Doing Science, Writing Science*

Jutta Schickore†‡

This article identifies a fundamental distinction in scientific practice: the mismatch between what scientists do and what they state they did when they communicate their findings in their publications. The insight that such a mismatch exists is not new. It was already implied in Hans Reichenbach's distinction between the contexts of discovery and justification, and it is taken for granted across the board in philosophy of science and science studies. But while there is general agreement that the mismatch exists, the epistemological implications of that mismatch are not at all clear. Philosophers, historians, and sociologists of different stripes have expressed widely different views about how one should understand and interpret the relation between what scientists do and what they state they did. This article surveys a number of approaches to the mismatch. Based on this survey, I offer an assessment of the epistemological significance of the mismatch and identify the major meta-epistemological challenges that it poses for the analysis of scientific practice.

1. Introduction. There are two complementary aspects to scientific activity: Scientists observe, perform experiments, make measurements, interpret data, draw up models, and formulate theories; and they communicate their findings and insights to others beyond their immediate research groups. In their presentations—papers in proceedings, journal articles, text books, and more—researchers usually do not tell their readers in any detail how they obtained their results. Rather, they present their findings embedded in a web of arguments and reasons, thereby changing the order and justification of their research steps. In short, there is a mis-

*Received May 2007; revised April 2008.

†To contact the author, please write to: Department of History and Philosophy of Science, 1011 East Third Street, Goodbody Hall 130, Indiana University, Bloomington, IN 47405; e-mail: jschicko@indiana.edu.

‡A version of this paper was presented in the HPS Reading Group at the Department of HPSC, IU Bloomington, in spring 2007. I thank the audience as well as Domenico Bertoloni Meli, Jordi Cat, Uljana Feest, Jan Frercks, John Johnson, and two anonymous referees for this journal for their critique and helpful suggestions. The final version of this article was produced while I was a member of the Institute for Advanced Study (Princeton). Generous support from the Mellon Foundation is gratefully acknowledged.

Philosophy of Science, 75 (July 2008) pp. 323–343. 0031-8248/2008/7503-0004$10.00
Copyright 2008 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/592951 Published online by Cambridge University Press

match between what scientists do and the presentation of results in scientific publications.

The insight that there is such a mismatch is not new. For instance, it was a driving force of the famous (or rather infamous) distinction between the contexts of discovery and justification Hans Reichenbach put forward in his 1938 book *Experience and Prediction*. Until today, it remains unclear whether this distinction should be kept or abandoned.[1] In contrast, the mismatch between what scientists do and what they state they did is never questioned. On the contrary, it is taken for granted across the board in philosophy of science and science studies. But as I show in my article, the epistemological implications of that mismatch are not at all clear, and philosophers, historians, and sociologists of science have expressed widely different views about it.

My essay has three parts. I begin with a brief characterization of Reichenbach's position. The main part of the article surveys a number of approaches to the mismatch between scientific practice and its presentation and draws out the differences between the various interpretations of it. Based on this survey, I then offer an assessment of the epistemological significance of the mismatch. I conclude with suggestions as to how to approach the mismatch in epistemological analysis.

2. Reichenbach's Distinction. In *Experience and Prediction*, Reichenbach motivates the distinction between the contexts of discovery and justification by highlighting the "well-known" difference between scientists' actual thinking processes ("context of discovery") and the presentation of a theory before a public ("context of justification"). In the course of their communication, scientific results are reformulated in such a way that logical analysis can be performed on them (Reichenbach 1938, 6). Reichenbach ties this mismatch between what scientists do and the reconstruction presented in the paper to a disciplinary distinction and at the same time to the distinction between normative and descriptive attitudes to science. According to Reichenbach, it is the task of psychology, sociology, and other empirical sciences to study scientists' actual thinking processes. This descriptive project aims at empirical accuracy. Epistemology, by contrast, is concerned with the context of justification.[2]

Two tasks arise for epistemology, one descriptive, one critical. Epistemology's descriptive task is to provide an account of the scientists' think-

1. On the history, context, and significance of the context distinction, see Schickore and Steinle 2006.

2. For details on Reichenbach's distinction and its contexts, see Howard 2006; Richardson 2006; Schiemann 2006.

https://doi.org/10.1086/592951 Published online by Cambridge University Press

ing processes. This is a *descriptive* task because it is concerned with thinking that has actually happened, but not necessarily in the sense that a comprehensive account of the order of events is given. Providing a descriptive account of the scientists' thinking processes is a task *for epistemology* in that only the epistemic aspects of the actual practices are considered and presented in an ordered fashion. Reichenbach regards this reformulation as a "rational reconstruction" of research results. It is the task of descriptive epistemology to reconstruct these thinking processes in such a way that logical analysis can be performed on them. Scientists and philosophers both contribute to this task, the scientists providing the initial reconstruction, the philosophers improving on it. The scientists' part is to present their thinking processes with an eye to validation. Epistemologists bring the reconstruction to completion. According to Reichenbach, epistemology also provides the principles of rational reconstruction. In accordance with his general philosophical position, the logical principles for the rational reconstruction he laid out in his work were inductivist. Revealing in their light the logical flaws and shortcomings of scientists' reconstructions of their activities is the critical task of epistemology.

Finally, Reichenbach identifies an "advisory task" for epistemology, which consists in recommendations concerning decisions about conceptual frameworks and conventions. Reichenbach calls these decisions "volitional"; it is up to the scientist to make them and epistemology can merely help clarify the options (see Richardson 2006, 44). In effect, the advisory task collapses into the critical task. The philosopher cannot do more than identify the complete set of possible decisions open to the scientists as well as the entailed decisions that follow from each. The actual decisions must be left to the scientists. Reichenbach's ultimate concern is with the critical and advisory tasks of epistemology. According to Reichenbach, philosophy of science should lay down the guiding principles for the exposition and validation of knowledge claims and perform a rational reconstruction of scientists' actual claims to knowledge. The rational reconstruction of scientific reasoning processes is the epistemologically significant part of scientific practice, and the epistemological aim is to improve the presentation of scientific ideas.

Many twentieth-century philosophers adopt a distinction similar to Reichenbach's, but usually in a somewhat simplified version, collapsing Reichenbach's manifold distinction into a twofold one. Disregarding Reichenbach's point about scientists' reconstructive efforts as an integral part of descriptive epistemology, most philosophers in Reichenbach's wake have concentrated on the regulative function of philosophy of science. The basic assumptions tied to this kind of distinction are the following: The scientists' actual reasoning processes to develop a theory have no bearing

https://doi.org/10.1086/592951 Published online by Cambridge University Press

on its appropriate conceptual structure and testing. According to advocates of this view, the knowledge claims scientists expose when they communicate their final results are the only part of scientific practice that is amenable to philosophical assessment. The philosophical reconstruction considers only the fully developed theory and its relation to the facts.

This article is not concerned with the fate and significance of the discovery-justification distinction. Instead, I consider one of Reichenbach's motives for conceiving it, namely the insight that the scientists' published accounts of their research provide an idealization, a "rational reconstruction" of what happened but never correctly represent the actual events that led up to the results obtained.

Since *Experience and Prediction*—and usually without reference to Reichenbach—numerous historical, philosophical, and sociological studies of science have produced strong evidence that there is indeed a mismatch between the actual process of research (what is going on, say, in the lab) and the accounts by which the results of that research are exposed and defended as valid (say, in an article in a scientific journal). There is no doubt that the published accounts of scientific activities do not in fact report these activities exactly as they occurred but recast the events, replacing the actual steps that were undertaken with operations that can be demonstrated as valid. Such reconstructions may change the temporal order of the events, "rationalize" the motives why certain experiments were performed or substitute entirely new reasons, reorganize the data sets, and recast evidential relations between data and hypotheses. But is this distinction epistemologically significant, and if so, what is its significance?

In the next section, I survey a wide range of approaches to the mismatch.[3] The survey shows general agreement among philosophers, historians, and sociologists of different stripes about the existence of the mismatch, but fundamental disagreement about how one should understand and interpret the relation between what scientists do and what they state they did. The interpretation of this relation raises major meta-epistemological questions, which I draw out and address in the final section of the article.

3. I should add that my survey is not comprehensive: I picked characteristic representatives of laboratory anthropology, history of scientific practice, and so forth. Of course, I do not deny that there are considerable differences between the positions of, say, different laboratory anthropologists, for instance, between Karin Knorr-Cetina's and Bruno Latour's views (or between the views of the earlier and later Latour, for that matter). But here I am more concerned with highlighting differences between such projects on a larger scale.

https://doi.org/10.1086/592951 Published online by Cambridge University Press

3. Approaches to the Mismatch.

3.1. Medawar's Distinction. Since the days of Reichenbach, many scholars of different persuasion have addressed the mismatch between actual scientific activity and its presentation in the scientific paper. Twenty-five years after Reichenbach, Peter Medawar again highlights the mismatch between scientists' thinking processes and their communication to the public in a famous BBC talk tellingly titled "Is the Scientific Paper a Fraud?" It is this article rather than Reichenbach's context distinction that has been most frequently quoted as a source of inspiration for thinking about the mismatch. Medawar argues that scientists purposely misrepresent their research activity in their publications. He claims that they do so in order to adhere to the methodological demands made, inappropriately so, by philosophers of science. The scientist Medawar states that to comply with philosophical demands, scientists commonly render their research in inductive terms. According to Medawar, this reflects the conception of scientific discovery that we owe "the work of a great and wise but in this context, I think, very mistaken man—John Stuart Mill" (Medawar [1963] 1996, 34). In their actual research practice, however, scientists do not proceed inductively but follow the hypothetico-deductive model (37). For Medawar, the inductivist framework is a "ludicrous" straitjacket for the scientist. The scientific paper is "a fraud in the sense that it does give a totally misleading narrative of the processes of thought that go into the making of scientific discoveries" (38). Medawar does not tell his readers why, in his opinion, philosophers have been able to assume a position from which they can dictate scientists' reconstructions. He does claim that the "inductive format of the scientific paper should be discarded" (38). The scientific paper should be organized in such a way that it reflects the actual thinking procedures, and philosophical accounts of scientific practice should also capture those thinking processes. Medawar notes that scientists begin with hypotheses, guesses, and expectations that are subsequently put to test. Not surprisingly, he urges that instead of Mill or other inductivists, William Whewell and Popper assume the role of the scientists' philosophical muse.

The comparison between Medawar's and Reichenbach's positions is instructive. Both agree that there *is* a mismatch between actual thinking processes and their expositions and that scientists put together the account of scientific results with an eye to philosophical models of testing and confirmation. Both characterize the dominant philosophical models as inductive. They differ, however, in the evaluation of the nature and role of the reconstruction that encapsulates the results. For Reichenbach, who endorses inductivist philosophy, inductive principles are the norm for epistemological reconstructions, and this norm overrides, as it were, both

https://doi.org/10.1086/592951 Published online by Cambridge University Press

the scientists' actual thinking practice and the scientists' own exposition of their thinking practice in the scientific paper. The scientific paper does not, and indeed should not, reflect that thinking practice. It should, ideally and ultimately, stand up to philosophical scrutiny.

Medawar, by contrast, not only is no inductivist but also reverses the relation between scientists' argumentative practices, the scientific paper, and philosophical reconstructions. For Medawar, scientists' actual thinking practice overrides the philosophical principles for reconstruction that inform the scientific paper. Contrary to Reichenbach, Medawar appears to believe that the gap between actual thinking practice and its presentation should be bridged. He assumes that scientists conduct their research in accordance with an identifiable underlying method: Scientists propose hypotheses and proceed to test them. Scientific papers can and should provide an accurate portrayal of these methods. The implicit assumption is that philosophy of science should epitomize the essential features of scientific practice. Medawar acknowledges that available philosophical frameworks may come in handy as tools for devising an adequate descriptive account of scientists' thinking practices. He turns to Whewell and Popper for conceptual resources. But the scientists, not the philosophers, judge whether the philosophical account of scientific reasoning practices is adequate; and if there is a tension, the philosophical account is the one that requires adjustment and improvement.

3.2. Lab Anthropology and the Mismatch. About two decades after Medawar's BBC talk, sociologists of science began to carry out empirical studies of laboratory life. These studies of scientific practice reinforce the point that had long been made by traditional philosophy of science, namely that philosophical methodologies had no grounding in actual laboratory practice. Notably, the earliest sociological studies of scientific research practice are explicitly motivated by Reichenbach's distinction between the contexts of discovery and justification. In his 1976 article "The Transformation of Research Findings into Scientific Knowledge," for instance, sociologist of science Nigel Gilbert laments that Reichenbach's distinction confined sociological study to the domain of discovery (1976). He states that it was time for sociologists to reclaim the domain of justification and to study "the procedures actually used by natural scientists to decide on the validity of claims to scientific knowledge" (282). He then sets out to reconstruct the career of a claim from astronomy about the velocity of meteors, using contemporary philosophy of science, in particular, Popper's and Lakatos's models of scientific testing and Kuhn's conception of paradigms.

The lab anthropologists who publish their studies of laboratory life around 1980 no longer link their work to Reichenbach. But they often

https://doi.org/10.1086/592951 Published online by Cambridge University Press

explicitly refer to Medawar's paper. Empirical studies of scientists' lab activities provide ample evidence for the fact that scientists' reasoning and experimental practice do not follow any recognizable pattern. "Lab anthropologists" like Karin Knorr-Cetina describe lab work as messy, highly contextualized, opportunistic tinkering with all kinds of resources that happen to lie about and contrast this activity with the neat, orderly, and systematic devising and testing of hypotheses (1981, Chapters 2–3). For Knorr-Cetina, the lab is the principal site of validating and credentialing scientific work (1981, 8–9). Judged by scientists' lab activities, scientific reasoning appears entirely pragmatic and opportunistic and driven by occasioned selections.

Such accounts of laboratory life follow Reichenbach and Medawar to the extent that the mismatch between scientific reasoning practices and their presentation is also regarded as the result of a conscious effort of scientists to reconstruct their research. But the reconstruction is not done with an eye to philosophy (for good or evil) as Reichenbach and Medawar had it. Rather, it is done with an eye to the fellow scientists who provide a more or less skeptical audience that needs convincing. Laboratory research and the production of research papers proceed "*with a view toward* potential criticism or acceptance (as well as with respect to potential allies and enemies!)" (Knorr-Cetina 1981, 7). Explicitly taking the cue from Medawar, Knorr-Cetina argues that the research paper that scientists produce transforms through several stages of revisions "laboratory reasoning" into literary reasoning. These transformations include the reversal of the order of the research steps and the recontextualization and motivation of research efforts (Knorr-Cetina 1981, Chapter 5).

Bruno Latour and Steve Woolgar's *Laboratory Life* ([1979] 1986) and Latour's sequel *Science in Action* (1987) underline the distinction between laboratory practice and the presentation of knowledge claims in scientific articles that Knorr-Cetina draws. Latour and Woolgar also quote Medawar (among others) to make the point that scientists' presentations "systematically conceal the nature of the activity which typically gives rise to their research reports" (Latour and Woolgar [1979] 1986, 28), and that the observation of scientific activity is necessary to elucidate the nature of scientific facts. Like Knorr-Cetina, Latour and Woolgar analyze the processes by which laboratory activities are transformed into reports. Their focus is on the linguistic activity of "performing operations on statements." For Latour and Woolgar, this is the principal feature of laboratory activity, regardless of whether these statements are part of a paper or part of an oral communication. They distinguish five types of statements with respect to their degree of "facticity" with conjectures and speculations at one end of the spectrum and assertions of commonly known facts at the other. Each type of statement is associated with a

https://doi.org/10.1086/592951 Published online by Cambridge University Press

particular context of communication: Speculations and conjectures, for instance, typically occur in private conversations, while assertions of accepted facts ("type 4 statements") appear primarily in textbooks (Latour and Woolgar [1979] 1986, 75–78).

In *Science in Action*, then, Latour stresses that scientists write their papers, first and foremost, with an eye to winning over an audience. He presents this audience as skeptical, belligerent, and ready to attack the author. The literary strategies employed in scientific papers are means to gain allies and build strongholds. They include references to authorities and other texts as well as the presence of numbers, equations, and figures. They bolster the claim to be advanced because the attacker must also overpower the authorities and challenge the numbers and figures. In this sense, literary strategies help the authors draw up accounts that can "withstand the assaults of a hostile environment" (Latour 1987, 45).

3.3. Historians of Science and the Mismatch. Like the laboratory anthropologists, many advocates of the "practical turn" of history of science—the turn away from the history of ideas toward the history of experimental practice—refer to the mismatch between research activity and the reconstructions offered in scientific publications. In doing so, they warn against mistaking the study of published papers for a source of information about how knowledge is generated in experiments. Larry Holmes, for example, reminds us—again, with a bow to Medawar—that the insights into scientific practice garnered from reading the scientists' published work are fundamentally incomplete precisely because these publications are retrospective reconstructions of research activities (Holmes 1987, 220). Since the research reports do not preserve the temporal order or logic of the thinking processes nor the steps of the experimental pathway, they do not provide any dependable information about the nature of scientific practice.

Larry Holmes makes this point most forcefully with his studies on Antoine Lavoisier and Claude Bernard. He also makes an additional point that calls into question one tacit assumption pertaining to the mismatch, namely that the research process is neatly divided into two successive steps. The generation of results and the construction of the argument do not follow one another. Rather, they are two aspects of one and the same research activity. For instance, Holmes's archival work on Lavoisier shows that writing a scientific article was, for Lavoisier, an integral part of the creative research process: Lavoisier's attempts to put together a draft of a paper changed his investigative pathway (Holmes 1987, 222–223).

Studies of scientific papers in the early modern period also can be read as variations of the theme of the mismatch, and they add to the discussion in significant ways. In their famous study of seventeenth-century experimental philosophy *Leviathan and the Airpump*, Steven Shapin and Simon

https://doi.org/10.1086/592951 Published online by Cambridge University Press

Schaffer (1985) find a rather close connection between research activities and published reports. They point out that the papers published by experimental philosophers working in the context of the Royal Society did provide particulars of the scientists' activities. These reports described in detail concrete experiments performed at specific times and places. However, Shapin and Schaffer do not deny that there is an element of reconstruction even in the most circumstantial report, and like the lab anthropologists, they suggest that the reconstruction is made with a view to peer audiences to gain their approval. In writing a circumstantial report, researchers employ a literary strategy: Providing minute details serves as a tool to convince an audience that the experiment had indeed taken place as reported. The identification of the witnesses present, the depiction of the instruments used, the detailed account of the circumstances of the experimental procedure, and the account of failed experiments are additional literary strategies for credentialing the experimental results.

In response to Shapin and Schaffer's work, Holmes and historian Peter Dear have drawn attention to other kinds of early modern research papers that were not detailed reports of particulars. Research papers produced at the French Academy of the Sciences around 1700, for instance, combined elements of arguments and elements of narrative in varying proportions. Some papers do describe in great detail the course of an investigation; in others, the investigation is "all but buried in an argument" (Holmes 1991, 177). Holmes suggests that these differences reflect the structure of the community that is addressed (the more background is shared, the less details need to be reported), the intellectual context (the more contested the claim, the more argumentative the paper), and so on.

Dear's work on the literary structure of scientific arguments in the early modern period gives the subject a slightly different twist. According to Dear, the way in which experiments are presented in the experimental report reflects deep-seated assumptions about the nature of experience.[4] Galileo, for instance, utilized an Aristotelian-scholastic framework to account for his experiments with the inclined plane. He did not describe specific experiment carried out at a particular time and place. Rather, he claimed that he had "found the results to agree exactly with his theoretical assumptions—having, he says, repeated the trials "a full hundred times." In doing so, Galileo sought to establish that falling bodies usually behave in the way described. According to Dear, this assertion reflected the con-

4. Dear insists that the distinction between doing experiments and reporting is misleading. Only the experimental report brings the experiment to completion. Still, the point of interest for my argument is that he also insists on the difference between what scientists do and what they state they did.

https://doi.org/10.1086/592951 Published online by Cambridge University Press

temporary scholastic notion of experience: "the memory of *many instances*—a multiplicity of unspecified instances adding up to experimental conviction" (1991, 138). Dear thus reminds us that the epistemological concepts scientists use to communicate their experimental activities have a history.

3.4. Philosophy of Science, Again. In the late 1990s, then, Medawar and the mismatch return to philosophy of science. The 1998 issue of this journal contains a set of four articles on the structure of a scientific paper. In the opening essay, Frederick Suppe refers to Medawar as well as to subsequent studies of scientific writings by lab anthropologists and sociologists of science more generally. Suppe emphasizes that the refereed scientific journal article is the premier site of disseminating, evaluating, and credentialing scientific knowledge, and he goes on to express his astonishment about the fact that no philosopher has embarked on structural analyses of scientific papers. Because scientific papers reflect what scientists consider the valid logic of scientific research, it is the task of philosophy to provide an adequate account of that logic.

Suppe claims that philosophical analyses of scientific articles need to fulfill two requirements. They "ought to apply at the microlevel and be able to take into account everything in the write-up" (Suppe 1998, 382). According to Suppe, no available philosophical model of testing and confirmation can meet these requirements. Standard inductivism and hypothetico-deductivism as well as the more recent Bayesian inductive model and the inference to the best explanation (IBE) model "seriously garble" the logic of the scientific paper (Suppe 1998, 382).

Suppe supports his verdict with a study of W. Jason Morgan's seminal text on plate tectonics, "Rises, Trenches, Great Faults and Crustal Blocks" (Morgan 1968). Notably, this article was published only a few years after Medawar's BBC talk. Suppe offers a microanalysis of the architecture of the paper and draws particular attention to the ways in which Morgan attempted to impeach alternative hypotheses or interpretations of data (ultimately, Morgan ignores the data that does not fit). Suppe claims that both the hypothetico-deductive and the Bayesian inductive model cover only 30% of the paragraphs of the paper, and that all three models present only inadequate accounts of the parts of the argument that they do cover. Moreover, none of them provides an empirically adequate account of Morgan's treatment of those portions of data that are incompatible to his hypothesis. In particular, the hypothetico-deductive model does not accommodate the fact that Morgan's actual data disconfirms the hypothesis. The Bayesian inductive model disregards that the reasoning in the paper is not probabilistic; and the IBE model disregards the fact that only one hypothesis is considered and no comparative evaluation is made.

https://doi.org/10.1086/592951 Published online by Cambridge University Press

The critics do not agree with Suppe's verdict. In his response to Suppe, advocate of the IBE model Peter Lipton reinforces Suppe's claim about the importance of exculpatory explanations in science. But he argues that the IBE model illuminates rather than obscures the role of impeachment of recalcitrant data (Lipton 1998). In the second response, two Bayesians, Franklin and Howson, refer to their previous work to argue that Bayesian logic can capture the argument of an experimental paper and that Suppe's rendering of Bayesianism is in fact faulty (1998).

This is not the place to discuss and decide whether the IBE model and the Bayesian inductive model of testing and confirmation are indeed inadequate for the purpose of reconstructing Morgan's paper, or whether Suppe's adversaries succeed in salvaging their models from his critique. Rather, I wish to highlight what is not at stake in this debate: All four philosophers agree that the (peer-reviewed) scientific paper is the site of disseminating, evaluating, and credentialing scientific knowledge;[5] that the paper is composed with a view to an audience; and that the relevant audience is the immediate peer group. It is assumed that the scientific paper is the best representation of the scientists' notions of valid argumentation because it is directed to the authors' immediate peers and thus likely to contain the best and strongest arguments. It is the epistemologist's task to draw out the basic structure of the arguments presented in the scientific paper and to render them in a systematic fashion.

In shifting the focus back to the scientific publication, these four authors side with traditional philosophy of science. However, all four also agree that it is the task of epistemology to offer an empirically adequate account of the practices of evaluating and credentialing scientific knowledge as they are manifest in the scientific paper, not to provide independently justified norms to improve on these arguments. No effort is made to sew straitjackets for scientists. On the contrary, at one point, Suppe notes that philosophers would be "scandalized" by Morgan's argument. But because it was never challenged by Morgan's peers, we need to consider this move legitimate and "call into question conventional philosophical understanding of the relations between data and theory, suggesting that adequate philosophical analyses of testing and conformation should accommodate such reasoning" (Suppe 1998, 389–390). The evidence of the scientific paper serves as a test case for the philosophical account.

5. In the context of his dispute with Harry Collins's on gravity waves, Franklin explicitly states that there are "good reasons to believe that these are the arguments believed by scientists. After all, a scientist's reputation for good work is based primarily on the published record, and it seems reasonable that they would present their strongest arguments there" (Franklin 1994, 465).

https://doi.org/10.1086/592951 Published online by Cambridge University Press

3.5. Science Education and the Mismatch. Lastly, one can find traces of the issue of the mismatch in recent discussions about science education. Science educators like the former chemist William Harwood are once again complaining that scientists are required to adhere to philosophical norms when writing their papers, and that these norms force them to misrepresent their research. Ironically, Harwood stresses that the prevailing model of "the scientific method" requires scientists to utilize the hypothetico-deductive scheme to represent their research although actual scientific practice does not follow it, and that hypothetico-deductivism is an impediment for science education if it is taught as the norm of proper scientific behavior (Harwood 2004).

Reading such complaints in light of Medawar's essay, one might come to think that scientific practice has changed dramatically since the 1960s and that presently all philosophers of science are staunch Popperians. But of course, what Harwood means by "the scientific method" are those guidelines and methods sections that one finds in scientific journals and science textbooks. Such regulations may, to an extent, reflect philosophical fashions. But those that are present in the textbook and journals guidelines at best display the fashion of yesteryear.

4. The Significance of the Mismatch. Evidently, historians, philosophers, and sociologists of science of different stripes acknowledge that there *is* a mismatch between what scientists do and what they state they did. This mismatch is the common core of a large set of approaches to science, ranging from laboratory anthropology to traditional philosophy of science. There is also general agreement as to why the mismatch occurs: it is the result of scientists' conscious attempts to present their research with a view to audiences. But there is no agreement (*a*) about what the relevant audiences are; (*b*) whether the mismatch is epistemologically significant; and (*c*) what the best way to analyze it would be.

Reichenbach, Medawar, and my chosen representative of science educators Harwood agree that the relevant audiences for the scientists' reconstructions are philosophers of science. They disagree, however, as to whether it is a good or bad thing that scientists seek to meet philosophers' demands. For Medawar (as well as for Harwood), the attempt to please philosophers is an impediment for scientific practice. For Reichenbach, the orientation to philosophy is certainly not a bad thing but actually it does not matter, as philosophy will have the last word in the reconstruction.

Most everyone else discussed in my survey regards the scientific peers as the relevant audience for scientific papers. Scientists construct their arguments with a view to their immediate peers; it is they who need to be convinced (or won over, or defeated). Latour (the Latour of *Science*

https://doi.org/10.1086/592951 Published online by Cambridge University Press

in Action) and, to an extent, Latour and Woolgar imply that even audiences beyond the scientists's immediate peers matter for the establishment of a knowledge claim, and that forms of communication may change relative to the target audiences.

Views about the epistemological significance of the mismatch are widely different. At first blush, it appears that there are two opposing camps: Reichenbach, Medawar, as well as Suppe and his critics agree that the mismatch *is* epistemologically significant; the other scholars discussed here deny its epistemological significance. But this dual grouping is quite misleading because it does not take into account that the there is no general agreement among the scholars about the very tasks of epistemology.

Take Reichenbach and Suppe and his critics. For Reichenbach, epistemology has a threefold—descriptive, critical, and advisory—task. The central function is the critical (rational) reconstruction of the scientists' presentation of their thinking processes. So for Reichenbach, the distinction between actual thinking processes and their reconstruction has great epistemological significance. The mismatch marks the contrast between logic and thinking processes that may or may not be rational, thereby outlining the domain of epistemological analysis.

The kind of epistemological analysis exemplified in the articles by Suppe and his critics is quite different from the conception of epistemology commonly associated with Reichenbach's context distinction. Their ultimate aim is to accommodate scientists' practices of argumentation, not to improve them. If there is a discrepancy between philosophical frameworks and scientific practices, the frameworks must be improved. This analysis is an *epistemological* analysis because it is concerned with the epistemic features of knowledge production. It focuses on scientists' practices of evaluating and credentialing knowledge claims. Suppe, Franklin, Howson, and Lipton consider the article published in peer-reviewed journals the closest representation of scientists' notions of valid arguments and make it alone the focus of their analysis; it is in this sense that the mismatch has epistemological significance. In their attitude to science (if not in the target) the four philosophers side with Medawar, not with orthodox advocates of Reichenbach's version of the context distinction.

Arguably, for Medawar, the mismatch also marks an epistemologically significant distinction between two domains: the domain of scientific rationality and the domain of pretense or even (as he provocatively puts it) deliberate deceit. The relevant site of scientific rationality is the scientists' reasoning processes involved in the generation of knowledge. Precisely because the scientific paper is not a "true" representation of these reasoning processes, it is a fraud. The aim of the analyst is to draw an appropriate image of scientific rationality, but this is a descriptive, not a critical task.

https://doi.org/10.1086/592951 Published online by Cambridge University Press

Knorr-Cetina and Latour and Woolgar agree with Medawar's (implicit) assumption that the principal site of validation and credentialing is the laboratory (or the field or the drawing board) and the interactions among the scientists at the bench. In contrast, Latour (the Latour of *Science in Action*), Shapin and Schaffer, and Dear shift the focus to the interactions between scientists and their audiences. The negotiations between these social groups are the relevant site of scientific argument and validation procedures.[6]

All these scholars (and Harwood, too) implicitly or explicitly deny that epistemological analysis as a separate enterprise is useful or even possible. Historical and sociological analysis of the generation, presentation, and negotiation of knowledge claims is all one can reasonably pursue; there is nothing specifically "epistemological" to add to the historical and sociological insights. But these scholars do have critical agendas, and their agendas are not so different from Medawar's. Of course, lab anthropologists do not describe the scientific paper as a fraud. Yet they do aim to unmask, as it were, the research report and to show that it is, in fact, a product of various literary strategies, which contextualize and give reasons for decisions made (for other reasons) during the research process. Studies of laboratory practice aim to be sensitive to the actual thinking processes, the experimental activities, and the processes by which these processes and practices are recast into arguments and communicated to peers. If a normative force outside the scientists' immediate practices is acknowledged, then it is the force exerted by the scientific community or relevant parts of it. Literary output is not so much constrained by philosophical norms, as Medawar assumed, but rather by scientific journals' guidelines and conventions about how scientific papers must be organized[7] as well as by expectations about the audiences and *their* expectations. So if there is a critical impetus in the lab anthropologists' approaches it does not concern the ways in which scientific research is conducted and validated. Rather, the critique is aimed at common prejudices about science as the epitome of rational pursuit that, according to some sociologists of science, are manifest also in philosophical reconstructions of scientists' practice.[8]

Obviously, the analytic approach to the mismatch hinges on the epistemological significance attached to it. For Reichenbach, the distinction between the scientists' reasoning processes and their communication marks the boundary between two domains of analysis: the empirical study

6. Larry Holmes is mainly interested in scientific creativity, not in validation, so he does not appear in this section.

7. See, in particular, Knorr-Cetina 1981, Chapter 5.

8. This is Knorr-Cetina's motivation, for example.

https://doi.org/10.1086/592951 Published online by Cambridge University Press

of informal reasoning processes (a task for psychology and sociology) and the descriptive and critical analysis and reconstruction of arguments (a task for epistemology). For Medawar, it is ultimately the task of the scientists to deal with the mismatch. The important issue is that scientists should be liberated from the unreasonable demands of philosophers.

Shapin and Schaffer, Dear, and others agree with the anthropologists of laboratory life that the adequate analytic approach to scientific knowledge production should be restricted to historical and sociological means and all-encompassing in scope: It should comprise studies of laboratory life—contemporary laboratories, early modern "houses of experiment" (Shapin) and everything in between—the transformation of this practice into a literary product, as well as the structure, organization, and attitude of the intended audiences and the negotiations between scientists and the different groups of audiences they address. To achieve this goal, the analysts turn to sociology, rhetoric, anthropology, and related fields for useful conceptual tools and models of interaction and negotiation between group members.

Suppe and his critics bring the focus of the analysis back to the detailed structure of the scientific paper. Their analysis is epistemological in that it is devoted to the study of how scientists evaluate and sanction scientific knowledge. The mismatch is epistemologically significant because the scientific paper is the site of disseminating, evaluating, and credentialing scientific knowledge. But to put the matter in Reichenbach's terms, Suppe and his critics pursue the descriptive task of epistemology. The goal is to provide an adequate account of the arguments presented in the peer-reviewed paper. To achieve this goal, the analysts turn to philosophy for useful conceptual tools and models of testing and confirmation.

5. Conclusion: Epistemological Implications. With the demise of traditional normative philosophy of science and the expectations that were once tied to the context distinction, the mismatch no longer marks a division between "epistemologically relevant" and "epistemologically irrelevant" domains of science or between philosophical evaluation and empirical study. Most philosophers of science today agree with Suppe and his critics that it is not the analysts' task to regulate science but to understand and account for the relevant features of scientific practice. Epistemology thus becomes a specific part of science studies, the part that is concerned with the reconstruction of scientific procedures of evaluating and credentialing scientific knowledge. Epistemology retains its critical (if not prescriptive) impetus in that the reconstruction may contribute to the clarification of scientific concepts and arguments.

But the take-home message from my survey is that there is another dimension to epistemological analysis, which one may call *meta-episte-*

https://doi.org/10.1086/592951 Published online by Cambridge University Press

mology: the selection of the focus of the analysis and the appropriate analytic framework for the reconstruction of practices of evaluation and credentialing require decisions that are necessary to begin the analysis yet cannot be made solely on empirical grounds. The questions about the appropriate focus and analytic framework are, in fact, intimately intertwined. We saw that one principal requirement for an epistemological framework is that they yield an empirically adequate account in the sense that they accommodate the scientists' strongest arguments.[9] To evaluate whether a certain framework is empirically adequate, we need to identify the scientists' strongest arguments. My survey of approaches to the mismatch has shown that this is no easy matter.

Suppose one agrees with Suppe and his critics that scientific knowledge is validated in the arena of interacting peers and that therefore the most significant site of evaluating and credentialing scientific arguments is the peer-reviewed scientific paper. Even then, important decisions remain to be made. The points of contention in the exchanges between Suppe and his critics illustrate this nicely. All authors agree that the framework should accommodate the scientists' practices of evaluation and credentialing, but they disagree about most everything else. Not only do they disagree profoundly about the overall merits of the IBE model and Bayesian inductivism but they also disagree about the focus of the analysis: the exact portion of the text that encapsulates the practitioners' relevant practice of evaluation and credentialing. This, in turn, affects their appraisal of the proposed frameworks for reconstruction.

The issue of scope cannot be resolved just by reading the scientific documents. Some selection criteria are required, and the selection criteria call for validation. Suppe is explicit about the scope of his analysis and offers reasons for his selection. He aims to reconstruct as much of the scientific paper as possible; in fact, for him, it counts against philosophical reconstructions of scientific arguments if the exposition of the scientific topic is not accommodated in its entirety. He argues that because of severe space restrictions and high page charges, scientists will only include in their papers what they consider absolutely necessary to their argument. Lipton indicates that he is not happy with Suppe's assumption but his response is inconclusive in this respect. Franklin and Howson demand that the philosophical account "adequately explain the *argument* offered in the paper"—if this is only a small part of the text, then so be it (1998, 413, emphasis added). But these two philosophers do not offer any criteria

9. Of course, there are other requirements besides empirical adequacy that models for testing and confirmation must fulfill, such as consistency and generality.

https://doi.org/10.1086/592951 Published online by Cambridge University Press

for the identification of the argument.[10] None of these meta-epistemological decisions is fully justified.

Actually, I doubt that full justification of this meta-epistemological decision is possible prior to the analysis of the scientific paper, and I doubt it is necessary. All that is needed is a preliminary decision about what portion of a scientific paper one wants to analyze and what framework appears suitable for the purpose. The preliminary analytic framework is then applied to those portions of the scientific documents that initially appear relevant. The analytic framework serves as a guideline for the reconstruction and the empirical data serves as a corrective for the initial framework. The process of mutual adjustment continues until a cogent account is reached, much in the sense of obtaining "reflective equilibrium" between the analytic framework and the empirical record.[11]

The exchange between Suppe and his critics brings another meta-epistemological issue to the fore, namely, the problem of picking the right paper for the analysis. As Franklin and Howson point out, there are different textual genres of peer-reviewed papers. Important experimental findings are often published in "letters-type journals" but these short communications are usually followed by a more detailed account (Franklin and Howson 1998, 413). But which paper should be the focus of the epistemological analysis, the compressed letter announcing important new findings to a journal like *Physical Review Letters* or the detailed elaboration and fully developed argument in the follow-up paper? Franklin and Howson imply that the short communication may not provide complete justification for the claim to be advanced, but that it would be absurd to require that epistemological analysis accommodate the follow-up paper in its entirety.

Again, a meta-epistemological decision must be made and validated: In today's scientific enterprises, there may be a string of research papers that emerge from one project. The first research reports in conference proceedings may be followed by press releases, letters to *Nature*, research articles, synoptic articles, review articles, synthetic expositions in handbooks, and new research proposals building on previous results. All these

10. The problem is familiar from Imre Lakatos's work on rational reconstructions of science. Lakatos stresses that every historical narrative of science necessarily encapsulates a theory of scientific rationality, a theory of "good" science. He suggests comparing historiographical frameworks to establish how big a portion of science they could accommodate. For him, the best framework is the one that reconstructs the most of good science as rational (Lakatos 1974). But as his critics have pointed out, this evaluation hinges on tacit assumptions about what to count as "good" science: how can one distinguish "good" from "bad" or "pseudo" science?

11. John Rawls coined the term "reflective equilibrium" in the context of his theory of moral judgements in Rawls 1971, thereby drawing on Nelson Goodman.

https://doi.org/10.1086/592951 Published online by Cambridge University Press

texts are written for peers—perhaps not exclusively and not for immediate peers, but still, they are written with a view to practicing scientists in the field. And one may expect that the findings presented in each of these communications are credentialed in different ways. Are any of these communications superior representatives of the best and strongest scientific arguments, and why?

Suppe and Franklin and Howson would probably agree that the most immediate outcomes of lab research (be it the brief communication to *Physical Review Letters* or the more detailed follow-up report) reflect best what scientists consider to be good and valid arguments, because they are directed at the immediate, and often competing, co-workers in a scientific field. On the other hand, it also appears reasonable to argue that those knowledge claims in publications more remote from the immediate context of their generation are the strongest, because they have proved to be the most robust and the most resistant to the critique advanced by other peers beyond the small circle of referees for a journal, distinguished as these referees may be.[12]

If one compares the scope of epistemological analysis as it is presented in the set of papers on Morgan with the scope of the analysis in sociological studies such as Latour's *Science in Action*, another meta-epistemological problem comes to the fore, namely, the role of the audiences for the establishment of claims to knowledge. What role, if any, do (peer) audiences play in the evaluation and credentialing of scientific knowledge claims? If the public arena is the site of evaluating and validating scientific knowledge claims, the (peer) audiences' responses too are an integral part of the evaluation procedures.

In order to engage with the meta-epistemological issue of the site of the scientists' "best and strongest" arguments, one may want to circumvent the decision and instead trace the *transformation* of the scientists' claims to knowledge and supporting arguments through different textual genres. We saw that the earliest sociological studies of scientists' procedures of validating knowledge claims like Nigel Gilbert's 1976 paper on the transformation of research findings into scientific knowledge are in

12. Ludwik Fleck made this point in *Genesis and Development of a Scientific Fact*. Fleck introduced the distinction between esoteric and exoteric forms of presentation in science to highlight the structural differences between journal and vademecum or handbook science (1981, 118–124). While the esoteric journal article provides at least a glimpse of the experimental production of results—if only in the highly regulated and schematic sections on 'methods' and 'observations'—exoteric handbooks (and, one might add, popular books and articles) are more remote from actual scientific practice. According to Fleck, these remote representations of lab science are the most important in the establishment of scientific knowledge because the scientist gains reassurance from the glance into the textbook.

https://doi.org/10.1086/592951 Published online by Cambridge University Press

rather close conceptual proximity to contemporary philosophy of science. Recent work in sociology of science, on the other hand, while retaining the goal of studying scientists' negotiations among themselves, has moved far from Gilbert's concern with the concrete structure of scientific argumentation. The scientists' activities are reconstructed as attempts to gain allies and build strongholds to defend knowledge claims. Such models of scientific practice are also rather remote from Suppe's and his critics' reconstructions of scientific arguments. So in a way, epistemologists interested in the procedures used by scientist to validate knowledge need to reclaim some of the ground now covered by sociological analysis. Instead of tracing, say, the formation of alliances between social groups and the strategies to enroll and control others, one would need to re-direct attention to the changing structure of the argument in different textual genres, the changing ways of marshaling the evidence and so forth with the help of more recent models of confirmation and testing.

In doing so, one should not ignore two lessons that can be learnt from sociological and historical analyses. First, those scholars who insist that audiences' responses play a role in scientific procedures of evaluating and credentialing do have a point. It is, in fact, a reasonable extension of the generally accepted notion that scientists write their research papers with a view to convincing their audiences. But available studies of scientists' negotiations with their (peer) audiences usually focus on the social forces driving the negotiations between scientists and their audiences.[13] Instead, what is needed are sustained attempts to reconstruct the argumentative structure of these negotiations in terms of philosophical models of testing and confirmation.

Secondly, historians of scientific practice remind us that the structure and organization of audiences (including peer audiences) change over time and even from local context to local context. Depending on the contexts, audiences of professional scientists may or may not share the author's background knowledge. Authors and their audiences are more likely to have a common background in institutions like the early modern Royal Society or the French Academy of the Sciences, where public demonstrations and discussions were frequent, than, say, in the German university system around 1800 with its many far-flung universities, where researchers mostly communicated via letters and articles in journals and magazines. Presentations of results reflect these structures and forms of organization of the scientific community as well as deep-seated but nev-

13. Franklin's study on the detection of gravity waves is an exception. This work is mainly a response to Harry Collins's notion of the experimenter's regress. Franklin's main aim here is to show that the controversy about gravity waves was ultimately resolved in a rational fashion, using a set of epistemological strategies (1994).

https://doi.org/10.1086/592951 Published online by Cambridge University Press

ertheless historically variable epistemological conceptions of evidence, proof, and so on.

We see that several meta-epistemological decisions regarding scope and means are required for the analysis of scientific practices of evaluating and credentialing scientific work. Generally speaking, I do not think that there is a principled and ultimate answer to the question of the types of scientific communication that are likely to contain the scientists' best and strongest arguments or the portions of scientific communications that need to be accommodated. However, my survey has shown that it is crucial to compare alternative reconstructions and consider their respective merits and demerits. This comparison reveals the meta-epistemological decisions involved in the reconstructions, and the different tacit assumptions about the epistemological analysis of scientific practice are brought to the fore.

REFERENCES

Dear, P. (1991), "Narratives, Anecdotes, and Experiments: Turning Experience into Science in the Seventeenth Century", in P. Dear (ed.), *The Literary Structure of Scientific Argument: Historical Studies.* Philadelphia: University of Pennsylvania Press, 135–163.

Fleck, L. (1981), *Genesis and Development of a Scientific Fact*. Chicago: University of Chicago Press.

Franklin, A. (1994), "How to Avoid the Experimenters' Regress", *Studies in History and Philosophy of Science* 25: 463–491.

Franklin, A., and C. Howson (1998), "Comment on 'The Structure of a Scientific Paper' by Frederick Suppe", *Philosophy of Science* 65: 411–416.

Gilbert, G. N. (1976), "The Transformation of Research Findings into Scientific Knowledge", *Social Studies of Science* 6: 281–306.

Harwood, W. S. (2004), "A New Model for Inquiry: Is the Scientific Method Dead?" *Journal of College Science Teaching* 33 (7): 29–33.

Holmes, F. L. (1987), "Scientific Writing and Scientific Discovery", *Isis* 78: 220–235.

——— (1991), "Argument and Narrative in Scientific Writing", in P. Dear (ed.), *The Literary Structure of Scientific Argument: Historical Studies*. Philadelphia: University of Pennsylvania Press, 164–181.

Howard, D. (2006), "Lost Wanderers in the Forest of Knowledge: Some Thoughts on the Discovery-Justification Distinction", in J. Schickore and F. Steinle (eds.), *Revisiting Discovery and Justification: Historical and Philosophical Perspectives on the Context Distinction*. Dordrecht: Springer, 3–22.

Knorr-Cetina, K. (1981), *The Manufacture of Knowledge*. Oxford: Pergamon.

Lakatos, I. (1974), "History of Science and Its Rational Reconstructions", in Y. Elkana (ed.), *The Interaction between Science and Philosophy*. Atlantic Highlands, NY: Humanities Press, 195–241.

Latour, B. (1987), *Science in Action. How to Follow Scientists and Engineers through Society*. Cambridge, MA.: Harvard University Press.

Latour, B., and S. Woolgar ([1979] 1986), *Laboratory Life: The Social Construction of Scientific Facts*. Beverly Hills, CA: Sage.

Lipton, P. (1998), "The Best Explanation of a Scientific Paper", *Philosophy of Science* 65: 406–410.

Medawar, P. ([1963] 1996), "Is the Scientific Paper a Fraud?", reprinted in *The Strange Case of the Spotted Mice and Other Classic Essays on Science*. Oxford: Oxford University Press, 33–39.

Morgan, W. J. (1968), "Rises, Trenches, Great Faults and Crustal Blocks", *Journal of Geophysical Research* 73: 1959–1982.

https://doi.org/10.1086/592951 Published online by Cambridge University Press

Rawls, J. (1971), *A Theory of Justice*. Cambridge, MA: Harvard University Press.
Reichenbach, H. (1938), *Experience and Prediction: An Analysis of the Foundations and the Structure of Knowledge*. Chicago: University of Chicago Press.
Richardson, A. (2006), "Freedom in a Scientific Society: Reading the Context of Reichenbach's Contexts", in J. Schickore and F. Steinle (eds.), *Revisiting Discovery and Justification: Historical and Philosophical Perspectives on the Context Distinction*. Dordrecht: Springer, 41–54.
Schickore, J., and F. Steinle (2006), *Revisiting Discovery and Justification: Historical and Philosophical Perspectives on the Context Distinction*. Dordrecht: Springer.
Schiemann, G. (2006), "Inductive Justification and Discovery: On Hans Reichenbach's Foundation of the Autonomy of the Philosophy of Science", in J. Schickore and F. Steinle (eds.), *Revisiting Discovery and Justification: Historical and Philosophical Perspectives on the Context Distinction*. Dordrecht: Springer, 23–39.
Shapin, S., and S. Schaffer (1985), *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, NJ: Princeton University Press.
Suppe, F. (1998), "The Structure of a Scientific Paper", *Philosophy of Science* 65: 381–405.

https://doi.org/10.1086/592951 Published online by Cambridge University Press