CHAPTER 1

The Metatheory of the Project

First and foremost, this work is a bit of history and philosophy of science. It uses historical context and philosophical concepts, such as constitution, implementation, and abduction, to develop a theory of the scientific interpretation of experimental results in the recent primary experimental literature. It attempts to make explicit what is often merely implicit in published scientific research.

On the surface, this is a thoroughly pedestrian project, but its methodological starting point leads my theory to depart from many common expectations. There are so many departures, in fact, that the methodological stage setting in this chapter makes for one of the longest chapters of the book.

As described in Section 1.1, the theory introduces philosophical vocabulary and concepts, unconstrained by vocabulary and concepts familiar to scientists. One point of setting aside such a constraint is that the goal is not merely to report what scientists claim to be doing or think they are doing in one particular scientific domain. To get beyond describing scientific practice and on to a theory of the practices, it is often helpful to have new concepts and terminology.

Section 1.2 describes the exclusively descriptive character of the theory. Many projects in the philosophy of science propose to draw metaphysical and epistemological conclusions from scientific practice. Stated simplistically, a metaphysical argument might observe that scientists postulate X's, and from there infer that there are X's. Stated simplistically, an epistemological argument might observe that scientists use method Y to draw conclusions about the world, and from there conclude that method Y is justified and justificatory. Philosophers, of course, hedge these inferences, but I propose not to take either the ontological or the epistemological steps, even with hedges added. A principal part of the rationale for this is that the descriptive work is challenging enough.

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

Section 1.3 describes the theory's focus on the scientific interpretation of experimental results presented in the recent experimental literature. This is important, since a theory that focuses on what is happening at the moment of scientific insight, when a scientist first conceives an answer to a scientific question, would likely be quite different. A philosopher's theory of the scientific genesis of "happy thoughts" should probably draw on results from cognitive science. By contrast, the theory developed here will rely upon a close, granular reading of arguments in the recent scientific experimental literature. This work also backgrounds what is compiled into review articles or textbooks. It makes sense to bracket the reasoning in review articles since these often include reasoning that is not limited to that about experimental results. Further, textbooks, for their part, focus on communicating scientific conclusions rather than supporting scientific conclusions.

Section 1.4 describes the explanatory and abductive pluralism adopted here. Scientists provide many types of explanations leading them, in principle, to provide many types of abductive inferences. This section foreshadows the plurality of explanations and abductions discussed more concretely in Chapters 2–6.

Section 1.5 attempts to address some of the challenges facing the philosophical use of case studies, such as concerns about "cherry-picking" and overgeneralization. These are serious concerns that cannot be met simply by beginning with actual cases. To address the issue of generalization, the proposal developed here is that to generalize reasonably from a handful of case studies one should identify specific features of the historical cases that can guide generalizations.

1.1 Invoking Philosophical Concepts

By invoking "philosophical concepts" and "philosophical terminology," I reject a common assumption about how historians and philosophers of science should proceed, namely, that they should limit themselves to the concepts or vocabulary that scientists use. For my part, this is overly restrictive. Although I am sensitive to scientific terminology and will often draw attention to it, there are times when a departure from that terminology is in order. My goal is to develop a theory that accurately describes some of the scientific abductive reasoning in the experimental literature. Sometimes new terminology helps.

Sometimes, scientific terminology is well-suited to describing actual scientific reasoning. So, for example, scientists often use the terms "data"

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

and "results" to describe at least a subset of the products of their experimental work. Further, in the case studies I examine, scientists use "account for" and "explain" interchangeably. I generally use "explain," because it is more commonly used by philosophers. So, in some instances, I largely embrace scientific terminology – fleshing out some of the scientific presuppositions – because the scientific terminology does the theoretical work I want to do.

As a point of clarification, I try to thread a needle in using the words "explain" and "explanation." On the one hand, I do *not* intend for the words "explain" and "explanation" to capture an English language, or even scientific language, use of the words "explain" and "explanation."[1] Thus, I do not mean to do lexicography. Nor, on the other hand, do I propose merely to stipulate what I mean by "explain" or "explanation." To repeat, my goal is to develop a theory of some of the abductive reasoning scientists use in connecting experimental results to compositional hypotheses regarding the world. This requires a theory of what I will call "explanation." My theory embraces certain ideas that philosophers have recently had regarding explanation, so it is not entirely detached from what has been popularly, scientifically, or philosophically meant by "explanation." Then again, it may have some features that are not found in some theories of explanation in the philosophical literature.

Although there are times when scientific terminology is well-suited to my project, there are also, unfortunately, times when it is not. This happens for very many reasons. Sometimes scientists do not adopt a consistent vocabulary, even within the space of a single paragraph or two. Günter Baumgartner wrote of *explaining* perception when viewing white bars on a black background but *understanding* perception when viewing black bars on a white background.[2] One philosophical interpretation of this is that Baumgartner thinks there is some important difference between the two cases, whereas the interpretation I adopt is that the terminological shift is merely stylistic. Semmelweis wrote of atmospheric-cosmic-terrestrial influences/conditions causing puerperal fever, operating on individuals, and occasioning an epidemic.[3] One might assume that Semmelweis attaches some significance to the difference

[1] For a sample of philosophers who are concerned with ordinary language use of "explain," and its cognates, see Jenkins (2008), Craver (2014), Şerban (2017), Wright and Van Eck (2018), and Craver and Kaplan (2020).

[2] Baumgartner (1960, p. 22). [3] Semmelweis (1983, pp. 65–66).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

between causing, operating, and occasioning. Alternatively, one might think, as I do, that the differences are merely stylistic.

Sometimes scientists have not developed a fixed, standard terminology that is useful for philosophical purposes. Scientists never use "explanans" or "explanandum" in the cases I examine, but those terms are extremely useful in developing a philosophical theory of explanation. In the cases I study, scientists never use "constitution," "realization," "implementation," and most notably "abduction," but I think that those terms/concepts are essential for understanding certain features of scientific reasoning. Just as one can speak in prose, reason using *modus ponens*, or commit a microaggression without having articulated the concepts of prose, *modus ponens*, or a microaggression, so scientists can invoke implementation or reason abductively without having those concepts.

There are also more subtle reasons for departing from extant scientific terminology. Sometimes there is a scientific term, but that term does not capture certain features that I take to be important. Here, I have in mind scientists writing about the interpretation of results. In general, "interpretation" is a term of ordinary and scientific language that can sometimes include the singular compositional explanation of results. As I read science, scientists sometimes interpret results using statistical methods, but sometimes they interpret them using singular compositional abduction. Thus, the interpretation of results covers more cases than does the singular compositional explanation of results. So, while I take it to be fine to discuss the interpretation of results, I also need more specific terminology that captures the additional features of the cases that I have in mind.

Another reason that some scientific terminology is sometimes not suited to my purposes is that it is used only within one area of research. I want to develop a theory that covers multiple distinct areas of research. Hodgkin and Huxley proposed that sodium ions carried the initial inward current of the action potential. This term "carried" describes a relation between an activity instance of an axon and the collective activity instances of multiple sodium ions. While there continue to be uses of this term for this purpose, it has not been exported to other areas of scientific research.[4] When Günter Baumgartner proposed that retinal ganglion cells (RGCs) provide part of the biological basis of the Hermann grid illusion, he did not say that retinal ganglion cells carry the Hermann grid illusion. Insofar as one wants a single term to describe both the Hodgkin–Huxley case and the Baumgartner case, one will have to improvise. My improvisation is to write

[4] Hammond (2014, p. 35) mentioned ions carrying currents.

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

of implementation. This term covers a relation I take to be found in multiple areas of scientific investigation.

Perhaps the most problematic cases where scientific terminology does not suit my purposes is when scientists in my case studies use "cause" or "effect," where I would describe what they are doing in terms of composers and composed. So, for example, Hodgkin and Huxley claim that "the rising phase of the action potential is caused by the entry of these ions" (Hodgkin & Huxley, 1952a, p. 449). A common presupposition is that if scientists call something a "cause" or an "effect," then these are the terms that philosophers should use. This is a defensible view, so for that reason I spend a fair amount of time supporting an alternative. Chapter 2 has a substantial discussion that motivates drawing a distinction between causal relations and compositional relations. Here, I suppose that this is simply another case where scientists do not make a distinction that is important for accurately describing what they are doing.

For what it is worth, at least some New Mechanists also implicitly reject the "stick to the scientific language" restriction. Many New Mechanists focus on something like characterizing the extension of the scientific use of the term "mechanism," so in this case try to abide by scientific terminology. Nevertheless, some New Mechanists rely on concepts and terminology, such as mechanistic constitution, constitutive relevance, mutual manipulability, and constraining a hypothesis space, that are not native to scientific discussion.[5] These are not naturally occurring concepts/terminology in scientific discussion. This is not to say that there are no scientists who have used these philosophical concepts or terminology. Some scientists have read the New Mechanist literature and embraced parts of it.[6]

Perhaps what motivates the restriction of philosophical concepts and terminology to scientific concepts and terminology is the sense that going beyond the scientific concepts and terminology amounts to an unwarranted philosophical interpretation. Here, my counterproposal is that there are interpretations that are warranted and interpretations that are not warranted. Indeed, at many points in my discussion, I will present objections to other philosophical interpretations of scientific reasoning. My objections will not, however, be that these interpretations are interpretations. Instead, the objections will be that the rejected interpretations somehow do not correctly describe the cases they may be thought to describe. The entire crux of determining which interpretations are

[5] Craver (2007), Craver and Darden (2013). [6] Japyassú and Laland (2017) is an instance.

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

warranted and which are not is an *a posteriori* issue that has to be settled on a case-by-case basis.

My method is not merely to report what scientists write in the experimental literature; I offer a theory of why they write what they do. They sometimes write as they do because they are making use of singular compositional abduction. Here, an analogy might be helpful. Sociologists and social psychologists sometimes distinguish between ingroups and outgroups. An ingroup is roughly a group with which an individual psychologically identifies, such as a race, religion, or culture. In contrast, an outgroup is a group one does not psychologically identify with. Members of the groups may not be aware of the ingroup–outgroup distinction. Still, sociologists might use the distinction to explain the observation that a comment in the mouth of a member of one's ingroup provokes a different reaction in the mouth of a nonmember. Other sociological concepts, such as "implicit bias" and "racial resentment," might not be familiar in a given human population but might nevertheless be scientific concepts that provide illuminating explanations of various sociological phenomena in that population.

Like the sociologist or social psychologist, a historian and philosopher of science might develop concepts such as realization, implementation, constitution, singular compositional explanation, singular compositional abduction, etc., to illuminate the work of scientists. Moreover, just as individuals in a society may not have articulated concepts of ingroup and outgroup, so the community of scientists may not have articulated concepts of realization, implementation, and so forth. The goal of this work is in this spirit. It attempts to render explicit the commitments often implicit in actual scientific theorizing and argumentation.

The analogy here is perhaps not perfect. Scientists conduct research, and are trained to conduct research, on entities in the natural world, such as action potentials. In addition, in their research and training, they develop a sense of what they are trying to do and how they are trying to do it. Roughly speaking, they have views about the natural world as well as views about science, its methods, and its products. They have views about science and, perhaps, a philosophy of science. Scientists may, therefore, be more self-conscious about what they do in their scientific work than average individuals are about what they do in their social interactions.

A potentially more worrisome observation is that scientists might have a philosophy of science that conflicts with what a historian and philosopher of science might have to offer. Anecdotally speaking, Popperian falsificationism

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

is quite popular among scientists.[7] For another example, Hodgkin and Huxley once claimed that "the rising phase of the action potential is caused by the entry of these [sodium] ions" (Hodgkin & Huxley, 1952a, p. 449). However, on the theory proposed here, the relation between the rising phase of the action potential and the entry of sodium ions is not causation but implementation. The justification for my proposal is that there are features of the implementation relation that differ from features of the causation relation. Most notably, the rising phase of the action potential is contemporaneous with the influx of sodium ions into the axon.

In the face of such conflicts, I do not assume that the historian and philosopher of science must be mistaken. The scientist could be mistaken. I do not suppose that this is because philosophers have some special insight into the nature of reality. Nor do I suppose that philosophy or metaphysics somehow "sets the limits of scientific theorizing."[8] Instead, the philosopher of science could get it right about how science works simply by devoting more effort to figuring it out. Where scientists might focus their attention on the rising phase of the action potential, historians and philosophers of science might devote their attention to what is implicit in the relation between ion fluxes and action potentials. Such detailed consideration has led many philosophers to distinguish causal relations from compositional relations, even though scientists may sometimes say that lower level activities cause higher level activities. In the end, the best that one can hope for, perhaps, is some measure of agreement between what scientists say about science, its methods, and its products, and what historians and philosophers of science have to say. The possibility of such agreement, at least, is presupposed in this book.

1.2 An Exclusively Descriptive Account of Scientific Reasoning

My project is to provide a descriptively adequate account of published scientific reasoning. In selecting this project, I intentionally set aside metaphysics and epistemology. I will not make claims about the fundamental nature of reality or what scientists may, or may not, justifiably believe true. Instead, I offer an account of what I take to be some scientific commitments and practices. The simple and primary reason for this is that

[7] Raman and Ferster (2022, p. 69) interpret Hodgkin and Huxley (1952a) as engaging in falsificationism at one point.
[8] Heil (2012, p. 11).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

the descriptive project is challenging enough without trying to support metaphysical and epistemological conclusions.

1.2.1 Some Metaphysical Presuppositions of Science

My account assumes that scientists are ontologically committed to, among other things, the existence of individuals, property instances, activity instances, and ontological dependence relations among them. The account proposes that certain ontological, hence metaphysical, commitments or presuppositions are found in science. The account does not, however, offer ontological or metaphysical commitments *simpliciter*. For, it is one thing to say that a scientist thinks that X exists and quite another to say that X exists. Metaphysics, as I understand it, concerns the latter existence claims.

One can, of course, forge a link between scientific metaphysical commitments and metaphysics. Tim Maudlin offers the following argument:

> Metaphysics is ontology. Ontology is the most generic study of what exists. Evidence for what exists, at least in the physical world, is provided solely by empirical research. Hence the proper object of most metaphysics is the careful analysis of our best scientific theories (and especially of fundamental physical theories) with the goal of determining what they imply about the constitution of the physical world (Maudlin, 2007, p. 104).

I happen to share Maudlin's view that our best-supported metaphysics is the metaphysics in science, so that philosophers should accept the ontology found in science. This, however, is not a claim I support in this book. The careful analysis of scientific theories Maudlin alludes to is challenging enough, so I limit myself to that.

As a note of clarification, I might add that Stavros Ioannidis and Stathos Psillos agree with the idea that one can describe scientific metaphysical commitments without endorsing them as *bona fide* metaphysical commitments.[9] They may be alone among New Mechanists in adopting this view. Nevertheless, they go too far when they deny that there are arguments such as Maudlin's: "Concerning the question of metaphysics, one can, surely, give reasons to prefer some particular metaphysical picture over others. Our point is that whatever else these reasons are, they are not related to scientific practice" (Ioannidis & Psillos, 2022, p. 6). Whatever objections one might have to Maudlin's argument, one should not include the claim that it is unrelated to scientific practice. At the very least, I part ways with Ioannidis and Psillos on this.

[9] Ioannidis and Psillos (2022, p. 6).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

Although many philosophers are inclined to dismiss an exclusively descriptive metaphysical project, this inclination underestimates its difficulty. An accurate description of scientific ontological commitments is not a trivial matter. Let me provide an illustrative example concerning the scientific commitment to property instances, something about which New Mechanists have had relatively little to say. My official ontology for singular compositional explanation involves property instances, rather than properties. I will justify this in Chapter 3, but for the next few paragraphs I will run the discussion in terms of properties. This will simplify matters. Moreover, the point I make regarding properties will carry over to property instances.

Sometimes the "New Mechanism," as opposed to one or another form of "old mechanism," is traced to Stuart Kauffman's idea that "Typical explanations in biology exhibit the manner in which parts and processes articulate together to cause the system to do some particular thing" (Kauffman, 1970, p. 257).[10] New Mechanists have, accordingly, for decades, largely limited themselves to two ontological categories corresponding to Kauffman's "parts" and "processes." Properties, however, are mentioned only on the periphery.[11]

The discussion in Machamer et al. (2000) is illustrative. They famously claim that "Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions" (Machamer et al., 2000, p. 3). There is no mention of properties here. Quickly thereafter, however, they propose that "Mechanisms are composed of both entities (with their properties) and activities. Activities are the producers of change. Entities are the things that engage in activities. Activities usually require that entities have specific types of properties" (Machamer et al., 2000). Glennan and Illari (2017) have more recently illustrated this same pattern. They propose that "A mechanism for a phenomenon consists of entities (or parts) whose activities and interactions are organized so as to be responsible for the phenomenon" (Glennan & Illari, 2017, p. 1), but also quickly thereafter note that "Metaphysical mechanists ... likely see mechanisms as central to understanding some traditional metaphysical questions, e.g., about the nature of

[10] For a brief history, see Wimsatt (2018).

[11] One notable exception is Piccinini (2020), which often refers to properties. Piccinini's approach, however, is apparently driven by an attempt to integrate the philosophical literature on realization (which often focuses on properties) and the literature on New Mechanism (which largely ignores properties). Piccinini's approach is not driven by fidelity to scientific practice. He does not argue that a descriptively adequate account of scientific practice should recognize a role for properties.

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

objects and properties, parts and wholes, causal relations and laws of nature" (Glennan & Illari, 2017, p. 2).

What scientific practice, one might ask, indicates a role for properties? Craver and Darden (2013) rightly mention that properties enable activities. So, they observe that "[e]ntities have properties, such as structure and orientation, that enable them to engage in specific activities" (Craver & Darden, 2013, p. 16).[12] Suppose that properties do enable activities. *Prima facie* there should be a vast number of property-invoking explanations of why certain entities engage in certain activities. There should be a vast number of "enabling explanations." So, one might ask, "Why does water expand upon freezing?" The answer has to do with the polarity of water molecules. Why does the movement of sodium ions into the axon generate a current? The answer has to do with sodium ions having positive charge. There is, however, more to be said than Craver and Darden mention. Given that water molecules are polar, one should expect there to be a scientific explanation of their polarity. Why, one might ask, are water molecules polar? Although New Mechanists have rarely considered such cases, there is a compositional explanation of this property of water molecules in terms of the properties of its parts. The explanation of the polarity of a water molecule is based on the relative electron affinities of hydrogen and oxygen. An oxygen atom is more electronegative than a hydrogen atom – electrons tend to cluster on the oxygen side of the molecule – thereby making the oxygen side of the water molecule more negative. This line of thought suggests that (a) for every explanation of an activity of a composite individual, there will be a property that enables that activity and (b) that there will be a scientific compositional explanation of the activity-enabling property of the composite individual. Maybe this is mistaken, but it is a line of thought that merits some philosophical consideration. I will return to enabling in Chapter 3.

To summarize this subsection, this book offers a purely descriptive theory of what scientists take to be involved in the singular compositional abductive interpretation of scientific results in the experimental literature. This departs from a common move in the philosophy of science to infer from scientific commitments regarding the way the world is to metaphysical conclusions regarding the way the world is. In this book, I do not take this further step, since I think the accurate description of actual scientific commitments is a worthy project in its own right. As an illustration of this challenge, I have drawn attention to the decades-long New Mechanist

[12] See, also, Craver and Darden (2013, pp. 17–18, 78–79, 152).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

practice of relative inattention to properties. On my theory, individuals, property instances, and activity instances are "equal ontological players" in scientific reasoning. They are ontologically intertwined in scientific thinking. New Mechanists have not accurately described this "ontological egalitarianism."

1.2.2 *Some Methodological Presuppositions of Science*

As my focus is on describing scientific compositional abduction as found in the primary experimental literature, I will set aside any number of issues that have occupied epistemologists and philosophers of science for decades. Thus, I do not consider whether abduction or IBE is a "fundamental rule of inference."[13] I will not be concerned with whether warranted enumerative induction is a special case of warranted abduction.[14] I do not consider whether any specific abductive inference is justified or whether abduction, in general, is justified.[15] One topic I will completely avoid is Bayesian confirmation. Although I find Bayesianism implausible as an account of the scientific reasoning I examine, I find that the pros and cons of Bayesianism are so well-worked that I do not see that I have anything to add.[16]

Focusing exclusively on the descriptive project sets this book apart from many other works in the philosophy of science that propose to use actual scientific reasoning as a guide to how scientific justification proceeds. New Mechanists often combine the study of the descriptive and the normative, but they are not alone in this. Lindley Darden has proposed that the history of science might serve as a source of "compiled hindsight."[17] William Bechtel and Robert Richardson propose to look at actual scientific reasoning, seeing how it succeeds and fails, for a guide to how scientists should reason.[18] Craver provides a lengthy defense of providing both a descriptive and normative account of mechanistic explanations in neuroscience.[19] Among logicians of abduction, Dov Gabbay and John Woods have proposed an "*Actually happens rule*. To see what agents should do, look first to what they actually do. Then repair the account if there is

[13] Cf., for example, Harman (1965, 1968), Ennis (1968), Fumerton (1980, 2017), Lipton (2003), Mackonis (2013), Khalifa et al. (2017), and Williamson (2016).
[14] Cf., for example, Harman (1965, 1968), Ennis (1968), and Fumerton (1980).
[15] For a small sample of this literature, see Van Fraassen (1980), Psillos (1996), Ladyman et al. (1997), Okasha (2000), Weber (2004), Carter and Pritchard (2017), and Douven (2022).
[16] For a very tiny sample, see Glymour (1980a), Earman (1992), Howson and Urbach (2006), and Norton (2021).
[17] Darden (1987, p. 39), Cf., Craver (2007, p. x) and Craver and Darden (2013, p. 10).
[18] Bechtel and Richardson (2010, pp. 10–11). [19] Craver (2007, pp. vii–xi).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

reason to do so" (Gabbay & Woods, 2005, p. 5). John Norton also looks closely at actual science and uses that as a basis for conclusions about warrant or justification.[20]

One reason not to pursue the descriptive and normative projects simultaneously is that it is a lot of work to put into practice. Although Gabbay and Woods advocate the *actually happens rule*, they do not discuss any empirical evidence about how agents actually reason. To flesh this out a bit with an example, Gabbay and Woods follow Peirce in assuming that abduction is "ignorance preserving." They do not, however, cite any empirical work in support of the view that scientists, or any others, treat abduction as ignorance preserving. I have proposed that Hodgkin and Huxley reasoned abductively about the sodium hypothesis. After they performed all those experiments, did they not know that sodium carries the initial inward current of the action potential? There are times when they write about what they have supported or established regarding how the action potential works.[21] Nevertheless, this is not a question Gabbay and Woods try to address.

A second reason not to pursue the descriptive and normative projects simultaneously is that matters quickly become complicated. Peirce proposes that abduction provides no confirmation. Suppose, as will be argued in Chapter 5, that Hodgkin and Huxley understood abductive reasoning to confirm the hypothesis that the initial inward current was generated by sodium. In other words, suppose that Hodgkin and Huxley assumed that abduction provides confirmation. Then there is a *prima facie* conflict. Peirce apparently had one assessment of abductive reasoning; Hodgkin and Huxley apparently had another. What is to be made of it? There are many possibilities. Perhaps Peirce made a descriptive mistake about how scientists reason. Or, maybe he did not mean to describe how scientists reason, but how scientists should reason. Maybe then Hodgkin and Huxley were drawing an unwarranted conclusion. But, if one concludes that Hodgkin and Huxley were in some sense wrong to draw this conclusion, then why were they wrong? What basis is there for saying that Hodgkin and Huxley were making a mistake? And, what if one finds that many scientists break with a normative interpretation of Peirce's proposal, as by awarding Hodgkin and Huxley the Nobel Prize for their work? Does this mean that there is a lot more scientific "malpractice" than one might have anticipated?

[20] Norton (2021).
[21] See, for example, Hodgkin and Huxley (1952a, p. 471) and Hodgkin and Huxley (1952d, p. 541).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

The point here is not that these questions cannot be resolved. It is that they involve complexities that I propose to bracket in this book. Indeed, they are questions that have typically been set aside in the literature. There are numerous discussions of Peirce's idea of abduction as merely generating hypotheses, but many do not consider what to say should descriptive-normative conflicts arise.[22] And Peirce is not the only one to have advanced apparently descriptive and/or normative claims. Indeed, part of what enables many philosophers of science to think they can simultaneously address both descriptive and normative issues is that they do not examine cases in which proposed normative claims appear to depart from scientific practices. Matters are a lot simpler if one does not entertain descriptive-normative differences.

A third reason to begin with an exclusively descriptive account of scientific reasoning is that some philosophers working on abduction are inclined to begin with the normative. Many philosophers begin with Harman's picture of warranted abduction and picking a best explanation on the basis of simplicity, plausibility, etc. Inspired by this picture, they might assume that, in using their first experimental results, Hodgkin and Huxley ruled out rival hypotheses by appealing to sundry explanatory virtues such as simplicity, plausibility, etc. If one reads through the text of Hodgkin and Huxley's paper discussing the sodium hypothesis experiments, however, one sees that Hodgkin and Huxley consider no rivals and there is no explicit mention of simplicity or plausibility.[23] One might then propose that Hodgkin and Huxley are *implicitly* considering rival hypotheses and *implicitly* ruling them out by appeal to simplicity and plausibility. But what is the evidence that Hodgkin and Huxley implicitly rule out rivals? Is this just an *ad hoc* presupposition one makes in order to apply to Harman picture?

Why should a historian and philosopher of science take this path? Why begin with Harman's picture of warranted abduction – a picture developed for a project in epistemology – as a starting point for a project in the history and philosophy of science? Why begin with such a preconception? A casual familiarity with scientific research suggests that, at any given time, scientists believe that some rival hypotheses have been eliminated, but that

[22] See, for example, Hoffmann (1999), Niiniluoto (1999), Magnani (2001), Gabbay and Woods (2005), Schurz (2008), Campos (2011), Plutynski (2011), Mackonis (2013), McAuliffe (2015), Misak (2017), and Mohammadian (2021). None of these accounts explain why Peirce thought that abduction merely introduces hypotheses (worthy of pursuit). Nor do they consider whether Peirce was, in fact, correct in thinking this.

[23] More specifically, see the discussion at Hodgkin and Huxley (1952a, pp. 450–452).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

others have not. So, there is some sense in which some rival hypotheses have implicitly been eliminated. Then again, a large part of actual science involves entertaining rival hypotheses and designing and performing experiments that will rule them out. In an important sense, many scientific inferences, such as those by Hodgkin and Huxley, may not be warranted to Harman's standards at the time they are made. Some rival hypotheses may have yet to be addressed. Instead, such warrant as scientists may have for some hypothesis may emerge, if it ever does, through a sustained diachronic process of scientific research. Indeed, Huxley himself had something like this thought. In autobiographical reflections some forty plus years after the publication of the Hodgkin–Huxley model, he wrote,

> Most of my life has been spent in posts where I was teaching medical students, and I am very glad to have had even that single year of clinical study since it enabled me to see how different the attitude of the clinician has to be from that of a scientist: the clinician has to make immediate decisions, often on slender evidence, while the scientist has (almost) unlimited time. (Huxley, 2004, p. 293)

While there is plenty of room to interpret what Huxley meant here, there is enough to suggest that historians and philosophers of science take seriously the idea that scientific justification often arises, when it does, not merely from the results of a single experiment.[24] Thus, rather than beginning with the Harman picture, I propose to look at actual published science without beginning with Harman's preconception. This is part and parcel of my science-first approach.

1.3 The Recent Primary Scientific Experimental Literature

Whereas this book is focused, first and foremost, on a descriptively adequate theory, it does not aspire to a descriptively adequate account of the whole of science. Nor does it aspire to a theory of the whole of scientific reasoning or even scientific inductive reasoning. To a first approximation, it focuses on a descriptively adequate account of how scientists sometimes justify compositional hypotheses in the recent primary scientific experimental literature. The qualifier "first approximation" is meant to cover a number of issues. For one thing, I do not have a precise

[24] Huxley's comment also invites a serious examination of whether abductive inference in medicine might be different than abductive inference in science. Maybe on limited evidence, the clinician must typically turn to simplicity, plausibility, etc., whereas the scientist typically has time to do another experiment.

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

characterization of "recent." Post-1930 is in the ballpark, but there is no specific event that suggests 1930, rather than 1925 or 1935 or 1945. Moreover, there are times when nineteenth-century science provides some helpful illustrations and helpful background. Finally, there are times when it is helpful to stretch the boundaries of "experimental literature" a bit.[25]

This book draws attention to abduction as scientists use it in the scientific literature. My engagement with science is, thus, what N. R. Hanson once derisively referred to as a "Logic of the Finished Research Report" (Hanson, 1960, p. 92). Of course, the engagement here is not so much with what I would call a "logic" *per se* but, more modestly, with a characterization of the justifications – the reasons – scientists sometimes provide for their hypotheses regarding what is compositionally related to what in the primary experimental literature. Adopting this focus, I, therefore, set aside abductive inference as it may be used in other domains, such as computer science, legal reasoning, and philosophical arguments for scientific realism.[26]

In focusing on finished research reports, this book focuses on a fragment of the reasoning found in "unsettled, dynamic, research sciences" (Hanson, 1958, p. 1). It focuses on some of the initial interpretations of experimental results for publication in the experimental literature. It, therefore, sets aside what is sometimes described as the "logic of discovery," understood as the psychology of individual scientists at the moment of scientific insight.[27] The present work is not concerned with the likes of what was passing through Archimedes's mind as he sat in his bathtub nor through Kekulé's mind as he sat by his fire. Nor is it concerned with the thinking that precedes those insightful moments. This book also sets aside the "packaging" of scientific theory and reasoning into review articles and textbooks. Nor is it focused on a somewhat more abstract concept of theory confirmation. The rationale for this focus merits articulation.

1.3.1 The Moment of Insight

Many philosophers will suppose that a philosophical theory of abduction will focus on scientists' initial conception of hypotheses. Peirce is widely credited with having distinguished abductive reasoning from inductive and

[25] Here I have in mind Hermann (1870), Baumgartner (1960), and Schiller and Carvey (2005).
[26] Magnani (2001), Psillos (2005).
[27] New Mechanists sometimes use the term "discovery" to cover both the initial advancement of a hypothesis and its subsequent confirmation. See, for example, Craver and Darden (2013). For discussion of this usage, see Schickore (2018).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

deductive reasoning. He believed that abduction was the only means by which new ideas enter science. Once abduction has introduced a new hypothesis, deductive and inductive reasoning take over to confirm that hypothesis.[28] Hanson embraced this role for abduction, and Darden apparently carries it over into her brand of New Mechanism.[29]

Although linking abduction to the "moment of insight" will be natural to some, that is not a linkage I pursue. I take seriously the possibility that the process, or processes, underlying the initial psychological formulation of hypotheses might be distinctly different from the reasoning that goes into the interpretation of the results in the experimental literature. Since New Mechanists first proposed to study scientific discovery some decades ago, psychologists have developed an extensive body of research on human reasoning. As the result of experimental work on, for example, statistical and syllogistic reasoning, psychologists have offered multiple versions of what is often called a System 1/System 2 distinction. In one popularization of the idea, Daniel Kahneman briefly describes these:

> System 1 operates automatically and quickly, with little or no effort and no sense of voluntary control.
> System 2 allocates attention to the effortful mental activities that demand it, including complex computations. The operations of System 2 are often associated with the subjective experience of agency, choice, and concentration. (Kahneman, 2011, pp. 20–21)

System 1 operates quickly, as it apparently acts on various heuristic assumptions – biases – that violate principles of valid statistical and logical reasoning. One should, therefore, take seriously the possibility that the reasoning involved in the initial formulation of a scientific hypothesis might be performed by System 1, whereas the reasoning involved in the interpretation of the results in the experimental literature might be performed by System 2.

I do not here mean to explain and defend any version of this analysis; that is why I propose instead merely that the possibility needs to be taken seriously. Taking the proposal seriously would mean reviewing a large psychological and philosophical literature on human reasoning and explaining and defending one or a family of analyses. While undoubtedly an important philosophical project, it is not this project. It is a project that should rely quite heavily on psychology but less heavily on the close reading of scientific texts. It makes sense to separate these two projects.

[28] Peirce (1992). [29] Hanson (1960), Craver and Darden (2013, p. 81), Darden (2017, p. 255).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

Here is another way of emphasizing the importance of the distinction between laboratory life and journal articles. In the IBE tradition, it is common to think of IBE in so general a fashion as to cover both what happens in the spontaneous scientific inferences in the laboratory and what is published more formally in scientific journal articles. Harman and Lipton apparently do this. In the history and philosophy of science, however, it is commonly held that what appears in scientific journal articles is not a faithful representation of what takes place in the lab. In a famously provocative paper, "Is the scientific paper a fraud?," Peter Medawar argued that scientific papers are not faithful representations of laboratory experimentation.[30] Hodgkin had his own version of this idea, touching on themes of chance and design in science:

> I believe that the record of published papers conveys an impression of directness and planning which does not at all coincide with the actual sequence of events. The stated object of a piece of research often agrees more closely with the reason for continuing or finishing the work than it does with the idea which led to the original experiments. In writing papers, authors are encouraged to be logical, and, even if they wished to admit that some experiment which turned out in a logical way was done for a perfectly dotty reason, they would not be encouraged to 'clutter-up' the literature with irrelevant personal reminiscences. (Hodgkin, 1976, p. 1)

I think there are likely to be differences between reasoning in the lab and reasoning in publications. We should not assume that they are the same. It might turn out that some general theory of abduction/IBE applies to both laboratory reasoning and experimental reasoning, but, for the space of this book, I will remain noncommittal about laboratory reasoning and focus exclusively on the experimental literature. In adopting this focus, I do not mean to weigh in on a debate about the relative importance of lab life versus publication. As I see it, laboratory experimentation and publication both have essential roles in science. It is, therefore, perfectly reasonable to focus on an account of the reasoning scientists offer in the experimental literature without trying to assess its importance relative to laboratory work or laboratory reasoning.

1.3.2 *Scientific Reviews*

What about later scientific work in reviews? In review articles, many important details of scientific reasoning are omitted for the sake of

[30] Medawar (1963).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

characterizing the experimental and theoretical lay of the land regarding a particular topic at a particular time. To give an example, review articles will sometimes not mention experimental controls that are commonly employed in the primary experimental literature. In the many studies of rat spatial navigation through mazes – to provide a concrete example – steps are taken to ensure that the rats cannot rely upon olfactory cues from the surfaces of the maze to find their rewards. Such steps might involve wiping down the surfaces of the maze. As this is common knowledge and practice in the area, review articles will often not mention them. Nevertheless, the fact that scientists regularly use experimental controls is important for the abductive approach, since one rationale for an experimental control is that it enables one to rule out alternative explanations of experimental results. With appropriate controls in place, one cannot explain a rat's navigation abilities by appealing to olfactory cues, since olfactory cues would not have guided the rats to their rewards.

1.3.3 Textbook Science

Textbooks also often omit much of the reasoning found in the primary experimental literature, as their burden is not so much to convince students of the truth of scientific hypotheses as it is to transmit to them what the hypotheses are. To return to our example of Hodgkin and Huxley's 1952 experimental work, Hodgkin and Huxley performed three experiments that involved varying the concentration of extracellular sodium and the potential at which the axon was clamped. Textbook accounts of their work may not describe any of these experiments or the reasoning behind them on the assumption that this level of historical detail is unlikely to be needed by neuroscience students.[31]

As a second illustration, consider the textbook discussions of the action potential in Shepherd (1988) and Kandel et al. (2000), alongside some of the reasoning offered in support of the theory circa 1950. Around 1950, it was known that sodium ion concentrations were much higher outside of the cell than inside and that potassium ion concentrations were much higher inside of the cell than outside. Hodgkin and Huxley, among others, proposed that features of the action potential could be explained by postulating fluxes of these ions. As will be spelled out in more detail in

[31] Cf., for example, Shepherd (1988, chapter 6), Kandel et al. (2000, p. 151), and Hammond (2014, chapter 4). In recent years, however, neuroscientists have taken greater interest in these papers. See Raman and Ferster (2022) and Brown (2020).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

Chapter 5, abductive methods were at work here. Hodgkin and Huxley's so-called indirect electrophysiological methods were complemented by "direct" methods, reported in Keynes (1948, 1949), Rothenberg (1950), and Keynes and Lewis (1951a, 1951b). In these latter papers, scientists argued that radioactively labeled sodium and potassium crossed the membrane during action potentials. For sodium, the experiments stimulated neurons in an external medium containing radioactively labeled sodium ions. When the cell was subsequently transferred to a sodium-free medium, it was found that the cell contained radioactivity that it must have carried from the external sodium ions. For potassium, the cells were loaded with radioactive potassium by keeping them in an external medium rich in radioactive potassium. After transferring the cells to a non-radioactive medium, then stimulating them, it was found that radioactive potassium was released into the medium. None of this experimental detail is reported in either Shepherd (1988) or Kandel et al. (2000). Hammond (2014) mentions the radioactive tracer work, but cites Hodgkin and Keynes (1955). Without examining the original research, the philosopher of science might overlook important instances of scientific reasoning. Chapter 5 will review a number of "minor" scientific inferences that do not make their way into textbooks, but that illustrate compositional abductive reasoning in action.

1.3.4 Theory Confirmation

I propose singular compositional abduction as part of an account of the scientific interpretation of certain experimental results. This focus distinguishes my account from one way in which accounts of abduction – generally understood as some form of IBE – may be related to scientific practice. Sometimes IBE is treated as an account of hypothesis confirmation, but it is also sometimes treated as an account of theory confirmation. Paul Thagard's work provides an early illustration, although it is framed in terms of theory choice. Thagard observes that,

> The phrase "inference to the best explanation" is relatively new, but the idea is old. Inference to scientific hypotheses on the basis of what they explain was discussed by such nineteenth-century thinkers as William Whewell and C. S. Peirce, and earlier still by David Hartley, Leibniz, and Descartes. To put it briefly, inference to the best explanation consists in accepting a hypothesis on the grounds that it provides a better explanation of the evidence than is provided by alternative hypotheses. We argue for a hypothesis or theory by arguing that it is the best explanation of the evidence. (Thagard, 1978, p. 77, cf., p. 76)

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

The discussion, up until the final sentence, is in terms of hypotheses. Thagard's subsequent discussion, however, ranges over Charles Darwin's theory of evolution by natural selection, Antoine Lavosier's oxygen theory of combustion, and Christian Huygen's wave theory of light. Further, in working out his account of theory choice, Thagard proposes that scientific theory choice is a function of, among other things, a theory's explanation of "facts."[32] This shift between hypotheses explaining and theories explaining is common in the literature.[33]

Singular compositional abduction is more plausible as an account of hypothesis confirmation than it is as an account of theory confirmation. Without venturing a characterization of the difference between hypothesis confirmation and theory confirmation, I appeal to the idea that theory confirmation is a function of – in a deliberately broad construal of the phrase – hypothesis confirmation. Theory confirmation, thus, to speak metaphorically, must cover more ground. In covering more ground, one is likely to find a much broader range of considerations than one finds in covering less ground.

To make this point more concrete and less metaphorical, consider the Hodgkin–Huxley–Katz papers circa 1950. These papers brought many experimental results to bear in support of the Hodgkin–Huxley theory of the action potential. In these papers, they mention simplicity from time to time. So, for example, Hodgkin and Katz comment that "A simpler type of hypothesis has recently been worked out, in collaboration with Mr Huxley, and forms the theoretical background of this paper" (Hodgkin & Katz, 1949a, p. 37).[34] Maybe there is some theory according to which the simplicity of a hypothesis counts in favor of the simplicity of an explanation.[35] And maybe this could be incorporated into a version of IBE. Thus, a complete account of how Hodgkin, Huxley, and Katz made their case for their theory of the action potential would have to include an account of the role they thought simplicity to have played. In contrast, if one examines, say, Hodgkin and Huxley's interpretation of the first

[32] Thagard (1978, p. 79f). For further, discussion of Thagard's account of theory choice, see Chapter 8.

[33] Okasha (2002, 22f), McCain (2019, 55f), Lange (2022, p. 85). The shift is more diffuse in Lipton (2003) and Bird (2022).

[34] See, also, Hodgkin and Katz (1949a, pp. 53, 55, 66), Hodgkin and Huxley (1952d, p. 505). Hodgkin, Huxley, and Katz also use "simple" in other ways that appear to have nothing to do with explanation. So, for example, there is this: "There is no simple way of estimating the error introduced by this approximation" (Hodgkin & Huxley, 1952d, p. 505).

[35] Thagard's theory of simplicity, however, does not appear to be such a theory, as it is based on the size and "nature" of the set of auxiliary hypotheses used in an explanation (Thagard, 1978, p. 86).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

experiment of Hodgkin and Huxley (1952a), one will find no mention of simplicity. This provides *prima facie* reason to think that, while a singular compositional abductive account that does not involve simplicity will not have sufficient theoretical resources to account for the whole of the confirmation of the Hodgkin–Huxley theory of the action potential, it might be more at home in the narrow context of the interpretation of the results of Hodgkin and Huxley's first experiment.[36] These are themes to which I will return in Chapters 8 and 9.

1.3.5 *The Primary Experimental Literature*

In focusing on the argumentation in the primary experimental literature, I do not mean to deny an important role for moments of insight. What would science be without them? Nor do I deny the important role of review articles and textbooks. Each of these parts of the scientific enterprise has a role to play, but those roles are not the same. Putting matters quite roughly, publication of experimental results is the empirical input to the scientific community's theorizing; review articles attempt to package experimental and theoretical work into more intellectually manageable frameworks; textbooks are meant to communicate current scientific views to a new generation of scientists. Moreover, coming to understand these different roles may require different methods. As proposed below, a close reading of scientific texts – be they in the primary experimental literature, reviews, or textbooks – can plausibly give philosophers some insight into how scientists make their case for empirical hypotheses. However, such methods might not be adequate for understanding the psychology of scientific insight. As I have mentioned, to understand Archimedes's reasoning in the bathtub or Kekulé's reasoning before his fire, philosophers of science might be helped by cognitive psychology.

The fine granular attention to the primary scientific experimental literature, I believe, sets my work apart from many – but by no means all – studies by New Mechanists. Rather than surveying discovery or justification in whole scientific disciplines or fields, such as biochemistry, cell biology, genetics, neuroscience, or psychology, my case studies focus on the reasoning in the primary scientific literature wherein scientists provide arguments for some specific compositional relations.[37] The discussion of Hodgkin and Huxley's first argument in their 1952 paper regarding the

[36] NB: Saying that the singular compositional abductive account is "more at home" is intentionally vague. A more detailed account will be developed in Chapters 2–4.

[37] My granular approach, thus, differs from that found in Bechtel (2006, 2008) or Darden and Maull (1977). It is closer in approach to Glymour (1980a).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

initial inward current is a prime example. It would be impractical for a single human to review any significant portion of an entire scientific discipline or field at the level of detail at which I will be working.[38] Indeed, some scientific topics have so many experimental publications dedicated to them that reviewing even a small fraction of them at the level of detail I adopt would be impractical. Here, I have in mind a fixture of the New Mechanist literature, namely, long-term potentiation (LTP).[39] So, rather than review the evidence for all of neuroscience or all of neuronal function or even all of LTP, I focus on the physiological research on the action potential circa 1950. Rather than review the whole body of justifications for compositional hypotheses concerning vision, I focus on efforts to explain the Hermann grid illusion.

1.4 Explanatory and Abductive Pluralism

Universalism about X in the philosophy of science is the view that there is one universal, static scientific conception of X. Thus, universalism about explanation might hold that there is one universal static conception of explanation. Universalism about abduction might hold that there is one universal static conception of abduction. Pluralism, in contrast, holds that there are many distinct conceptions of X that may vary over time, place, and scientific discipline.

The view of explanation and abduction developed in this book is decidedly pluralist, for which there are precedents. Psillos, Gerhard Schurz, Igor Douven, and Jonah Schupbach have each proposed a disunity to abduction/IBE. Psillos argues that

> IBE should be considered as an inferential genus. The several species of the genus IBE are distinguished, among other things, by plugging assorted conceptions of explanation in the reasoning schema that constitutes the genus. For instance, if the relevant notion of explanation is causal, IBE becomes an inference to the best causal explanation. Or, if the relevant notion of explanation is subsumption under laws, IBE becomes a kind of inference to the best nomological explanation, and so forth. ... Many things can be explanations and the explanatory relation is not fixed and immutable. (Psillos, 2007, p. 444)

[38] See Norton (2021, pp. 256–267) for comments on this topic.

[39] Although it is inessential to my point to assign a precise number of publications, there is this observation: "By 1989, the U.S. National Library of Medicine listed some 312 articles with the term 'long-term potentiation' in the title, and, in the 1990s alone, over 1,000 additional articles have appeared" (Shors & Matzel, 1997, p. 597). Considering the additional articles that have since been published, it would be impractical to try to review all the articles in the whole of this literature.

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

Schurz proposes that there will not be "*one most general* schema" for abduction, as abductions will depend on details of a specific "abduction scenario" (Schurz, 2008, p. 205). In support of this account, Schurz provides a detailed taxonomy of multiple patterns of abduction. Douven proposes that the phrase "inference to the best explanation" is a mere slogan for similar reasons: "we do best to conceive of IBE as the aforementioned slogan, a slogan that not only permits different fleshings-out, but that may even rationally require different fleshings-out, depending on what our interests are" (Douven, 2017b, p. 23). Schupbach proposes that, "If one is a pluralist about the nature of explanation itself, then varieties of IBE may further multiply, with Inference to the Best Causal-Mechanical Explanation, for example, differing from Inference to the Best Covering Law Explanation, and so on" (Schupbach, 2017, p. 41, fn. 43).

Chapters 2 and 3 will articulate a small fraction of what I take to constitute the explanatory plurality Psillos, et al. refer to. It describes a plurality of singular compositional explanations. Of course, in point of logic, scientific explanatory pluralism does not entail scientific abductive pluralism. In point of logic, there can be scientific explanations that scientists do not use in abductive inferences. What combinations of explanation and abduction scientists use is, as is often said, an empirical matter.

This last point about scientific explanatory pluralism is meant, in the first instance, to be a descriptively accurate account of scientific practice. It does, however, have a secondary purpose. Norton (2021) objects that earlier conceptions of abduction/IBE overlook the diversity of philosophical accounts of explanation.

> Philosophical analysis of explanation has failed even to find a univocal sense of explanation at work in science. Instead, it has found multiple, competing senses of explanation. This multiplicity indicates that the notion is a loose one – an umbrella concept covering several disparate notions. They have no common core. (Norton, 2021, p. 248)

Norton is quite right to point this out, but, as Psillos implies, this is not a conclusion that an abductive approach needs to resist. As I shall argue more concretely in Chapters 2 and 3, there are at least three species of singular compositional explanation. In Chapter 4, and in the case studies, I will argue that there are at least three species of singular compositional abduction. In Chapter 6, I will broach the idea of "singular interlevel abductions," an idea that will be developed in greater detail in Chapter 10.

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

1.5 The Use of Case Studies

As the use of case studies is such a prominent part of this work, I have to address some of the perennial worries about their use. Joseph Pitt has given the issue a trenchant formulation in what he calls "the dilemma of case studies":

> What do appeals to case studies accomplish? Consider the dilemma: On the one hand, if the case is selected because it exemplifies the philosophical point, then it is not clear that the historical data hasn't been manipulated to fit the point. On the other hand, if one starts with a case study, it is not clear where to go from there – for it is unreasonable to generalize from one case or even two or three. (Pitt, 2001, p. 373)

My response to the dilemma is to tackle it head on. In truth, neither horn of the dilemma is insuperable. Regarding the first horn, it is an *a priori* logical possibility that a case study manipulates the historical data. (The use of "data" in this discussion should not be confused with the use of "data" articulated in Chapters 4–11.) Case studies should not, however, be dismissed simply based on this *a priori* possibility. Clarity regarding possible manipulation of a case must be achieved *a posteriori* by comparing the account with the historical record. Regarding the second horn, it is indeed unclear, *a priori*, how one might generalize from cases. Nevertheless, I provide guidance *a posteriori* on how to generalize. I do this by drawing attention to features of the cases that are likely to be applicable to other cases.

This last point merits elaboration. What ties together the cases of compositional abduction that I examine is the use of intralevel experiments. (I depart from this in Chapters 6 and 10 with the discussion of singular interlevel abduction.) In many cases, scientists manipulate some activity or property instance of an individual and then measure some response of that individual, but do not directly measure or manipulate the activity or property instances thought to underlie the connection between the manipulation and the measurement. Recall the situation with Hodgkin and Huxley's 1952 experiments regarding the sodium hypothesis. Hodgkin and Huxley voltage-clamped axons and measured the consequent total currents. They could perform this intralevel experiment. They did not, however, track the movements of individual ions. Indeed, they could not even track the movements of distinct ion species, such as potassium or sodium. Instead, they could only postulate such ionic movements as they thought would explain the interactions between the voltage clamping and the subsequent currents in different media.

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

This scientific use of intralevel experiments to inform conclusions about lower level goings on is a guide that suggests that scientists are using compositional abduction. Indeed, at this abstract level, one can see that Tolman had engaged in such reasoning about maze experiments with rats. Tolman, or his colleagues, placed rats in start boxes of mazes and measured the rats' movements through the maze, but he did not measure what was going on in a rat's brain when it was navigating a maze. In interpreting such experiments, Tolman relied on compositional abduction. In the child development literature, the preferential looking paradigm is an intralevel method. It involves having an infant view various sequences of events and measuring their responses, which are then used to make inferences about what is going on in the infant's mind. So, merely starting with a single case study may not, in general, make it clear how to proceed, but one can at least entertain a reasoned hypothesis about how to proceed.

De Regt (2017) takes up Pitt's dilemma, but interprets it differently than I do. After citing the passage above, he writes, "In other words, the philosopher who uses case studies is either guilty of 'cherry-picking' or of overgeneralization" (De Regt, 2017, p. 7). As I read the first horn of Pitt's dilemma, manipulation of data is a matter of interpreting a scientific case in a prejudicial manner. In contrast to data manipulation, I understand "cherry picking" as selecting individual cases, while ignoring similar cases that contradict one's preferred account. One might get the case cited correct, but one ignores relevant near neighbors. Further, as I read the second horn of Pitt's dilemma, it is a matter of not knowing how to generalize, rather than overgeneralizing. Thus, as I read things, there are two dilemmas in play. There is Pitt's dilemma of case studies and de Regt's dilemma of case studies. Both are serious and must be addressed.

As with Pitt's dilemma, it is indeed *a priori* logically possible that the cases are cherry-picked. Whether a case is cherry-picked must be determined *a posteriori* by looking to nearby cases. One precaution I take against cherry picking is to use a case favored by the New Mechanists, the Hodgkin–Huxley theory of the action potential. In this case, there is at least the common ground that scientists are dealing with compositional explanations. That is not all. In Chapter 5, I do not cover just one experiment from Hodgkin and Huxley. I review the interpretation of many experiments in mid-twentieth-century physiology. In Chapter 6, I review several different experiments regarding the biological basis of the Hermann grid illusion. *A priori* one cannot tell if these many examples are cherry-picked. Skeptics will have to review the relevant history for themselves.

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

As for overgeneralization, my general strategy is to err on the side of caution. My default proposal is to claim merely that there exist cases of a particular description, namely, the ones demonstrated. An existential conclusion is, of course, a weak conclusion, but the existence claims are "beachheads" for further theoretical exploration. Such a cautious default is warranted as many of the proposals I offer barely register on the philosophical landscape. Peirceans rarely, if ever, entertain the possibility of abduction being used for confirmation. Compositional abduction is barely noticed in the IBE literature. New Mechanists have rarely commented on abduction or IBE. Perhaps one can first agree that there are cases as I describe, and then turn to how pervasive they are. That is my strategy.

As an aside, I would add that I accept both Pitt's and de Regt's dilemmas. They are genuine challenges that the historian and philosopher of science must face. De Regt offers a different reply. He proposes to "dissolve" the dilemma on the grounds that it rests "on a misguided view of the relation between philosophy and history of science. It assumes what Jutta Schickore (2011) calls the 'confrontational model of HPS,' on which philosophical theories of science have to be confronted with historical data" (De Regt, 2017, p. 7).[40]

I do not embrace de Regt's dissolution for two reasons. First, for my project, I am very much concerned to draw attention to features of scientific reasoning that have largely been ignored by philosophers of science. One may read this as testing extant philosophical theories against the historical record. This does not commit me to drawing an analogy between this philosophical method and a scientific method of testing theories by experimental data. Nor does this commit me to claiming that confrontation is *the* relation between philosophy of science and the history of science, but it is one important relation. Nor does this commit me to a history and philosophy of science model in which the *discipline* of the history of science provides data for the *discipline* of the philosophy of science. In my experience, historians of science have their own intellectual agendas, which generally do not include being aides to philosophers of science. Historians of science are rarely interested in, say, whether Hodgkin and Huxley used abduction to introduce the sodium hypothesis (a la Peirce), rather than to confirm the sodium hypothesis, as I contend.

Second, de Regt's preferred model of integrated history and philosophy of science does not avoid the dilemma(s). He proposes that

[40] Schickore (2011).

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press

> philosophical analysis and historical research should continuously and mutually interact. . . . The attempt to construct a coherent narrative may require articulation or modification of the employed philosophical concepts, which can subsequently be applied in other historical case studies. (De Regt, 2017, p. 9)

This is a sensible reply. There should indeed be a continual interaction between one's philosophical account and the historical record. Starting with particular historical episodes is reasonable as well. My science-first approach embraces it. I would say that the account of compositional abduction in Chapter 2 constitutes a relatively neat package of ideas, but that close contact with scientific cases precipitated a less organized package in Chapter 3. Indeed, throughout the text I will draw attention to cases where historians and philosophers of science should distinguish singular compositional abductive inferences from other "neighboring" inferences they are often combined with. Such neighboring inferences may be simple enumerative inductive inferences.

Nevertheless, even using an iterative method and starting with a particular historical episode, Pitt's and de Regt's dilemmas remain. One must take care not to distort a case to fit one's preconceptions. One must figure out *a posteriori* how, if at all, one might generalize from a particular historical episode. One must take care not to cherry pick one's examples. One must be careful not to overgeneralize. De Regt does not speak to the matter of how one addresses these concerns with an iterative method, whereas I have at least given it a try.

1.6 Summary

This book offers a theory of how scientists confirm compositional hypotheses. It is driven by a close reading of scientific experimental work. Because it is offering a theory of these practices, it uses philosophical concepts that are often unfamiliar to scientists themselves. To what extent the scientific ontological commitments described here conform to reality is the topic for another project; to what extent the reasoning described here is rational or warranted is also the topic for another project. A psychologically informed account of the "moment of insight" is another important project set aside for future work. While there are no *a priori* guarantees that our case studies are correct and representative of larger patterns in science, this must be assessed largely on an *a posteriori* basis.

https://doi.org/10.1017/9781009435710.002 Published online by Cambridge University Press