

## Negotiating History: Contingency, Canonicity, and Case Studies

Agnes Bolinska and Joseph D. Martin

Department of History and Philosophy of Science, University of Cambridge

*Abstract:* Objections to the use of historical case studies for philosophical ends fall into two categories. Methodological objections claim that historical accounts and their uses by philosophers are subject to various biases. We argue that these challenges are not special; they also apply to other epistemic practices. Metaphysical objections, on the other hand, claim that historical case studies are intrinsically unsuited to serve as evidence for philosophical claims, even when carefully constructed and used, and so constitute a distinct class of challenge. We show that attention to what makes for a *canonical* case can address these problems. A case study is canonical with respect to a particular philosophical aim when the features relevant to that aim provide a reasonably complete causal account of the results of the historical process under investigation. We show how to establish canonicity by evaluating relevant contingencies using two prominent examples from the history of science: Eddington's confirmation of Einstein's theory of general relativity using his data from the 1919 eclipse and Watson and Crick's determination of the structure of DNA.

*Note:* This essay was awarded the 2019 IUHPST Essay Prize in History and Philosophy of Science.

### 1. Introduction

A century ago, Arthur Stanley Eddington made his renowned eclipse observations. According to the best-known version of the story, his measurements of starlight bending in response to the sun's gravitational field validated predictions of Albert Einstein's general theory of relativity, rocketing both men to fame and cementing relativity's superiority over Newtonian mechanics (see Missner 1985). This would appear to be a clear case from which to draw more general conclusions about the scientific process. It seems to exhibit many salient features beloved by philosophers: two theories that make different predictions; an experiment carefully designed to test those predictions; and a result favoring one theory over another.

But a shadow of doubt has darkened Eddington's analysis. In 1980, John Earman and Clark Glymour published a reanalysis of Eddington's published results. Eddington and his collaborator Frank Watson Dyson, they concluded, had fudged the data: "the eclipse expeditions confirmed the theory only if parts of the observations were thrown out and the discrepancies in the remainder ignored; Dyson and Eddington, who presented the results to the scientific world, threw out a good part of the data and ignored the discrepancies" (Earman and Glymour 1980, 85). The reasons, others have speculated, were political and ideological: Eddington, already convinced that general relativity was correct, wanted to see scientific ties to Germany restored after the damage of the Great War. Promoting the theory of a German

scientist and fellow pacifist advanced that end. Perhaps these considerations, conscious or not, compelled him to put his finger on the scales (Collins and Pinch 1993, ch. 2; Waller 2002, ch. 3).

Such a case might worry philosophers seeking historical support for their theories. If historical reinterpretation can render such an ostensibly clear case irrelevant for philosophical questions, then are all historical cases similarly vulnerable? Here, we address the challenge posed by this and other critiques of the uses of historical case studies in philosophy of science. Although previous commentators have outlined various reasons philosophers might regard historical case studies skeptically, their objections have rarely been systematized.<sup>1</sup> We identify and characterize the objections that have been leveled against the use of case studies as evidence for philosophical claims. Systematizing these objections reveals that most of them amount to admonitions to exercise methodological caution—the same sort of caution that scientists should observe when interpreting data or that philosophers should employ when constructing thought experiments. In section 2, we describe and respond to these challenges, treating them not as rigid constraints, but as components of a roadmap for applying historiographical sense in service of better philosophical reasoning.

However, one family of challenges to using historical case studies in support of philosophical conclusions, which we confront in section 3, has to do with the nature of history itself: for one, history is *contingent* in the sense that things could have gone otherwise, and, on the face of it, this raises doubts about whether we can draw philosophical conclusions from case studies. Building on Hasok Chang's (2012) analogy between the historical process of science and the narrative sweep of a television program, we show how attention to contingency in the historical process can establish the *canonicity* of certain cases, relative to particular philosophical aims. Canonical cases, we argue, *can* serve as evidence for philosophical claims.

Before embarking on this project, we must delineate the scope of our analysis. Philosophy of science addresses a range of questions, which can be divided roughly into metaphysical and epistemological questions. Metaphysical questions are concerned with what the world is like: what kinds of entities and processes does it contain, and what are their characteristic features? We do not ask about the nature of time or causation from the armchair; instead, we consult our best physical theories, supplementing the answers these theories provide with interpretations and argumentation that are beyond their remit. Metaphysical questions address the nature of reality; science informs these answers insofar as it furnishes evidence for this.<sup>2</sup> It thus seems that, in order to answer metaphysical questions, we should consult our best *current* science.

---

<sup>1</sup> One exception is Rafael Scholl and Tim Rätz (2016), who distinguish between foundational and methodological critiques as a preamble to their more thorough taxonomy of case studies. We follow a similar approach, with the aim of developing a more fine-grained typology of the critiques themselves.

<sup>2</sup> Here, we take an implicit stance on debates in the metaphysics of science (see Ladyman et al. 2007; Slater and Yudell 2017) in order to constrain the scope of our analysis. This stance, however, does not bear on the argument presented in the rest of the paper.

Epistemological questions, on the other hand, aim to illuminate the nature of scientific knowledge: What is it? How is it acquired? These questions concern things like the nature of explanation, the consequences of theory-ladenness of observation, the rationality of theory change, and the relationship between theory and experiment. In contrast to metaphysical questions, which ask what the world is like, the subject matter of epistemological questions is the practice of science and the relationship between this practice and what we can know about natural phenomena. Since the history of science gives us information about the very practices whose epistemic value we aim to understand, we address how case studies can help answer this latter type of question.<sup>3</sup> The overarching aim of this paper, then, is to show how historical case studies can best provide evidence for answers to questions in the epistemology of science.

## 2. Problems of Method

Existing critiques of the use of historical case studies in philosophy of science fall into the broad categories of problems of method and problems of metaphysics.<sup>4</sup> We consider them in turn, suggesting ways that each critique can be either obviated or turned to the philosopher's advantage.

The first sort of critique is that the use of historical case studies in support of philosophical conclusions is prone to bias. First, historical case studies are narratives that historians must *construct*; because narrative construction is underdetermined, historians' values and philosophical commitments can influence this process. Second, philosophers tend to *select* those case studies that support their theories and ignore those that contradict them. Third, philosophers' aims can influence how historical cases are *interpreted*. Finally, even on the assumption that we agree on a particular interpretation, how we *apply* it may differ: which conclusions we draw from cases will depend on prior philosophical commitments.

### 2.1 Construction Bias: The Theory-Ladenness of Historical Accounts

The first sort of bias arises because historical case studies are not foraged from the wild, but must be constructed on the basis of various types of evidence. The historian furnishes an account of events: what happened; who was involved; which factors were more or less salient? Why did, say, an experiment or discovery unfold the way it did, rather than some other way— who was responsible, what enabled them, and which social, economic, and other factors laid the groundwork?

How does the historian build an account that answers these kinds of questions? Drawing on Hayden White's (1973) account of historical narrative building, Katherina Kinzel (2015a; 2016) suggests that each stage of historical-case construction is theory-laden.<sup>5</sup> Historical cases are *selected* from a chronicle—the entire space of past events—and molded into a story with

---

<sup>3</sup> Schindler (2018, 188) makes a similar point.

<sup>4</sup> Discussions about the role of history of science in the philosophy of science date to Kuhn (1962). For summaries of the history of these discussions, see Nickles (1995), Schickore (2011), and Kinzel (2015a).

<sup>5</sup> This aligns with Lakatos's view that "history without some theoretical 'bias' is impossible" (1978, 107).

a beginning, middle, and end; certain aspects of that story are *emphasized*; and the way in which events are *emplotted*—the precise way in which the story is told—affects the meaning of the information the case contains. For instance, Kinzel (2015a) argues that Steven Shapin’s (1975) sociological approach to history disposes him to emphasize the relationship between science and society, historical actors’ social classes, and social transformations in his account of phrenology in nineteenth-century Edinburgh. Moreover, emplotment requires making numerous decisions, each of which is theory-laden. How thickly or thinly should events be sliced? How much and what kinds of evidence must be included in the plot for it to be considered satisfactory? What do key terms mean?

Kinzel argues that although case studies can serve a limited evidential function, they cannot adjudicate between competing philosophical theories. It is always possible to construct alternative narratives for a given historical episode, and it is often impossible to choose between them in a theory-neutral way. For instance, Geoffrey Cantor (1975) criticizes Shapin’s account of the Edinburgh phrenology debates on the grounds that Shapin fails to explain *how* the social forces he emphasizes shape scientists’ beliefs. Kinzel argues that this criticism reflects a difference in methodological standards, and that any choice between standards is theory-laden. For Shapin, stating the sociological facts is sufficient to account for differences between scientists’ theoretical approaches; Cantor insists that for these factors to be explanatory, we would also need to know how they influence individual beliefs and attitudes. According to Kinzel, the fact that criteria for choosing among alternative historical narratives are theory-laden implies that these narratives cannot arbitrate neutrally between competing philosophical theories.

## 2.2 Selection Bias

Even given an array of theory-neutral narratives to choose from—the possibility of which the above objection denies—we would still encounter another problem. Philosophers, addressing philosophical questions, naturally proceed with philosophical agendas. And the history of science offers a cornucopia of examples; historical grist exists for almost any philosophical mill. The utility of case studies might then be undermined by philosophers’ psychological predisposition to cherry-pick those cases that support their preexisting views (Pitt 2001).

This critique has a long history. Thomas Nickles points to a legacy dating to the 1960s of historians criticizing philosophers for “yank[ing] specially selected cases out of time and historical context to support a favorite thesis” (1995, 141). More recently, Jutta Schickore worries that “philosophers sifting through historians’ works trying to find suitable data would ... do so with an eye to the conceptual framework they sought to support” (2011, 467). This is a concern about how philosophers tend to behave when they wade into historical waters. Rather than being led by the content of historical examples, they tend to select case studies that match their philosophical presumptions. If this critique is right, then we should be skeptical of any historical evidence presented in favor of philosophical claims: equally potent examples might negate them.

## 2.3 Interpretation Bias: The Priority of Philosophical Commitments

The influence of philosophers' agendas, according to Anjan Chakravartty (2017, ch. 1), runs even deeper: not only might our theoretical commitments bias the construction and selection of case studies; they can also affect how case studies are *interpreted*. After the pessimistic meta-induction undermined scientific realism, realist philosophers sought to salvage some of its parts. Although realism about our theories *tout court* was no longer a tenable position, we might nonetheless still be *selective* realists, that is, believe that some aspects of our best theories are (approximately) true. Chakravartty argues that case studies cannot adjudicate between various selective realisms and their critics because how we use case studies depends on prior philosophical commitments, in the sense that we are inclined to interpret them so as to support our positions. This is a similar claim to construction bias, except that here, it applies to philosophers' interpretations of these narratives—which may well bear significant resemblances to the process of narrative construction that Kinzel (2015a) outlines.

Stathis Psillos (1994), for example, argues that even though scientists in the late eighteenth and early nineteenth centuries accounted for thermal phenomena by reference to the absorption or emission of caloric, they were only modestly committed to its existence. Moreover, the empirical success of caloric theory did not depend on caloric's existence: thermodynamics, which replaced caloric theory, retained the mechanisms involved in explaining thermal phenomena in caloric theory. Thus, Psillos contends, we can believe those parts of theories that are necessary for explaining their empirical successes; since caloric was not necessary for explaining these successes, belief in caloric's existence was not justified, and realism survives the critique.

However, Chang (2003) and Kyle Stanford (2003) argue that Psillos's interpretation of the history of caloric theory is subservient to his realism. An alternate interpretation of this history reveals that scientists *were* committed to the existence of caloric, and that this commitment was central to the empirical success of caloric theory. In fact, these authors argue, the examples Psillos cites as evidence of scientists' lack of commitment to the existence of caloric ignore the influence of contemporary rhetorical norms. Such norms required modesty about metaphysical commitments. Given these norms, we *cannot* conclude that these scientists were not committed to the existence of caloric.

For Chakravartty, this is a case where our philosophical ends might affect how we interpret historical case studies. Even assuming that historians share historiographical standards that guide interpretations, disagreements are sure to arise. Given how prior philosophical commitments sway our interpretations, it is unclear how we can adjudicate between competing interpretations in anything like a theory-neutral way.

#### 2.4 Application Bias: The Necessity of Prior Philosophical Commitments

Chakravartty (2017, ch. 1) also suggests that we cannot avoid prior philosophical commitments even if we agree on how to interpret a case study. They still color our analysis when we try to decide what philosophical view the case supports. Consider entity realism, the view that we should believe that certain unobservable entities exist because we can putatively use them to manipulate other systems. On this view, although what we believe *about* these entities changes when theories change, warranting a skeptical or agnostic attitude toward how these entities are

characterized, we are nevertheless justified in believing *that* they exist (Hacking 1983; Cartwright 1983).

Even if we agree on the facts about historical cases and how to interpret them, we might still disagree about what we ought to conclude on that basis. For instance, we can agree about when the term “electron” was coined and which conclusions scientists drew from central experiments, such as J. J. Thomson’s observations of cathode rays in discharge tubes or Millikan’s measurement of the charge-to-mass ratio of oil drops suspended between electric plates. However, because the properties scientists ascribed to electrons changed significantly over the course of the twentieth century, we might disagree about whether the early-twentieth-century electron *is the same entity* as the mid- or late-twentieth-century electron: it depends on which theory of reference we adopt (see also Arabatzis 2006, 32–35).

On a causal theory of reference, an entity is “baptized” with a name, after which that name continues to refer to that entity because of the causal chain leading back to the baptism (Kripke 1980; Putnam 1975). A causal theory of reference favors entity realism: even when properties of the electron radically change, we can maintain our belief in the electron’s existence. On a descriptivist theory of reference, in contrast, part of a term’s reference depends on how the entity it refers to is described. Thus, when our characterization of electrons changes to a significant degree, we are no longer warranted in our assumption that we refer to the same entity. Entity realism is undermined. We might still be talking about “electrons,” but this is a linguistic accident; “electron” refers to different entities across theory change.

### 2.5 Responses to Problems of Method

Case studies, it seems, are beset by methodological challenges on every side. Historians operate from (often tacit) philosophical presumptions when they conduct historical inquiry and so, to begin with, any case studies available to philosophers are freighted with philosophical baggage. Philosophers then select, interpret, and apply historical cases in light of preexisting agendas. Thus, even if we ignore the problems with how historical case studies are constructed, their ability to sustain any reliable inferences remains dubious. In this section, we demonstrate that these methodological challenges are not special ones—they reflect the potential for bias that threatens any epistemic practice.

Kinzel (2015a; 2016) argues against the use of case studies to arbitrate philosophical disagreements on the grounds that they are theory-laden, casting theory-ladenness as an insurmountable problem. If criteria are theory-laden, then we cannot say which criteria for historical case construction are best; all cases constructed in line with prevailing historiographical standards stand on equal footing. However, although individual historians will indeed tend to favor particular methodologies, and methodologies differ between historians, this does not imply that there is no way to resolve the tension. Instead, we can make arguments for one approach and against another, and we can evaluate these arguments from a metahistorical perspective. For instance, we might criticize an intellectual history for not taking sufficient account of social forces shaping historical actors’ motivations and actions. We could do this both by pointing to evidence about the case itself—for instance, showing that the “purely intellectual” factors are insufficient to explain a particular episode (see also section

4.3)—or by reference to more general arguments in support of the important role that social factors play in the history of science. The fact that we can and often do argue in this way undermines Kinzel’s claim that theory-ladenness prevents the successful arbitration of competing historical claims.

But are these metahistorical claims not also theory-laden? This is a prudent worry; theory-ladenness certainly can enter at the meta level, or at any level of analysis, but the possibility or even the actuality of its presence does not imply that it is *impossible* to adjudicate between positions. Instead, we do what we always do: we evaluate arguments on the basis of coherence, cogency, and other evaluative criteria—criteria that can be articulated on independent grounds. In other words, the fact of theory-ladenness does not render us wholly incapable of evaluating arguments. Theory-ladenness does indeed pose a challenge to the use of historical case studies for philosophical claims, as it does in science, or in any other epistemic pursuit, but that does not mean that any claim made in the course of these pursuits is as good as any other.

We can reply in similar ways to the selection-bias critique and to Chakravartty’s charge that philosophers’ interpretations of case studies are biased by the conclusions which the case studies are used to support. Just because it is *possible* to select and interpret case studies in light of one’s philosophical aims doesn’t mean one *must* do so! One should be wary of potential biases, and thus approach case-study selection and interpretation scrupulously.<sup>6</sup> Moreover, part of an academic community’s job is to point out biases that individual authors have missed, show that the cases they have selected are not representative, point to counterexamples, or argue that their interpretations are misguided.

What about the problem of application bias? Chakravartty argues that case studies cannot adjudicate between different views on entity realism because these views are underdetermined by history; we need a theory of reference to supplement them. This puts the cart before the horse: although it is true that we need a theory of reference to decide whether the history supports entity realism, it does not follow that we should choose our theory of reference to favor our desired outcome. Rather, we should adopt a theory of reference on independent grounds, viz. by carefully considering arguments for and against it, and only *then* look to the history to see what it tells us. That is, we *can* use historical case studies to adjudicate between positions on entity realism *with respect to a theory of reference*. One might object that we can be more confident in our knowledge of the historical facts than about which theory of reference is correct. But if Chakravartty is right that a theory of reference has to fill the gap between positions on entity realism and historical cases, then it will have to enter the picture somewhere. Our key claim is that our theory of reference should not be selected *with the aim of* supporting a particular view on entity realism, but rather on independent grounds.

For example, if we examine the history of science and see that whenever theories change, the new theory supposes entirely different entities to exist, we have a strong case

---

<sup>6</sup> Such biases might be explicit or implicit. In either case, the first step to addressing them is to make oneself aware of them. (Indeed, implicit bias training, designed to sensitize people to the ways in which subtle racial and gender biases influence our behavior, aims to do just this.)

against entity realism. But if we see that theories can change while beliefs in certain entities are retained, then we have a prima facie case for entity realism. Whether that case is strong does, indeed, depend on our theory of reference. But on the assumption that we have established a theory of reference, we can use the case to support or deny entity realism. For instance, if we establish a causal theory of reference on independent grounds, then we can argue that scientists really were talking about the same thing across theory change when using the word “electron”; *mutatis mutandis* for a descriptivist theory. But we should not choose our theory of reference to suit our needs. Doing so would be akin to cherry-picking data to support a favored scientific hypothesis.

This is not to say that disagreements about how a case ought to be constructed, selected, or interpreted, or about the nature of reference, are easy to settle; on the contrary. But this situation is not unique: the fact that some philosophical debates last centuries does not itself entail that these debates will never be resolved. Rather, arguments for different positions in these debates must be carefully considered, and whatever stand we take, we should admit its fallibility. And a resolution might involve a nuanced approach, reflecting the complexity of the issues at hand. Similarly, the presence of divergent views, and the psychological difficulty of avoiding bias at every juncture, does not imply that we should give up. History, philosophy, and indeed most academic disciplines rely on careful, critical analysis to answer difficult questions, even if a firm answer is not immediately forthcoming. The response to methodological critiques, in short, is to exercise methodological care.<sup>7</sup>

### 3. Problems of Metaphysics

The methodological problems described above are, at core, problems of underdetermination. Historical narratives are underdetermined by historical data; philosophical positions are underdetermined by historical cases. But what if history itself is just inherently unsuited to providing evidential support for philosophical claims? This concern runs deeper than worries focusing on how we construct and argue from case studies. If the utility of case studies is curtailed by virtue of features intrinsic to history, then the situation looks much grimmer. These critiques fall into two different, but related categories: Heraclitianism and contingency.

#### 3.1 *Heraclitianism*

Joseph Pitt suggests a reason to dismiss the case-study method that runs deeper than the methodological problems outlined in the previous section: the Heraclitian variability of history. Pitt claims that any responsible use of history requires a) appropriate contextualization, and b) consideration of the whole sweep of extended historical processes, rather than the isolated incidents that compose those processes. In the process of deploying a contextualist approach,

---

<sup>7</sup> It is beyond the scope of this paper to develop those standards of care here—a task that requires a long-term disciplinary commitment. However, the fact that we do not abandon other epistemic enterprises when they face similar challenges, but rather seek to mitigate those challenges, suggests that developing such mitigation strategies is the most appropriate approach to the methodological worries discussed in this section.

however, we find that our selection of the relevant features of historical context—for example, which individuals and practices to track—is both hopelessly arbitrary and reveals an essential instability of our analytical categories, as a result of which “all of the concepts we use to discuss science are in constant flux” (2001, 381). Scientific concepts, experimental methods and standards, and even the notion of science itself shift from one historical moment to another. On the strongest reading of Pitt’s view, history simply has no general features, and so case studies cannot support philosophical claims, which tend to be at least somewhat general.

Such a view comes with its own hefty metaphysical baggage about the nature of the historical process—assuming, for instance, that it varies in a stochastic way or according to indiscernible patterns—on the grounds of which some (e.g., Burian 2001) have challenged Pitt. Let us set that rejoinder aside for the moment, however, and focus on the deep differences Pitt observes between historical and philosophical standards of reasoning. Heraclitianism suggests that the sort of contextualization that is the gold standard of contemporary historical inquiry is incompatible with the philosophical impulse to generalize—and that history is therefore unable to play an evidential role for theses in the philosophy of science.

### 3.2 Contingency

In the final line of his condemnation of case studies, Pitt writes: “We [philosophers] seek precision, definitional clarity, analytic sophistication. These are good—but there is more to understanding: depth, flexibility, and a sense of the give and take and contingency found in history” (2001, 381). Pitt presents contingency here as it often appears in discussions of case studies: as an intrinsic feature of historical processes that tends to gum up the philosophical works. To the extent that philosophers are interested in history, this view suggests, they attend to processes that are deterministic (or nearly so) with respect to rational procedures and principles.

Contingency indeed presents serious problems if we accept this view. To any philosopher seeking to found normative claims about science in history, the critic can respond that any historical episode, or set of episodes, was either chancy or depended upon factors tangential to the philosopher’s normative aims. The critique might be formulated as follows: Because history is not governed by strict, deterministic rules (at least not rules that limited beings such as ourselves can discern), it might in some meaningful sense have gone differently. Because history might have gone otherwise, we have ample reason to doubt whether historical examples can constitute firm evidence for philosophical claims that seek to generalize about scientific practice and process.

### 3.3 Ungeneralizability: The Consequence of Heraclitianism and Contingency

The challenge for case studies, then, is this: Because of its fundamental variability and contingency, history might not be the type of thing that can sustain general claims. If so, it cannot support philosophical conclusions, which aim at (some level of) generality. We might worry that history is simply so variable that it cannot support even highly localized general claims. We can, if we look hard enough, find many counter-instances for any possible generalization. And even our best instances might have unfolded otherwise. It is therefore

doubtful that any number of historical cases can sanction universal or even relatively general claims about the scientific process.

Further, generalizing from case studies requires delineating cases, a process that creates additional opportunities for our biases to color our judgment and hampers our confidence in such generalizations (Pitt 2001, 374). How do we know when we are looking at something distinct enough to qualify as a case? What criteria should we apply to cordon off a case from adjacent events? We need cases before we can form case studies from which we might generalize, and the difficulty of carving history into cases hampers our ability to do so.

Schickore (2011, 469) notes that philosophers are often ambiguous about the target of their generalizations. It will remain unclear what a general philosophical claim is really about unless we agree on a well-defined scope for the philosophy of science—but such a consensus hardly seems forthcoming. And in its absence, it will be uncertain both what range of cases studies is germane to a given philosophical question and what the domain of applicability for any putative generalization is.

The problem of ungeneralizability is a concern for any underdetermined reasoning process, even those relying on highly consistent and reliable evidence. But Heracliticism and contingency might give us reason to believe that history is the sort of thing that is particularly ill-suited to generalization, above and beyond the problems of theory-ladenness. In addition, the fuzziness of cases and the uncertain scope of philosophy of science further suggest that examples from history might be fundamentally unsuitable for the philosophical project. Even philosophers seeking modest descriptive accounts of scientific reasoning and practice will find history too unruly to be systematized. The stronger, normative claims about right scientific practice other philosophers seek on the basis of historical evidence will then, *a fortiori*, be hamstrung from the outset.

Metaphysical critiques in their strongest form suggest that the process of history itself is simply inappropriate for the philosopher's evidentiary needs. These critiques, then, are distinct from the methodological critiques discussed above, and so have to be addressed somewhat differently.

#### **4. Responses to Problems of Metaphysics: Confrontation, Contingency, and Canonicity**

Metaphysical criticisms suggest that case studies cannot bear on philosophical problems *at all*, leading some to abandon talk of case studies altogether. Schickore (2011), in her own reaction to metaphysical criticisms, challenges what she calls the *confrontation model*, in which philosophical claims are confronted with historical evidence. Considering this model too naïve an analogy between the way data informs theory in the natural sciences, Schickore suggests that history and philosophy are partners in a shared interpretive exercise, bringing complementary resources to bear on a common object of inquiry.

We agree with Schickore's contention that philosophy of science is an interpretive exercise, but stop short of discarding the analogy with the natural sciences entirely. One does not need to adopt the confrontation model in order to draw useful parallels between philosophical and scientific inquiry. As we have seen, many concerns about case studies center on how biases inform interpretive processes—problems that have received a great deal of

attention from natural scientists. To the extent that reasoning from experimental scientific data is like reasoning from historical data, we can learn much from how natural scientists manage potential biases. We therefore follow Kinzel in suggesting that “historical case studies might provide some kind of evidence for philosophical claims even if the confrontation model is misguided,” and that we should not discard any babies in our haste to drain the bath water (2015a, 50).

Chang proposes another way forward: “I prefer to speak of historical ‘episodes’ rather than ‘cases’. When we have an episode of *The Simpsons*, or *Buffy the Vampire Slayer*, or what have you, the episode is not really a case or an example of whatever the general idea of the show might be. Rather, the episode is a concrete instantiation of the general concepts” (2012, 110–11). In Chang’s picture, history and philosophy interact iteratively, such that historiographical puzzles can improve philosophical frameworks, which in turn clarify the significance of historical episodes. Below, we take up and extend Chang’s proposal.

One limitation with the existing literature is that it focuses, with a few exceptions (e.g. Scholl and Räs 2016), on the broad question of using case studies, rather than on the particulars of what makes a *good* case study. This approach suggests that the conceptual validity of case studies needs to be established *before* we decide how to deploy them. But this is not obviously the best way forward. Instead, we might begin by thinking about what features a case study would need to have in order to be suitable for drawing inferences about philosophical claims. We undertake that task by showing how attending to the practices that establish what constitutes a *canonical* case study in philosophy of science obviates the objections we sketched above.

#### 4.1 Case Studies and Canonicity

Chang’s analogy to fictional television programs is compelling, and provides a useful alternative to the confrontation model. But it fails to reflect some pertinent aspects of the historical story we have. A single television show, despite employing many writers, is typically driven by a clear directorial vision. Although it might have many sub- and side-plots, it largely follows a single narrative arc. A similar claim with respect to science would be contentious at best. As a result of these characteristics, each episode of a television program instantiates the show as a whole to a similar degree. The same is not true of historical case studies.

Science, therefore, does not resemble a television show so much as a more sprawling fictional world, the product of many (often competing) authorial visions, with diverse and interlocking, but not necessarily unified, sub-plots, that evolves through multiple media. Roy Cook (2013) calls these *Massive Serialized Collaborative Fictions* (MSCFs): multi-author fictional universes with so many component works that no individual can reasonably consume them all.<sup>8</sup> An analogy between the history of science and an MSCF preserves the benefits of Chang’s analogy to television shows, while also capturing the essential variability of scientific practices.

---

<sup>8</sup> Examples would include the *Star Trek* and *Star Wars* universes, the DC and Marvel comic book universes, the *Buffyverse*, and the worlds of Sherlock Holmes and many daytime soap operas.

Cook's analysis identifies a distinctive feature of MSCFs. Because of the practical limitations to any individual gaining a synoptic view of these fictional universes, detailed practices must emerge to negotiate *canonicity* within them. Canonicity, for Cook, is a property of "a privileged subfiction that constitutes the 'real' story regarding what is fictionally true in the MSCF, whereas noncanonical stories are 'imaginary' or are de-legitimized in some other sense" (2013, 272). The canonical status of particular works within an MSCF is established through negotiation, a process that involves both producers and consumers of the work, responds to political and commercial pressures as well as matters of internal consistency, and results in a provisional judgment that is subject to renegotiation.

If we accept Chang's picture of the history of science as like a fictional world insofar as each case instantiates general concepts, and incorporate the insight that the history of science is more like an MSCF than a single television program, then we should accept the inevitability that some cases will instantiate general concepts better than others. Some historical cases are canonical for particular philosophical questions, others are non-canonical. And canonicity is a negotiated property of a historical episode: it must be established through a give-and-take between historians and philosophers.

This characterization has three consequences. First, canonicity is a provisional designation. As with works of fiction, the canonicity of historical cases is always open to renegotiation. As we discuss below, this feature nicely captures the actual practice of the history of science. History is a complex system that provides messy data. Using it requires tools suited to that data. Previous commentators have tended to view historians' and philosophers' disagreements over how to interpret case studies as evidence that a multiplicity of competing interpretations undermines their utility (e.g. Kinzel 2016). But if we conceptualize this as a process of negotiating canonicity, then we can view it instead as aiding the utility of history for philosophy of science by sharpening our understanding of particular cases and forcing us to examine and defend their canonical status.

Second, canonicity is a local, purpose-relative property of case studies. Here, our notion of canonicity differs somewhat from Cook's, which establishes canonicity with respect to a maximal fictional universe. It would make little sense to say that a historical case study is canonical for all of science; we must take seriously the provision that a case is only ever canonical with respect to a particular philosophical aim. Thus, case studies that we might want to consider non-canonical with respect to questions about good scientific practice—such as Lysenkoism—might still be canonical with respect to philosophical questions about the influence of ideology on scientific practice. When we negotiate canonicity, we must be clear about the domain of applicability and the level of generality of our cases.

Third, canonicity implies a particular relationship between canonical case studies and the philosophical disputes upon which they bear. If we establish a case as canonical, can that case *settle* a philosophical dispute? In principle, yes; in practice, it is not so straightforward. Before we can deem a case study canonical, we require a thoroughgoing understanding of its relevant features. However, we can expect ignorance of some those features to linger, for at least two reasons. First, we face the methodological challenges outlined in the previous section. Second, both history and philosophy are complex, so the sort of complete knowledge required

is unattainable for limited beings such as ourselves. That ignorance provides one opportunity for philosophical disagreement, even about cases that seem to motivate a stance on a philosophical problem particularly strongly. Nevertheless, if we have a case that would decide a philosophical question in the limit of our knowledge of its features, our improved knowledge of that case will more strongly motivate a particular stance.

This point resembles various discussions of the relationship between the abstract and the concrete (see, e.g., Cartwright 1989). An abstract concept *instantiates* a concrete one: something cannot be a game without also being football, or chess, or bridge, or some other specific game. Similarly, a case study cannot, *in principle*, be canonical for a philosophical problem without being an instantiation of some position or other. But the “in principle” here is key: actual cases deviate significantly from in-principle cases, due both to our difficulties in overcoming various biases and to the complexity of history and philosophy. In practice, there is room for disagreement even once a case is established as canonical.

To conclude, a clarification: when we try to establish canonicity, our first concern is whether the example is *relevant for deciding* a particular case. The fact of its relevance might or might not motivate a particular perspective on the philosophical problem. The discussion here, therefore, is concerned with the suggestion that variability and contingency entirely undermine the value of history for philosophy of science. We might agree that a case is canonical and nevertheless still have reasonable disagreements about what the case tells us; but this is a problem of method and subject to the considerations discussed above.

#### 4.2 Contingency Makes the Canon

We contend that, far from bedeviling philosophers’ attempts to use historical cases, contingency can make history matter for philosophy. To see how, we must first distinguish between three senses of contingency. John Beatty (2006) makes the most critical distinction between types of contingency. The first is unpredictability contingency (or “contingency *per se*”), which presumes fundamental randomness or indeterminacy in the historical process. We can understand Pitt’s Heraclitianism as implying a claim of this sort: history is unsuited for supporting philosophical claims because it could have been otherwise, even given indistinguishable starting parameters. The second is causal-dependence contingency (or “contingency upon”), which does not presume indeterminacy, but rather identifies the antecedent factors upon which historical outcomes crucially depend. To these, we can add explanatory-insufficiency contingency (Martin 2013). This can be understood as a weaker version of unpredictability contingency. A claim of this type takes the form: “X is contingent with respect to a particular set of explanatory factors.” That is, the underlying processes might or might not be fully determined, but if we only consider certain antecedent factors, then it will look to us as though the process is indeterminate.<sup>9</sup>

Critiques of the use of historical case studies in support of philosophical claims tend to invoke either unpredictability contingency or explanatory-insufficiency contingency. The

---

<sup>9</sup> The contingency concept has been decomposed further (Martin 2013; Kinzel 2015b), but for our purposes it will suffice to focus on these three general senses in which it is employed.

former kind of critique is implausible because it comes with deep metaphysical commitments about the historical process—commitments that are suspect. When we look at physical processes, we find some of them to be highly stable with respect to external perturbations. Think of the semiconducting circuits in your computer that are designed to perform identically across a wide range of temperatures. Others, like weather systems, are more chaotic. Small changes in initial conditions, like temperature, can lead to vastly different outcomes. Chemical and biological systems also run the gamut from highly robust to highly chaotic, depending on the type of system and the factors against which we assess their stability. (Your computer's circuits are robust with respect to temperature changes, for example, but less robust with respect to the action of a sledgehammer. Vice versa for the weather.)

We cannot experiment with historical systems, and so we cannot say with certainty whether they are robust or chaotic, but the most responsible assumption is that they lie across a similarly broad spectrum. Our understanding of processes that make up history gives us good reason to believe that some historical processes are robust, and others chaotic (Ben-Menahem 1997). Historical processes are constituted by physical, biological, psychological, and other processes that we know to have variable robustness. In short, we should be suspicious of any account of history that suggests that it is either always inherently stable or always inherently chaotic. One of the tasks of historical inquiry, then, is to offer arguments for why we should regard a particular system as either robust or chaotic, to what extent, and with respect to what.

Next, let us consider explanatory-insufficiency contingency, the other principal way in which contingency functions in critiques of historical case studies. Such critiques do not insist that a particular process is random or infinitely variable, but rather that the outcome of that process is not explicable in terms of a specific set of explanatory factors. The most familiar example comes from Stephen Jay Gould and Richard Lewontin (1979), who famously suggested that evolutionary outcomes are contingent (unpredictable) with respect to selection effects alone. This does not imply a commitment to the impossibility of evolutionary outcomes being fully determined when we introduce other factors—and, in fact, Gould and Lewontin's purpose was to draw attention to certain additional factors, such as genetic drift and pleiotropy, as necessary components of more complete evolutionary explanations.

Similarly, explanatory-insufficiency arguments in history suggest that a set of factors, typically a set that previous historians have favored, do not explain the outcome of the historical process and that we need to invoke other factors to develop a satisfactory account. In the Eddington case (which we discuss in more detail in 4.3), Earman and Glymour suggested that responsible data-handling practices do not provide an adequate historical explanation of the reported results of the eclipse expedition: their reconstruction of Eddington's data yielded a different result. Thus, we need to invoke something more than the data themselves and best practices for interpreting them to explain why Eddington got the result he did. And subsequent historians argued that ideological factors were best suited to fill the explanatory gap. Consequently—and critically—the argument that the outcome of the eclipse expedition derived from ideological factors threatens to render the case moot for a whole swath of philosophical questions. If we are interested in the rational process of theory choice, for instance, and we

believe that Eddington arrived at his results in an irrational way, the case will not help us resolve the philosophical debates in question.

Here, we should clarify the notion of rationality that underpins this contention. We follow Schindler in understanding rational processes or methods, first, to be justifiable independently of the historical facts describing how they were deployed (2013a), and second, to be those that are directed toward promoting truth or empirical adequacy (2018, 192–93). Schindler acknowledges that these are utopian goals, but argues that such a conception of rationality nevertheless enables us to judge when an agent is irrational, for instance, if she adopts a theory that contradicts all available evidence. This notion of rationality does not exclude social factors—indeed, rational decisions are often made communally—but it does exclude decisions that are made primarily on the basis of, for instance, political considerations. That is, had Eddington and Dyson interpreted their data to align with their political aims, failing to take into consideration whether their interpretation accorded with the evidence, we would take this to be irrational.

We pointed out above that explanatory-insufficiency contingency can be understood as a weaker form of unpredictability contingency. But another way to understand it is as the inverse of causal-dependence contingency. Whereas a causal-dependence claim singles out a factor or set of factors as potent determinants of a historical outcome, an explanatory-insufficiency claim instead highlights the causal impotence of certain factors. For example, the explanatory-insufficiency claim in the Eddington case is that rational factors alone were insufficient to explain why Eddington got the results he did. In other words, contrary to how historians traditionally presented the case, rational factors do not provide a causally complete account of Eddington’s results. The consequence, if such an argument holds, is that the case is immaterial for assessing scientific reasoning.

In order to argue that the Eddington case is indeed canonical with respect to philosophical questions about rational processes, we would need to show that rational factors can provide an adequate causal account of his reported results. In other words, to establish a case as canonical with respect to a philosophical aim, we necessarily invoke causal-dependence contingency, which consists of three related, but distinguishable claims:

- (1) *The causal claim*—a causal connection exists between a historical outcome and a particular antecedent factor;
- (2) *The counterfactual claim*—that antecedent factor might plausibly have been different;
- (3) *The sensitivity claim*—the historical outcome was non-robust with respect to changes in that antecedent factor (see Ben-Menahem 1997).

The plausibility of the contingency claim depends on each of these holding. The argument is undermined if the causal connection fails, if the relevant antecedent factor was in some sense inevitable, or if the outcome was robust with respect to other factors and therefore insensitive to changes in the relevant antecedent condition. For instance, we cycled to work this morning, and so perhaps we could claim that our presence at work is (causal-dependence) contingent upon our bikes being in working order. The bikes were causally relevant for our arrival at work (satisfying the causal claim) and bikes are the sorts of things that sometimes break (satisfying the counterfactual claim). However, if we live within walking distance of the

office, or on a reliable transit line, then we might challenge this contingency claim, because in the event of a flat tire, we could simply walk or take the bus. Our arrival at work is robust with respect to the condition of our bikes. For the causal-dependence contingency claim to hold in this case, we would have to be in circumstances such that the bikes were the only reliable means of getting to work.

How, then, does attending to contingency help us establish the canonicity of a case study? A canonical case is one that can be explained by factors relevant to the philosophical question at hand—and in which the outcome is sensitive to those explanatory factors. Establishing a case's canonicity, then, requires demonstrating that those explanatory factors are not washed out by other contingencies inherent in the process. That is, it involves making judgments about the relevant robustness relationships.

That is all well and good, a skeptic might argue, but how do we do that? We can understand how a well-controlled experimental system responds to small perturbations in initial conditions by running the experiment over and over again with the necessary modifications. History is anything but a well-controlled experimental system! But all is not lost. Two examples can illustrate how good historiographical sense can ground claims about the robustness of historical systems.

#### 4.3 Reviving the Canonicity of Eddington's Eclipse Expedition

Recall again the charge that Eddington and Dyson were selective in which data they factored into their analysis. Had they taken it all into account, Earman and Glymour argued, the result they obtained for the sun's effect on passing light rays would have been much closer to the classically expected value. But more recent historical work that takes a broader view of the relevant experimental practice has shown that the judgments Eddington and his collaborators made were based on a finely tuned *experimental* sense (Kennefick 2012; 2019). They worked closely with the finicky instruments they used to take the photographic plates that provided the eclipse data, and so they had a good idea of which data were reliable and which were questionable. They were not being led down the garden path by prior theoretical or political commitments; they were inching ahead carefully based on sound experimental sensibilities and intimate knowledge of their instruments.

How can attending to the contingencies involved in the Eddington example help us assess its canonicity with respect to philosophical questions about theory and evidence? First, consider the history of the example itself. We have a historical claim,  $C_1$ , that Eddington's observation of light deflection during the 1919 eclipse confirmed Einstein's general theory of relativity. This is a case that we might use to support a philosophical argument, for instance about the importance of novel prediction for theory choice (*pace* Brush 2015).

We can understand further historical work as arguing that this example is in fact non-canonical with respect to those questions. Earman and Glymour reexamined Eddington's publications, tried to redo his analysis, and developed a new historical claim,  $C_2$ , that Eddington's data could have been interpreted—maybe even could have been *better* interpreted—as supporting the classical theory.

This claim involves two contingency arguments. First is an explanatory-insufficiency argument: Eddington and Dyson's conclusion that Einstein's theory was right was *contingent with respect to* the eclipse data and the contemporary best practices for interpreting it. We must instead invoke other factors to account for their conclusion. And we can do so via a causal-dependence contingency claim. Subsequent historians advanced such a claim by suggesting that Eddington's conclusion was *contingent upon* his politics. In the wake of World War I, he had an interest in supporting the work of a German scientist and fellow pacifist in order to heal the damage done to the European scientific community (Collins and Pinch 1993; Waller 2002). In other words: (1) Eddington's politics were causally relevant to his results; (2) individual political commitments are not the sorts of things we consider inevitable, and so a differently inclined Eddington, or a scientist with different political sensibilities, might have behaved differently; (3) the results were sensitive to changes in the interpreter's political commitments; other factors would not robustly produce the same result.

But Daniel Kennefick's (2019) recent work represents another historiographical shift. He conducts a detailed analysis of Eddington and Dyson's reasoning process, including not only the photographic plate data, but also a close consideration of the instruments used to collect it. Interpreting data, as we all know, is not straightforward. It requires tacit knowledge and instrumental sense. Experience with specific instruments informs experimenters' judgments about which data is reliable and which should be discarded.  $C_2$ , that is, rests on an assumption that Eddington and Dyson discarded data without good methodological cause. But that claim must be defended with respect to the experimenters' interactions with their tools. Kennefick shows why, on these grounds, they had good reason to discard the data they did. That gives us a new claim,  $C_3$ , that Eddington's interpretation is (causal-dependence) contingent upon his data *and* his detailed knowledge of the instruments that produced them. On  $C_3$ , the Eddington case regains its canonical status with respect to questions about theory and evidence.

A case-study skeptic might suggest that philosophers lack the tools to judge these different interpretations, and so will be inclined to cherry-pick the one that best suits their preconceptions. But this would be cynical. Kennefick's analysis suggests a prudent principle of charity: it makes good methodological sense to assume that scientists broadly act in accordance with prevailing methodological norms. The burden of proof falls on those who propose deviations from them. On that basis, we have strong grounds for provisionally accepting Kennefick's contention that the eclipse data can count as a canonical example of responsible reasoning from data.

This is not to say that Kennefick's interpretation is the final word. It has, in fact, been challenged by Samuel Schindler (2013b, 96–98), who suggests that Eddington and Dyson did not have sound *experimental* reasons to selectively discard data. However, Schindler argues that they were nonetheless methodologically justified in looking to theory for guidance in pruning their dataset. This point, then, is aligned with Kennefick in seeking an explanation of scientists'

behavior in rational terms before reaching for contextual explanations tangential to the aims of the scientific enterprise.<sup>10</sup>

We can make an analogy here between scientists using their instruments and philosophers using case studies. Just as a good scientist ought to cultivate tacit knowledge that promotes good judgment about the data her instruments generate, a good philosopher ought to acquire the sensibilities that promote good judgment when using the tools of the trade. Good science involves selecting data with care, together with a keen know-how about how this selection ought to take place. Similarly, good philosophy of science, when it involves the use of historical case studies, must select and interpret these case studies judiciously using both philosophical and historiographical tools.

Some practical considerations come into play here. Most philosophers are not trained in history, and so need to develop a sense of how to manipulate it in the course of their philosophical work. One consequence of the view we present here is that philosophers who use historical case studies should seek out the sorts of experiences that confer historiographical sense. It is beyond our scope to suggest what such an effort would entail, but we would suggest that analyzing the contingencies at play in historical cases, as the above example demonstrates, is one of the most valuable tools for judicious selection of case studies.

Of course, we will never have a final, complete historical account. Prevailing historical judgments change, and so the interaction between history and philosophy must be iterative. But close contact between history and philosophy can ensure that this iterative process is a productive one—a point that comes through also in a second example.

#### *4.4 Could the Discovery of DNA Structure Have Been Otherwise?*

James Watson and Francis Crick's discovery of the structure of DNA marked a milestone in biology, but was subsequently mired in controversy. As is well known, a key piece of evidence supporting Watson and Crick's double-helical structure was Rosalind Franklin's Photo 51, whose sharp, distinctive "X" pattern was known to be characteristic of helical structures. Without Franklin's permission, her colleague Maurice Wilkins showed this photo to Watson, raising questions about how credit for the discovery of the DNA structure ought to be distributed. What role did this photo play in Watson and Crick's postulation of their structure? Would they have been so sure about it had Watson not seen the photo? If not, would they have announced their structure when they did? Might Franklin have arrived at the correct structure herself first?

These are questions about causal-dependence contingency, about how sensitive Watson and Crick's discovery was to their having seen the photo. Answers to these questions have implications for who deserves credit for the determination of the DNA structure, as well as for

---

<sup>10</sup> This point is similar, but not identical, to Lakatos's (1978) views on rational reconstruction, as explicated by Schindler (2018, 194–96). Lakatos contends that better accounts of scientific methodology can generate rational reconstructions that explain more historical facts in rational terms. The different, but related version of this claim that we defend is that it behooves us to see how far rational explanations can take us before we reach for other sorts of explanations.

larger issues about reasoning and inference in science. We acknowledge that discoveries do not happen in isolation; science is a collective enterprise, and those who are credited with discoveries rely on expertise, evidence, and input from other sources. To the extent that other contributions were essential or central to a particular discovery, they ought to receive credit. And one way of assessing centrality is by considering whether the discovery would still have been made had the contribution been absent.

Thus, although Watson and Crick did, in fact, find and announce the structure of DNA first, the extent to which they should share credit for this discovery with Franklin depends on how central her photo was to this discovery. And that depends on how contingent Watson and Crick's determination of the structure was upon Watson's seeing the photo. If they would have determined the structure anyway, then Franklin deserves less credit; conversely, had they not been able to find it at all without her photo, then Franklin deserves ample credit.<sup>11</sup>

We might even attempt a stronger claim: had Franklin been a man, she would have determined the structure herself, thereby receiving *full* credit for the discovery of the structure of DNA. In addition to how we answer the questions above, this stronger claim depends also on our evaluation of other contingencies. In general, how significant were barriers to women scientists' productivity in the mid-twentieth century? That is, were Franklin a man, would the absence of critical barriers have enabled her to get the structure first? In particular, how significant was Franklin's being a woman in Wilkins's lab at King's College? Had she been a man, would Wilkins have reacted to her presence at King's so negatively, and would they have had the toxic relationship they did, which was no doubt detrimental to their progress? Would he have shown Watson her photo?

Answering questions about contingency with certainty is impossible: in order to do so, we would have to alter the features of interest—for instance, by creating a world in which Franklin is a man, holding all else fixed, and then “replaying the tape” (Gould 2000; Beatty 2006). However, our inability to answer such questions *with certainty* does not imply that we cannot address them *at all*. Just as with any historical question, we must view our answers to questions about contingency as probabilistic and fallible; but we nevertheless have tools that can allow us to make educated inferences on the basis of the most plausible interpretation of all available evidence.

Often, in fact, we *can* assess claims about historical contingency using such techniques, just as we can in ordinary life. We can say, for instance, that had we removed the pizza from the oven ten minutes sooner, it would not have burned, that had we not prepared for our lectures, we would not have been able to deliver them, that had our bicycles not gotten flats, we would have arrived at work on time. We can say these things because of our past experience making pizza, delivering lectures, and riding bicycles. Similarly, we can assess historical

---

<sup>11</sup> It is outside our remit to defend the methodological wisdom of counterfactual reasoning here. For a lucid defense of their appropriateness—and, indeed, necessity—in the history of science, see Radick (2016). For a discussion of counterfactual reasoning in light of questions about contingency, see Soler (2015).

contingencies on the basis of what we know about particular historical actors and their contexts. For instance, if we find out that Wilkins had female collaborators with whom he worked well, and that King's was generally a good place for female scientists in the mid-twentieth century, then we have evidence that Franklin's gender alone was likely not an overriding hinderance to her research. Similarly, if we have evidence that, although seeing Franklin's photo lent further support to Watson and Crick's structure, they had already submitted their paper to *Nature*, the significance for their discovery of Watson's seeing the photo is undermined.

Asking such questions is relevant not only to the apportioning of credit, which might be regarded as a historical question, but also to answering philosophical questions. For example, one of us (Bolinska 2018) has considered the question of which research strategy for determining the structure of DNA—that of building the model up from component parts (Watson and Crick's) or that of deriving it from X-ray diffraction photographs (Franklin's)—was most promising. The contention is that, given certain features of available evidence in the context of mid-twentieth century biochemistry—how well-confirmed it was and how many structural candidates it enabled one to rule out—Watson and Crick's bottom-up strategy was, all else being equal, most likely to lead to the correct structure in the shortest period of time.<sup>12</sup>

We can use the historical case to assess this philosophical claim if we can establish that it is canonical with respect to this claim. In order to show that it is canonical, we would need to show that the salient features of the case are robust with respect to the relevant contingencies. And in order to do this, we would need to ask similar sorts of questions to those we would ask if we wanted to determine how to appropriately allocate credit for the discovery of the DNA structure. First and foremost, we would have to ask questions about contingency with respect to research strategies: had Watson and Crick opted for a different strategy, would they still have determined the DNA structure? Had Franklin adopted Watson and Crick's strategy, might she have determined the structure instead?

The answers to these questions can be informed by parallel cases: were similar strategies adopted in relevantly similar cases, and were they successful? Indeed, the race for the determination of the structure of protein a few years before the race for DNA is just such a case, with Linus Pauling adopting the model-building strategy (and subsequently inspiring Watson and Crick to do so) and competitors Sir Lawrence Bragg, John Kendrew, and Max Perutz working toward the structure from X-ray diffraction photographs. The results in this case were similar: Pauling, rather than Bragg et al., determined the structure first.

Second, we can also ask the same questions we did when we were asking about credit for the discovery. If we can establish that the case was robust with respect to these other factors, but contingent upon research strategies, then we will have further supported the canonicity of the case with respect to choice of research strategy. If, for instance, by evaluating the relevant contingencies, we find that Franklin's being a woman did not significantly hinder her work

---

<sup>12</sup> For a discussion of the roles of data, models, and theories in the context of the discovery of the DNA structure, see Schindler (2008).

toward the structure, this makes it more likely that it was instead her reliance on an inferior research strategy.

A few clarifications are in order. First, the cases above and their interpretations are meant only as examples of how contingency can be assessed in light of historical evidence in order to establish the canonicity of a historical case study with respect to a philosophical aim. The particular cases and their interpretations are of course contentious, and arguing that they *should* be interpreted as we suggest is beyond the scope of this paper. Second, and relatedly, establishing canonicity is a negotiation, and is by no means straightforward. Thus, disagreement about how to assess contingency claims is to be expected. Nevertheless, as we stressed, we should not take this disagreement to indicate that answers are not forthcoming or, worse, that there is no fact of matter. Making such an inference would be analogous to inferring from the absence of scientific consensus to the conclusion that such a consensus will never be reached, or that there is no fact of the matter (!) about the issue in question.

## 5. Conclusion

Recent work on the use of historical case studies as evidence for philosophical claims has resulted in several objections to this practice. This paper began by systematizing these objections. We found that they fall largely into two categories: methodological objections and metaphysical objections. The former, we argued, do not identify special challenges that do not also apply to other epistemic practices. Case studies demand responsible handling, but this is unsurprising. History is messy and philosophy is difficult. But the need for care is hardly the mark of a hopeless endeavor. Rather, attention to the *ways* in which history is messy and in which philosophy is difficult can be resources for developing better historiographical and philosophical practices.

Metaphysical objections do, however, raise special problems for the philosophical use of historical case studies. We showed that attention to what makes for a *canonical* case can address these problems. A case study is canonical with respect to a particular philosophical aim when the features of the historical system relevant to that philosophical aim provide a reasonably complete causal account of the results of the process under investigation. We showed how to establish canonicity by evaluating relevant contingencies using two prominent examples from the history of science: Eddington's confirmation of Einstein's theory of general relativity using his data from the 1919 eclipse and Watson and Crick's determination of the structure of DNA. These examples suggest that the analogy between philosophical inquiry and the natural sciences, although imperfect, has important elements that make it worth retaining. This is not to say that we should think of philosophy as modeled on scientific practice, but rather that both succeed by virtue of something more general: their reliance on shared principles of sound reasoning.

Taking seriously the practices necessary to establish the canonicity of case studies makes clear that some examples of the historical process of science are more salient to particular philosophical aims than others. With historiographical sense, we can pick these examples out. Doing so requires attention to the contingencies of history. Rather than undermining the use of historical cases, philosophical attention to contingency aids the development of case studies

as resources by making explicit otherwise tacit assumptions about which features of them are most salient and why.

It is possible, perhaps even easy, to use irresponsibly the rich resources that history provides to make a predetermined point. But that is not a genuine case of history of science informing philosophy of science—in part because it proceeds in the absence of historiographical sense. By outlining the practices that render particular cases canonical for certain philosophical aims, we have offered a route by which such sense can be integrated into standard philosophical practices.

### **Acknowledgments**

For insightful comments and conversations that much improved this paper, we thank Michael Barany, Mary Brazelton, Andrew Buskell, Jeremy Butterfield, Cecilie Erikson, Dan Kennefick, Josh Nall, Karoliina Pulkkinen, Darrell Rowbottom, Samuel Schindler, Jim Secord, and a lively audience at the Cambridge Philosophy of Science seminar. Finally, are grateful to the IUHPST for proposing the prize that inspired us to write this essay, and to the prize committee, Rachel Ankeny, Theodore Arabatzis, Hasok Chang, and Takehiko Hashimoto, for awarding it their 2019 prize.

## References

- Arabatzis, Theodore. 2006. *Representing Electrons: A Biographical Approach to Theoretical Entities*. Chicago: University of Chicago Press.
- Beatty, John. 2006. "Replaying Life's Tape." *The Journal of Philosophy* 103, no. 7: 336–62.
- Ben-Menahem, Yemima. 1997. "Historical Contingency." *Ratio* 10, no. 2: 99–107.
- Bolinska, Agnes. 2018. "Synthetic versus Analytic Approaches to Protein and DNA Structure Determination." *Biology and Philosophy* 33, no. 3–4: 26.
- Brush, Stephen G., with Ariel Segal. 2015. *Making 20th Century Science: How Theories Became Knowledge*. Oxford: Oxford University Press.
- Burian, Richard M. 2001. "The Dilemma of Case Studies Resolved: The Virtues of Using Case Studies in the History and Philosophy of Science." *Perspectives on Science* 9, no. 4: 383–404.
- Cantor, Geoffrey N. 1975. "The Edinburgh Phrenology Debate: 1803–1828." *Annals of Science* 32, no. 3: 195–218.
- Cartwright, Nancy. 1983. *How the Laws of Physics Lie*. Oxford: Oxford University Press.
- Cartwright, Nancy. 1989. *Nature's Capacities and Their Measurement*. Oxford: Oxford University Press.
- Chakravartty, Anjan. 2017. *Scientific Ontology*. Oxford: Oxford University Press.
- Chang, Hasok. 2003. "Preservative Realism and Its Discontents: Revisiting Caloric." *Philosophy of Science* 70, no. 5: 902–12.
- Chang, Hasok. 2012. "Beyond Case-Studies: History as Philosophy." In *Integrating History and Philosophy of Science*, edited by Seymour Mauskopf and Tad Schmaltz, 109–24. Dordrecht: Springer.
- Collins, Harry M., and Trevor Pinch. 1993. *The Golem: What Everyone Should Know about Science*. Cambridge: Cambridge University Press.
- Cook, Roy T. 2013. "Canonicity and Normativity in Massive, Serialized, Collaborative Fiction." *The Journal of Aesthetics and Art Criticism* 71, no. 3: 271–76.
- Earman, John, and Clark Glymour. 1980. "Relativity and Eclipses: The British Eclipse Expeditions of 1919 and Their Predecessors." *Historical Studies in the Physical Sciences* 11, no. 1: 49–85.
- Gould, Stephen Jay. 2000. *Wonderful Life: The Burgess Shale and the Nature of History*. London: Vintage.
- Gould, Stephen Jay, and Richard 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, Biological Sciences* 205, no. 1161: 581–98.
- Hacking, Ian. 1983. *Representing and Intervening*. Cambridge: Cambridge University Press.
- Kennefick, Daniel. 2012. "Not Only Because of Theory: Dyson, Eddington, and the Competing Myths of the 1919 Eclipse Expedition." In *Einstein and the Changing Worldviews of Physics*, Einstein Studies 12, edited by Christoph Lehner, Jürgen Renn, and Matthias Schemmel, 201–32. New York: Springer.
- Kennefick, Daniel. 2019. *No Shadow of a Doubt: The 1919 Eclipse That Confirmed Einstein's Theory of Relativity*. Princeton: Princeton University Press.

- Kinzel, Katherina. 2015a. "Narrative and Evidence: How Can Case Studies from the History of Science Support Claims in the Philosophy of Science?" *Studies in History and Philosophy of Science Part A* 49: 48–57.
- Kinzel, Katherina. 2015b. "State of the Field: Are the Results of Science Contingent or Inevitable?" *Studies in History and Philosophy of Science Part A* 52: 55–66.
- Kinzel, Katherina. 2016. "Pluralism in Historiography: A Case Study of Case Studies." In *The Philosophy of Historical Case Studies*, Boston Studies in the Philosophy and History of Science 319, edited by Tilman Sauer and Rafael Scholl, 123–49. Switzerland: Springer.
- Kripke, Saul A. 1980. *Naming and Necessity*. Oxford: Blackwell.
- Kuhn, Thomas. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Ladyman, James, and Don Ross, with Don Spurrett and John Collier. 2007. *Everything Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.
- Lakatos, Imre. 1978. *The Methodology of Scientific Research Programmes*. Cambridge: Cambridge University Press.
- Martin, Joseph D. 2013. "Is the Contingentist/Inevitabilist Debate a Matter of Degrees?" *Philosophy of Science* 80, no. 5: 919–30.
- Missner, Marshall. 1985. "Why Einstein Became Famous in America." *Social Studies of Science* 15, no. 2: 267–91.
- Pitt, Joseph. 2001. "The Dilemma of Case Studies: Toward a Heraclitian Philosophy of Science." *Perspectives on Science* 9, no. 4: 373–82.
- Psillos, Stathis. 1994. "A Philosophical Study of the Transition from the Caloric Theory of Heat to Thermodynamics: Resisting the Pessimistic Meta-Induction." *Studies in History and Philosophy of Science Part A* 25, no. 2: 159–90.
- Putnam, Hilary. 1975. *Philosophical Papers*, vol. 2, *Mind, Language and Reality*. Cambridge: Cambridge University Press.
- Radick, Gregory. 2016. "Presidential Address: Experimenting with the Scientific Past." *British Journal for the History of Science* 49, no. 2: 153–79.
- Schickore, Jutta. 2011. "More Thoughts on HPS: Another 20 Years Later." *Perspectives on Science* 19, no. 4: 453–81.
- Schindler, Samuel. 2008. "Model, Theory, and Evidence in the Discovery of DNA Structure." *British Journal for the Philosophy of Science* 59, no. 4: 619–58.
- Schindler, Samuel. 2013a. "The Kuhnian Mode of HPS." *Synthese* 190, no. 18: 413–54.
- Schindler, Samuel. 2013b. "Theory-Laden Experimentation." *Studies in History and Philosophy of Science Part A* 44, no. 1: 89–101.
- Schindler, Samuel. 2018. *Theoretical Virtues in Science: Uncovering Reality Through Theory*. Cambridge: Cambridge University Press.
- Scholl, Raphael, and Tim R az. 2016. "Towards a Methodology for Integrated History and Philosophy of Science." In *The Philosophy of Historical Case Studies*, Boston Studies in the Philosophy and History of Science 319, edited by Tilman Sauer and Rafael Scholl, 69–91. Switzerland: Springer.

- Shapin, Steven. 1975. "Phrenological Knowledge and the Social Structure of Early Nineteenth-Century Edinburgh." *Annals of Science* 32, no. 3: 219-43.
- Slater, Matthew, and Zanja Yudell, eds. 2017. *Metaphysics and the Philosophy of Science*. Oxford: Oxford University Press.
- Soler, Léna. 2015. "The Contingentist/Inevitabilist Debate: Current State of Play, Paradigmatic Forms of Problems and Arguments, Connections to More Familiar Philosophical Themes." In *Science as It Could Have Been: Discussing the Contingency/Inevitability Problem*, edited by Léna Soler, Emiliano Trizio, and Andrew Pickering, 1-42. Pittsburgh: University of Pittsburgh Press.
- Stanford, P. Kyle. 2003. "No Refuge for Realism: Selective Confirmation and the History of Science." *Philosophy of Science* 70, no. 5: 913-25.
- Waller, John. 2002. *Fabulous Science: Fact and Fiction in the History of Scientific Discovery*. Oxford: Oxford University Press.
- White, Hayden. 1973. *Metahistory: The Historical Imagination in Nineteenth-Century Europe*. Baltimore: John Hopkins University Press.