Philosophy of Science & the Curse of the Case Study[1]

Forthcoming in *The Palgrave Handbook of Philosophical Methods*

1. Introduction

We can divide the philosophy of science into two projects. Let's call the first *philosophy-directed*. Here, we seek to describe, systematize and explain scientific practice, and draw on this to argue for philosophical positions. Science informs philosophy.[2] The second could be named *science-directed*. We aim to clarify, clean-up and unify scientific concepts. Philosophy informs science[3]. Both projects lean on generalizations about scientific method, practice, development and so on. Frequently, such generalizations are made in reference to *case studies*, particular, detailed descriptions of scientific activity. Here, I defend the use of case studies in both philosophy-directed and science-directed contexts.

My defence is pluralistic: case studies play a variety of roles in philosophical methodology, and particular case studies act in more than one. These roles fall into three categories. First, we take case studies as inductive evidence. As we shall see, due to the heterogeneity of scientific practice, this can lead to a 'natural history' of science: philosophers ought to restrict themselves to detailed studies of particular scientific practices and refuse to posit generalizations. I resist both that science is so heterogeneous, and that philosophy of science need be so particularist. Second, case studies play a non-justificatory, pragmatic or perhaps rhetorical role. For example, referring to the same case study can centre dispute, providing philosophers common ground. I argue that over-emphasizing the heuristic use of case studies can undermine them. The third

[1] An early version of this paper was presented at PBDB7. Drafts were read by Marc Ereshefsky, Kirsten Walsh, Brandon Holter, Chris Daly and Alex Prescott-Couch. Their feedback is greatly appreciated.

[2] Alex Prescott-Couch has suggested to me that philosophy-directed projects be divided again: one project aims to explain scientific practice, where the other draws on science to inform philosophical debate.

[3] Godfrey-Smith (2009) makes a similar distinction, naming the 'science-directed' philosophy 'philosophy of nature'.

route is somewhere in between: philosophical analyses of sciences stand and fall on their capacity to explain, illuminate and unify scientific practice, and the case study is an important part of this success. Here, philosophers design concepts to solve either philosophical questions about scientific practice or to delineate and unify scientific concepts. Case studies play a role in justifying this by demonstrating, and in part establishing, the explanatory capacity of our theoretical work.

Rather than enquiring about why we use case studies, I defend the idea that we ought to use them. Case study methodology could be explained historically: that is, our method was driven by quirks in philosophy's development, rather than its being underwritten by rational methodology. I think it plausible that historical contingency takes its part of the blame, but I am not in the business of providing that kind of explanation. Lakatos (1970) distinguished between *internal* philosophy of science, which understands science rationally (in terms of some logic of science) and *external* philosophy of science, which understands science in terms of its contingent, irrational development. I am interested in an internal examination of philosophical method. Following Lakatos' lead, then, my project could be thought of as a *rational reconstruction* of the use of case studies in the philosophy of science.

Throughout the paper I sketch a moderate position on science's 'disunity'. I agree that science is heterogeneous—and heterogeneous in a way that matters—but I will suggest it is nonetheless 'patterned'; that is, claims about scientific practice will apply in some, but not all, contexts. This will underwrite the 'ahistorical' methodology I defend in 3.3.

In section 2 I provide a few examples of the use of case studies before articulating the 'curse': that case studies, if read evidentially, are a poor basis to build inductive generalizations about science. In section 3 I detail and critique the three broad roles discussed above. Finally, I conclude in section 4 by briefly discussing the mutual inclusivity of those roles and pointing to some methodological upshots.

2. The Curse

Here's a paper-schema philosophers will be familiar with. First, introduce some philosophical issue; second, launch into a detailed description some scientific endeavour; third, draw some general lessons on the original issue. In essence, I'm going to argue in support of that strategy. Why does it need support? Because *prima facie,* there's something problematic about how philosophers of science use case studies. At least one of our tasks is to generalize across scientific practice, and case studies don't help. The 'curse of the case study', then, is that case studies, by their nature, are peculiar and individual, but the generalizations philosophers seek are broad and unitary. On the face of it, we should switch our aims to suit our evidence, or switch our methodology to suit our aims.

I'll start by providing some examples of the practice, before describing the problem. Note that I will be drawing on examples of philosophy-directed work rather than science-directed, but will assume the lessons carry over.

2.1 Case Studies

Case studies drawing on historical and contemporary science are common in philosophy, anchoring individual papers, books, and entire debates. Although I am primarily interested in the philosophy of science, scientific case studies also play their role in philosophy more broadly. For instance, bee 'waggle' dances (Devitt 2013, Millikan 2005) and bacterial quorum signalling (Sykrms 2008) inform naturalistic accounts of intentionality, while the philosophy of perception frequently draws on psychological pathology. Here I will focus on the philosophy of science.

Carl Craver's *Explaining the Brain* (2007) has been influential. The book's aim is to provide an account of mechanistic explanation, a mode of explanation which is taken to be distinctively scientific. A mechanistic explanation targets some phenomenon by decomposing it into parts, and then understanding its behaviour in terms of the causal relationships between them. This results in both a constitutive explanation of the phenomenon in terms of its components, and a causal explanation of the phenomenon's behaviour in terms of the components' causal structure. In the book, important conceptual work is undertaken: Craver links mechanistic explanation with manipulationist accounts of causation and he undertakes an extensive analysis both of 'levels' talk and issues of explanatory reduction. It is clear that Craver's work is intended to uncover a relatively prevalent and moreover *explanatory* aspect of scientific practice. Throughout *Explaining the Brain*, we are treated to a detailed case study: explanations of action potential in neuroscience. Neurons 'communicate' via electrical discharges fired by the opening of voltage-gated ion channels embedded in their membranes. As Craver details, neuroscientists have built increasingly sophisticated mechanistic explanations of such behaviour.

And so, Craver made conceptual strides in our understanding of a general approach to scientific explanation and drew on this to discuss important philosophical issues. These general claims were made against the backdrop of the history of the quest to understand action-potential. Here, we draw on an episode in the history of cognitive neuroscience, rather than a particular study; case studies can differ in breadth. Let's turn to a wider debate.

The role and justification of idealized models has become a central issue in the philosophy of science. Philosophers ask what role such idealizations play, what the relationship is between model and world, and so on… In this debate, ecological Lokta-Volterra predator/prey models are frequently mentioned (Strevens 2003, Levy & Currie 2014, Weisberg 2007, 2013, Weisberg & Reismann 2008, Matthewson & Calcott 2011). These models consist of two coupled differential equations, one representing prey population and the other predator population. The populations' dynamics are linked and exhibit distinctive, out-of-phase oscillations. Predator-prey models have become a touch-stone for many philosophers of science interested in the role of abstraction and idealization in scientific theorizing.

And so, philosophers have theorized about the role of models (and idealizations generally) across science, often in reference to a set of models from community ecology, which are presumably taken to be representative of a broader set of scientific practices.

Philosophers, then, appeal to case studies frequently in their work. Sometimes case studies anchor a paper or monograph, sometimes they become paradigm examples, and sometimes structure entire debates. So, what's the problem?

2.2 The problem

The problem, in a nutshell, is this: *prima facie*, philosophers of science appeal to case studies to justify general claims about scientific practice, methodology or development, but case studies appear ill-suited for this role. To explain why, I'll take a quick foray into ampliative inference.

One way of justifying a claim about a general class is via *sampling*: we examine a sub-set of the class and then extrapolate. Let's take a simple example, caching behaviour amongst scrub-jays. Scrub-jays put aside food for winter, apparently employing rather remarkable episodic memory to do so (see Clayton & Dickinson 1998). One worry, if you are a scrub-jay, is pilfering: given the chance, one of your conspecifics will steal your food. Daly et al (2007) investigated whether scrub-jay caching behaviour differed when observed by other scrub-jays. They found that *experienced* scrub-jays, that is, those who have pilfered themselves, will shift caches if they have been observed, while naïve scrub-jays will not. How was this conclusion reached? They used a simple experiment:

> … in our 'barrier' experiment, the cachers were given the opportunity to cache food in two different, visuospatially distinct trays that were placed in the bird's home cage. A second bird, the observer, was placed in a cage opposite to that of the cacher such that it could easily see the caching bird. One of the trays was placed behind a barrier such that it was out of view of the observer, but the other was in full view so that the observer could clearly see the location of caches hidden in the full-view tray (511).

Unsurprisingly, scrub-jays preferred to cache out of sight of conspecifics, typically caching behind barriers. This behavior, however, does not distinguish between the hypothesis that scrub-jays exploit what their conspecifics cannot see, or simply prefer to cache when *they* cannot see conspecifics (the 'out of sight out of mind' hypothesis). One way Daly et al discriminated between these hypotheses utilized shading. They found that scrub-jays, particularly when observed, cached in shadowy rather than well-lit locations. They see this as reason to think that scrub-jays will actively exploit limitations in their conspecifics' visual field. The 'shade' experiment was carried out 21 times with 6 birds under observed conditions (there were 8 birds under private conditions). An inference is then drawn on this basis[4]. Is this sufficient?

Two things which determine the justification of the ampliative inference is (1) the sample size, and (2) the expected heterogeneity of the class[5]. If I only examine a few scrub-jays, then the behaviour I observe could be fluky: my subject scrub-jays might be particularly vigilant, or perhaps scrub-jay re-caching behaviour just happened to express in a patterned fashion on those occasions. To mitigate these possibilities, I need to see the behaviour in a sufficiently large group of scrub-jays. How many? The answer to this depends on how heterogeneous I expect scrub-jay behaviour to be. If I have reason to expect scrub-jays to be reasonably uniform in their behaviour, then a small sample size will do. If, however, scrub-jays are plastic, random or highly context-dependent in their behaviour (if the class is heterogeneous) then I need a much larger subset.

Case studies of scientific episodes are problematic as data-points, that is, evidence for ampliative inferences, as both the sample-size is low *and* we should expect the class to be heterogeneous. The second point deserves discussion. If science is fairly homogenous, and we have good reason to think our case study is representative, then low sample sizes could be

[4] I have simplified things somewhat: Daly et al's inference is not merely on the basis of the 'shadow' experiments, but also relied on field observations and various other experiments.

[5] This is certainly not the whole story here, but it is all I need to set up the curse.

mitigated. So, why think that science is a heterogenous phenomenon? Because of the appeal of a *disunified* picture of science (Wylie 1999, Wimsatt 2007, Dupre 1993, Cartwright 1999).

Science is a social activity which targets a wide variety of different phenomena and incorporates a wide variety of interests. Although such breadth does not guarantee disunity, it does at least put pressure on the intuition that science can be happily captured by unified generalizations. This point is made frequently by historically-minded philosophers. Pitt (2001), for instance, has this to say about definitions of 'scientific observation':

> … a serviceable universal account of scientific observation is not possible because the activity of making a scientific observation depends on, among other things, the sophistication of the technology available at the time, hence what we mean by a scientific observation changes. What is allowed as an observation varies in time, place and with respect to changing criteria influenced by technological innovation (Pitt, 374).

Because what counts as an 'observation' changes through time, we cannot provide a unitary story of that concept: the target shifts too much. I would add differences in explanatory target to Pitt's historical and technological changes. The scientist interested in understanding scrub-jay behavior has a very different task to the scientist testing physical theories, or one developing new pharmaceutical products, or one examining action potential in neural systems, or one trying to identify and account for morphological changes in the fossil record. It is undeniable that science is heterogeneous, but for this to bite it must be *relevantly so,* that is, it must be disunified in a way which undermines the philosopher's goals. For now, let us assume that is the case. How science might be relevantly heterogeneous will be clarified as we go. I suspect that some aspects of scientific practice will be more patterned and uniform than others, but nonetheless we should approach science expecting it to be heterogeneous, and thus ought to be suspicious of small data-sets. . This worry is not restricted to universal generalizations about scientific practice, but also effects more restricted general claims. Science is not only heterogeneous between domains, but *within* domains. Even if Craver is right about the nature of explanations of action-potential, this might not be true of ‘nearby’ areas of cognitive neuroscience.

And so, because any particular aspect of science is likely to be heterogeneous and a case study is a single data point, if case studies are inductive bases, we are making a mistake: this is the curse. Note that I am not denying that science is not unified, albeit partially, rather case studies do not grant epistemic license for making general claims about science. Before shifting to solutions, I turn to two weaker evidential roles which case studies could play.

Philosophers could appeal to case studies as *actuality proofs* or as *counterexamples*. I’ll discuss these in order. A restriction on the acceptability of an inductive generalization is there being at least one instance of it. That Craver is right (assuming that he is) about explanations of action-potential shows that his account of mechanistic explanation is on the table.

Science’s heterogeneity undermines actuality proofs. First, note that it is possible for a class to be heterogeneous, but still patterned. For instance, scrub-jays might change their re-caching behaviour when observed by conspecifics 95% of the time, while 5% of the time performing a wild variety of activities (adding food, ignoring it, putting a collection of colored pebbles around the cache, putting a leaf on their head and going ‘tweet’). In this situation, as there are examples of nearly any particular scrub-jay behaviour, unless there is some systematicity in the behaviours, actuality proofs for hypotheses of scrub-jay behaviour are boring.

Science could be analogous: it could be that for almost any faintly plausible claim about scientific methodology, there is some scientist somewhere doing it (and, perhaps doing it for good reasons). Call this the *patterned heterogeneity thesis*. If the thesis is true, then demonstrating that a theory is realized by some scientific field is boring. This type of heterogeneity in no way undermines drawing general lessons, but undermines the point of providing actuality proofs. In 3.3 the patterned heterogeneity thesis will underwrite an 'ahistorical' vindication of case studies.

Another use of case studies is as counterexamples. After all, philosophers make general statements about science, and case studies could be used to falsify those claims. One worry about this is that, in the face of heterogeneity, philosophers shouldn't be making claims *that generally* in the first place. Another is that, as Baumslag (2000) points out, most philosophical work is partly normative. Showing that scientific practices diverges from the prescriptions of philosophers is not alone a good reason to deny the prescription. The scientists in question must be both going against our normative claims and also succeeding.

> Information about how scientists behave can be used to test normative principles. Where the principles and the behavior diverge, we must examine the behavior carefully to see whether a justification can be given for it. If the behavior is justified, then we should reject our principles (Baumslag pp 267).

If the patterned heterogeneity thesis is true it is possible that, for any epistemic principle *p*, there will be a scientist somewhere doing *~p*, and succeeding. This doesn't show that scientists who *do* obey the principle are doing something wrong, or that *p* is mistaken in anything other than breadth. Our response to 'counterexamples' should not be to abandon our principles, but to restrict their domain. This kind of approach is implied by the methodology sketched in 3.3.

Here's a better thing to say: appeal to case studies is not important because it provides actuality proofs or tests *per se*, but because it anchors philosophical analysis. That is, given the abstractness of philosophical work, grounding it in actual science is better than nothing. Sure. I think there is much more to be said, however.

3. Lifting the curse

So, the curse of the case study can be summarized:

(1) If case studies play a legitimate role in the philosophy of science, they must *justify* our claims.
(2) Case studies cannot justify claims in the philosophy of science.
(3) Therefore, the use of case studies in the philosophy of science is illegitimate.

The discussion in 2.2 assumed that case studies, if legitimate, are inductive evidence of philosophical claims. The defence of the second premise claimed that the nature of science (its heterogeneity) undermines ampliative induction. There might be some general truths to be had about science, but case studies do not licence us.

How do we lift the curse? In the following, I consider three strategies. By the first, we accept the terms of the argument, and change our method, that is, pull back our ambitions. Philosophers of science, by this line, should be *natural historians* of science: we either restrict our scope to particular cases, or wait until extensive surveys are in before drawing any general conclusions. The second strategy denies the first premise by adopting an 'anti-realist' position. Case studies do not justify philosophical positions (as they are not inductive evidence), but play heuristic or rhetorical roles. The final strategy denies the second premise, claiming that case studies *do* play a justificatory role, but not as evidence for inductive generalizations. Rather, they are illustrative of conceptual tools. As I shall argue, the success of the conceptual machinery employed by philosophers of science, both on their case studies and in other scenarios, gives us reason to believe their claims.

As I spell out in the conclusion, although I prefer some roles more than others, they are not mutually exclusive. Philosophers of science incorporate case studies for a variety of reasons, and the curse is lifted via a combination of these strategies.

3.1 A 'Natural History' of Science?

Let's commit to the curse: if philosophers of science appeal to case studies, they had better be as evidence. Given that case studies cannot support general claims, then, we should either stop making general claims, or stop using case studies. Call the first kind of natural historian (the one who thinks we ought not make general claims) 'particularist', and call the other (one who thinks we should abandon case studies) 'baconian'.

One way of being a natural historian of science is to embrace particularism. Recall Lakatos' distinction between rational internal philosophy of science, and irrational external philosophy of science. On the 'particularist' approach, internal philosophy will be restricted to the 'logic' of some scientific debate (or field, or whatever). The best that philosophers of science can do is provide detailed, longitudinal studies of particular scientific events and investigations. Some historically minded philosophers argue forcefully for this.

James Lennox (2001), like myself, denies that '… the history of science is a sort of 'inductive data base' to be used as confirmation for various philosophical views about science (656).' He reacts by restricting the scope of philosophy, taking himself to be drawing on historical cases in order to understand the foundational issues of a particular science in terms of its origins:

> … the activity of understanding foundational problems in biology through a study of the historical origins and development of those problems. 657.

I think this is important: understanding foundational issues in a scientific domain in terms of its history can be a sophisticated and illuminating approach. This kind of natural historian takes scientific debate to be, for all intents and purposes, unique, and provides a highly detailed, contingent explanation of the debate's unfolding[6]. It is undeniable that historical case studies can do this. It is my contention that this is not *all* they do.

[6] In other work (Currie 2014) I have called this kind of explanation a 'complex narrative explantion'. These can be contrasted with 'simple narrative explanations', which account for events by unifying them

The 'baconian' natural historian of science sees herself as a fact-gatherer. Before our claims about science are justified we need a large body of data about scientific practice on the table. The hypotheses we have about scientific practice might be important for how we go about gathering these facts, but until the data are in they are strictly for heuristic purposes. Philosophers of science should not feign hypotheses. To some extent, Larry Laudan could be read as pushing this kind of agenda:

> Nothing resembling the standards of testing that these very authors [post-positivist philosophers of science] insist upon within science has ever been met by any of their theories about science. Those of us who claim some modest expertise in the logic of empirical inference have been notably indifferent about subjecting our own theories to empirical scrutiny, even though our own philosophies of science suggest that without such scrutiny we might well be building castles in the air (Laudan et al 1986, pp 142).

Laudan et al go on to reformulate several 'big picture' views of scientific change into a common language, in order to better test them. Once this cleaning up is achieved, and locations of disagreement identified, bodies of historical cases must be carefully and painstakingly constructed in order to test the theories across their appropriate domains.

Another broadly baconian view is Burian's (2001). His response to the curse is to recommend that we create collections of relevant case studies in order to test and support philosophical claims. For Burian, we can 'group' case studies in two ways. First, we unify them around an effort to understand a particular scientific problem: '… studies that follow the evolution of the problem and of scientists' ways of dealing with it' (387). Second, we can compare different case studies from different contexts: '… such studies need to take account of the multiple settings within which scientific work takes place—theoretical, technical, instrumental, institutional, political, financial, national...' (387). The second route approaches the kind of ahistorical project I sketch later, although I am more optimistic of our capacity to abstract from practice than Burian.

with others: they are treated as event-types. Natural historians of science could be read as denying that simple narrative explanations are available to philosophers of science. I disagree.

And so, accepting the curse offers us two routes: we can go *particularist* and retain the use of case studies but stop making generalizations, or go *baconian*, and lose the case studies in favor of a more methodical examination of science. I should point out that there is nothing wrong with taking either of these two routes: interesting and important work can be done (as I think the last twenty years of work in the history and philosophy of science can attest to). However, I argue that these routes should not be over-emphasized: it is not *all* we should be doing.

I have two complaints against baconian philosophy of science. First, adopting a baconian approach might involve simply changing our attitude, as opposed to our actual method. There's a way of thinking about the baconian project that vindicates the case study methodology. Philosophy of science is an essentially social activity. Even though the vast majority of philosophical papers have a single author, philosophers are not carving truth from the cliff-face of knowledge unaided. Rather, philosophers uncover truth slowly, and socially. Perhaps the curse of the case-study is no curse at all, but rather we have been mistaken in taking individual case studies, and individual papers and books, as stand-alone pieces. Rather, justification comes from a series of works—layers of dialogue—which tests philosophers' claims against each other[7]. The point here is that philosophical positions are not justified *in a* paper, but rather *throughout a debate. If* that is right, then demanding case studies provide inductive evidence within the context of a single paper, argument or book, is a mistake.

The second returns to the patterned heterogeneity thesis: although science is heterogeneous, it has *patchy unity*. This is to say that there are partial generalizations to be had, but when these generalizations hold will be, at best, difficult to ascertain and at worse random. If science is like that, then it seems to me that the kind of methodical data-gathering which the baconian prefers is not obviously better than a kind of 'scatter-shot' method.

[7] There is also an individualist, psychological version of this position. As a philosopher of science, I have developed a tacit understanding of scientific practice through my study of it. The case study has been picked as an exemplar of this tacit understanding.

Here is my general complaint against the particularist natural historian: science is heterogeneous but it isn't *that* heterogeneous! Consider Lennox:

> … the foundations of a particular scientific field are shaped by its history, and to a much greater degree than many of the practitioners of a science realize. There is more conceptual freedom in the way theories—even richly confirmed theories—may be formulated and revised than is usually realized (657).

Even if there is generous wiggle-room in how fields turn out, that doesn't show that there are not patterns in scientific practice: it's just that which sciences exhibit which patterns is not fully determined by their respective epistemic situation. To simplify matters, imagine there are two possible ways that a richly confirmed theory about some phenomenon *x* might be formulated, *a* and *b*. Imagine also that there are two analogous ways that a richly confirmed theory about some other phenomenon *y* could be formulated, *a** and *b**. Even if, let's say, the choice between *a* and *b,* and *a** and *b** is not rational (that is, epistemically speaking, picking one or the other doesn't make a diffe*rence*) it could still turn out that *y*-scientists and *x*-scientists end up picking analogous theories. At which point there is a pattern to be discovered and explained. Moreover, this pattern is partly rational: although epistemic reasons are not the whole story about why theories of *x* and *y* converge, that *a* and *a** (let's say) are both richly confirmed, legitimate theories of their respective phenomena is part of the explanation of their use. That is to say, even though on this picture science is heterogenous, and even though rational response to epistemic situation does not alone explain the convergence, there is still a legitimate 'ahistorical' project which abstracts from those contingent details. Lennox's pessimism is too quick.

I agree that the shape taken by a field is highly context dependent, and for some questions one must take the kind of highly detailed, constrained approach that Lennox recommends. We mustn't prejudge the way science, and philosophy of science, will turn out. In 3.3 I will articulate an ahistorical approach which happily draws from both historical and contemporary cases, but first I will consider the idea that case studies do not play a justificatory role.

3.2 Anti-Realism

Anti-realists deny the first premise, that is, they claim the point of case studies is not to justify philosophical positions. Rather, they play a heuristic role. We appeal to case studies to help structure our thinking, facilitate communication and articulate our commitments.

Joseph Pitt (2001) argues that the use of case studies is either unjustified in the sense discussed in 2.2 (the sample size is insufficient) or that '… if the case is selected because it exemplifies the philosophical point being articulated, then it is not clear that the philosophical claims have been supported, because it could be argued that the historical data was manipulated to fit the point' (373). Either we have an insufficient sample size, or we are cherry-picking. I am less convinced by the latter worry. Recall that philosophy is a social activity and we should be *both* examining the case studies philosophers discuss directly seeing how they go in other contexts. Problematic cherry-picking should be quickly detected. Nonetheless, the curse of the case study can lead to pessimism:

> … even very good case studies do no philosophical work. They are at best heuristics. At worst, they give the false impression that history is on our side… (Pitt, 373)

I will briefly outline three senses in which case studies can play heuristic roles.

They could provide something like a 'psychological model' for philosophical enquiry. That is, they give the philosopher something to 'hang her mind on'. Nersessian (1999) has analyzed conceptual shifts in science in terms of 'mental models': psychological objects which structure scientific thinking. I think case studies playing this role for philosophers is perhaps inevitable, at least reflecting on my own experience. It is difficult for me to call scientific models to mind without considering predator/prey dynamics.

Case studies also play a more social role: they are 'paradigms' insofar as they structure both debate and practice. Understanding the same case studies—sharing the same psychological models—is part of how philosophers of science learn the same skills, explanatory approaches, and so on. That philosophers interested in scientific models are familiar with the same cases (that they are using similar mental models) provides common ground for, and thus facilitates, debate.

Finally, case studies could act as ‘intuition-pumps’, they ‘prime’ the reader for the views which the philosopher pushes. Indeed, in the philosophy of science, drawing intuition pumps from actual scientific cases is surely more convincing, and gives us reassurance that our intuitions are being pumped legitimately.

While case studies undoubtedly play these non-justificatory roles, this is not alone sufficient to justify their central role in our methodology. Here is a shopping list of anti-realism’s dangers.

First, it can narrow debate. If we become focused on a particular scientific example, then we can find ourselves circling unhelpfully in arguments about how to understand that case. Moreover, philosophers are then knowledgeable about only a few scientific cases. This both means they are missing out on further inspiration, and that the debate becomes hopelessly parochial.

Second, antirealism can encourage caricaturing science. Scientific practice is rich, varied and sophisticated. Typically, having a few mental models of how testing works, which include, say, how Newton’s early optical experiments went, leads to an overly simple model of scientific testing. Scientists will test theories using natural experiments, abstract simulations, enormous data sets, not just ‘crucial experiments’[8]. If anti-realism encourages conservatism in the application of case studies, then it is likely that it will also encourage caricatures.

Third, and relatedly, anti-realism about case studies can encourage what Ladyman & Ross (2007) have delightfully called *domesticated science*. As opposed to understanding cutting-edge science, which is difficult and sometimes inconvenient, caricatured ‘pet’ scientific cases can be

[8] Naturally, this is a danger for the realist as well, but it is more pressing for the anti-realist.

wielded as rhetorical bludgeons. This leaves the philosopher dangerously out of touch with actual science.

Fourth, if the purpose of case studies is to structure debate, then we should select case studies based on their potential for that, not about whether they justify our claims. If the purpose is rhetorical, then it seems like cherry picking is the right thing to do: we should locate the case study which makes our argument look the most forceful or convincing. We should be able to distinguish between legitimate and illegitimate cherry-picking. We might pick a case because it makes some distinction starkly, or one which is particularly representative—this seems fine; whereas cherry picking which misrepresents scientific practice, or non-rationally manipulates the reader is undesirable. On the face of it, anti-realism lacks the resources to distinguish kosher from illegitimate cherry-picking.

Fifth and finally, if the purpose of case studies is purely heuristic, then it is no longer obvious why we should bother getting the case studies *right*. That is, who cares about the actual details of scientific practice if case studies are dialectical props? I think the vast majority of philosophers of science genuinely care about reaching the truth about science, and work hard at mastering complex scientific ideas. Reading case studies in merely heuristic terms does not do justice to this.

All this is not to suggest that case studies cannot, and ought not, play these kinds of roles. However, that had better not be the end of the story. As I stated in the introduction, my aim here is to provide a multi-faceted defense of the use of case studies in the philosophy of science, and surely heuristic uses play their part.

3.3 Illustrating Conceptual Tools[9]

[9] The term 'conceptual engineering' has appeared in an epistemological context in relation to *constructionism*, that is, the view that an agent has knowledge when they are able to construct a

The second premise of the argument claimed that case studies could not justify our claims. This assumed that if case studies justify, they do so as data-points for ampliative inference. Some of our claims, the descriptive ones, should certainly be read in these terms. However, epistemic justification is not exhausted by ampliative inference. Knowledge is not merely a matter of unified description: we also want to explain, understand and illuminate scientific practice and reasoning. Moreover, we want our work to be fruitful, to extend both philosophy and science in surprising ways, and to be normative: the best philosophy of science can guide us in both theory and practice. The dichotomy between natural history and anti-realism is false, as there are more routes to epistemic justification than the curse allows. I will argue for such a non-inductive epistemic justification for case study methodology. The idea is this: philosophers of science are in the business of explaining science, and success in this grants us epistemic licence for belief. At least some of the time, a case study's role is to demonstrate the potential of some piece of conceptual machinery to explain scientific practice (philosophy-directed) or unify or clarify scientific work (science-directed). As we shall see, this picture opens the door to what I will call an 'ahistorical' philosophy of science.

This involves neither anti-realism nor taking case studies as data points for inductive inferences. Rather, we should take case studies as being *illustrative of conceptual tools* which are used to explain, critique, and solve scientific problems and practice. The thought is that the proof of such tools is in the pudding, and case studies play two roles: first, an *indication* of the tool's use; its success in one case is a good start. Second, and more importantly, it plays a pedagogical role in helping us see the *potential* for the machinery in question. Indeed, a good case study can illustrate a concept or point much more effectively than the abstract details alone. I will draw a capricious analogy.

representation of it, and apply it to the relevant information networks (see Floriadi 2011). I don't think we need to commit to this epistemic claim in order to drive the point that *conceptual engineering*, in addition to conceptual analysis, is an important part of the philosopher's aim.

Consider traditional lego sets. These consisted of a collection of various interlocking blocks of different sizes which could be used to build various structures, the only restrictions being the number of blocks and the user's imagination. The sets came with instructions: guides on how to build a few structures, say, a house. In a step-by-step way, one could follow these instructions to build a house with your lego set. The primary purpose of the instructions was not to show the child how to build a house. They were intended to show the child both how to use lego, and its potential, via building the house. The function of lego sets is to facilitate imaginative construction, and the function of the house-building instructions was to help build practical skills and demonstrate the kind of things these blocks could do.

Case studies play a role analogous to lego instructions. Assuming it is true that Craver's story about mechanistic explanation captures action-potential in neurons, the work the case study is doing is *not* primarily to provide evidence that the conceptual machinery applies to science generally. Rather, it is to both illustrate how the machinery operates and demonstrate its potential. The function of lego is to facilitate creative play, and its success in doing this is facilitated by a good set of lego instructions. The function of conceptual engineering in philosophy of science is to explain, articulate and unify scientific concepts, and case studies do indeed provide a single example of this, but more importantly they demonstrate the potential of the machinery to do that work. They therefore play a role in justifying the conceptual work in question[10].

This view is in some ways akin to Lakatos' (1970) and distinguishing us is helpful. Lakatos' aim is to understand the relationship between the history and philosophy of science. The picture, roughly, is this: the external (that is, detailed and complex) history of science is *made sense of* by the internal (that is, abstracted and normative) philosophy of science. In short, we apply some

[10] Of course, the lego analogy is not perfect. There is a big difference in aim: science targets truth (or, if you must, empirical adequacy), while lego aims for enjoyment, developmental enrichment, and commercial success. In virtue of this, our attitudes towards the two are quite different, I might *believe* scientific claims, but lego doesn't really make 'claims' as such.

logic of scientific methodology to rationally reconstruct an aspect of scientific history. The external picture explains the growth of scientific knowledge, while the internal picture plays dual roles as explanandum and raw data. My picture differs in three crucial respects. It is ahistorical, it is unambitious, and it has a wider sense of the philosopher's job.

I think that conceptual engineering, with a fairly minimal claim about patterns of scientific reasoning, grounds an 'ahistorical' project in the philosophy of science. Recall the patterned heterogeneity thesis, that is, different aspects of different scientific work will be unified to greater or lesser extents. If that is right, then we might be able to abstract from the contingent details of sciences to draw general lessons. For instance, I have argued that mechanistic explanation is illuminative of some work in the historical sciences (geology or paleontology say), and not in others (Currie 2014). Sometimes, historical scientists' explanations of past events unify the target with other events through some common mechanism. For instance, the (rejected) 'Nemesis' hypothesis explains the apparent periodicity of mass extinctions in the fossil record as the result of a cosmological mechanism (see Bailey 1984). The sun, according to this hypothesis, has a twin star 'Nemesis'. Every so often, Nemesis' orbital path takes it through the Oort cloud, sending vast quantities of debris towards our solar system. This leads to periodic increases in extra-terrestrial impacts on earth, and thus mass extinctions. This explanation is, I argue, broadly mechanistic: mass extinctions are explained as a *type* of event unified by a common cause. The system is thought of in terms of components (the earth, nemesis, the oort cloud, etc…) and its behaviour (mass extinctions) is explained in terms of the causal relationships between those components. Mechanistic explanation, then, helps us understand some work in historical science. However, not all work. Sometimes past targets are too disjoint, too complex, and too contingent to be captured in mechanistic terms. Here, historical scientists treat their target as if it is unique, and do not unify it as a type of mechanism (I suggest this is true of explanations of sauropod gigantism). And so, Craver's work is indeed illuminative of some aspects of historical science, but not all of it: its application is *patchy*.

Ian Hacking drew inspiration from Crombie's 'styles' of scientific reasoning (Crombie 1981). 'Styles' are different ways in which scientists go about justifying themselves (Hacking 1992). One style, for instance, is '…the deployment of experimental techniques both to control postulation and to explore by observation and measurement', another the '… hypothetical construction of analogical models' (Hacking 4). Although the styles of reasoning have a history, they may be described abstractly and, in my terms, they apply 'patchily'. Paleontologists and geologists, for instance, have been characterized as paradigmatically following a 'non-experimental' mode of reasoning (most clearly by Carol Cleland 2002, 2011), however in some circumstances they are experimentalists: 'The paleontologist uses experimental methods to carbon date and order the old bones' (Hacking, 5—also see Jeffares 2008).

The possibility of science being 'patchily unified' opens the door to an ahistorical philosophy of science driven by case-study methodology. *Contra* Lakatos, (and other natural historians of science) we need not tie our conceptual work to particular debates, or particular histories, but rather we can abstract the relevant aspects for explanatory purposes. This is what I mean by 'ahistorical'. I have given two examples of this: one is the structure of explanations, the other is styles of reasoning. I suspect there are many more.

For Lakatos, philosophers of science are largely in the business of explaining how scientists generate knowledge. I think this is indeed a very important part of the business of philosophy of science, but it is certainly not the whole picture. We can also engineer concepts for scientific purposes. Here, the philosopher of science is actively involved in the generation of scientific knowledge via unifying and discriminating between scientific concepts. For instance, many philosophers of biology have concerned themselves with the unity or otherwise of species-concepts. Part of the concern here is philosophy-directed: the ontological question of *whether there are species or not*. But another aspect is science-directed: ought scientists use *the same* concept of species, is concept pluralism acceptable for scientific practice, and so on. By engineering species-concepts, and answering these methodological questions, philosophers are not merely explaining science, they are active participants (if rather abstract ones).

Finally, the ‘logics of scientific methodology’ which Lakatos considers are *unitary frameworks* for understanding science. My approach is both more flexible and less ambitious. It is less ambitious because the philosopher does not, in engineering concepts, apply some ‘one size fits all’ model of scientific practice. Rather, they specify some conceptual machinery which explains some scientific work in certain contexts, or unifies some scientific concepts. It is more flexible because the success of these conceptual tools don’t turn on unifying or explaining large swathes of scientific practice, but rather patchily distributed practices[11].

The applicability of this defence turns on the heterogeneity of science. We shouldn’t prejudge how unified science is or ought to be, and presumably some practices will be common, while others will be rare. For an ahistorical approach to work, science needs to be at least patchily unified. If it is not, then it fails. For instance, if it is true that science can only be understood in terms of historical context, that there are no meaningful continuities between fields across time, then conceptual engineering is not applicable outside of the case study in question. I find this picture unlikely generally, but there may be some aspects of scientific practice where we must be so restricted.

4. Conclusion

The curse of the case study is not broken in a unified manner, but rather via a plurality of applications. In a sense, overcoming the curse has provided a backdrop for discussing a set of uses which philosophers of science put case studies to. These have included:

[11] Chang (2012), again in an historical context, has argued that we can avoid reading the use of case studies inductively by thinking of case studies (the history) in terms of being ‘concrete’ and our conclusions (the philosophy) being ‘abstract’. On his view, we require abstract philosophical machinery to do history in the first place, and must construct such machinery if some does not already exist. And so doing history can lead to new philosophical concepts. I’m on board with Chang’s suggestion, however, my account is not restricted to historiography.

- Inductive testing of general claims;
- 'Restricted' natural history of science: building a picture of a particular scientific field in terms of its foundational history;
- 'Baconian' natural history of science: fact gathering about scientific practice;
- Providing 'mental models' to facilitate the philosopher's thinking;
- Providing 'paradigms' which facilitate debate by providing common ground;
- Rhetorical intuition pumps;
- Illustrating patchily-applicable conceptual tools.

These are not mutually exclusive: indeed, much of the time a case study might be playing a multitude of roles.

One important upshot from this discussion, to my mind, is the defence of an ahistorical philosophy of science. Philosophers of science are right to emphasize the heterogeneity of scientific practice, but this only goes so far: heterogeneity only undermines partial unity if the heterogeneity is unpatterned. The ahistorical project involves engineering conceptual tools which are widely, if patchily, applicable in explaining and unifying scientific practice and concepts. I don't see this as being in competition with more historical projects: both have something to add to our understanding of science. Rather, it is wrong to think that *all* we can do is natural history of science. We can also be conceptual engineers.

I have also emphasized the *social* aspects of philosophy. Like science itself, philosophy of science is primarily a cooperative group activity (despite appearances). The function which case studies are playing is important here. If they are heuristically structuring a debate, providing a paradigm and thus common ground, then sticking to the same case can be beneficial[12]. However, if we are in the process of applying conceptual tools, or testing inductive generalizations, relying

[12] Indeed, this is Levy & Currie (2014)'s reason to use the Lokta-Volterra model to illustrate scientific models.

on too small a subset of cases can retard progress. Overall, it seems to me that the picture I have sketched ought to encourage philosophers of science to cast a very wide net indeed when searching for case studies. In many contexts, rehashing the same old examples will not do.

And so, the curse of the case study is no curse after all. Philosophers of science are justified in using detailed single examples to drive their conceptual work.

Bibliography

Bailey, M (1984). Astronomy: Nemesis for Nemesis? Nature 311, 602-603

Baumslag, D (2000). How to test normative theories of science. Journal for General Philosophy of Science 31 (2):267-275.

Burian, R (2001). The Dilemma of Case Studies Resolved: The Virtues of Using Case Studies in the History and Philosophy of Science. Perspectives on Science, Vol. 9, No. 4 pp 383-404

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

Chang, H. 2011. Beyond case-studies: History as philosophy. In: S. Mauskopf and T. Schmaltz (eds.), Integrating History and Philosophy of Science: Problems and Prospects (pp. 109-124). Dordrecht: Springer.

Clayton, N & Dickinson, A. Episode-like memory during cache recovery by scrub jays. Nature 395, 272-274

Cleland, C. E. (2011). Prediction and explanation in historical natural science. The British Journal for the Philosophy of Science, 62, 551–582.

Cleland, C. E. (2002). Methodological and epistemic differences between historical science and experimental science. Philosophy of Science 69 (3):447-451.

Craver, C (2007). Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience. Oxford University Press, Clarendon Press

Crombie, A (1981). Philosophical perspectives and shifting interpretations of Galileo. IN: Theory change, ancient axiomatics and Galileo's methodology: proceedings of the 1978 Pisa conference on the history and philosophy of science

Currie, A (2014) Narratives, mechanisms and progress in historical science. Synthese Volume 191, Issue 6, pp 1163-1183

Dally, J. Emery, N. Clayton, N (2007) Social cognition by food-caching corvids. The western scrub-jay as a natural psychologist Phil. Trans. R. Soc. B (2007) 362, 507–522

Devitt, M (2013). The 'Linguistic Conception' of Grammars. Filozofia Nauki 2.

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

Godfrey-Smith, Peter (2009). Darwinian Populations and Natural Selection. OUP Oxford.

Jeffares, B (2008). Testing times: Regularities in the historical sciences. Studies in History and Philosophy of Science Part C 39 (4):469-475.

Ladyman, J & Ross, D (2007). Every Thing Must Go: Metaphysics Naturalized. Oxford University Press.

Lakatos, I (1970). History of Science and Its Rational Reconstructions. PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association

Laudan, L. Donovan, A. Laudan, R. Barker, P. Brown, H. Leplin, J Thagard, P. Wykstra, S. (1986). Scientific change: Philosophical models and historical research. Synthese 69 (2):141 - 223.

Lennox, J (2001). History and Philosophy of Science: a Phylogenetic approach. Historia, Ciencias, Saude—Manguinhos VIII(3): 655-69

Levy, A & Currie, A (2014). Model Organisms are Not (Theoretical) Models. Br J Philos Sci (online first) doi: 10.1093/bjps/axt055

Luciano Floridi (2011). A defence of constructionism: Philosophy as conceptual engineering. Metaphilosophy 42 (3):282-304.

Matthewson, J & Calcott, B (2011). Mechanistic models of population-level phenomena. Biology and Philosophy 26 (5):737-756.

Millikan, R (2005). Language: A Biological Model. Oxford: Clarendon Press.

Nersessian N (1999) Model-based reasoning in conceptual change. In: Magani L, Nersessian N,Thagard P (eds) Model-based reasoning in scientific discovery. Kluwer/Plenum, New York, pp 5–22

Pitt, J (2001). The dilemma of case studies: Toward a heraclitian philosophy of science. Perspectives on Science 9 (4):373-382.

Skyrms, B (2008). Signals. Philosophy of Science 75 (5).

Strevens, M (2003). Bigger than Chaos: Understanding Complexity through Probability. Harvard University Press.

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

Weisberg, M & Reisman, K (2008). The Robust Volterra Principle. Philosophy of Science 75 (1):106-131.

Weisberg, M (2007). Who is a Modeler? British Journal for the Philosophy of Science 58 (2):207 - 233.

Wimsatt, W (2007). Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality. Harvard University Press.

Wylie, A (1999). Rethinking unity as a ‘working hypothesis’ for philosophy: How archaeologists exploit the disunities of science. Perspectives on Science 7 (3):293-317.