

# Justifying the method of historical case studies: a phylogenetic approach

## Abstract

Philosophers use historical case studies to support general claims about science. Such inferences are *prima facie* problematic and are widely criticized. In this paper we argue that inferences based on case studies are no more problematic than phylogenetic inferences from model organisms in biology.

## 1 Introduction

The use of historical case studies in the philosophy of science is ubiquitous: they can be found in classical debates concerning theory appraisal, scientific realism, explanation, and many more. Crucially, case studies are usually taken to be representative of a broader class of cases. For example, Lakatos took his discussion of development of the Bohr model of the atom to be representative of research programmes *in general*. When Kuhn argued that the meaning of the term ‘mass’ is different in Newtonian and relativistic mechanics, Kuhn believed to have lent support to the thesis that paradigms *in general* are incommensurable. Scientific realists have argued on the basis of historical cases like the Fresnel wave theory of light that realism as a *general* position about science is sustainable despite radical theory change, and not just as a position about particular cases.<sup>1</sup> Whether the Hodgkin-Huxley model is explanatory is believed to have implications not only for this particular model, but for explanations in biology *in general*.<sup>2</sup>

Despite its widespread use in the philosophy of science, the case study approach has long been criticized for its perceived overreach. In his famous “marriage of convenience” paper on the difficult relationship of the history and the philosophy of science, Giere (1973) criticized the case study approach as being “without a

---

<sup>1</sup> See Worrall (1989). There are also realist of a particularist bent (see Magnus and Callender (2004)). For lists of historical cases that have been discussed in the realism debate see (Laudan 1981, Vickers 2013).

<sup>2</sup> See (Weber 2008), (Craver 2008) and (Levy 2013).

conceptually coherent programme” because it did not address the question of how “philosophical conclusions may be supported by historical facts” (292). Later writers have shared this skepticism. Nickles (1995) summarized a widespread sentiment that “historical case studies can be too much like the Bible in the respect that if one looks long and hard enough, one can find an isolated instance that confirms or disconfirms almost any claim” (141). Similarly, Pitt (2001) worried that “it is unreasonable to generalize from one case or even two or three” and even believed that case studies run the risk of being “manipulated to fit the point” (373). Subsequent contributions have continued to address these concerns, among them Burian (2001), Chang (2011), Schickore (2011), Schindler (2013)), and Scholl and Rätz (2016). Yet even as the use of case studies in the philosophical literature is booming, there is no consensus view on why they permit general conclusions. In that respect, at least, the history-philosophy-relationship remains as opaque as it was when Giere wrote in the early 1970s.

Here we defend the case study approach against skepticism by drawing analogies to a well-established practice in biology: the use of model organisms. We submit that the extrapolation from individual case studies to broader philosophical claims can be understood in the same way as the extrapolation from model organisms to broader biological claims. In the philosophy of biology, such extrapolation is usually taken to be warranted by phylogeny: the fact that populations and species are related by descent can ground our extrapolations in facts about similarity. We will outline an analogous phylogenetic justification for extrapolation from historical case studies. This will reveal a number of deep similarities between the biological and the philosophical practices.

In Section 2 we review the philosophical literature on model organisms with an emphasis on their epistemic role. In Section 3 we spell out the analogies between case studies and model organisms concretely, by reference to a widely-used case study in the history and philosophy of science: Semmelweis’ investigation of the cause of childbed fever. In Section 4, we articulate our phylogenetic approach to the justification of case studies in more detail, and we defend the view against criticisms. Section 5 concludes our discussion.

## 2 Model organisms and case studies

Biological model organisms are intriguing scientific objects: they promise inferences from a very limited set of instances to an indefinite one. As Ankeny and Leonelli (2011) note, “model organisms are always taken to represent a larger group of organisms beyond themselves” (318). *Drosophila melanogaster*, for example, was used extensively in the modern synthesis in the early twentieth century to enable inferences to higher level organisms such as moths, pigeons, cats, silkworms, rabbits and even humans (Levy and Currie 2014). It is striking just how limited the number of model organisms is in biology. As Weber (2004) put it felicitously in his chapter on model organisms in his *Philosophy of Experimental Biology*, “molecular biology laboratories are extremely impoverished in biodiversity” because “most laboratories work on only a single species, and a large number of laboratories work on the same species” (155).

The use of model organisms defines the landscape of experimental biology. Perhaps most famously, Thomas H. Morgan and his research group laid the foundation of modern genetics with their experiments on *Drosophila melanogaster* in the 1920s. But both before and after Morgan, many other organisms were established as models in particular fields and for particular research questions: the sea urchin in early developmental biology, *E. coli* in the study of bacterial conjugation, squid for the study of nerve cells, rats for the study of metabolic pathways, mice for the study of the immune system, baker’s yeast in the study of eukaryotic cells, *C. elegans* for the study the molecular basis for behavior and development, and *Arabidopsis thaliana* for the study of plants (154).

Weber suggests three related questions about this seemingly peculiar practice of the use of model organisms: (i) why do biologists choose particular species as their model organisms?, (ii) why do biologists keep using the same model organisms instead of diversifying their induction base?, and (iii) how is it possible to extrapolate from model organisms to other organisms such as humans?

With regard to question of why biologists choose model organisms, Weber suggests that pragmatic reasons play an important role in addition to epistemic ones. For example, the organism must be easy to breed in the laboratory, its generation

time must be short, and its features must be suitable for specific research questions (e.g., the size of the squid giant axon, or the size of chromosomes in *Drosophila*'s larval salivary glands, 176ff.). With regard to the question of why biologists return to a limited set of models, Weber argues that standardization has a positive cumulative effect. Once experimental techniques and procedures have been developed, it is reasonable not to shift to different organisms where the known experimental techniques might not work as well and where new techniques might have to be developed (175f.). A closely related advantage of returning to known model organisms is that this makes it easier to reproduce results.

With regard to the question of extrapolation, Weber points out that inferences from model organisms to other organisms are grounded in phylogeny, that is, in their evolutionary history and in their shared genetic code (180f.). An inference from a model organism (such as fruit flies) to a target organism (such as humans) is thus justified because both the model organism and the target organism share a common ancestor, which in turn possesses features which they both share. Levy and Currie (2014 333) develop the phylogenetic grounding of model organisms in more detail. In particular, they distinguish theoretical modeling from "empirical extrapolations" involving model organisms and argue that whereas in theoretical modelling one must always check whether the target is actually similar in the relevant aspects to the model in order for the model-inferences to be justified, in empirical extrapolations involving model organisms „the relatedness of the lineages licenses inferring from one to another, *without the need to explicitly compare the underlying traits*" (330, our emphasis). Inferences on the basis model organisms, according to them, are thus justified phylogenetically.

Levy and Currie also point out that model organisms undergo modification and genetic standardization in the laboratory with the aim of increasing reproducibility and comparability (333), and, presumably, of lending greater stability to inductive inferences. There are arguably various other non-epistemic functions of model organisms (Levy and Currie 2014 333). Our interest here, however, lies squarely with their epistemic function.

We believe that the use of historical case studies in the philosophy of science can be understood along the lines laid out by Weber for model organisms in biology. That is, we believe that (i) historical case studies are selected in part for pragmatic reasons such as simplicity and comprehensibility, (ii) they are used repeatedly in part because it is efficient for philosophers to re-use the relevant resources, and (iii) case studies allow inductive inferences to broader classes of cases because each case is embedded in scientific research traditions which relate it to other cases. Therefore, we can expect the practices studied in one case to be representative of the practices in other cases. In other words, inferences from historical case studies are (at least to some extent) justified by their phylogenetic relationships.

In the next section, we will articulate these notions by considering in detail Ignaz Semmelweis' discovery of the cause of puerperal fever between 1844 and 1848. We will structure our discussion along the three questions raised by Weber about model organisms.

### **3 Choosing, stabilizing and learning from a case study: Semmelweis on puerperal fever**

Semmelweis' discovery of the cause of puerperal fever began its life as a philosophical case study in Hempel's *The Philosophy of Natural Science* (1966). Hempel explained that Semmelweis, a physician working in Vienna around the middle of the nineteenth century, was motivated by a puzzle. The mortality rate of so-called childbed fever differed markedly between two divisions of the same maternity clinic: in the first division, the mortality rate was near 10%, but in the second division it was comparatively low at 3%. On Hempel's account, Semmelweis demonstrated the cause of the difference using the hypothetico-deductive method. He framed a number of hypotheses to explain the difference: He suspected differences in weather conditions, hospital crowding, birthing positions, and examination techniques, among others. But he found that each hypothesis yielded false predictions. Eventually Semmelweis hit upon a more successful hypothesis. The first division was run by physicians, who conducted autopsies before examining pregnant patients, while the second division was run by midwives, who performed

no autopsies. Semmelweis surmised that the physicians transferred some kind of infectious matter from autopsies to patients. This hypothesis yielded correct predictions since the institution of thorough hand-washing measures reduced the mortality in the first division to levels below those of the second. On the hypothetico-deductive reconstruction, we can understand this as cycles of conjecture and refutation followed by an eventual confirmation.

In the succeeding decades, the Semmelweis case was revisited by numerous historians and philosophers of science. The most extended treatment was by Peter Lipton (1991/2004), who used Semmelweis as a case study in his seminal *Inference to the Best Explanation*. Lipton aimed to show in detail how the hypothetico-deductive account fails, while inference to the best explanation succeeds, at capturing and justifying Semmelweis' actual scientific reasoning. Another important contribution was by Donald Gillies (2005), who studied Kuhnian factors in the case in order to explain why Semmelweis' findings were initially rejected. Alexander Bird (2010) argued that Semmelweis' reasoning should be understood as an instance of inference to the *only* explanation. Later writers continued to find new aspects in Semmelweis' reasoning. Scholl (2013, 2015) found Semmelweis' inferences to correspond closely to J. S. Mill's (1843) methods of experimental inquiry and argued that an explanationist framework was not necessary to recover Semmelweis' inferences. Tulodziecki (2013) argued that Semmelweis' reasoning was often careless and should not be held up as a paradigm of scientific inference at all.

We can profitably ask the same three questions about the Semmelweis case that Weber asked about model organisms: Why was the case considered suitable in the first place? Why have philosophers and historians of science returned to it repeatedly instead of expanding their empirical basis? And why do they think that concepts that are useful for understanding the Semmelweis case speak to the question of scientific discovery and confirmation more broadly?

### 3.1 Why are particular episodes chosen as case studies?

We suggest that case studies, like biological model organisms, are chosen in part for pragmatic and in part for epistemic reasons. *Drosophila*, for example, is a preferred model organism because it is easy to breed, has a short life cycle, and is rich in

phenotypically traceable mutants (Weber, 2004, 177). Likewise, philosophers select certain historical case studies because they are straightforward to present and to understand (we may say that they offer “cognitive ease”), and because they are rich in philosophically informative detail. Hempel, for example, chose the Semmelweis case in part because it is “a simple illustration of some important aspects of scientific inquiry” (1966, 3).

Cognitive ease, however, comes with tradeoffs: cases that are easy to present and understand are not necessarily typical of many aspects of science—just as the plant *Arabidopsis thaliana* is relatively quick and easy to breed but perhaps not representative of many aspects of the long-lived *Sequoia sempervirens*. But atypicality need not be an obstacle to inductive reasoning. Some atypical traits are unrelated to the traits under investigation. A high breeding rate need not distort Mendelian ratios or impede the production of chromosomal maps. Even when atypical traits themselves are investigated, this need not be a problem for extrapolation. For example, the discovery of the mechanisms of action potential propagation was enabled by the giant squid axon, which, as its name suggests, is atypically large. But circumstantial evidence indicated early on that that the same mechanisms are shared by more typically sized nerve cells (Levy and Currie, 334f.). The situation is similar for the Semmelweis case. Some of its atypical features are irrelevant to the philosophical research questions it has served to illuminate. An example of this is the case’s eventual co-option in the service of Hungarian nationalism: This is atypical in the sense that few medical discoveries were co-opted in this way, but it is also irrelevant to the practices of discovery and confirmation in mid-nineteenth-century medicine. Other atypical features of the case, by contrast, are relevant to our philosophical concerns. An example of this is the striking difference in mortality between two distinct hospital wards. Such ready-made research questions are both atypical and relevant to philosophical interests such as discovery and confirmation. But does this aspect of the case distort our results or impede extrapolation? We would argue that it is only an accentuated version of the kinds of unexplained contrasts that often drive research. It is just that, more typically, such contrasts occur in mixed populations rather than pre-sorted into hospital divisions. Thus, in spite of

its atypical features, the Semmelweis case may be highly representative of a broad range of cases of successful scientific reasoning.

Another requirement for a good model organism is that it must present insightful variation (Weber, 2004, 177). The use of *Drosophila* for genetic analysis, for example, was helped by the fact that its populations included mutants with phenotypic effects that were reasonably easy to discern. Similarly, the Semmelweis case contains philosophically insightful variation. It instructs us about both discovery *and* justification, experiment *and* observation, refutation *and* confirmation, and so on. For instance, on Hempel's reconstruction the Semmelweis case shows that narrow inductivism is false (because discovery follows no rules), that some hypotheses are refuted by observation and some by experiment (but that this is logically the same), and that even successful theories are fallible (because of the fallacy of affirming the consequent). Lipton, by contrast, stresses that the Semmelweis case gives us a handle on the nature of hypothesis generation (since we search for explanatory differences between contrasting groups) and that it elucidates why some explanatory failures speak against a theory while others are simply irrelevant to it (it depends on the relative strengths of competing hypotheses). The Semmelweis case is thus suitable in part because of its richness. It allows a wide range of philosophically interesting questions to be raised and examined.

In sum, the choice of a historical case study is analogous to the choice of model organisms: the goals are fundamentally epistemic, but the choice of models is partly pragmatic.

### 3.2 Why are case studies used repeatedly?

We saw in section 2 that there are at least two good reasons for reusing the same model organism instead of expanding our empirical basis: standardization and reproducibility. Once laboratory techniques have been adapted to a particular model organism and have become productive, it is costly to switch to a different experimental system. What is more, switching may make it harder to compare, contrast and integrate results from different studies. If our goal is to study a particular system in depth, it makes no sense to switch exemplars.

The reuse of historical case studies relies on similar considerations. Once editions of source documents and other secondary works have been written, it is efficient to keep studying the now well-documented episodes. Starting anew requires disproportionate effort. But even if resources were unlimited, reusing case studies would have the advantage of allowing our analyses to build on each other in a cumulative manner. We will see below that there is much to be learned by comparing different reconstructions of the *same* episode.

The Semmelweis case illustrates the first advantages of the repeated use of a single case study: the relevant sources and background materials are easily available and often edited, so that research on conceptual questions can proceed from a rich foundation. By the time of Lipton's (1991/2004) use of the Semmelweis case as an extended study of inference to the best explanation, Semmelweis' main work, the *Etiology, Concept and Prophylaxis of Childbed Fever*, had already appeared in a new and accessible English translation by K. Codell Carter (Semmelweis 1983). The translation had been written expressly in order to facilitate philosophical and historical study of the case, particularly in the context of introductory courses in the philosophy of science (see Carter's introduction to the translation). Most writers from then on used the new translation: both the "Kuhnian" take on Semmelweis by Gillies (2005) and the "Holmesian" take by Bird (2010) rely on it. Carter also provided further historical material on Semmelweis, his work, and his predecessors, which proved invaluable for the continuing study and reassessment of Semmelweis' reasoning. Thus, a fair amount of research on the historical sources and the context of the Semmelweis case contributed to its standardization and further use as a case study.

However, standardization does not come without risks. Carter's translation of Semmelweis' *Etiology* made editorial choices that reflected Hempelian preconceptions about Semmelweis' goals and methods (Scholl 2013). For instance, many pages of numerical tables are left out of the translation because they appear repetitious from the hypothetico-deductive point of view. In truth, however, the numerical tables attest to Semmelweis' use of methods akin to Mill's methods of agreement and concomitant variation. Without these tables, the methodological core

of the work is obscured. Similarly, Carter omitted an account of animal experiments, which do not feature in the hypothetico-deductive account. Yet Semmelweis' contemporaries considered these animal experiments to be among the clinching evidence for Semmelweis' case. Such omissions in secondary works can obviously reinforce existing biases. Thus, standardization produces efficiencies, but not without risk. Luckily, oversights resulting from standardization are amenable to correction.

### 3.3 How do we learn from individual case studies?

We now proceed to what is perhaps the most interesting and most puzzling question concerning case studies: how can single cases enable extrapolation to a broader group of cases? In Section 3 we saw that inferences from model organisms to a broader group of organisms are justified phylogenetically. Insights about mechanisms in *Drosophilidae* may be taken to reflect mechanisms in *Elephantidae*, at least in part: even though the two species belong to distinct genera, some of their biological mechanisms will be shared because of common descent.

What is the equivalent of phylogeny for case studies? We suggest that it is historical influence: any episode which we isolate in the form of a case study bears relations to research practices and traditions before and after. Researchers learn from each other. They take up ideas from their colleagues and predecessors, develop them, modify them, and pass them on. Even innovative findings rest on such a foundation. Semmelweis' discovery may have been a breakthrough for our understanding of infectious diseases, but its methodology is continuous with earlier work. Creating contrasts with control groups, excluding confounders: these are concerns we find in earlier clinical research, for instance by James Lind in Britain or P. C. A. Louis in France. Although many of the details of the methods for causal inference were in flux, these researchers contributed to a shared methodological tradition. Semmelweis is interesting to us in part *because* he is representative of this methodological tradition, as one link in a long chain. Today's randomized controlled trials are the distant offspring of the tradition in which Semmelweis worked. Thus, the lines of influence between researchers provide a basis for extrapolation from case

studies: because of their historical connections, studying one case can be expected to teach us something about others.

The notion of historical connections between cases should be understood broadly. Beyond the classical lines of influence between major figures that historians of science used to be obsessed with, broader methodological currents can be discerned, of which individuals philosophers and scientists are representatives. For example, on the reconstruction of Scholl (2013), Semmelweis used Mill's (1884) methods of experimental inquiry. But there is no claim that Semmelweis actually read Mill, since the same methodological ideas may have come to him by different paths – for instance, by way of his Viennese teachers and their connection, in turn, to Parisian exponents of the numerical method in the early nineteenth century. The Parisian physicians, of course, did their research decades before the writings of Mill, whom we need not understand as the inventor nor even as an indispensable contributor to that branch of methodological thought. Mill merely developed earlier proposals that attempted to capture principles of scientific reasoning – principles which were probably quite commonplace among practicing scientists. Both Mill and Semmelweis must be understood as representatives of, rather than indispensable links in, a methodological tradition that extends far and wide in the history of science. We believe that similar broad currents of thought characterize many facets of science. Scientists learn from each other, directly and indirectly, what it takes to explain, to measure, to intervene. Such widespread methodological principles may be less tractable historically than research traditions and scholarly influence *sensu stricto*, but they exist and are important for the justification of extrapolations from case studies.

#### **4 The phylogenetic approach to case studies: articulation and defense**

##### 4.1 Recapitulation

We have argued that historical case studies in the philosophy of science work in close analogy to model organisms in biology. The ultimate justification for extrapolation from model organisms or cases is a phylogenetic relationship between them: since they derive from common historical sources, one can be expected to

represent the other at least to some extent. In biology, this relationship is descent with modification: we can expect aspects of model organisms to match aspects of related species because they share common ancestors. In the case of history and philosophy of science, there is a comparable phylogenetic relationship. It is grounded in research traditions that transmit forms of practice between scientists: Individual scientists rarely invent out of thin air the standards by which theories are assessed, by which experiments are conducted, or by which explanations are constructed. The more usual situation is for such standards and practices to be transmitted by various routes from scientists to scientist. Just as in the biological case, there is thus good reason to assume – at least as a starting point of our investigation – that a detailed study of practices in one case will be transferable to a greater or lesser extent to other cases.<sup>3</sup>

Neither Weber nor Levy and Currie specify the *extent* to which inferences from model organisms to a larger class of organisms are justified by phylogeny. Our view is that the inductive support provided by phylogenetic relationships is already reasonably strong, but that it can be improved decisively by a comparison between the model organisms and the target organisms. Likewise, extrapolations from case studies to other episodes are reasonably strong on phylogenetic grounds, even when the decisive examination of the target episodes is impossible for lack of time or resources.

Beyond the question of inductive support, even skeptics about the use of historical case studies should admit that cases minimally constitute something like existence proofs: they show that a piece of philosophical theorizing *actually* corresponds to scientific practice. Case studies thus ensure that our philosophy of science is one of actual science, and not one merely concocted in the comfort of the armchair. Case studies can show that particular processes and concepts are

---

<sup>3</sup> We should note that although Lennox (2001) too speaks of a „phylogenetic approach“ to history and philosophy of science, Lennox pursues an altogether different project. In contrast to us, Lennox is not interested in and rejects viewing the history of science / philosophy of science relationship as one of epistemic support of philosophical claims by historical evidence. Secondly, Lennox describes his approach as “understanding foundational problems in biology through a study of the historical origins and development of those problems” (657), which strikes us as purely historical approach.

operational in at least some episodes of the history of science, laying the groundwork for their application and extension to other cases.

#### 4.2 Efficiencies and blind spots

Construing the methodology of case studies in this way gives insight into their efficiencies. Once an historical episode has been studied in depth, we possess editions of published and often unpublished primary works as well as secondary literature on the details of the episode. Scholars can thus hone in on philosophically salient aspects of the episode from an advantageous position. What is more, once a case is reasonably well known in a community, scholars who refer to it enjoy the benefit of the cognitive ease with which their audience can assimilate new claims and arguments about the material already known. Moreover, previous work will also provide a baseline against which new arguments can be made: there will be lacunae in previous accounts that new work can productively close, as well as tensions that it can profitably resolve. It is thus sensible to return to cases repeatedly, just as new biological questions are easiest to ask about organisms we already know how to breed and study in the laboratory.

#### 4.3 Transferring the phylogenetic analogy to the history of science

There are differences as well as analogies between the biological use of model organisms and the historico-philosophical study of cases. An important difference concerns the topology of the descent-relationships in the two domains. The paradigmatic evolutionary case is one where descent with modification produces strict branching: lineages are connected through common ancestry, but lineages do not exchange genetic material once separated. The same is not true in the history of science: the paths of learning and imitation are reticular rather than branching. Different research traditions can cross-fertilize each other even after lengthy separation. Fortunately, however, we need not expect this to be detrimental to our approach, for two reasons. First, branching in biology is often reticular as well, for example in bacteria (O'Malley 2014). Nevertheless, biologists use bacteria as models not only for other bacteria but also for species that are phylogenetically far removed. Second, the crucial point is that the system we observe in one model or case is related to other models or cases; it does not matter how complicated the topology of

the relationship is. What matters is that we can trace relationships sufficiently to inform our extrapolations, and this seems to be possible both in biology and in the history of science.

In both the biological and the historical cases, we can trace lineages at different resolutions, but finer resolutions will often be more useful. For instance, the *overall* phylogenetic relationship between current human and cow populations may not matter greatly if we wish to compare a *particular* aspect of these organisms – for instance, when we study bovine mitochondria in order to learn about human mitochondria. For understanding mitochondrial genetics, the gene tree may matter more than the species tree. The same is true in history and philosophy of science: If we wish to learn about experimental approaches from the Semmelweis case, we require a close kinship between Semmelweis' experimental methodology and the methodology deployed in the other cases we are interested in. However, there is no need for the cases to stand in a close relationship or to belong to the same broader research tradition with regard to *other* aspects. Perhaps Semmelweis was in many ways quite idiosyncratic and atypical, standing outside the main medical and biological research traditions of the middle of the nineteenth century. Nevertheless, we can extrapolate from his experimental methodology to other cases so long as the methodologies are appropriately related. Thus, the phylogenetic relationship between different historical cases is at the core of our views on the justification of extrapolation. However, the relevant phylogenetic relationship must be understood as fine-grained rather than coarse-grained: our approach is grounded in the study not of broad research traditions, but of lineages of methodologies, concepts and practices. These may be meaningfully related to each other in ways that cut across research traditions in a broader sense.

#### 4.4 Philosophical progress by historical means

The Semmelweis case discussed in section 3 demonstrates that case studies allow philosophical progress by historical means. Consider how Lipton (1991; 2004) improved on Hempel's account of induction by careful attention to the details of the Semmelweis case.

Lipton noted that one would expect the main difficulty of understanding scientific inference to lie in the justification of induction. Why are particular principles fit for purpose? In actual fact, however, the mere description of the principles by which we make inductive inferences has proved to be elusive. Lipton wrote: “[It] is not merely that we have yet to capture all the details, but that the most popular accounts of the gross structure of induction are wildly at variance with our actual practice” (2004, 12). He proceeded to show how, precisely, Hempel’s hypothetico-deductive account was both too permissive and too restrictive to capture Semmelweis’s actual inductive inferences.

Consider, first, an example of the hypothetico-deductive account’s excessive permissiveness. In the course of his investigation, Semmelweis rejected the hypothesis that childbed fever was caused by overcrowding of the hospital ward. Hempel reconstructed this in terms of a logical contradiction between the hypothesis and the observation that the two divisions were equally crowded. However, Semmelweis’ rejection of the hypothesis would not have been licensed on hypothetico-deductive grounds, since the hypothesis that overcrowding causes childbed fever is perfectly compatible with equal crowding of the divisions. Overcrowding may cause childbed fever, for example, only in conjunction with other conditions that were only realized in the physicians’ division. Overcrowding may also be only one of multiple causes of childbed fever, of which *others* were realized in the first division. Thus, argued Lipton, the hypothetico-deductive account cannot account for Semmelweis’ rejection of the overcrowding hypothesis and is thus too permissive.

The second example, along similar lines, shows that the hypothetico-deductive account is also too restrictive, as it would not have licensed Semmelweis to accept the hypotheses he actually accepted. Hempel took the cadaveric hypothesis to be confirmed by its observable consequences, in particular by the experiment showing that childbed fever decreased when hand-washing measures were instituted. If the cadaveric hypothesis is joined with appropriate auxiliary hypotheses, it may indeed entail the contrast between the experimental and the control group. However, there were many other relevant contrasts that the cadaveric hypothesis did not entail.

While most women who delivered on the way to the hospital still did not contract childbed fever, some did, even though they were far removed from autopsies. Similarly, even in the midwives' division, where no autopsies were performed, childbed fever occurred sometimes. The cadaveric hypothesis did not entail these contrasts. If it was justified to reject the overcrowding hypothesis because it did not entail the contrast between the physicians' and the midwives' divisions, then it would equally have been justified to reject the cadaveric hypothesis because it did not entail other salient contrasts.

In summary, Lipton made considerable headway on a descriptive track, by showing that the hypothetico-deductive account of confirmation cannot recover the inferences that Semmelweis actually made. By building on Hempel's previous writings on the Semmelweis case, Lipton made a powerful case that the hypothetico-deductive account did not reflect actual scientific reasoning. In a next step, he naturally argued that his account in terms of inference to the best explanation is able to pick out and motivate those inferences, and only those inferences, that Semmelweis did make. Even then, much remained to be said about the justification of the explanationist proposal. Nevertheless, considerable philosophical work was done by a close scrutiny of the details of the case study. Lipton made philosophical progress by historical means.

#### 4.5 The phylogenetic approach and historical epistemology

It may be objected that the phylogenetic approach to the justification of case studies is ahistorical. Scientific methods and practices related to confirmation, explanation, reduction, and so on, change over time. It is thus presumptuous to assume that studying experiments in one time period will reveal much about experiments half a decade earlier or later: it ignores the historical dimension of epistemology. Even those who maintain that scientific epistemology must ultimately be reduced to fixed and perhaps extra-empirical principles will agree that actual methodologies and practices have changed over time (Worrall 1988, 1989, Laudan 1989). Semmelweis' understanding of causal inference was surely different from Claude Bernard's; Bernard's was different from R.A. Fisher's; and Fisher's was different from Austin Bradford Hill's. Such developments are easy to find throughout the history of

science. We do not think, however, that a historicized understanding of epistemology and scientific practice more generally undermines the phylogenetic approach. On the contrary, the phylogenetic approach invites us to trace the development of scientific concepts and procedures over time – the descent *and* modification of practices. By recognizing scientific practices as part of a phylogenetic tree, we become more rather than less sensitive to their historical change. A historicized understanding of philosophical topics is inherent in the phylogenetic approach.

#### 4.6 The normativity objection

The phylogenetic approach does not initially distinguish between successful and unsuccessful scientific practices. It gives us license to extrapolate from practices studied in one case to appropriately related other cases, irrespective of whether those practices are fit to their intended scientific purpose such as inferring causes, giving explanations, or any other. Some will see this as a weakness of the phylogenetic approach: while it may be a great tool for a descriptive project, it fails to further the normative goals that the philosophy of science is generally taken to aim for. It cannot help with the task of justification, or so the objection goes.

We think that this objection lacks force. Far from impeding the project of normative assessment, the phylogenetic approach provides the empirical data for it. On a naturalistic conception, history and philosophy of science is normative not in the sense that it provides extra-empirical (e.g. logical) justification for individual scientific practices, but simply because it can discern, with appropriate distance, which practices achieve their goals and which do not, and ideally, how to explain success when it occurs (Laudan 1987). The phylogenetic approach helps this project because it is intrinsically comparative: it invites us to compare different instances of particular practices both diachronically and synchronically. Such data allows us to determine the conditions under which those practices succeed or fail, and thus to advance the task of justification.

## 5 Conclusion

The justification for extrapolating the lessons learned in historical case studies rests on a material assumption about scientists and scientific practices. The assumption is that few scientists invent methodologies, standards or techniques out of thin air.

Usually these practices are transmitted by teaching and imitation, and we can learn something about entire lineages of practices by studying representative episodes.

We have argued that this is analogous to the way in which biologists justify inferences from model organisms to larger groups by appeal to phylogenetic relationships.

The analogy between model organisms and case studies extends beyond the surface. Not only is the justification for inductive inferences from individual cases similar to the justification for inferences from model organisms, but the approaches also share strengths and weaknesses. It is at first glance surprising that biologists focus on a limited set of model organisms instead of expanding their empirical basis. However, this restriction allows researchers to standardize both organisms and techniques so that rapid progress can be made in the study of particular systems. This procedure is also typical of the history and philosophy of science, where intricate accounts of explanation, confirmation, and many other topics, are usually honed by reference to a small set of cases that are discussed in waves – we have called this philosophical progress by historical means. There is a danger in this, since model systems or case studies may be atypical in crucial respects but not recognized to be so. In such cases, extrapolations may (but as we argued, need not) eventually prove to be erroneous. Luckily such inductive errors are corrigible through further historical research which in turn can lead to important differentiations. For example, if we understand the ribosome of the bovine myocardium in great detail, this will presumably give us the theoretical and instrumental foundation for determining whether other mitochondria function in similar or different ways. Similarly, understanding neuroscientists' explanation of the action potential will make it easier to determine whether other instances are of the same type. Having achieved an in-depth understanding of one case, we can more quickly assess the range of applicability of our results.

On reflection, the similarities between historical case studies and model organisms should be unsurprising. Both biology and the history and philosophy of science have similar research objects: complex historical entities marked by branching and reticular descent with modification. Naturally their empirical methods will converge to some extent.

## References

- Ankeny, Rachel A, and Sabina Leonelli. 2011. "What's so special about model organisms?" *Studies in History and Philosophy of Science Part A* 42 (2):313-323.
- Bird, Alexander. 2010. "Eliminative abduction: examples from medicine." *Studies in History and Philosophy of Science Part A* 41 (4):345-352.
- 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 (4):383-404.
- Chang, H. 2011. "Beyond Case-Studies: History as Philosophy." In *Integrating History and Philosophy of Science*, edited by Seymour Mauskopf and Tad Schmaltz, 109-124. Heidelberg: Springer.
- Craver, Carl F. 2008. "Physical law and mechanistic explanation in the Hodgkin and Huxley model of the action potential." *Philosophy of Science* 75 (5):1022-1033.
- Giere, Ronald N. 1973. "History and philosophy of science: Marriage of convenience or intimate relationship." *British Journal for the Philosophy of Science* 24 (3):282-297.
- Gillies, Donald. 2005. "Hempelian and Kuhnian approaches in the philosophy of medicine: the Semmelweis case." *Studies in history and philosophy of science part C: Studies in history and philosophy of biological and biomedical sciences* 36 (1):159-181.
- Hempel, Carl G. 1966. "Philosophy of natural science."
- Laudan, Larry. 1981. "A confutation of convergent realism." *Philosophy of science* 48 (1):19-49.
- Laudan, Larry. 1987. "Progress or rationality? The prospects for normative naturalism." *American Philosophical Quarterly* 24 (1):19-31.
- Laudan, Larry. 1989. "If it ain't broke, don't fix it." *The British Journal for the Philosophy of Science* 40 (3):369-375.
- Lennox, James G. 2001. "History and philosophy of science: A phylogenetic approach." *História, Ciências, Saúde-Manguinhos* 8 (3):655-669.
- Levy, Arnon. 2013. "What was Hodgkin and Huxley's achievement?" *The British Journal for the Philosophy of Science*:axs043.
- Levy, Arnon, and Adrian Currie. 2014. "Model organisms are not (theoretical) models." *The British Journal for the Philosophy of Science*:axt055.
- Lipton, Peter. 1991/2004. *Inference to the best explanation*. London: Routledge.
- Magnus, PD, and Craig Callender. 2004. "Realist Ennui and the Base Rate Fallacy." *Philosophy of Science* 71 (3):320-338.

- Mill, John Stuart. 1884. *A system of logic, ratiocinative and inductive: being a connected view of the principles of evidence and the methods of scientific investigation*: Longmans, green, and Company.
- Nickles, Tom. 1995. "Philosophy of science and history of science." *Osiris* 10:139-163.
- O'Malley, Maureen. 2014. *Philosophy of Microbiology*. Cambridge: Cambridge University Press.
- Pitt, Joseph C. 2001. "The dilemma of case studies: toward a Heraclitian philosophy of science." *Perspectives on Science* 9 (4):373-382.
- Schickore, Jutta. 2011. "More thoughts on HPS: Another 20 years later." *Perspectives on Science* 19 (4):453-481.
- Schindler, Samuel. 2013. "The Kuhnian mode of HPS." *Synthese* 190 (18):4137-4154.
- Scholl, Raphael, and Tim Rüz. 2016. "Towards a methodology for integrated history and philosophy of science." In *The Philosophy of Historical Case Studies*, edited by Till Sauer and Raphael Scholl. Boston Studies in the Philosophy and History of Science.
- Semmelweis, Ignaz. 1983. "The etiology, concept, and prophylaxis of childbed fever, translated by K." *Codell Carter (Madison, WI: University of Wisconsin Press, 1983)*.
- Vickers, Peter. 2013. "A Confrontation of Convergent Realism." *Philosophy of Science* 80 (2):189-211.
- Weber, Marcel. 2008. "Causes without mechanisms: Experimental regularities, physical laws, and neuroscientific explanation." *Philosophy of Science* 75 (5):995-1007.
- Worrall, John. 1988. "The Value of a Fixed Methodology." *The British Journal for the Philosophy of Science* 39 (2):263-275.
- Worrall, John. 1989. "Fix it and be damned: a reply to Laudan." *The British Journal for the Philosophy of Science* 40 (3):376-388.