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

**Abstract**

The use of historical case studies in philosophical theorizing about science is inherently problematic: single cases are claimed to allow inferences about large parts of science. On the face of it, such inferences are entirely ungrounded. Nevertheless, it seems that similar inferences are not only possible, but even routine in many parts of the biological sciences: geneticists, for instance, successfully reason from very limited sets of organisms to indefinitely many. In this paper we explore whether the philosophical use of historical case studies could work analogously to the use of model organisms in biology.

## 1 Introduction

The use of historical case studies in the philosophy of science is ubiquitous. A prominent example is the current debate about scientific realism, which revolves around long lists of cases tabulated by authors such as (Laudan 1981, Vickers 2013). These cases are shared touchstones for widely disparate views, and they constitute part of the foundation of that community's research project. Any successful account of scientific realism (or anti-realism) will eventually have to make sense of Fresnel's wave theory of light and its successful prediction of the bright spot ((Worrall 1989); (Saatsi 2005); (Psillos 1999)), the caloric theory of heat (Chang 2003; Psillos 1999), and the phlogiston theory (Carrier 1991; Schurz 2011). In other philosophical debates, historical cases are less conspicuous but no less central. Topics include theory appraisal and theory choice, observation and experiment, biological explanation, and many more.[1] In debates large and small, historical case studies are used to support or to challenge theses in the philosophy of science.

[1] For theory appraisal and choice, see Kuhn (1962), Lakatos (1978) and Laudan (1977). For observation and experiment see (Hacking 1982), (Pickering 1984) and (Shapere 1982). Concerning biological explanation, see for instance the debates surrounding the Hodgkin-Huxley-model of the action potential by (Weber 2008), (Craver 2008) and (Levy 2013).

The philosophical findings from a case study are usually taken to reach beyond the case itself. They are routinely extrapolated to a wider class of cases or to other parts of science. Realists view studies showing the conservation of success-fuelling elements of past theories in successor theories as evidence for realism per se (and not just for realism about particular cases). Conversely, antirealists take the lack of conservation to be evidence for antirealism per se (and not just antirealism about the cases in question).[2] Participants in debates concerning biological explanation view the determination of whether the Hodgkin-Huxley model is explanatory as having implications not only for this particular model, but for explanations in biology in general.

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 conceptually coherent programme" because it did not address the question of how "philosophical conclusions may be supported by historical facts" (292).[3] 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 and to answer these concerns, among them Burian (2001), Chang (2011), Schickore (2011), and Scholl and Räz (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 (if at all). In that respect at least, the history-philosophy-relationship remains as opaque as it was when Giere wrote in the early 1970s.

---

[2] Magnus and Callender (2004) have criticised both realists and antirealists for ignoring the "base rates", which are needed to determine how representative the discussed cases really are.

[3] Although Giere, in this contribution, has been interpreted as criticising history and philosophy of science for violating the norm-fact divide, Giere already back then made clear that he did not want to restrict his criticism to this interpretation: "I would argue that all norms have their roots in facts" (292). His concerns were thus of a much broader nature.

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 we understand the extrapolation from individual model organisms to broader biological claims. We will outline a phylogenetic justification for extrapolation that closely mirrors the methods of biology. By adopting the view of cases as model organisms, we will see how extrapolation from individual cases is possible, why philosophers frequently return to previously used case studies instead of expanding their empirical basis, and how philosophical progress is ultimately achieved by historical means.

We proceed as follows. In Section 2 we delineate the object of our discussion. In Section 3 we review the philosophical literature on model organisms with an emphasis on their epistemic role. In Section 4 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's investigation of the cause of childbed fever. Section 5 concludes our discussion.

## 2 Historical case studies characterized

What is a historical case study? Although we do not pretend to possess any authoritative definition, we think it will be helpful to bring the object of our discussion into sharper focus by identifying some of the features we believe generally characterize historical case studies.

We can identify two main functions of historical case studies in philosophical arguments: a clarificatory and an evidential function. The former concerns the attempt to clarify some concept by means of investigating closely some episodes of scientific practice, for example concepts such as causation, explanation, experiment, confirmation, and discovery. We think it is important that philosophers engage in this kind of work, because it allows them to establish a relevance of their deliberations to actual scientific practice. Case studies have an evidential function, because philosophers also use case studies in order to lend support to, or to challenge, a philosophical thesis. This, we think, is for example apparent in the

realism debate, where antirealists have presented historical case studies in order to undermine realist claims about continuity and approximation of the truth over time. Realists, in turn, have faced the challenge by arguing that important historical episodes allow for less overt continuity, for example, structural continuity. Realists and antirealists wouldn't be having this debate, if case studies had no evidential function.

We see no clean separation between case-study-based philosophy of science and the history of science. In both, certain conceptions about science, theories, and experiment will influence the way in which philosophers and historians structure the historical facts at their disposal. So it is not the case that historians would merely just describe 'things as they are' and that philosophers would be particularly biased by virtue of their trade, as it were. Where we do see a difference, however, is in the kinds of questions philosophers and historians ask about history. Whereas philosophers tend to be more interested in epistemic questions such as "was method X used in the discovery of Y?", "was theory T sufficiently supported by evidence E?", "were any elements of theory T1 conserved in the shift to theory T2?", etc., the contemporary historians' main concerns centre around political and cultural contextualization, and questions such as "did individual I influence individual J in her research?", "was individual I influenced by institution G or the political climate at the time?", "how powerful was individual I in research community C?", etc.

Different questions require different criteria for selecting historical facts. Whereas historians often constrain their analyses by the choice of particular historical dates or by the focus on particular protagonists or institutions,[4] such constraints are absent from the philosopher's use of historical case studies. Even though any historical analysis will of course have to include certain protagonists and their actions within a certain time frame, they are not themselves criteria for individuating a case study. Instead, the choice of historical protagonists and periods is often secondary to the choice of historical material for reasons to do with the philosophical questions asked. As a result of different selection criteria, it is then no

[4] For example: "The Unmaking of a Modern Synthesis: Noam Chomsky, Charles Hockett, and the Politics of Behaviorism, 1955–1965" (G. Radick, *Isis*, Volume 107, Number 1, March 2016).

surprise that the research outputs end up looking very different: whereas historical analyses tend to stretch over longer historical periods and follow the trajectories of certain protagonists, philosophical case studies tend to be historically more 'limited' in that they focus on shorter time periods (such as instances of theory change) and on what might be considered more 'internal', non-contextual issues, such as cogency of a theory and evidential support.

We saw that three problems have been raised to challenge the use of case studies in the philosophy of science. We call these the problems of distortion, bias, and inference, and we will discuss them in turn.

### 2.1 The problem of distortion

Pitt (2001) has accused philosophers of distorting the facts when using historical case studies. A plausible example of this may be Imre Lakatos's rational reconstructions, his "radically improved versions" of actual history (1970). But these were declared openly and so could not do too much damage. The challenge must be directed against something more insidious: unrecognized and perhaps subtle distortions in the service of one argument or another. In order to assess this threat, we would have to canvass a large sample of suspected distortions and study them in detail. We would also have to distinguish between willful distortion (which is a matter of research ethics) and inadvertent distortion (which may point to a deeper methodological difficulty for the case studies project). Such a survey is beyond the scope and purpose of the present paper.

Some might accuse philosophers of distortion on the basis of the very fact that philosophers *select* historical facts with a certain philosophical agenda in mind, i.e., with the aim to argue for or against a certain philosophical thesis. One might want to contrast this with the work of the historian, whose main goal is not to argue a certain point, but rather to accurately represent a certain historical period by paying due attention to all the relevant causal factors (such as political and cultural ones). First of all, we would like to shed doubt on this rather innocent portrayal of the historian's work: also historians use history to make arguments. For example, historians have argued that scientists 'sleepwalk' into discoveries and are often unaware of what they actually discover (Koestler 1990), that the "historical

development of our current world view" cannot be understood without paying due credit for the long dismissed steady state theory in cosmology (Kragh 1996), or that the caloric theory of heat was given up prematurely (Chang 2012). In so far as distortion is a threat whenever historical facts are used to make an argument for a particular view, historians therefore seem to be no less subject to it than philosophers. Second, we must emphasize that selection does not imply distortion. Selecting historical facts with a certain philosophical (or historical) agenda in mind does not imply that those facts will automatically be distorted. This is of course not to say that the selection of facts can never lead to distortion. In particular, bias must be avoided.

### 2.2 The problem of bias

Even when facts are presented without significant distortion, they may mislead if they are presented in an unbalanced, biased way. This is apparent in our main case study below: the philosophical reception of Semmelweis's work on puerperal fever. When Carl Hempel (1966) presented his hypothetico-deductive account of the episode, he did not, when viewed with minimal charity, misrepresent facts about Semmelweis's work. Nevertheless, his selection of facts was such that the hypothetico-deductive account seemed natural even though a broader view of the historical facts easily casts doubt on that interpretation. A closer look suggested instead that the overall methodology driving Semmelweis's work was concerned with causal inference and mechanistic reasoning, concepts that were neglected in the tradition of logical empiricism. Thus, the effects of a biased selection of facts are significant. However, the problem again affects all historical scholarship, not just philosophical case studies. Beard (2015) argues that our view of ancient Rome has changed considerably since we began to consider more than the writings of aristocratic men: archeological findings allowed us to redress an imbalance by revealing facts about, for instance, the lives of plebeian women. In sum, the problem of selection is real, but it is not confined to case studies in philosophy of science. It is the nature of history that an ever more pluralistic succession of perspectives allows us to understand historical episodes in greater and greater depth. Of course there is a risk that philosophers might willfully ignore those facts that could challenge their

own views (although, as in the case of willful distortion, this seems more like a problem of research ethics than of methodology). But this problem is in principle suitably addressed by communal control mechanisms such as peer-review and in particular by scholarly dispute. In the Semmelweis case discussed below, we will see that such dispute, conducted over decades, enabled philosophical progress by historical means.

A particular concern has been that philosophers often fail to take into account a certain *kind* of historical context (Burian 2001, Pitt 2001): that they focus on scientific epistemology at the expense of the institutional, social and cultural context of science in which historians are often interested. This is a legitimate concern in so far as philosophical case studies may fail to capture the dynamics of scientific practice if such contextual factors are omitted. Surely, philosophers need to be sensitive to this issue. However, we must emphasise once more, selectivity is ever present: in order to be tractable at all, historical research must always focus on some aspects of an episode and neglect others. Here, too, we are dealing not with a problem of philosophical case studies, but with a general feature of historical scholarship. And the issue cuts both ways: To say that a historical case study is inadequate unless it focuses on social and cultural context (as some historians might) is as tendentious as to say that a case study is inadequate unless it focuses on scientific epistemology (as some philosophers might). In sum, selectivity is a fact of life in history and most other endeavors. But selection does not imply distortion, nor does it necessarily constitute a vicious bias so long as the criteria for inclusion and exclusion of facts are open to debate.

There is a stronger version of the problem of bias, implicit in the aforementioned quote by Nickles, which we consider even less threatening to the historical case study approach. According to this stronger version, the historical record is so rich and diverse that basically *any* philosophical claim can be supported. We regard this as obviously wrong. Consider for example claims like "The electron was and was not discovered in the late 1890s", or "Fresnel's equations were and were not conserved in Maxwell's equations". However rich the historical record, we do not believe—contrary to what Nickels quote seems to imply—that such contradictory

claims can be supported. Instead we do believe that there is good support for either one or the other view (even though philosophers might disagree about the strength of the support). More broadly speaking, we do believe that the historical facts can constrain philosophical theorizing and that they are not as nearly as malleable as Nickels' quote seems to suggest.

### 2.3 The problem of inference.

We do not believe that either the problem of distortion or the problem of bias pose any serious threat to the method of case studies – or if they do, then only in so far as they are a threat to all historical scholarship. However, the problem of inference gives us stronger grounds for concern. How can philosophers justifiably believe that demonstrating the "validity" of a philosophical account by means of a case study can have any bearing beyond this case? In particular, how are wide-ranging inferences possible given the restricted number of case studies that philosophers return to again and again? It seems almost as if they are resistant to increasing their empirical basis. Despite the urgency of this problem for historically informed philosophy of science, it remains unaddressed.

In this paper we propose that the use of historical case studies can be justified in a similar way as the use of model organisms in biology. This approach avoids some of the criticisms that have been raised against case studies.[5] We hope to shed light not only on the problem of inference, but also on the reasons for which philosophers select certain historical case studies and why they tend to keep reusing them.

## 3 Model organisms

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 20th century to enable inferences to higher level organisms such as moths, pigeons, cats, silkworms, rabbits and even

[5] There are other important problems the historical case study approach faces, such as the norm-fact divide. See Schindler (2013) for a discussion.

humans (Levy and Currie 2014). It is striking just how limited is the number of model organisms in biology. In his chapter on model organisms in his book *Philosophy of Experimental Biology*, Weber (2004) writes felicitously that "molecular biology laboratories are extremely impoverished in biodiversity" (155). He adds: "most laboratories work on only a single species, and a large number of laboratories work on the same species". 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 model organisms in particular fields and for particular research questions: the sea urchin in early developmental biology, *E. coli* in the study of bacterial conjunction, 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 mustard seed *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 higher organisms such as humans?

With regard to question of why biologists choose model organisms, Weber suggests that biologists select particular model organisms mostly for pragmatic reasons. For example, the organism must be easy to breed in the laboratory, its generation time should be short, there must be viable mutants with noticeable phenotypic effects, and its features must be suitable for specific research questions (e.g., the size of the neurons in giant squid, and 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 there is a positive cumulative effect of the development of experimental techniques and procedures over time, which makes it 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.). With regard to the question of extrapolation,

Weber argues 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 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, whereas 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). Inferences on the basis model organisms, according to them, are thus fully justified phylogenetically.[6]

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, in order to lend greater stability to inductive inferences from model organisms. 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 and with possible relations to historical case studies. Let us now explore those in the rest of the paper. Suitably, we will do this by means of a case study, namely Ignaz Semmelweis's discovery of the cause of puerperal fever during his work as an obstetrician at the Vienna General Hospital from 1844 to 48. We will structure our discussion along the three questions raised by Weber about model organisms: why do we choose them, why do we return to them, and how do we learn from them?

[6] See section 4 for our more modest view of this sort of justification for historical case studies.

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

Semmelweis's discovery began its life as a case study with Hempel's *The Philosophy of Natural Science* (1966). Hempel explained that Semmelweis was motivated by a puzzle (which was in fact well known at the time) about different mortalities from childbed fever in two divisions of the same maternity clinic: in the first division, the mortality rate was near 10%, while it was comparatively low at 3% in the second division. On Hempel's account, Semmelweis discovered why this was so hypothetico-deductively. He framed a number of hypotheses to explain the difference between the divisions – he suspected differences in weather conditions, crowding, examination techniques – but found each to yield false predictions. But eventually Semmelweis hit upon a better hypothesis. The first division was run by physicians, who conducted autopsies before examining pregnant patients, while the second division was run by midwives, who followed no such autopsy practice. Semmelweis surmised that the physicians transferred some kind of infectious matter from autopsies to patients. This hypothesis yielded correct predictions since, as one would expect, the institution of hand-washing measures removed the difference. On the hypothetico-deductive reconstruction, we can understand this as cycles of conjecture and refutation followed by an eventual corroboration.

In the succeeding decades, Semmelweis was revisited by numerous historians and philosophers of science. Among them is Peter Lipton (1991/2004) who used Semmelweis as an extended case study in his seminal *Inference to the Best Explanation* in order to show in detail how the hypothetico-deductive account fails, while IBE succeeds, at capturing and justifying actual scientific reasoning. Another contribution is by Donald Gillies (2005), who studied Kuhnian factors in the case in order to explain why Semmelweis's findings were initially rejected. Alexander Bird (2010) has argued that Semmelweis's reasoning should be understood as an instance of inference to the *only* explanation. Later writers continued to find new aspects in Semmelweis's reasoning: Scholl (2013, 2015) found Semmelweis's inferences to mirror closely the whole range of methods of experimental inquiry formulated by J. S. Mill (1843), thus casting additional doubt on the hypothetico-deductive

reconstruction, while Tulodziecki (2013) has argued that Semmelweis's reasoning was often careless and should not be held up as a paradigm of scientific inference.

The whole range of questions introduced above must be asked about the Semmelweis case: Why was the case considered suitable in the first place? Why have philosophers and historians of science returned to it countless times since? Why do they think that concepts that are useful for understanding the Semmelweis case speak to the question of scientific discovery and confirmation in general?

### 4.1 What makes a case study suitable?

Why do philosophers pick certain case studies, such as Semmelweis's discovery? We suggest that the situation is similar as in the case of biological model organisms, discussed above: the same issues can in principle be investigated using many different case studies, and cases are chosen in part for pragmatic rather than epistemic reasons. *Drosophila*, for example, is a preferred model organism because it is easy to breed and it has short life cycles (Weber, 2004, p. 177). Likewise, philosophers select certain historical case studies because they are cognitively easily accessible: easy to present and understand. Hempel, for example, chose the Semmelweis case, because it is "a *simple* illustration of some important aspects of scientific inquiry" (1966, p. 3; our emphasis).

Ease of cognitive access, however, comes with tradeoffs: cases that are easy to present and understand are not necessarily typical of many aspects of science—just as *Arabidopsis thaliana* is relatively quick and easy to breed but perhaps not representative of many aspects of the long-lived *Sequoia sempervirens*. But as in the case of model organisms, atypicality need not be an obstacle to inductive reasoning. Atypically high breeding rates, for example, are a feature that may well be irrelevant to a question of interest, such as the genetic map of chromosomes. And even if the feature of interest is indeed atypical in the model organism, this need not be a problem. For example, the axons of giant squid are exceptionally large, and yet are a good choice for an object of investigation because their atypicality makes it easy to study them and to obtain clean results. To this day, mechanisms discovered in neurons in giant squid are taken to be representative of a very broad class of neurons in other living beings, sometimes on the basis of only circumstantial evidence (Levy

and Currie 2014 333). Likewise, Semmelweis's discovery may well be atypical with regards to features that are not relevant to philosophical interests. An example of this is the case's eventual co-option in the service of Hungarian nationalism: it is interesting historically, but irrelevant to discovery and confirmation in the mid-19th-century. Similarly, the case may yield insights despite features that are more exaggerated than is commonly the case. Just think of the astonishing difference in mortality between two otherwise similar parts of the same hospital, which is atypical but neutral with regard to many philosophical questions. In spite of these atypical features, the Semmelweis case is, arguably, highly representative of a broad range of cases of successful scientific reasoning.

As we saw in section 3, another requirement for a good model organism is that it presents insightful variation. A reason for pursuing the study of heredity in *Drosophila* is that many viable phenotypic variants could be observed that were of interest to geneticists. Similarly, the Semmelweis case contains intrinsic variation that instructs us about both scientific success *and* scientific failure: before finding the correct explanation of the phenomenon, Semmelweis investigated several unsuccessful hypotheses, such as the influence of weather conditions or hospital crowding. Thus, there are instructive contrasts to be found within the model system that invite exploration: What were the different sources of Semmelweis's hypotheses? What factors explain that Semmelweis rejected overcrowding as a cause of puerperal fever but accepted the transfer of cadaveric matter?

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 model is pragmatic. That is, historical case studies, just like model organisms, are chosen because they exhibit certain features, possibly accidental, which are conducive to the philosophical inquiry at hand. While some may worry that this sorty of cherry-picking introduces an unacceptable selection bias, we do not believe that a selection of historical cases led by philosophical concerns is inherently problematic. As we argued above, selection in no way implies distortion. While Hempel, Gillies, Lipton, Bird and Scholl agree that the Semmelweis case is in many ways a representative

instance of important aspects of scientific reasoning,[7] they draw very different conclusions from it. Philosophers may have cherry-picked the Semmelweis case because for pragmatic reasons, but the case nevertheless fulfills a genuine epistemic role in philosophical debates.

### 4.2 Why are case studies used repeatedly?

As we saw in Section 2, Weber argues that good reasons for continuing to work with the same model organism are standardization and reproducivility: it is laborious and risky to change model organisms once laboratory techniques have become productive. In a very similar way, the repeated use of particular historical case studies enables philosophers to engage with history in a focused way and with low expenditure of resources. It is useful for philosophers to be able to develop and test their views using well-known cases. We do not always have to start anew.

The Semmelweis case illustrates one of the advantages of the repeated use of a single case study: the relevant sources and background materials are easily available, 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's main work, the *Etiology, Concept and Prophylaxis of Childbed Fever*, had already appeared in a new 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's reasoning. Thus, a fair

[7] Tulodziecki (2013) is the only philosopher writing on the case who takes exception to this. She thinks, for instance, that Semmelweis lacked care in ruling out alternatives to his favored "monocaual" explanation of childbed fever. Although there were without a doubt other factors modulating the observed effect, we do believe that Semmelweis's methods were apt to identify the main causal drivers.

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. As Scholl (2013) shows, Carter's translation of Semmelweis's *Etiology* makes editorial choices that reflect Hempelian preconceptions about Semmelweis's goals and methods. For instance, many pages of numerical tables are left out of the translation, even though they attest to Semmelweis's 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. Semmelweis's contemporaries considered these to be a particularly crucial part of the evidence. Thus standardization *can* lead to biases. However, even if this happens, the situation is not without redress: biases can be fixed by a closer scrutiny of history. Increased attention to the historical detail of cases often challenges the philosophical claims that have been based on them, as the Semmelweis case demonstrates. When this happens, we consider it an instance of philosophical dispute by historical means.

### 4.3 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 inductive inferences to a broader class of cases? In Section 3 we saw that inferences from model organisms to a broader class of organisms are justified phylogenetically. Insights about developmental mechanisms in *Arabidopsis* may be taken partly to reflect development in *Sequoia*: 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 phlogeny 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's discovery may have been a breakthrough for our understanding of infectious diseases, but its methodology is continuous with earlier

work. Creating contrasts with a 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 their methods for causal inference were in flux, none of these researchers reinvented the wheel. There is a methodological tradition to which all of them belong, and Semmelweis himself was, similarly, only one link in a long chain. Today's randomized controlled trials are the distant offspring of the methodological tradition in which Semmelweis also worked. Thus, the reticular 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.

Beyond the classical lines of influence between major figures that historians of science have traced for many decades, broader methodological currents can be made out. We have already mentioned that Semmelweis used John Stuart Mill's (1884) method of experimental inquiry. But we need not understand Mill as the inventor or even an indispensable contributor to that branch of methodological though: Mill merely developed earlier proposals that themselves attempted to capture quite general principles of scientific reasoning. Mill must be understood as a well known exponent of a tradition that extends far and wide in the history of science. We believe that there are other general principles which characterize science throughout: 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 may be just as important for the justification of inductive inferences from case studies.

As we noted in Section 3, Levy and Currie claim that inferences from model organisms to a larger class of organisms is *fully* justified by phylogeny. Although we do believe that such inferences have inductive support, we do not believe that a detailed comparison between the model organism and the target is entirely dispensable. On the contrary, we believe that the inductive support would be very much strengthened by such comparisons. Thus, the inference from neuro-toxicological effects of certain drugs in fruit flies to the same effects in humans has

some inductive support by virtue of phylogeny, but it is better practice to check whether the effect is also found in humans. Likewise, we believe that inductive inferences involving case studies do have some justification, but would nevertheless be strengthened by further investigation of the target. However, time constraints and limited resources often prevent such further investigations. What we would like to stress, though, is that we should not conclude in such cases that inductive inferences involving case studies have no justification whatsoever.

Even the sceptic about inductive inferences should admit that historical case studies 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 arm-chair. Case studies can show that particular processes and concepts are operational a least in some episodes of the history of science, laying the groundwork for their well-motivated application and extension to other cases. This is an important reason why historical case studies have assumed such prominence in the arguments made by philosophers of science.

# 5 Conclusion

We have explored the analogy between historical case studies as used by philosophers and model organisms in biology. We conclude that there are crucial similarities with regards to the choice and the repeated use of the same case studies and model organisms. Moreover, a case study's conceptual and practical relatedness to precedents and successors provides a justification for inductive inferences similar to the phylogenetic justification for the use of model organisms.

**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, R.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.
Chang, Hasok. 2012. *Is Water H2O? Evidence, Realism and Pluralism*. Vol. 293, *Boston Studies in the Philosophy of Science*: Springer.
Craver, Carl. 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* 24:282-297.
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.
Hacking, I. 1982. "Experimentation and scientific realism." *Philosophical Topics* 13:71-87.
Hempel, Carl G. 1966. "Philosophy of natural science."
Koestler, A. 1990. *The sleepwalkers: a history of man's changing vision of the Universe*: Arkana.
Kragh, Helge. 1996. *Cosmology and controversy: The historical development of two theories of the universe*: Princeton University Press.
Lakatos, Imre. 1970. "History of science and its rational reconstructions." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1970:91-136.
Laudan, Larry. 1981. "A confutation of convergent realism." *Philosophy of science*:19-49.
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, Thomas. 1995. "Philosophy of science and history of science." *Osiris* 10:139-163.
Pickering, Andy. 1984. "Against putting the phenomena first: The discovery of the weak neutral current." *Studies in History and Philosophy of Science* 15 (2):85-117.
Pitt, Joseph C. 2001. "The dilemma of case studies: toward a Heraclitian philosophy of science." *Perspectives on Science* 9 (4):373-382.
Psillos, Stathis. 1999. *Scientific realism: How science tracks truth*. London: Routledge.
Saatsi, Juha. 2005. "Reconsidering the Fresnel–Maxwell theory shift: How the realist can have her cake and EAT it too." *Studies in History and Philosophy of Science Part A* 36 (3):509-538.

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.
Semmelweis, Ignaz. 1983. "The etiology, concept, and prophylaxis of childbed fever, translated by K." *Codell Carter (Madison, WI: University of Wisconsin Press, 1983)*.
Shapere, Dudley. 1982. "The concept of observation in science and philosophy." *Philosophy of science*:485-525.
Tulodziecki, Dana. 2013. "Shattering the myth of Semmelweis." *Philosophy of Science* 80 (5):1065-1075.
Vickers, Peter. 2013. "A Confrontation of Convergent Realism." *Philosophy of Science* 80 (2):189-211.
Weber, Marcel. 2004. *Philosophy of experimental biology*: Cambridge university press.
Weber, Marcel. 2008. "Causes without mechanisms: Experimental regularities, physical laws, and neuroscientific explanation." *Philosophy of Science* 75 (5):995-1007.
Worrall, John. 1989. "Fresnel, Poisson and the 'White Spot': The Role of Successful Prediction in Theory-acceptance." In *The Uses of Experiment*, edited by David Gooding, Trevor Pinch and Simon Schaffer. Cambridge: Cambridge University Press.