To appear in: Tilman Sauer and Raphael Scholl, eds., *The Philosophy of Historical Case Studies*. Boston Studies in the Philosophy and History of Science, Springer.

Towards a Methodology for Integrated History and Philosophy of Science

Raphael Scholl and Tim Räz

> A glaring asymmetry, obvious at this meeting, is that historians dress better than philosophers – historians always being interested in the details, sartorial and otherwise, while philosophers seem concerned only with dressing in general.
>
> – Richards (1992)

Abstract We respond to two kinds of skepticism about integrated history and philosophy of science: foundational and methodological. Foundational skeptics doubt that the history and the philosophy of science have much to gain from each other in principle. We therefore discuss some of the unique rewards of work at the intersection of the two disciplines. By contrast, methodological skeptics already believe that the two disciplines should be related to each other, but they doubt that this can be done successfully. Their worries are captured by the so-called dilemma of case studies: On one horn of the dilemma, we begin our integrative enterprise with philosophy and proceed from there to history, in which case we may well be selecting our historical cases so as to fit our preconceived philosophical theses. On the other horn, we begin with history and proceed to philosophical reflection, in which case we are prone to unwarranted generalization from particulars. Against worries about selection bias, we argue that we routinely need to make explicit the criteria for choosing particular historical cases to investigate particular philosophical theses. It then becomes possible to ask whether or not the selection criteria were biased. Against worries about unwarranted generalization, we stress the iterative nature of the process by which historical data and philosophical concepts are brought into alignment. The skeptics' doubts are fueled by an outdated model of outright confirmation vs. outright falsification of philosophical concepts. A more appropriate model is one of stepwise and piecemeal improvement.

Raphael Scholl
University of Pittsburgh, Center for Philosophy of Science, 817 Cathedral of Learning, Pittsburgh, PA 15260, USA, e-mail: raphael.scholl@gmail.com

Tim Räz
University of Konstanz, FB Philosophie, 78457 Konstanz, Germany, e-mail: tim.raez@gmail.com

1 Introduction

Integrated history and philosophy of science must steer clear of two opposite methodological pitfalls. One is the unwarranted generalization of grand philosophical theses from a handful of historical case studies. The other is the outright imposition of philosophical concepts upon cases that serve as mere Potemkin villages. The task is difficult, and many have claimed it not to be worth the trouble. Some philosophers will argue that their discipline seeks a kind of systematic knowledge that does not need to be sensitive to the detailed study of cases from actual scientific practice. Conversely, some historians of science will argue that historical understanding is possible without the abstract concepts that contemporary philosophy of science has to offer.

Thus, the present contribution has a foundational and a methodological purpose. Section 2 articulates our view of what we can find at the intersection of history and philosophy of science (HPS) that each discipline on its own cannot offer. We expect some version of our "fundamental motivation" for an integrated HPS to meet with broad consensus – not only among those who self-identify as historians and philosophers of science, but also among the many philosophers of science who embrace case studies (from contemporary or historical science) as a routine part of their work.

Our remaining discussion takes up the greater difficulty: It outlines some best methodological practices for relating history to philosophy (and vice versa) in case studies. While a robust grounding in both historical and philosophical methods is indispensable, we believe that the integrated HPS project requires an additional set of skills that currently receive little discussion. We will begin with the so-called dilemma of case studies, which encapsulates the two methodological challenges we introduced above (section 3). First, how can we apply philosophical concepts to historical cases without selection bias – that is, without choosing our historical cases such that they are merely convenient illustrations of preconceived philosophical theses? We will outline how historical cases can play a *probative* role in philosophical investigations by allowing us to test and refine our best concepts about science (section 4). Second, how should the confrontation of philosophical concepts and historical theses proceed? We will suggest a model that emphasizes iterative, piecemeal conceptual engineering instead of generalizations from particular cases (section 5).

The view of integrated HPS we will be presenting is not new, and it is not our intention to educate colleagues. Rather, we see our contribution as an attempt at explicating a method already widely used by many practitioners in the field. If our contribution is successful, other scholars will accept this invitation to join us in a renewed, explicit reflection on the methodology of integrated HPS.

2 The fundamental motivation for an integrated history and philosophy of science

We recognize a plurality of different approaches within HPS.[1] However, for the purposes of the present contribution our interests mainly belong to a particular view of the history-philosophy interaction. On this view, philosophy of science provides us with a set of concepts and worthwhile questions about the history of science – questions that are particularly pertinent to science's core epistemological concerns. Conversely, the history of science grounds philosophical reflection in actual science – in brief, it enables a naturalistic approach to the questions of philosophy of science concerning such issues as confirmation, explanation, reduction, discovery, long-term theory change, and so on. Our flavor of integrated HPS derives its value from this unique interaction: it provides the historian with theoretically rich concepts and questions, and it grounds its philosophical concepts in actual scientific practice.[2]

2.1 Questions for history

Historical sources are powerful: they can bring into focus vague questions and diffuse notions about science. But it is also a well understood fact that the sources fail to speak for themselves. The sources speak because we ask specific, carefully crafted questions of them. It is at best a curiosity to learn that Darwin, in his *Beagle* notebooks, used various spellings of words like "occasion", "coral" and "Pacific" – but it becomes a revealing fact once we use these spelling variants to date Darwin's conversion to transformationism after the Galapagos (Sulloway, 1982). If the sources sometimes seem to almost force themselves on us, it is because we read them with well prepared minds.

The focus of most current historians of science is on questions about the cultural, social and material context of scientific inquiry. Three decades ago, this shift was welcome as a corrective to an earlier tradition which was overly focused on great men and great ideas, or theories and theoreticians. We now have a much richer understanding of science as a human activity than in the past.

However, other types of questions have been neglected or even rejected. In particular, the core philosophical concerns about science have not in recent years received much historical elucidation: discovery, justification, representation, explanation, and

[1] Jutta Schickore (2011) gives a good overview of the debates about integrated history and philosophy of science since the middle of the 20th century, and she cites many key works. Howard (2011) offers a more long-term history of the relation of the two fields and discusses some of the fundamental reasons for their separation in the 20th century. A good snapshot of the state of the field in the new millennium can be found in Arabatzis and Schickore (2012).

[2] While we focus on one flavor of HPS, we see others as complementary and equally valuable. For instance, a key project is to trace the origin and growth of modern concepts, theories and questions (Lennox, 2001; Schickore, 2011).

so on. It is time to reverse this trend. To echo an earlier generation of scholars, the historical neglect of philosophical issues is an opportunity missed and a responsibility avoided (cf. Laudan et al, 1986, p. 152). By reintegrating philosophical concerns with historical scholarship, we will gain richer histories.[3] The penalty for foregoing this reintegration is an inadequate picture of the scientific enterprise – a picture just as inadequate as one that pays no attention to cultural context, institutional frameworks, or societal organization.

A serious historical investigation of philosophical questions promises to teach us much about how scientists conceive of new hypotheses, and more broadly about what strategies they employ for solving empirical problems. We will be able to study in greater depth how actual scientific communities have debated different types of empirical evidence (not just rhetorically, but epistemologically) and how they have adjusted their judgments in accordance with it. We will finally find out whether individual scientific disciplines grow essentially cumulatively or by sharp discontinuities – which remains, in many ways, as open a question as it was when Kuhn put it on the scene in the 1960s. Similarly, once we go beyond the limiting notion that there is nothing to scientific consensus except for rhetoric, negotiation and power, we find ourselves confronted with the very real and pressing question of how groups – rather than mere individuals – conduct collaborative and competitive epistemic projects. For the most part, historically deep and philosophically informative historical studies of these and similar issues remain to be written.

In summary, philosophical issues are a rich resource for asking historical questions that are particularly pertinent to science, but this approach remains underappreciated both in theory and in practice. In consequence, we miss the chance of a deeper understanding of what makes science special as an epistemic enterprise. However, we do not envision the use of philosophical questions in historical scholarship as a one-way interaction, as we will discuss further in the next subsection and in the remainder of this paper: While philosophical questions should be asked of history, we also believe that history will allow us to refine our philosophical questions and indeed to answer them in unexpected and uniquely informative ways.

2.2 Naturalism for philosophy

Fifty years ago, philosophical skepticism about integrated history and philosophy of science centered around the issue of normativity: If the goals of philosophy of science are normative, then what does a descriptive project like the history of science have to do with it? For one may study Galielei's or Darwin's methodology to one's heart's content, but this will not shed any light on the normative question of whether their methods are justified. When we ask whether science *should* proceed in one way or another, the answer will not come from a descriptive historical study but

[3] An indication of this is the tradition of "historical epistemology" (see e.g. Rheinberger, 1997, Daston and Galison, 2007, and the special issue of *Erkenntnis* edited by Feest and Sturm, 2011).

from a normative philosophical argument. In effect, history was short-circuited out of philosophical discussions.

The normativity objection found its most famous expression in Ronald Giere's "marriage of convenience" paper, long a cornerstone of philosophical skepticism about integrated HPS. Giere wrote:

> If one grants that epistemology is normative, it follows that one cannot get an epistemology out of the history of science — unless one provides a philosophical account which explains how norms are based on facts. This ought to be a central problem for historically oriented philosophers of science, but few seem willing even to acknowledge the question, let alone attack it head on. (Giere, 1973, p. 290)

From the normative perspective, Giere's challenge is perhaps as close as one can get to an ironclad argument against the philosophical relevance of the history of science. On this view, history can only play a heuristic role by allowing us to identify problems and outline possible solutions that then require proper philosophical analysis. At most, normative philosophy of science needs a link to the history of science in order to be a philosophy *of science*, rather than of some logically possible state of affairs. Or as Norwood Russell Hanson (1962) put it, the pure philosopher's "analytical skill may be admirable, but it does not take us anywhere" (p. 586).

However, in the decade after the "marriage of convenience" paper Giere changed his mind and began to give a much more crucial role to history. He now argued that he had misconstrued the issue; that all known normative approaches had stalled; and that naturalism offered the best prospect for a successful philosophy of science (Giere, 1985). In a recent paper, he summarized his change of mind:

> I came to the conclusion that the philosophy of science should be transformed into something like the theory of science. That is, philosophers should be in the business of constructing a theoretical account of how science works. Philosophical claims about science would then have the status of empirical theories. In short, the philosophy of science should be naturalized. This means, among other things, giving up pretensions to finding autonomous standards for the practice of science. (Giere, 2011, p. 60–61)

In a naturalized philosophy of science, history of science ceases to be a heuristic crutch.[4] Instead, it turns into the indispensable empirical basis for a theoretical enterprise that is best understood in analogy to other theoretical enterprises in the natural and social sciences. Much as ecologists model interspecific competition, or as macroeconomists model the effects of monetary policy, philosophers of science model scientific confirmation, explanation, theory choice, or what have you.

The naturalistic turn deflects the normativity objection by reconstruing the HPS project along the lines of the empirical sciences. But even though it begins with and emphasizes description, the naturalized project shares many of the goals of the old

[4] We prefer not to draw a strong distinction between historical and contemporary scientific practice as an object of study. What would have been "contemporary" science to Giere in 1985 is "historical" now, but the theoretical questions we ask about our cases remain largely the same. The only difference are in the methods of study: How recent an episode is will partly dictate whether our tools will include archival studies, oral histories, laboratory notebooks, or questionnaires, not to mention "embedding" oneself in a research group. Depending on method, of course, some questions will be easier to answer than others.

normative project. HPS does not turn into a purely descriptive project because we understand it to have two tasks: description *and* justification (see also Lipton, 2004). That is, we wish to give an adequate description of past and present science, but we also wish to understand how the epistemic successes and failures of science can be accounted for. A theoretical and historical approach to science that cannot explain successes or recognize mistakes would be a poor one. Traditionally, however, the normative project took the descriptive task to be fairly trivial, while many naturalists would argue that adequate description may be the harder of the two tasks.

The task of justification has much in common with the normative project, yet the naturalized view has at least two advantages. First, a close engagement with the past and present of scientific practice can serve as an accelerator. Even if it were in principle possible to do normative philosophy of science from first principles (such as logic or probability calculus), the project is likely to advance more quickly if existing practice is taken as a guide. If we wish to understand epistemology, we should begin with the most successful epistemological enterprise that we know. A second, much stronger naturalistic argument holds that many issues in the philosophy of science cannot be tackled from first principles at all. This is because scientific practices such as induction or explanation may ultimately be grounded in facts about the world (on induction, see Norton, 2003). For example, the justification for biologists' interest in mechanistic explanations (Machamer et al, 2000; Bechtel, 2006) probably does not derive from any formal philosophical property of such explanations. More likely, biologists have learned in the course of research that mechanistic explanations are adequate to many parts of their area of inquiry. What counts as a "normatively" adequate explanatory standard in this case has a necessary empirical and historical dimension: it concerns what is the case in the world and how we have learned about it. Thus, while the old normative project aimed for some sort of extra-empirical justification for the methods of science, strong naturalists expect the task of justification to be continuous with empirical science itself. But it ultimately does not matter whether one takes a weak or a strong naturalistic stance: the consideration of history and actual science helps in either case.

Where the methodology of integrated HPS is concerned, a commitment to some sort of naturalism is the beginning and not the end of the discussion. It is far from trivial to see how the combined goals of descriptive adequacy and philosophical insight can be achieved in practice. Therefore, in the remainder of this contribution we will discuss what is now perhaps the most pressing methodological problem in the discipline: best practices for relating history and philosophy of science to each other.[5] We will begin our discussion at the skeptical extreme: with the "dilemma of case studies", which suggests that the project of integrated history and philosophy of science – whether construed naturalistically or not – may be doomed in principle.

[5] Pinnick and Gale (2000) commented that "despite the possibility of doing so, philosophers have not pursued a method of case-study design" (p. 116). They also observed that disciplinary consensus about method coincides with progress.

3 The dilemma of case studies

If we wish to use historical cases in order to test, refine and expand our best concepts about science, we are faced with two main problems: selection bias and inappropriate generalization from individual cases. In an issue of *Perspectives on Science* devoted to the legacy of Thomas Kuhn, Joseph C. Pitt (2001) labeled these twin dangers as "the dilemma of case studies". If we approach our project "top down" (proceeding from philosophy to history) then this leads into the first horn of the dilemma:

> [I]f 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. (p. 373)

Yet it is no solution simply to stick closer to the facts of history, since proceeding "bottom up" (form history to philosophy) only leads into the second horn of the dilemma:

> [I]f one starts with a case study, it is not clear where to go from there – for it is unreasonable to generalize from one case or even two or three. (p. 373)

Pitt's dilemma seems to us to capture the core worries of both philosophers and historians who are presently skeptical about the project of integrated HPS. Many historians are wary of philosophically motivated work since the philosophy might well dictate which historical cases are chosen and how they are interpreted – very much in line with the first horn of Pitt's dilemma. Conversely, many philosophers (even those with a broadly naturalistic outlook) worry that any conclusions drawn on the basis of historical case studies are ultimately unwarranted generalizations – and this mirrors the second horn of Pitt's dilemma.

A number of practitioners of integrated history and philosophy of science have responded to Pitt's challenge (Burian, 2001, 2002; Schickore, 2011; Chang, 2011). Chang in particular has articulated conceptual moves that may allow us to break free from the dilemma of case studies. He argues that the we should not think of philosophical concepts as *general* and historical facts as *particular* – for this would indeed lead into fruitless debates about how many white swans are needed to show that all swans are white. Instead, philosophy provides *abstract* concepts which are instantiated to various degrees by *concrete* historical cases. Chang compares this – using an admittedly imperfect metaphor – to the relationship between the setting of a TV series and its episodes:

> When we have an episode of *The Simpsons*, or *Buffy the Vampire Slayer*, or what have you, the episode is not really a case or an example of whatever the general idea of the show might be. Rather, the episode is a concrete instantiation of the general concepts (the characters, the setting, the type of events to be expected, etc.), and each episode also contributes to the articulation of the general concepts. (Chang, 2011, pp. 110–111)

Once the relationship between history and philosophy of science is construed in this way, we stop thinking in terms of working "top down" or "bottom up". Instead, we start thinking of a cyclical process: Just as abstract concepts help us to understand

concrete episodes, so concrete episodes help us to further elaborate our conceptual tools. On Chang's view, doing HPS consists in a repeated cycling between the concrete and the abstract (figure 1).[6]

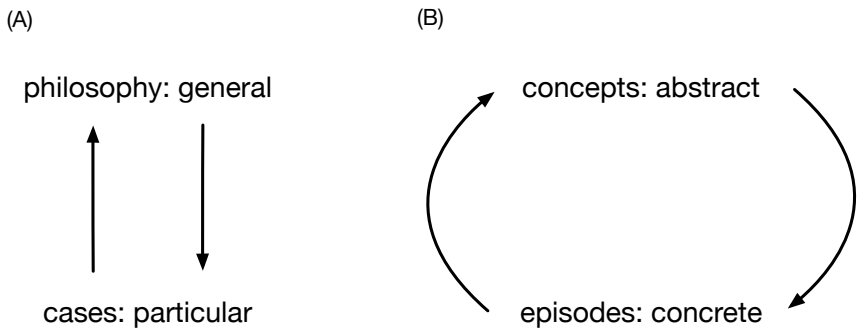

Fig. 1 **A**: The traditional view of a contrast between general philosophical concepts and particular historical cases, which are related to each other either top-down or bottom-up. The usual problems present themselves: How can we move top-down without selection bias? How can we move bottom-up without unwarranted generalizations? **B**: A schema of Chang's alternative view. Instead of contrasting the general with the particular, it contrasts abstract concepts with concrete instances in historical episodes. Instead of conceiving of either a bottom-up or a top-down confrontation of concepts and episodes, Chang proposes an iterative, cyclical movement between concepts and episodes. The entry into the cycle can occur either with concepts or with episodes: It does not require us to decide that either concepts or episodes are primary.

Chang's cyclical model fits better with the actual practice of HPS than Pitt's top-down/bottom-up model. It is a useful metaphor for the interaction between history and philosophy of science. However, it requires further elaboration if it is to provide the basis for a methodology of HPS.[7]

In the remainder of this contribution, we will discuss the particulars of both the downward arrow (from concepts to episodes) and the upward arrow (from episodes to concept), as well as their cyclical interaction. In section 4, we will discuss criteria for moving from concepts to episodes without incurring the risk of selection bias. In

[6] While we adopt Chang's general framework, we do not think that much hinges on whether we speak of "episodes" or "cases", and we will continue to use both terms.

[7] Schickore (2011) has argued that the history-philosophy relationship should not be understood in terms of a confrontational model, in analogy to the empirical sciences, but in terms of hermeneutics, or "the art of gradually reconciling provisional analytic concepts with a provisional reading of the historical record" (p. 459). However, we believe that the confrontational and the hermeneutic models can be reconciled. Certainly the confrontational model must be conceived, as we discuss, in cyclical and iterative terms. But this is no surprise, since the empirical sciences – on which the confrontational model is based – are similarly iterative in theory testing. Moreover, HPS is in part concerned with the beliefs and desires of human actors, the traditional domain of interpretive, hermeneutic approaches. But this has ample room in the confrontational model, which understands the study of human beliefs and motives in terms of empirical theses about cognitive states (how ever difficult these may be to ascertain).

section 5, we will discuss the handling of agreements and conflicts between concepts and episodes, which we call the dynamics of confrontation. In particular, we will discuss the cyclical movement between concepts and episodes which allows us to achieve a certain kind of generality for our concepts about science.

4 The selection of case studies

Many of the best works in integrated HPS explicitly discuss the merits of their chosen cases: they tell us why a particular case will not merely illustrate but investigate the worth of particular philosophical claims, or why conclusions reached for one case are likely to extend to others. However, for the most part the reasons for the choice of case studies remain implicit, and there is little discussion of a general methodology by which the selection of case studies should proceed. The aim of this section is to develop the outlines of such a general methodology. Like many systematic investigations, we begin with typology: There are a number of different purposes that case studies typically serve, and the issues of selection bias and generalization must be understood in relation to these purposes.

4.1 Hard cases

The basic idea of hard cases is to seek out challenges: instead of selecting cases that illustrate a philosophical thesis particularly well, we prefer those that are difficult to accommodate and that therefore put a thesis to the test. To use an engineering example: If we build a self-driving car that can navigate the busy and complex traffic of Beijing without accident, then it can probably handle the streets of Zurich as well. Hard cases demonstrate the power of a principle, and they show that the same principle can plausibly handle a host of similar but less difficult cases.

Analogies in the sciences are easy to find. Consider for example evolution by natural selection, for which the Giraffe's neck is a traditional and well-worn illustration. The illustration is excellent as a didactic tool: it is easy to understand and contrasts well with the alternative of evolution by use and disuse. However, those who are skeptical of the power of natural selection will not find the case particularly compelling: neck length may be a fairly trivial trait, and it would be easy to accept its origin by natural selection while denying that selection can produce more intricate and complex traits. It is no surprise, then, that evolutionists from Darwin onward have been particularly interested in hard cases for natural selection such as the human eye. If a trait as complex as the eye can be explained by natural selection, then selection has passed a high bar; and this success immediately makes plausible that more trivial cases can be explained in the same way.

To illustrate the point using an example from our own work, Scholl and Nickelsen (2015) investigated the genesis of Peter Mitchell's mechanism of oxidative phospho-

rylation. This is the main process by which mitochondria transform the energy in food stuffs into a chemical compound called ATP, which cells, tissues and organs then use to drive their various processes. Mitchell received a 1978 Nobel Prize for the formulation of the mechanism, and it has long counted as one of the most spectacularly original contributions to 20th century biology. Leslie Orgel once wrote that "[n]ot since Darwin and Wallace has biology come up with an idea as counterintuitive as those of, say, Einstein, Heisenberg and Schrödinger" (Orgel, 1999, p. 17). In our study we were able to show that the genesis of the theory can be explained using concepts from two recent strands in the philosophy of scientific hypothesis generation. One is interested in how the unknown causes of phenomena are sought (Graßhoff and May, 1995; Lipton, 2004); another is interested in how new mechanistic hypotheses are generated based on known entities and interactions (Darden, 2006). The first strand allowed us to see that Mitchell's process of hypothesis generation, however spectacular the result, occurred in a well-defined space of possible causal hypotheses. The second strand allowed us to see how this well-defined space of possible hypotheses was investigated by generating "how possibly" mechanisms.

Our study has special probative force because it deals with a hard case of scientific discovery. No one would claim that the mechanism of oxidative phosphorylation was a trivial extension of existing biochemical knowledge: It was a theory of acknowledged novelty and originality. If its genesis is intelligible in terms of a number of basic heuristics, then the power of those heuristics is credibly demonstrated. Moreover, that the hard case could be accommodated provides some warrant for the speculation that many further but less difficult cases of scientific discovery are amenable to similar analyses.

There is a general recognition that hard cases can be particularly telling. To pick just one example, many of the philosophically most interesting theses in *Inventing Temperature* (2004) rely on Chang first convincing his readers that the measurement of temperature is, against expectation, a hard case in the history of measurement techniques. Chang's notion of epistemic iteration becomes compelling precisely when we realize that it can illuminate a particularly difficult epistemic advance.

4.2 Paradigm cases

Many cases play a special role in HPS because they have become paradigms of some aspect of science. In a sense, these cases function like model organisms in biology: we study them not only as particulars, but as more or less typical instances of some aspect of science. Like model organisms, paradigm cases offer the advantage of pre-built resources. The relevant historical documents and the historical context are usually reasonably well understood, so that new conceptual studies can proceed rapidly.

A good example of a paradigm case is Semmelweis's discovery of the cause of puerperal fever around the middle of the 19th century. The case was introduced to the philosophy of science in Carl G. Hempel's *Philosophy of Natural Science*

(1966), where it served to illustrate aspects of the confirmation of theory by data. The case was later revisited by Peter Lipton (2004), who challenged Hempel's hypothetico-deductive reconstruction of Semmelweis's procedure and outlined an alternative in terms of inference to the best explanation (IBE). Lipton argued that a number of aspects of Semmelweis's investigation – including its context of discovery, the rejection of alternative hypotheses, and the confirmation of accepted hypotheses – remain obscure on a hypothetico-deductive reconstruction but become intelligible in the framework of IBE. Importantly, Lipton was able to draw on rich existing material concerning Semmelweis's discovery such as, for example, K. Codell Carter's translation of Semmelweis's main work (Semmelweis, 1983). The translation, in turn, was partly produced in order to facilitate the use of the Semmelweis case in a course in philosophy of science. The discussion of the Semmelweis case has continued in recent years: Donald Gillies (2005) has argued that a Kuhnian perspective is necessary for understanding the reception of Semmelweis's work; Alexander Bird (2010) sees Semmelweis as an instance of inference to the *only* explanation; and one of us has argued that Mill's four methods of experimental inquiry play an important role in the confirmation of Semmelweis's data (Scholl, 2013) – a fact which was previously overlooked because Carter's translation omitted Semmelweis's copious numerical tables, which seemed irrelevant from a Hempelian perspective. For the most part, these authors are not primarily interested in Semmelweis *qua* Semmelweis: the topic of interest is confirmation, of which Semmelweis is taken to be a representative instance.

Whether a case deserves the status of a paradigm is itself open to debate. For example, Dana Tulodziecki (2013) has recently argued that the discussion of Semmelweis proceeds from the false assumption that Semmelweis was an excellent reasoner. She discusses a number of flaws in Semmelweis's arguments which indicate that the case is not, after all, a representative instance of successful scientific reasoning. In our view, explicit arguments for and against the representativeness of a case are required when using paradigm cases.

Importantly, it remains an empirical question whether concepts can be transferred from the paradigm to other cases. The fact that paradigms are considered typical instances gives us reason for some optimism that many concepts, once developed and refined, can be transferred from them to other cases – but whether this is in fact the case must be checked in further detailed studies. This again mirrors the use of model organisms, where we also have the expectation but no guarantee of transferability to other organisms.

Many classical works of HPS use paradigm cases. Take for instance Shapin and Schaffer's *Leviathan and the Air-Pump* (1985). The authors' discussion of the construction of experimental knowledge is powerful precisely because the air-pump is emblematic of experimental science. What is true for the air-pump is plausibly true for countless other experiments. While we would not commit to many of Shapin and Schaffer's specific claims, from a methodological point of view the air-pump is a properly deployed paradigm case.

4.3 Big cases

The most traditional and straightforward reason for choosing a case study is that it concerns a big scientific achievement. It may be an achievement that served as a scientific paradigm for many further works; it may be the foundation for a large branch of present-day science; it may have yielded an understanding of a fundamental aspect of nature. In many of the most interesting cases, such as the works of a Newton or a Darwin, all of the above will apply.

Unlike paradigm cases, big cases cannot be expected to generalize particularly well. We often assume that big cases are also in some way typical of an aspect of science, and we may therefore be tempted to generalize from them in the same way as we do from paradigm cases (section 4.2 above). But of course typicality is something that cannot be assumed. It is possible that Newton's efforts to confirm his theories were quite atypical of how most confirmation in science happens; it may be that Darwin's standards for what is an acceptable explanation were atypical of scientific explanations at most times; and so on. That a big case is also typical of some aspect of science must be explicitly argued for (or at least stated as a premise) – and then these cases generalize in virtue of being paradigms.

Similarly, we should not be misled into thinking that all big cases are hard cases. Certainly influential scientists like Newton and Darwin solved hard empirical problems. But that does not mean that their work always qualifies as hard cases in the sense of section 4.1: Whether something is a hard case in our sense depends on the philosophical thesis under consideration. The genesis of Darwin's theory may have been particularly difficult, which makes it a hard case for those who argue that scientific discovery is explicable in terms of basic heuristics. At the same time, however, other aspect of Darwin's work may not constitute a hard test of relevant philosophical ideas – maybe there is little to be learned from finding that Darwin's concept of explanation conformed to the notion that most explanations are mechanistic. Our modest point is simply that whether big cases are also hard cases in the sense discussed here depends on the philosophical thesis under test.

Even though there is no reason to think that lessons learned from big cases are necessarily transferable to other cases, we do not think that selection bias is a major concern. A scholar may choose a big case specifically because it bears out his or her philosophical *idée fixe* – but this would nevertheless teach us something interesting about a case we already consider to be important, provided that the concepts actually apply. Simply put, it is inherently fascinating to understand the particulars of an important epistemic advance. Moreover, we certainly wish to know whether our best philosophical concepts can illuminate the epistemic breakthroughs we find most important. Thus, it is not only allowed but even necessary, sooner or later, to apply our conceptual tools to the big cases. The only mistake would be to think that inherent fascination is a substitute for carrying on with the broader program of HPS: Ultimately, the range of applicability of concepts must be checked using diverse cases. To know whether a Newton or a Darwin is typical of science at his time, or typical of key concepts from the philosophy of science, is part of understanding the episode.

Big cases are particularly prone to the underdetermination problem of HPS (see the introduction to this volume by Sauer and Scholl): the same historical episode is usually told again and again in different philosophical terms, which raises concerns that philosophical concepts hinder rather than help our understanding of science. Most influential scientists have had multiple careers in the literature: as good inductivists, as resourceful hypothetico-deductivists, as epistemically cautious Popperian falsificationists, perhaps as methodological anarchists, and finally as contingent products of mostly social forces. Not all of these accounts can be true, but deciding among them is hindered by the fact that the key questions often concern cognitive processes of past scientists – to which we have little access. The best defense against the mindless retelling of big cases according to prevailing philosophical fashion is not, however, to retreat to some form of dubious historical positivism, but to take the cyclical model of HPS seriously: We must consider a wide range of cases from the history of science, use them to improve our conceptual tools, and deploy these tools to understand episodes at different levels of importance. We will have more to say about how cases are used to evaluate and refine concepts in section 5.

4.4 Randomized cases

There already exists a widely accepted method for avoiding selection bias in the sciences: randomization. It is at least conceivable that case studies big and small could be chosen randomly from a database and submitted to philosophical analysis. If one had a particular hypothesis about, say, the steps by which model-building proceeds, it might be possible to ask the database for random instances of model-based science in order to check the applicability of the hypothesis. While we do not believe that this should (or could) replace historical judgment in the choice of case studies, it could be a valuable complement to the way in which historical scholarship traditionally proceeds.

Before the randomization of case studies becomes feasible, both practical and conceptual problems need to be addressed. On the practical side, no suitable database of case studies from the history or the philosophy of science currently exists. However, it is a reason for optimism that such a database would be desirable for any number of purposes apart from randomization. For instance, a database of case studies could restore unity to a field that has lately focused on historical micro-studies rather than grand narratives. On the conceptual side, the organization of the database would be a challenging issue. How are case studies to be individuated and classified in a way that is historically adequate, reasonably theory-neutral and useful for data retrieval? At minimum, something akin to Pubmed's Medical Subject Headings (MeSH) vocabulary would be required. Importantly, long-term institutional backing would be a prerequisite for the credibility of such a project.

4.5 First sketch of a typology

We have distinguished between four types of case studies. Each has its own conceptual relationship to key concerns such as selection bias and generalizability.

HARD CASES are chosen to be difficult for the philosophical concepts under study to handle. What counts as a hard case will thus vary depending on the philosophical concepts we are interested in. If a philosophical principle survives contact with a hard case, this speaks to its power. Hard cases circumvent selection bias by seeking challenges rather than convenient illustrations. They allow us to draw more forceful conclusions than individual cases normally do since they give us reason to think that the tested philosophical principle is powerful enough to handle less difficult cases as well.

PARADIGM CASES are the model organisms of HPS. We use them in teaching and research as typical instances of particular aspects of science. Because they are already accepted as typical, and because the relevant historical sources are usually easily available, paradigm cases are efficient tools for making new points and for revising existing concepts. Importantly, whether a case qualifies as paradigmatic is usually itself a point of debate. And whether concepts that apply to the paradigm case can be extended to further cases remains, as in the case of model organisms, an empirical question.

BIG CASES concern influential scientific achievements. They have particular appeal because we inevitably learn something we care about from them – big cases are important to us! However, big cases must not be assumed to be typical of some aspect of science without further argument; nor are they necessarily hard cases, since this depends on the philosophical concepts under study. Finally, big cases are particularly attractive targets for retellings according to prevailing philosophical fashions. This impulse must be resisted by committing in earnest to the cyclical model of HPS.

RANDOMIZATION of case studies to counteract selection bias is currently little more than a neat idea, but we think that it is coherent in principle. A reason to pursue the idea is that a database of case studies would be a useful tool with many additional uses.

Our typology is not intended as something static and upfront. Cases do not present themselves to us with a label identifying them as "paradigm" or "hard" cases. Whether we understand a case as paradigmatic, or as hard with respect to a philosophical thesis, and so on, should be allowed to develop as our studies progress. New historical evidence – the cyclical revisions we discuss in the next section – may well alter our assessment of a case study's status. Moreover, our typology is unlikely to be complete: we look forward to the critical discussion of additional types with their own particular function. Our main point is modest: As we develop research projects, we should pay greater attention to the ways in which our case studies relate to the philosophical theses we wish to demonstrate, criticize or test.

5 Dynamics of confrontation

Once philosophical theses are confronted with historical cases, agreements and disagreements between them are found. The temptation is to think of these as simple instances of the confirmation or rejection of general hypotheses by particular facts. But this temptation must be resisted on pain of two related philosophical sins: the universal ($\forall$) sin and the existential ($\exists$) sin. When we commit the universal sin, we assume that what is true for our case study is true for all of science. When we commit the existential sin, we assume that a single counterexample can serve to reject a philosophical approach. Both procedures are misguided, but how can we do better?

An example from our own research (Scholl and Räz, 2013) will serve to illustrate some of the moves we think are appropriate for iteratively evaluating and revising philosophical concepts and for interpreting historical episodes. We are not holding up our paper as particularly significant, much less do we think the paper should compel universal assent. Its usefulness as an example rests to a large extent on its ordinariness: The goal is to illustrate a methodological approach, but this is largely independent of the truth of the philosophical and historical theses defended.

The starting point of our project was a paper by Michael Weisberg (2007), who had proposed a distinction between different types of theoretical practices in science: modeling and "abstract direct representation" (ADR). In order to illustrate modeling, Weisberg used Volterra's work on predator-prey dynamics; to illustrate ADR, he turned to Darwin's work on the origin of coral atolls and Mendeleev's work on the periodic table of the elements. While we agreed that theorizing is heterogeneous, we were interested in how the scientists themselves understood the subcategories of their methodological practices, and so turned to the historical sources.

5.1 Agreement

To investigate Volterra's own views on method, we turned to Volterra and D'Ancona (1935), which presents a monograph-length exposition of the predator-prey model. Perhaps as a reaction to the critics of earlier publications of the basic model (Volterra, 1926a,b, 1928), the monograph contains a methodological discussion as a preface. These methodological reflections support the idea that there is a contrast between more "direct" and more "indirect" theoretical practices. Volterra and d'Ancona place their own work on predatory-prey dynamics in the latter subcategory. They label their own indirect approach as "deductive", but we argue (with Weisberg) that it should be understood in terms of our present-day notion of modeling. Thus, we found an agreement between our philosophical thesis and the historical case: a distinction between modeling and more "direct" theoretical practices was suggested by the scientists themselves.

The finding of an agreement is good news for our philosophical thesis (that model-based science is "indirect" theoretical practice that should be distinguished from more "direct" practices) and for our understanding of the historical episode.

Yet it is only a first step: We must immediately ask whether the agreement is the result of selection bias such that the historical episode is, in truth, a mere illustration (as opposed to a study) of a pet philosophical thesis. Is the case hard for our thesis? Is it paradigmatic of some aspect of science? (We would argue that the predator-prey model is indeed a paradigmatic case of model-based science.) Do we have reasons for believing that the thesis, although borne out by the case, might apply only locally? In general, agreement is nothing but a sign that we must now investigate the range of applicability of our thesis (see section 5.5).

5.2 Conflict

Next, we reexamined one of Weisberg's two examples of ADR: Charles Darwin's explanation of the origin and structure of coral reefs and atolls in the Pacific (Darwin, 1842). Weisberg understood this as a case of ADR, because "at all times, Darwin was talking about the actual atolls in the Pacific" (2007, p. 228). We disagreed with this assessment. Darwin had to examine the effects of the interaction between known processes: the subsidence of islands and the growth of corals, processes which he investigated in some detail during his journey on the *Beagle*. However, it was impossible for him to observe this interaction directly, because the process extends over centuries. Darwin thus had to resort to a mental model of the process and its consequences: This allowed for the indirect investigation of the genesis and structure of coral atolls that is typical of model-based science. We thus reclassified Darwin's explanation as a case of modeling, extending the category to an additional, less obvious case of non-mathematical modeling.

We must be careful about the conclusions drawn from this conflict. The first instinct is to conclude that the thesis of ADR has now been rejected, as practiced in the popular philosophical game of counterexamples. However, the appropriate thing to do is simply to reassign Darwin to the modeling category without making judgments about the existence or range of applicability of ADR, which the reevaluation of an individual case does not permit.

5.3 Incompleteness

After determining agreements and conflicts between existing concepts and cases, we may discover that there are some issues for which do not have any conceptual tools at all. This sort of incompleteness may of course indicate that our existing concepts need to be expanded or generalized by strenuous tinkering, but that strikes us as a problematical "winner takes all" mentality. Incompleteness may just as well indicate that a new concept needs to be developed in addition to those we already have.

In the example in hand, Volterra's methodological discussion confirmed to us the existence of an incompleteness in Weisberg's scheme of theoretical practices. Volterra and d'Ancona preferred to contrast model-based science with a more "direct" theoretical investigation in terms not of ADR but of experimental causal inference. The authors write that for their own investigation they would have *preferred* an experimental approach, which would have allowed for direct causal inferences in the system under scrutiny. However, this was unfeasible for practical reasons having to do with the size and time scales of ecological systems. In light of this, we added the subcategory of causal inference as an additional theoretical practice. While the historical case did not suggest the *existence* of this type of theorizing (since philosophers have studied it for a long time), the sources certainly urged that causal inference as a theoretical practice is a relevant *contrast* to model-based science.

Instances such as this one illustrate the *generative* role of historical cases. By turning to the historical sources, we were able to recognize a natural subcategory – one of considerable philosophical interest – that needed to be added to Weisberg's initial range of theoretical practices.

5.4 Redundancy

The complement to incompleteness is redundancy, where we may find that we have developed elaborate philosophical concepts that are not applicable to *any* part of science. There is a philosophical prejudice in such cases for thinking that the concepts are fine but that the search for good examples goes on. We would suspect, however, that a failure to find good instances of a concept is an indication that the concept is either not well developed or wholly misguided.

In the present example, the historical cases led us to suggest a redundancy. We were convinced that Darwin's work on coral atolls did not fit in the subcategory of ADR, while other scholars had already argued that Mendeleev's discovery of the periodic table should be understood differently as well.[8] We took this as an indication that the entire subcategory of ADR may be superfluous (unless, of course, additional instances can be presented).

5.5 Ranges of applicability

In all the steps discussed above, it is useful to keep thinking about the range of applicability of a thesis. An agreement may indicate not that a philosophical approach is correct overall, but only that it captures a subset of cases well, while other concepts may apply to other subsets. Similarly, a supposed counterexample may indicate not that a thesis is false, but only that our case is outside its range of ap-

[8] See Scerri (2012) for a critique of Weisberg's interpretation of Mendeleev's work.

plicability. Thus, we concluded that Volterra's work on predator-prey dynamics did indeed fit in Weisberg's subcategory of model-based science, while Darwin (as explained in section 5.2) had to be removed from the ADR subcategory and moved to modeling. An examination of the Mendeleev case led us to remove that case from the subcategory of ADR as well.[9] Finally, we immediately recognized cases that fit in the newly added subcategory of causal inference, as suggested by Volterra's own methodological discussions. For instance, Semmelweis's investigation of the cause of puerperal fever (Scholl, 2013) has long been discussed as a paradigm case of causal inference. Finding multiple instances of a single subcategory can be taken as an indication that a natural and effective subcategory has been found.

5.6 Summary

Just as the same historical episode can simultaneously be a hard case, a paradigm case and an big case (depending on the philosophical concepts under study), confrontation will rarely involve just one of the possibilities listed above: as we investigate the relationship between historical cases and philosophical theses, we will usually make several or all of the moves discussed. Some concepts will be in agreement with the historical case, while many will be in various degrees of conflict; some of these conflicts will be resolved by adjusting ranges of applicability, while others will require us to consider incompleteness or redundancy. As discussed above, we see the process of confrontation as cyclical, where a combination of the moves outlined above will occur over multiple iterations. For the case discussed here, we summarize the cyclical revisions in figure 2.

6 Conclusions

We have discussed a fundamental motivation for integrated history and philosophy of science, and for case-study-based philosophy of science more generally. Philosophical concepts are a useful resource for asking questions about the historical or contemporary practice of science: They provide particularly pertinent question that relate to science's core epistemic project. Conversely, historical and contemporary science provides the empirical basis for a naturalized philosophy of science, which should be properly conceived as a theory of science in analogy to theory construction in other empirical disciplines.

Moreover, we have examined the dilemma of case studies as a methodological challenge for integrated HPS. First, how can we minimize the danger of selection bias in choosing our case studies? How can we choose case studies that play a pro-

[9] For the time being we refrained from assigning Mendeleev to any of the other subcategories, although Scerri (2012) suggested "classification" – which we should presumably count among theoretical practices.

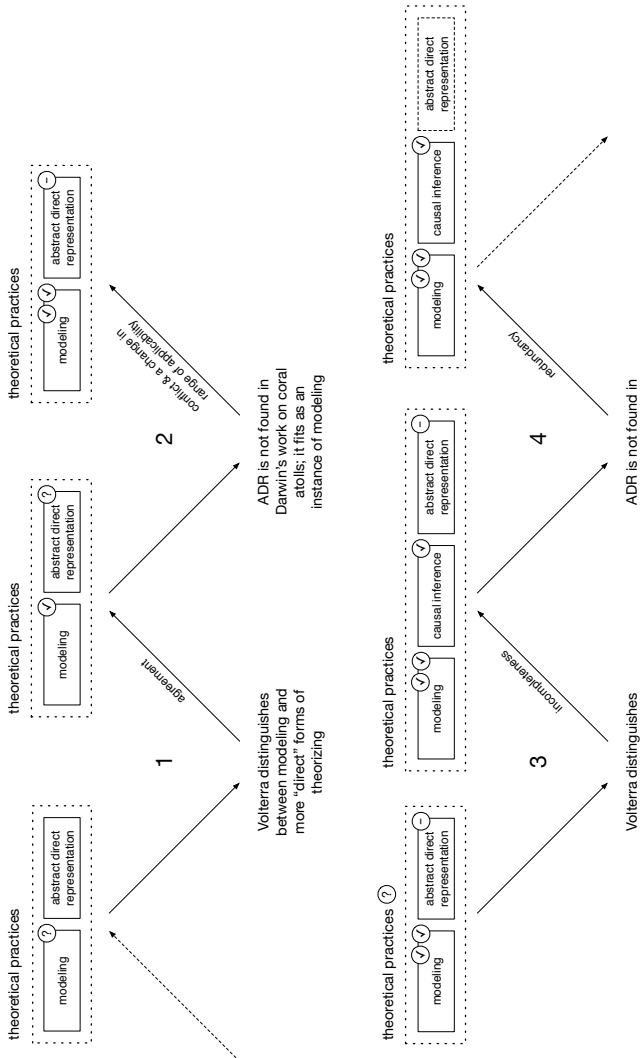

Fig. 2 Cyclical HPS and the exploration of theoretical practices. We begin with Weisberg's notion that two kinds of theoretical practices should be distinguished: modeling and abstract direct representation or ADR.

In cycle 1, we find an **agreement** between concepts and case: Volterra, in his work on predator-prey dynamics, recognizes a type of theorizing that corresponds to modeling.

In cycle 2, we test whether Darwin, in his work on coral atolls, follows a theoretical practice that corresponds to Weisberg's ADR. We find a **conflict** between concepts and case. Based on textual evidence, we adjust the modeling subcategory's **range of applicability** to include Darwin's work on coral atolls, giving modeling a second instance.

Next, in cycle 3, we recognize an **incompleteness**: Volterra's chosen contrast for model-based science is experimental causal inference rather than ADR.

Finally, cycle 4 suggests that the concept of ADR is **redundant** since it does not apply to Weisberg's second suggested instance either: Mendeleev's work on the periodic table.

bative instead of merely illustrative role? We have suggested that an initial remedy is the explicit discussion of criteria for choosing a case study, and of the function that the case study will play vis-à-vis the philosophical concepts under study. In particular, we have outlined four main types of case studies: Hard cases, which are chosen such that they challenge rather than illustrate the relevant philosophical concepts; paradigm cases, which are taken to be typical of some aspect of science; big cases, which concern particularly influential or otherwise important scientific findings; and randomized cases, where selection bias is minimized by leaving selection to chance.

Second, how should our conceptual engineering proceed when we are confronting philosophical concepts with historical cases? We have argued that the main mistake is to look for concepts that are too general, and to be content with their facile confirmation or rejection. A disciplined pluralism appears to pay dividends: We must ask about all the main categories of concepts about science (such as confirmation, explanation, and many more) whether they can be understood in terms of a number of subconcepts. Much of the work will consist in naming, elaborating and applying these subconcepts. In this procedure, individual cases do not usually confirm or disconfirm broad categories of concepts (such as "mechanistic explanation"). Instead, agreements and conflicts between concepts and cases urge us to assess the concepts' ranges of applicability. If our existing concepts prove inadequate to the understanding of historical cases, this incompleteness will challenge us to create new concepts. In this way, historical cases are more than just a testbed for philosophy. They also play a crucial generative role.

In 1992, David Hull asked philosophers of science for renewed ambition:

> Although grand theories about the nature of science are currently out of fashion, I think that we need to rehabilitate them. We need to construct theories about science the way that scientists construct theories about fluids, gene flow and continental drift. (Hull, 1992, p. 473)

We agree with Hull's ambition and with his naturalistic approach, but his notion of "grand theories" requires clarification. In line with the above discussion, we should not expect to find *the* grand theory of science, akin to Popper deriving all aspects of proper science from *modus tollens*. Rather, we should expect theories of the kind of we find in most special sciences: a series of overlapping models (to use Giere's term) whose interaction and integration with each other are themselves important questions for study. In brief, grand theories are not necessarily unified theories, and pluralists need not fear them.

Many important issues are left unanswered by our discussion. Our types of case studies are likely not exhaustive: We may well find a range of further types, with distinct functions in HPS, that we have not yet considered. And similarly, much of what we have said about the dynamics of confronting concepts and cases remains to be made more precise and to be extended by further procedures. Moreover, there are a number of important issues that we have only touched upon. For instance, case studies usually leave a certain room for interpretation. Different scholars may see different and contradictory concepts instantiated in the same episode, even when there is no malicious intent to misrepresent. The problem is defanged somewhat by

explicit criteria for choosing case studies, and by an awareness of the conceptual engineering involved in confronting concepts with cases. Nevertheless, the underdetermination problem of integrated HPS and its relationship to methodology deserves much further discussion (see also Franklin and Collins, this volume; Kinzel, this volume; and Räz, this volume).

To conclude, integrated history and philosophy of science promises a unique understanding of the scientific enterprise. However, we agree with Hasok Chang, who writes:

> I believe that the neglect to clarify the nature of the history-philosophy relationship in case-studies has contributed decisively to a widespread disillusionment with the whole HPS enterprise. (2011, p. 109)

We expect that a renewed methodological discussion can bring about the needed clarification. The effort promises to pay off on both the philosophical and the historical front: we stand to gain better general theories of science and a more adequate understanding of its history.

Acknowledgements For helpful comments and debates, which have shaped this contribution significantly, we thank first and foremost the participants of the workshop "The philosophy of historical case studies" at the University of Bern (November 21–22, 2013). In addition, we are indebted to Michael Bycroft, Allan Franklin and Jutta Schickore; the participants of the Fifth Conference on Integrated History and Philosophy of Science (&HPS5) at the University of Vienna (June 26–28, 2014); the Visiting and Postdoctoral Fellows at the University of Pittsburgh's Center for Philosophy of Science (2014–2015); and the members of the Lake Geneva Biology Interest Group (lgBIG). Raphael Scholl was supported by a grant from the Swiss National Science Foundation (SNSF), grant number P300P1_154590.

References

Arabatzis T, Schickore J (2012) Ways of integrating history and philosophy of science. Perspectives on Science 20(4):395–408

Bechtel W (2006) Discovering cell mechanisms: The creation of modern cell biology. Cambridge University Press

Bird A (2010) Eliminative abduction: examples from medicine. Studies in History and Philosophy of Science 41:345–352

Burian RM (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

Burian RM (2002) Comments on the precarious relationship between history and philosophy of science. Perspectives on Science 10(4):398–407

Chang H (2004) Inventing temperature: Measurement and scientific progress. Oxford University Press

Chang H (2011) Beyond case-studies: History as philosophy. In: Mauskopf S, Schmaltz T (eds) Integrating history and philosophy of science, Springer, pp 109–124

Darden L (2006) Reasoning in biological discoveries. Cambridge University Press

Darwin C (1842) The Structure and Distribution of Coral Reefs. Smith, Elder and Co., London

Daston L, Galison P (2007) Objectivity. Zone Books

Feest U, Sturm T (2011) What (good) is historical epistemology? Editors' introduction. Erkenntnis 75(3):285–302

Franklin A, Collins H (in press) Two kinds of case study and a new agreement. In: Sauer T, Scholl R (eds) The philosophy of historical case studies, Boston Studies in the Philosophy and History of Science

Giere RN (1973) History and philosophy of science: intimate relationship or marriage of convenience?

Giere RN (1985) Philosophy of science naturalized. Philosophy of Science 52(3):331–356

Giere RN (2011) History and philosophy of science: Thirty-five years later. In: Mauskopf S, Schmaltz T (eds) Integrating History and Philosophy of Science, Springer, pp 59–65

Gillies D (2005) Hempelian and Kuhnian approaches in the philosophy of medicine: the Semmelweis case. Studies in History and Philosophy of Biological and Biomedical Sciences 36(1):159–181, DOI 10.1016/j.shpsc.2004.12.003

Graßhoff G, May M (1995) Methodische Analyse wissenschaftlichen Entdeckens. Kognitionswissenschaft 5:51–67

Hanson NR (1962) The irrelevance of history of science to philosophy of science to philosophy of science. The Journal of Philosophy 59(21):574–586

Hempel CG (1966) Philosophy of natural science. Prentice Hall

Howard D (2011) Philosophy of science and the history of science. In: French S, Saatsi J (eds) The Continuum Companion to the Philosophy of Science, Continuum, pp 55–71

Hull D (1992) Testing philosophical claims about science. In: PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, pp 468–475

Kinzel K (in press) Pluralism in historiography: A case study of case studies. In: Sauer T, Scholl R (eds) The philosophy of historical case studies, Boston Studies in the Philosophy and History of Science

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 JG (2001) History and philosophy of science: A phylogenetic approach. História, Ciências, Saúde-Manguinhos 8(3):655–669

Lipton P (2004) Inference to the best explanation. Routledge

Machamer P, Darden L, Craver CF (2000) Thinking about mechanisms. Philosophy of science 67(1):1–25

Norton JD (2003) A material theory of induction*. Philosophy of Science 70(4):647–670

Orgel L (1999) Are you serious, Dr Mitchell? Nature 402(6757):17–17

Pinnick C, Gale G (2000) Philosophy of science and history of science: A troubling interaction. Journal for General Philosophy of Science 31(1):109–125

Pitt JC (2001) The dilemma of case studies: toward a heraclitian philosophy of science. Perspectives on Science 9(4):373–382

Räz T (in press) Gone till november: A disagreement in Einstein scholarship. In: Sauer T, Scholl R (eds) The philosophy of historical case studies, Boston Studies in the Philosophy and History of Science

Rheinberger HJ (1997) Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube. Stanford University Press

Richards RJ (1992) Arguments in a sartorial mode, or the asymmetries of history and philosophy of science. In: PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, pp 482–489

Sauer T, Scholl R (in press) Gone till november: A disagreement in Einstein scholarship. In: The philosophy of historical case studies, Boston Studies in the Philosophy and History of Science

Scerri ER (2012) A critique of Weisberg's view on the periodic table and some speculations on the nature of classifications. Foundations of Chemistry 14(3):275–284

Schickore J (2011) More thoughts on HPS: Another 20 years later. Perspectives on Science 19(4):453–481

Scholl R (2013) Causal inference, mechanisms, and the Semmelweis case. Studies in History and Philosophy of Science Part A 44(1):66–76

Scholl R, Nickelsen K (2015) Discovery of causal mechanisms: Oxidative phosphorylation and the Calvin-Benson-cycle. History and Philosophy of the Life Sciences 37(2):180–209

Scholl R, Räz T (2013) Modeling causal structures. European Journal for Philosophy of Science 3(1):115–132
Semmelweis IP (1983) The etiology, concept, and prophylaxis of childbed fever. University of Wisconsin Press, Madison, translated and edited by K. Codell Carter.
Shapin S, Schaffer S (1985) Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life. Princeton University Press
Sulloway FJ (1982) Darwin's conversion: The beagle voyage and its aftermath. Journal of the History of Biology 15(3):325–396
Tulodziecki D (2013) Shattering the myth of Semmelweis. Philosophy of Science 80(5):1065–1075
Volterra V (1926a) Fluctuations in the abundance of a species considered mathematically. Nature 118(2972):558–560
Volterra V (1926b) Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. Memorie della R Accademia dei Lincei 6(2):31–113
Volterra V (1928) Variations and fluctuations of the number of individuals in animal species living together. J Cons int Explor Mer 3(1):3–51
Volterra V, D'Ancona U (1935) Les associations biologiques au point de vue mathématique. Paris: Hermann
Weisberg M (2007) Who is a modeler? The British Journal for the Philosophy of Science 58(2):207–33