

Cambridge University Press

978-0-521-88008-4 - Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science

Edited by Deborah G. Mayo and Aris Spanos

Excerpt

[More information](#)

Introduction and Background

Deborah G. Mayo and Aris Spanos

I Central Goals, Themes, and Questions

1 Philosophy of Science: Problems and Prospects

Methodological discussions in science have become increasingly common since the 1990s, particularly in fields such as economics, ecology, psychology, epidemiology, and several interdisciplinary domains – indeed in areas most faced with limited data, error, and noise. Contributors to collections on research methods, at least at some point, try to ponder, grapple with, or reflect on general issues of knowledge, inductive inference, or method. To varying degrees, such work may allude to philosophies of theory testing and theory change and philosophies of confirmation and testing (e.g., Popper, Carnap, Kuhn, Lakatos, Mill, Peirce, Fisher, Neyman-Pearson, and Bayesian statistics). However, the different philosophical “schools” tend to be regarded as static systems whose connections to the day-to-day questions about how to obtain reliable knowledge are largely metaphorical. Scientists might “sign up for” some thesis of Popper or Mill or Lakatos or others, but none of these classic philosophical approaches – at least as they are typically presented – provides an appropriate framework to address the numerous questions about the legitimacy of an approach or method.

Methodological discussions in science have also become increasingly sophisticated; and the more sophisticated they have become, the more they have encountered the problems of and challenges to traditional philosophical positions. The unintended consequence is that the influence of philosophy of science on methodological practice has been largely negative. If the philosophy of science – and the History and Philosophy of Science (HPS) – have failed to provide solutions to basic problems of evidence and inference, many practitioners reason, then how can they help scientists to look to philosophy of science to gain perspective? In this spirit, a growing tendency is to question whether anything can be said about what makes an

Cambridge University Press

978-0-521-88008-4 - Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science

Edited by Deborah G. Mayo and Aris Spanos

Excerpt

[More information](#)

enterprise scientific, or what distinguishes science from politics, art or other endeavors. Some works on methodology by practitioners look instead to sociology of science, perhaps to a variety of post-modernisms, relativisms, rhetoric and the like.

However, for the most part, scientists wish to resist relativistic, fuzzy, or postmodern turns; should they find themselves needing to reflect in a general way on how to distinguish science from pseudoscience, genuine tests from ad hoc methods, or objective from subjective standards in inquiry, they are likely to look to some of the classical philosophical representatives (and never mind if they are members of the list of philosophers at odds with the latest vogue in methodology). Notably, the Popperian requirement that our theories and hypotheses be testable and falsifiable is widely regarded to contain important insights about responsible science and objectivity; indeed, discussions of genuine versus ad hoc methods seem invariably to come back to Popper's requirement, even if his full philosophy is rejected. However, limiting scientific inference to deductive falsification without any positive account for warranting the reliability of data and hypotheses is too distant from day-to-day progress in science. Moreover, if we are to accept the prevalent skepticism about the existence of reliable methods for pinpointing the source of anomalies, then it is hard to see how to warrant falsifications in the first place.

The goal of this volume is to connect the methodological questions scientists raise to philosophical discussions on *Experimental Reasoning, Reliability, Objectivity, and Rationality* (E.R.R.O.R) of science. The aim of the "exchanges" that follow is to show that the real key to progress requires a careful unpacking of the central reasons that philosophy of science has failed to solve problems about evidence and inference. We have not gone far enough, we think, in trying to understand these obstacles to progress.

Achinstein (2001) reasons that, "scientists do not and should not take . . . philosophical accounts of evidence seriously" (p. 9) because they are based on a priori computations; whereas scientists evaluate evidence empirically. We ask: Why should philosophical accounts be a priori rather than empirical? Chalmers, in his popular book *What is This Thing Called Science?* denies that philosophers can say anything general about the character of scientific inquiry, save perhaps "trivial platitudes" such as "take evidence seriously" (Chalmers, 1999, p. 171). We ask: Why not attempt to answer the question of what it means to "take evidence seriously"? Clearly, one is not taking evidence seriously in appraising hypothesis *H* if it is pre-determined that a way would be found to either obtain or interpret data as supporting *H*. If a procedure had little or no ability to find flaws in *H*,

Cambridge University Press

978-0-521-88008-4 - Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science

Edited by Deborah G. Mayo and Aris Spanos

Excerpt

[More information](#)

then finding none scarcely counts in H 's favor. One need not go back to the discredited caricature of the objective scientist as “disinterested” to extract an uncontroversial minimal requirement along the following lines:

Minimal Scientific Principle for Evidence. Data x_0 provide poor evidence for H if they result from a method or procedure that has little or no ability of finding flaws in H , even if H is false.

As weak as this is, it is stronger than a mere falsificationist requirement: it may be logically possible to falsify a hypothesis, whereas the procedure may make it virtually impossible for such falsifying evidence to be obtained.

It seems fairly clear that this principle, or something very much like it, undergirds our intuition to disparage ad hoc rescues of hypotheses from falsification and to require hypotheses to be accepted only after subjecting them to criticism. *Why then has it seemed so difficult to erect an account of evidence that embodies this precept without running aground on philosophical conundrums?* By answering this question, we hope to set the stage for new avenues for progress in philosophy and methodology. Let us review some contemporary movements to understand better where we are today.

2 Current Trends and Impasses

Since breaking from the grip of the logical empiricist orthodoxy in the 1980s, the philosophy of science has been marked by attempts to engage dynamically with scientific practice:

1. Rather than a “white glove” analysis of the logical relations between statements of evidence e and hypothesis H , philosophers of science would explore the complex linkages among data, experiment, and theoretical hypotheses.
2. Rather than hand down pronouncements on ideally rational methodology, philosophers would examine methodologies of science empirically and naturalistically.

Two broad trends may be labeled the “new experimentalism” and the “new modeling.” Moving away from an emphasis on high-level theory, the new experimentalists tell us to look to the manifold local tasks of distinguishing real effects from artifacts, checking instruments, and subtracting the effects of background factors (e.g., Chang, Galison, Hacking). Decrying the straightjacket of universal accounts, the new modelers champion the disunified and pluralistic strategies by which models mediate among data,

Cambridge University Press

978-0-521-88008-4 - Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science

Edited by Deborah G. Mayo and Aris Spanos

Excerpt

[More information](#)

hypotheses, and the world (Cartwright, Morgan, and Morrison). The historical record itself is an important source for attaining relevance to practice in the HPS movement.

Amid these trends is the broad move to tackle the philosophy of methodology empirically by looking to psychology, sociology, biology, cognitive science, or to the scientific record itself. As interesting, invigorating, and right-headed as the new moves have been, the problems of evidence and inference remain unresolved. *By and large, current philosophical work and the conceptions of science it embodies are built on the presupposition that we cannot truly solve the classic conundrums about induction and inference.* To give up on these problems, however, does not make them go away; moreover, the success of naturalistic projects demands addressing them. Appealing to “best-tested” theories of biology or cognitive science calls for critical evaluation of the methodology of appraisal on which these theories rest.

The position of the editors of this volume takes elements from each of these approaches (new experimentalism, empirical modeling, and naturalism). We think the classic philosophical problems about evidence and inference are highly relevant to methodological practice and, furthermore, *that they are solvable.* To be clear, we do not pin this position on any of our contributors! However, the exchanges with our contributors elucidate this stance. Taking naturalism seriously, we think we should appeal to the conglomeration of research methods for collecting, modeling, and learning from data in the face of limitations and threats of error – including modeling strategies and probabilistic and computer methods – all of which we may house under the very general rubric of the methodology of inductive-statistical modeling and inference. For us, statistical science will always have this broad sense covering experimental design, data generation and modeling, statistical inference methods, and their links to scientific questions and models. We also regard these statistical tools as lending themselves to informal analogues in tackling general philosophical problems of evidence and inference. Looking to statistical science would seem a natural, yet still largely untapped, resource for a naturalistic and normative approach to philosophical problems of evidence. Methods of experimentation, simulation, model validation, and data collection have become increasingly subtle and sophisticated, and we propose that philosophers of science revisit traditional problems with these tools in mind. In some contexts, even where literal experimental control is lacking, inquirers have learned how to determine “what it would be like” if we were able to intervene and control – at least with high probability. Indeed “the challenge, the fun, of outwitting and outsmarting drives us to find ways to learn what it would be like to control,

Cambridge University Press

978-0-521-88008-4 - Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science

Edited by Deborah G. Mayo and Aris Spanos

Excerpt

[More information](#)

manipulate, and change, in situations where we cannot” (Mayo, 1996, p. 458). This perspective lets us broaden the umbrella of what we regard as an “experimental” context. When we need to restore the more usual distinction between experimental and observational research, we may dub the former “manipulative experiment” and the latter “observational experiment.”

The tools of statistical science are plagued with their own conceptual and epistemological problems – some new, many very old. It is important to our goals to interrelate themes from philosophy of science and philosophy of statistics.

- The first half of the volume considers issues of error and inference in philosophical problems of induction and theory testing.
- The second half illuminates issues of errors and inference in practice: in formal statistics, econometrics, causal modeling, and legal epistemology.

These twin halves reflect our conception of philosophy and methodology of science as a “two-way street”: on the one hand there is an appeal to methods and strategies of local experimental testing to grapple with philosophical problems of evidence and inference; on the other there is an appeal to philosophical analysis to address foundational problems of the methods and models used in practice; see Mayo and Spanos (2004).

3 Relevance for the Methodologist in Practice

An important goal of this work is to lay some groundwork for the methodologist in practice, although it must be admitted that our strategy at first appears circuitous. We do not claim that practitioners’ general questions about evidence and method are directly answered once they are linked to what professional philosophers have said under these umbrellas. Rather, we claim that it is by means of such linkages that practitioners may better understand the foundational issues around which their questions revolve. In effect, practitioners themselves may become better “applied philosophers,” which seems to be what is needed in light of the current predicament in philosophy of science. Some explanation is necessary.

In the current predicament, methodologists may ask, if each of the philosophies of science have unsolved and perhaps insoluble problems about evidence and inference, then how can they be useful for evidential problems in practice? “If philosophers and others within science theory can’t agree about the constitution of the scientific method... doesn’t it seem a little dubious for economists to continue blithely taking things off

Cambridge University Press

978-0-521-88008-4 - Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science

Edited by Deborah G. Mayo and Aris Spanos

Excerpt

[More information](#)

the [philosopher's] shelf?" (Hands, 2001, p. 6). Deciding that it does, many methodologists in the social sciences tend to discount the relevance of the principles of scientific legitimacy couched within traditional philosophy of science. The philosophies of science are either kept on their shelves, or perhaps dusted off for cherry-picking from time to time. Nevertheless, practitioners still (implicitly or explicitly) wade into general questions about evidence or principles of inference and by elucidating the philosophical dimensions of such problems we hope to empower practitioners to appreciate and perhaps solve them. In a recent lead article in the journal *Statistical Science*, we read that "professional agreement on statistical philosophy is not on the immediate horizon, but this should not stop us from agreeing on methodology" (Berger, 2003, p. 2). But we think "what is correct methodologically" depends on "what is correct philosophically" (Mayo, 2003). Otherwise, choosing between competing methods and models may be viewed largely as a matter of pragmatics without posing deep philosophical problems or inconsistencies of principle. For the "professional agreement" to have weight, it cannot be merely an agreement to use methods with similar numbers when the meaning and import of such numbers remain up in the air (see Chapter 7). We cannot wave a wand and bring into existence the kind of philosophical literature that we think is needed. What we can do is put the practitioner in a better position to support, or alternatively, question the basis for professional agreement or disagreement.

Another situation wherein practitioners may find themselves wishing to articulate general principles or goals is when faced with the need to modify existing methods and to make a case for the adoption of new tools. Here, practitioners may serve the dual role of both inventing new methods and providing them with a principled justification – possibly by striving to find, or adapt features from, one or another philosophy of science or philosophy of statistics. Existing philosophy of science may not provide off-the-shelf methods for answering methodological problems in practice, but, coupled with the right road map, it may enable understanding, or even better, solving those problems.

An illustration in economics is given by Aris Spanos (Chapter 6). Faced with the lack of literal experimental controls, some economic practitioners attempt to navigate between two extreme positions. One position is the prevailing theory-dominated empirical modeling, largely limited to quantifying theories presupposed to be true. At the other extreme is data-driven modeling, largely limited to describing the data and guided solely by goodness-of-fit criteria. The former stays too close to the particular theory chosen at the start; the second stays too close to the particular data. Those

Cambridge University Press

978-0-521-88008-4 - Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science

Edited by Deborah G. Mayo and Aris Spanos

Excerpt

[More information](#)

practitioners seeking a “third way” are implicitly thrust into the role of striving to locate a suitable epistemological foundation for a methodology seemingly at odds with the traditional philosophical image of the roles of theory and data in empirical inquiry. In other words, the prescriptions on method in practice have trickled down from (sometimes competing) images of good science in traditional philosophy. We need to ask the question: *What are the threats to reliability and objectivity that lay behind the assumed prescriptions to begin with?* If data-dependent methods are thought to require the assumption of an overarching theory, or else permit too much latitude in constructing theories to fit data, then much of social science appears to be guilty of violating a scientific canon. But in practice, some econometricians work to develop methods whereby the data may be used to provide independent constraints on theory testing by means of intermediate-level statistical models with a “life of their own,” as it were. This is the key to evading threats to reliability posed by theory-dominated modeling. By grasping the philosophical issues and principles, such applied work receives a stronger and far less tenuous epistemological foundation.

This brings us to a rather untraditional connection to traditional philosophy of science. In several of the philosophical contributions in this volume, we come across the very conceptions of testing that practitioners may find are in need of tweaking or alteration in order to adequately warrant methods they wish to employ. By extricating the legitimate threats to reliability and objectivity that lie behind the traditional stipulations, practitioners may ascertain where and when violations of established norms are justifiable. The exchange essays relating to the philosophical contributions deliberately try to pry us loose from rigid adherence to some of the standard prescriptions and prohibitions.

In this indirect manner, the methodologists’ real-life problems are connected to what might have seemed at first an arcane philosophical debate. Insofar as these connections have not been made, practitioners are dubious that philosophers’ debates about evidence and inference have anything to do with, much less help solve, their methodological problems. We think the situation is otherwise – that getting to the underlying philosophical issues not only increases the intellectual depth of methodological discussions but also paves the way for solving problems.

We find this strategy empowers students of methodology to evaluate critically, and perhaps improve on, methodologies in practice. Rather than approach alternative methodologies in practice as merely a menu of positions from which to choose, they may be grasped as attempted solutions to problems with deep philosophical roots. Conversely, progress in

Cambridge University Press

978-0-521-88008-4 - Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science

Edited by Deborah G. Mayo and Aris Spanos

Excerpt

[More information](#)

methodology may challenge philosophers of science to reevaluate the assumptions of their own philosophical theories. That is, after all, what a genuinely naturalistic philosophy of method would require. A philosophical problem, once linked to methodology in practice, enjoys solutions from the practical realm. For example, philosophers tend to assume that there are an infinite number of models that fit finite data equally well, and so data underdetermine hypotheses. Replacing “fit” with more rigorous measures of adequacy can show that such underdetermination vanishes (Spanos, 2007). This brings us to the last broad topic we consider throughout the volume.

We place it under the heading of *metaphilosophical themes*. Just as we know that evidence in science may be “theory-laden” – interpreted from the perspective of a background theory or set of assumptions – our philosophical theories (about evidence, inference, science) often color our philosophical arguments and conclusions (Rosenberg, 1992). The contributions in this volume reveal a good deal about these “philosophy-laden” aspects of philosophies of science. These revelations, moreover, are directly relevant to what is needed to construct a sound foundation for methodology in practice. The payoff is that understanding the obstacles to solving philosophical problems (the focus of Chapters 1–5) offers a clear comprehension of how to relate traditional philosophy of science to contemporary methodological and foundational problems of practice (the focus of Chapters 6–9).

4 Exchanges on E.R.R.O.R.

We organize the key themes of the entire volume under two interrelated categories:

- (1) *experimental reasoning (empirical inference) and reliability*, and
- (2) *objectivity and rationality of science*.

Although we leave these terms ambiguous in this introduction, they will be elucidated as we proceed. Interrelationships between these two categories immediately emerge. Scientific rationality and objectivity, after all, are generally identified by means of scientific methods: one’s conception of objectivity and rationality in science leads to a conception of the requirements for an adequate account of empirical inference and reasoning. The perceived ability or inability to arrive at an account satisfying those requirements will in turn direct one’s assessment of the possibility of objectivity and rationality in science. Recognizing the intimate relationships between categories 1 and 2 propels us toward both understanding and making progress on recalcitrant foundational problems about scientific inference. If, for example, empirical

Cambridge University Press

978-0-521-88008-4 - Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science

Edited by Deborah G. Mayo and Aris Spanos

Excerpt

[More information](#)

inference is thought to demand reliable rules of inductive inference, and if it is decided that such rules are unobtainable, then one may either question the rationality of science or instead devise a different notion of rationality for which empirical methods exist. On the other hand, if we are able to show that some methods are more robust than typically assumed, we may be entitled to uphold a more robust conception of science. Under category 1, we consider the nature and justification of experimental reasoning and the relationship of experimental inference to appraising large-scale theories in science.

4.1 Theory Testing and Explanation

Several contributors endorse the view that scientific progress is based on accepting large-scale theories (e.g., Chalmers, Musgrave) as contrasted to a view of progress based on the growth of more localized experimental knowledge (Mayo). Can one operate with a single overarching view of what is required for data to warrant an inference to H ? Mayo says yes, but most of the other contributors argue for multiple distinct notions of evidence and inference. They do so for very different reasons. Some argue for a distinction between large-scale theory testing and local experimental inference. When it comes to large-scale theory testing, some claim that the most one can argue is that a theory is, comparatively, the best tested so far (Musgrave), or that a theory is justified by an “argument from coincidence” (Chalmers). Others argue that a distinct kind of inference is possible when the data are “not used” in constructing hypotheses or theories (“use-novel” data), as opposed to data-dependent cases where an inference is, at best, conditional on a theory (Worrall). Distinct concepts of evidence might be identified according to different background knowledge (Achinstein). Finally, different standards of evidence may be thought to emerge from the necessity of considering different costs (Laudan). The relations between testing and explanation often hover in the background of the discussion, or they may arise explicitly (Chalmers, Glymour, Musgrave). What are the explanatory virtues? And how do they relate to those of testing? Is there a tension between explanation and testing?

4.2 What Are the Roles of Probability in Uncertain Inference in Science?

These core questions are addressed both in philosophy of science, as well as in statistics and modeling practice. Does probability arise to assign degrees

Cambridge University Press

978-0-521-88008-4 - Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science

Edited by Deborah G. Mayo and Aris Spanos

Excerpt

[More information](#)

of epistemic support or belief to hypotheses, or to characterize the reliability of rules? A loose analogy exists between Popperian philosophers and frequentist statisticians, on the one hand, and Carnapian philosophers and Bayesian statisticians on the other. The latter hold that probability needs to supply some degree of belief, support, or epistemic assignment to hypotheses (Achinstein), a position that Popperians, or critical rationalists, dub ‘justificationism’ (Musgrave). Denying that such degrees may be usefully supplied, Popperians, much like frequentists, advocate focusing on the rationality of rules for inferring, accepting, or believing hypotheses. But what properties must these rules have?

In formal statistical realms, the rules for inference are reliable by dint of controlling error probabilities (Spanos, Cox and Mayo, Glymour). Can analogous virtues be applied to informal realms of inductive inference? This is the subject of lively debate in Chapters 1 to 5 in this volume. However, statistical methods and models are subject to their own long-standing foundational problems. Chapters 6 and 7 offer a contemporary update of these problems from the frequentist philosophy perspective. Which methods can be shown to ensure reliability or low long-run error probabilities? Even if we can show they have good long-run properties, how is this relevant for a particular inductive inference in science? These chapters represent exchanges and shared efforts of the authors over the past four years to tackle these problems as they arise in current statistical methodology. Interwoven throughout this volume we consider the relevance of these answers to analogous questions as they arise in philosophy of science.

4.3 Objectivity and Rationality of Science, Statistics, and Modeling

Despite the multiplicity of perspectives that the contributors bring to the table, they all find themselves confronting a cluster of threats to objectivity in observation and inference. Seeing how analogous questions arise in philosophy and methodological practice sets the stage for the meeting ground that creates new synergy.

- Does the fact that observational claims themselves have assumptions introduce circularity into the experimental process?
- Can one objectively test assumptions linking actual data to statistical models, and statistical inferences to substantive questions?

On the one hand, the philosophers’ demand to extricate assumptions raises challenges that the practitioner tends to overlook; on the other hand,