Dimensions of Inductive Risk: Prospects, Boundaries, New Facets

Elliott, Kevin and Ted Richards, eds. (2017). Exploring Inductive Risk: Case Studies of Values in Science. Oxford University Press, New York. ISBN: 019046772X, 304 pages, $44.00 (paperback).

Reviewed by:[1] Leuschner, Anna, DFG-Graduiertenkolleg 2073, Institute of Philosophy, Leibniz Universität Hannover, Germany, e-mail: anna.leuschner@philos.uni-hannover.de
Bueter, Anke, Institute of Philosophy, Leibniz Universität Hannover, Germany, e-mail: anke.bueter@philos.uni-hannover.de

This is the final draft of a book review forthcoming in *Science and Education.*

1 Background and Context

During the last two decades, the various challenges of inductive risk have been addressed by a number of philosophers of science in diverse contexts. Thus, it is fortuitous that there is now a volume available providing an overview of the status of the discussion as well as addressing hitherto unanswered questions. It provides useful reading material to classroom teachers who address all sorts of topics on science and values.

The volume is edited by Kevin Elliott and Ted Richards, both outstanding scholars in the field of socially relevant philosophy of science, and it contains a foreword by Heather Douglas whose work has been, as Torsten Wilholt put it, "vital for bringing the argument of [Richard] Rudner and others back into the current debate about science and values." (Wilholt, 2009, *Bias and Values in Scientific Research*, p. 94). Since Douglas' seminal paper "Inductive Risk and Values in Science" from 2000, the fruitfulness of the inductive risk argument has been established by a broad spectrum of philosophical analyses. This is then the starting point of the present volume.

Not only does this book thoroughly flesh out the variety of stages where values enter in scientific practices, it also makes clear that this concerns all scientific fields, regardless of whether they are more practically or more theoretically oriented. Although at some points slightly repetitive, as well-known arguments from Rudner, Hempel, Jeffrey, and Levi are reconstructed more than once, at other points the volume reveals surprising new aspects of

[1] Anna Leuschner is the corresponding author. Both authors contributed equally to this book review.

these classical arguments and also provides directions for future research. The volume's most important merit, however, lies in its presentation of new and original case studies. Let us now consider the volume in more detail by exploring its four parts.

2 Part One: Weighing Inductive Risk

The first part of the book starts with a contribution by Jacob Stegenga (chapter 2), who explores the role of inductive risk in drug regulation, a field highly relevant to the discussion on inductive risk because of the practical consequences of accepting or rejecting mistaken hypotheses. Moreover, false-positive and false-negative results come with different risks for different stakeholders, such as, e.g., financial costs for pharmaceutical companies or side-effects for patients. Many decisions in the design or assessment of clinical trials involve a trade-off between those different risks. Thus, as Stegenga points out, such risks have to be weighed by reference to non-epistemic values. Balancing the often-times competing demands of such values requires drug regulators to perform what he calls the "inductive risk calculus" (IRC).

Based on Philip Kitcher's account of well-ordered certification (Kitcher, 2011, *Science in a Democratic Society*), Stegenga argues that one should aim for a *balanced* state of the inductive risk calculus, which integrates the concerns of different stakeholders equally and accounts for the full range of non-epistemic values playing a role. He then considers the policies and standards of the U.S. Food and Drug Administration (FDA) and argues that these reflect an unbalanced state of IRC that is skewed towards avoiding mistaken drug rejections. He points out as problematic that the FDA relies on data provided by the respective drug manufacturers, which increases the risk for many kinds of bias in this data, such as *p*-value hacking, publication bias, or exploitation of the base-rate fallacy.

While these issues are mostly already well-known and thoroughly discussed in philosophy of medicine, Stegenga provides a very helpful and lucid summary. Moreover, he proposes some counter-measures that work on the institutional level (e.g., obligatory registering of all clinical trials, sticking to pre-defined primary outcomes, etc.) rather than calling on individual drug regulators' or scientists' ethical responsibility, which seems a more fertile approach. Whereas his argumentation that the current FDA standards reflect an imbalance in the inductive risk calculus is very convincing, however, the paper could have put more emphasis on grounding the presumed ideal of balance. Why should all stakeholders' interests be considered of equal importance? Why would it not be defensible to weigh

potential benefits for patients higher than those for drug manufacturers? Exploring this in more detail could have been very interesting and cast some light on the difficult questions on how to ethically weigh different non-epistemic values.

In chapter 3, Kent W. Staley offers a detailed reconstruction of Isaac Levi's account, referring not only to the well-known 1962 paper but also to his 1974 *Gambling with the Truth*, coming to the conclusion, in contrast to Douglas, that Levi's views have "a strong affinity with Douglas' own position" (p. 48). However, Staley apparently reaches this conclusion on the mere grounds of Levi's concession that the choice of a significance level "is a subjective factor that does in some sense reflect the investigator's attitudes," a "degree of caution" (p. 47). However, this overlooks the advancement offered by Douglas' position in that her argument goes much further, particularly as she reveals how values enter into decisions regarding experimental design and the characterization and interpretation of data, meaning that there is no such thing as "given data" or simply "data." This makes Staley's reconstruction of the inductive risk argument, "as discussed in the contemporary literature," problematic (pp. 41, 46).

In contrast, the case study presented in Staley's paper—the discovery of the Higgs boson at the LHC—is very thought-provoking. Staley describes how both scientific and public interest have exerted enormous pressure on the research project. The finding of the boson has been relevant for future agenda-setting in science as well as for future funding sources and the authority of science. This has led to a particular methodological cautiousness, especially evident in the emphasis given to the choice of the 5 Sigma standard.

3 Part Two: Evading Inductive Risk

Part two begins with a contribution by David Resnik (chapter 4) who provides a thorough discussion of the added complexity of inductive risk in dual-use research. He illustrates this by a case study of the debate on modifications of the H5N1 virus and the danger of their abuse by terrorists. He argues that scientists here have moral and professional obligations to consider the social and ethical consequences of mistaken inferences. This point is defended against new versions of the Jeffreyan defense of value-freedom, which he dubs the "clean-hands view": The idea that scientists can remain neutral with regard to the ethical dimensions of inductive risk, as long as they openly communicate uncertainties and probabilities to policy-makers. This, in fact, would leave it to the policy-makers to get their hands dirty in

risky, value-laden decision-making. Resnik makes the convincing point that deferring moral responsibility in this manner to people less well-equipped to assess the existing data and its uncertainties is ethically problematic.

However, his individualist solution to the problem of inductive risk in dual-use research has two problems, which arise exactly because of the added complexity in dual-use research that he himself lays out so clearly. First, this concerns the question of whether all the additional issues involved are adequately described in the terminology of inductive risk. This is because the concerned risks seem to refer to (at least) two different levels: (1) the level of making a (potentially wrong) decision on the acceptance or rejection of a hypothesis on, e.g., the contagion rate of a certain virus modification (a classic case of inductive risk); and (2) the meta-level of making a (potentially wrong) decision on whether to publish these results or not. Such a publication poses risks especially if the original decisions were *correct.* Thus, scientists would have to consider the potential harms of mistaken as well as of correct inferences.

Second, the estimations of risks on the meta-level (of whether and how to publish) involve additional dimensions of uncertainty, as Resnik makes clear. He shows that such estimates will be affected by the respective researchers' background assumptions and values, as, for instance, estimations of terrorists' lab equipment and skills obviously cannot be based on reliable, transparent reports of such matters. Here, it seems a lot to ask of the individual researcher to be capable to assess all these risks on both levels. This poses interesting questions about what it means to deal responsibly with dual-use research, and whom exactly this responsibility can be ascribed to.

The next chapter (5) by David M. Frank also addresses the Jeffreyan ideal of value-free science. Frank proposes three conditions he considers necessary for approximating the ideal: 1. "Decision-makers are able to understand the representations of uncertainty offered by scientists." 2. "Scientists are unlikely to manipulate representations of uncertainty to serve political or other ends." (Here, Frank refers to Oreskes and Conway's *Merchants of Doubt*, 2010.) 3. "Representations of uncertainty are produced by well-established methodologies that introduce minimal higher-order uncertainty" (pp. 90–91). Frank argues that, to the extent that these conditions are fulfilled, we come close to the Jeffreyan ideal in practice. He discusses Betz' Jeffreyan criticism (2013, In Defence of the Value-free Ideal) of Biddle and Winsberg's discussion of inductive risk in climate modeling (2010, Value-judgments and the Estimation of Uncertainty in Climate Modeling), finding it misguided since in climate modeling the three conditions are not fulfilled at all.

While the last point is certainly correct, the discussion seems to get somewhat off the track since one crucial point is neglected here, namely that in climate science it is not *scientists* who are manipulating results but rather stakeholders exerting pressure on climate science (and climate modeling) from outside of academia. This makes the problem of manipulation in the context of climate science much subtler and more complex. There is a growing body of evidence that the external pressure leads to a lopsided distribution of inductive risk in climate science, namely, to a systematic preference for false-negatives (e.g., Biddle et al., 2017, *Epistemic Corruption and Manufactured Doubt: The Case of Climate Science*; Brysse et al., 2013, *Climate Change Prediction: Erring on the Side of Least Drama*; Lewandowsky et al., 2015, *Seepage: Climate Change Denial and Its Effect on the Scientific Community*), and it is puzzling why this discussion was completely ignored by Frank, alongside the issue that Steven John's (2015, The Example of the IPCC does not Vindicate the Value Free Ideal: A Reply to Gregor Betz) reply to Betz (2013, In Defence of the Value-free Ideal) should at least have been mentioned.

Joyce C. Havstad and Matthew J. Brown (chapter 6) discuss the "pragmatic enlightened model" (PEM) proposed by Ottmar Edenhofer and Martin Kowarsch (2015, Cartography of Pathways: A New Model for Environmental Policy Assessments) as a tool for the IPCC's working group III (which focuses on climate change mitigation) to reconcile the challenge of the inductive risk argument with "the IPCC's mandate to adopt a model of climate science advising that produces relevant but neutral, non-prescriptive policy advice" (p. 108). According to this model, the costs, risks, climate change impacts, and co-effects associated with a 1.5° C, 2° C, and 3.5° C warming are discussed, leading to an exploration of "a range of different policy pathways" (pp. 107–108) for the purpose of well-informed policy-making. However, as Havstad and Brown convincingly analyze, the PEM falls short of meeting the challenge of the inductive risk argument, as value-laden decisions cannot be deferred for principled reasons, namely, the "endemic and pervasive uncertainties in science" (p. 109) which are actually aggravated by multiplying effects. This leads to such a vast complex of value-judgments in climate science that the PEM seems unworkable.

While Havstad and Brown's evaluation of the PEM is adequate, some details of their discussion appear problematic. First, the fact that an implementation of the PEM requires a pathway selection does actually not pose a problem for Edenhofer and Kowarsch's approach; clearly, such choices include value-laden decisions, but they could be handed over to policy-makers. The real challenge is that scientists have to make value-judgments throughout the scientific process, often occurring implicitly and tacitly (as is actually well explained by

Havstad and Brown). And, this is also why, secondly, the inductive risk argument does not apply only to certain "kinds of cases" (pp. 105, 109, 110). Rather, it applies universally, as empirical research is inevitably both empirically underdetermined and socio-economically embedded, thus, incorporating non-epistemic considerations into methodological choices (even, as demonstrated by Staley in chapter 3, the seemingly value-free search for the Higgs boson), a point to which Frank refers as "fallibilism" (p. 89). Therefore, lastly, stakeholder participation will not lead to a transformation of PEM into "a valuable tool" (p. 119) as this cannot reconcile the challenge of the inductive risk argument with the idea of "neutral, non-prescriptive policy advice" which is clearly committed to the value-free ideal. Rather, such an attempt is doomed to fail in principle.

4 Part Three: The Breadth of Inductive Risk

The first chapter of part three by Robin Andreasen and Heather Doty (chapter 7) provides a splendid contribution on the role of inductive risk in the social sciences by examining value-laden decisions in the process of measuring inequality. First, Andreasen and Doty point out that choosing between different significance tests, such as Chi-Square and Fisher, can make a difference, namely, in cases where "either test is appropriate, but they may output somewhat different *p* values" (p. 135). Second, they refer to different methods of disparity testing, where scientists have to choose between statistical significance and the four-fifth rule, which prescribes that a minority group must be selected (when it comes to hiring, promotion, etc.) at a rate of no less than 80 percent of the majority group. While significance tests "on a small unbalanced sample can run the risk of a false-negative—namely, a *p* value that is above the established threshold for significance even when the disparity is not due to chance," the four-fifth rule "can disadvantage small employers by raising the possibility of a false-positive because statistical significance need not be established" (p. 138). In all these methodological choices, Andreasen and Doty demonstrate that non-epistemic values play a legitimate role.

Lastly, they turn to the case of operationalizing scientific variables: "In study design, scientists sometimes reference variables, like promotion and retention, that are multifaceted and can be operationalized in a number of different ways" (pp. 141–142). This entails the problem that choosing a particular variable could lead to insufficient evidence, and, depending on how a specific variable is operationalized, there could be a preference for false-positives or for false-negatives. However, as long as it is impossible to know in advance

whether a certain choice could lead to a tendency to prefer one error type over the other, value-judgments cannot play a role in this decision process (p. 143).

Throughout the chapter, Andreasen and Doty illustrate their points with recourse to a 2012 *Science* article which reports that there were no significant gender disparities among STEM faculties at 14 U.S. universities. They reveal that, while the study is not wrong, its undifferentiated presentation is problematic given the methodological choices made by the researchers.

In the following contribution (chapter 8), Anya Plutynski explores the role of inductive risk (or of the broader category of epistemic risk, see below) with regard to cancer screening, especially mammography screening. Whether or not mammography screenings for breast cancer should be commonly done on women of different age groups is a highly controversial debate. It can also be difficult to assess due to its complexity and the multitude of issues discussed. Plutynski offers a very well-written analysis of this case, which is both accessible and thorough. She also traces this complexity back towards the pervasive role of epistemic risks involved in a clear and convincing manner: inductive or epistemic risks are relevant at multiple points and during different stages of trial design, data analysis, or communication of results. Particularly, it makes a big difference how one operationalizes "effectiveness"; that is, whether one focuses, e.g., on 5-year-survival rate, reduction in age-adjusted cancer mortality, or reduction in overall mortality. What exactly is measured and how, she points out, can differ greatly with regard to catching the different harms and benefits of screening measures.

According to Plutynski, this does require a greater transparency in reporting and communicating research results. Ideally, this would lead to better informed decision-making by patients. Rather than adopting a uniform program of mammography screening, she moreover calls for a more pluralistic approach that acknowledges uncertainties and grey areas. As this seems to be a fertile idea, it would be great to see it fleshed out in more detail in the future.

Roger Stanev (chapter 9) discusses inductive risk related to the use of composite outcome measures (that are comprised of a set of different outcomes) in clinical trials. According to Stanev, the use of composite outcome measures is usually defended by two kinds of rationales. First, it is argued that composites can increase trial efficiency, that is, increase statistical power given a fixed sample size or reducing trial size given a fixed statistical power. This decreases the financial costs of trials and, arguably, leads to more new interventions. Second, composite measures are used to include outcomes that are highly

relevant to patients, such as, e.g., hospital admissions. Stanev convincingly shows that the use of composite outcome measures increases the range of methodological decisions that affect the distribution of inductive risk, and that it is highly problematic if these added risks as well as layers of uncertainty get ignored.

In particular, he demonstrates through an analysis of different cases that composite outcome measures can be very misleading if the individual outcomes subsumed in the composite have different underlying causal mechanisms, display quantitative heterogeneity regarding effect size of the components, or qualitative heterogeneity regarding their importance to patients. This is because results on effectiveness of the composite are in practice often treated as applying to all the individual components, despite a lack of statistical power to make conclusions about the latter.

While this is obviously problematic, it is unclear whether this is best described as a situation of inductive risk (which usually involves a trade-off between false-positives and false-negatives) rather than some other kind of epistemic risk, e.g., that of making a clear mistake in statistical reasoning. Stanev concludes with giving some recommendations on how to deal with composites. The gist of his recommendations, namely, to make explicit and transparent decisions on why and how to use a composite, is certainly agreeable, and it seems a good direction for future work to think about ethical frameworks for such decision-making.

In her insightful and lucidly written contribution, Robyn Bluhm examines three ongoing debates on the design of clinical trials from the perspective of inductive risk (chapter 10). The first issue is the debate on the role of randomization in clinical trials and the decisive role it is ascribed to by proponents of evidence-based medicine. She shows that the arguments put forward in favor of the central role of randomized controlled trials (RCTs) in medicine can be understood as expressing the view that RCTs reduce inductive risk of both sorts, rather than involving a trade-off. Second, Bluhm reviews the debate on pragmatic trials versus "purer" explanatory trials. Pragmatic trials aim to enhance extrapolability of the generated data, e.g., by representing the target population more adequately or by allowing for flexible intervention management during a trial. This threatens to include all sorts of confounding factors; at the same time, it is epistemically advantaged with regard to the question of what will work in practice. Thus, pragmatic trials are more prone to false-positives and -negatives, whereas explanatory trials have an enhanced risk for false-positives, *if* the results are taken to be generalizable. The third issue concerns the use of placebo versus active controls. Bluhm here explores the different ethical and epistemological arguments given for either one; again, it turns out to be mostly an issue of what kind of question we seek an answer for.

Not only does Bluhm demonstrate the fertility of the inductive risk perspective here, she also uses the insights gained for a systematic contribution to the philosophical discussion. As pointed out, many of the issues that clinical trial design faces involve different kinds of risks, and many of these issues concern methodological design (cf. Douglas, 2000, *Inductive Risk and Values in Science*) or the prioritization of different kinds of research questions and, in consequence, of information. Bluhm then uses these results to argue against a normative framework that relies on distinguishing between direct (illegitimate) and indirect (illegitimate) roles of non-epistemic values (cf. Douglas, 2000, *Inductive Risk and Values in Science*; 2009, *Science, Policy, and the Value-Free Ideal*). As not all relevant questions are of the form "how much evidence is enough," but many rather point to what kind of evidence is prioritized, how data is characterized, etc., the underlying distinction between indirect and direct roles of values does not help much to answer them.

5 Part Four: Exploring the Limits of Inductive Risk

In the last part of the book, Justin Biddle and Rebecca Kukla provide the chapter (11) with the most systematic import. As they point out, Douglas (2000, *Inductive Risk and Values in Science*) has been seminal in arguing that not only the acceptance or rejection of a hypothesis faces inductive risk, but also earlier methodological choices such as the characterization and interpretation of data. This connects well with other contributions in this volume, for instance, by Bluhm, Plutynski, or Stanev, who all expand the range of decisions in the research process that face uncertainty and thus risk. Biddle and Kukla argue that not all of these decisions involve an inductive step from a certain amount of data to a general hypothesis. They propose to limit the use of "inductive risk" to such cases and introduce a broader covering notion of epistemic risk, which distinguishes different kinds of risks. Particularly salient here are "phronetic risks," which are risks arising during "activities that are preconditions for or parts of empirical […] reasoning" (p. 220) and require balancing in the light of non-epistemic values. As Biddle and Kukla argue, such a more fine-grained typology of risks would add much analytical clarity to the debate on inductive risk.

We concur partly, in particular with respect to the case study provided by Jack Powers (chapter 12). Powers argues for broadening the concept of inductive risk by pointing to value-laden scientific language choices implying inductive risk. He supports his idea with a case study of gendered terminology in endocrine-disruption research, i.e., research on the effects of

chemical pollution (such as herbicides in the ground and surface water) on the fertility of living creatures. As a case in point, he refers to a meta-study that reviewed "the available evidence on the effects of atrazine with respect to the hypothesis that atrazine 'demasculinizes' the gonads of male vertebrates" (p. 250).

Powers is doubtlessly right to point out that "demasculinization"—in contrast to the rather neutral alternative concept "gonadal lesions"—is a heteronormatively laden concept with a clearly negative connotation; moreover, here it is used in the context of a threat to human health: It "serves to reinforce a naturalized account of heteronormativity [… and capitalizes] on societal fears of demasculinization, feminization, and gender ambiguity" (p. 251). Powers claims that this is an "example of non-type I/II inductive risk" (p. 253). Scientists have to make a decision between a clearly value-laden and a more "value-neutral" conceptual framing.

However, inductive risk *always* entails a trade-off between false-negatives and false-positives. It may also entail considering the consequences of true-negatives and true-positives, as Elliott shows (p. 265, fn. 1). But it seems a stretch to conceive problematic language choices as a case of inductive risk.

Apparently, Powers begs to differ on this point, arguing that Wilholt and Douglas also broaden the concept of inductive risk, both uncovering subtle ways in which choices on acceptance of hypotheses can be biased (namely, via choices in experimental design and data characterization and interpretation). However, in both Wilholt's and Douglas' case studies scientists have to deal with the risk of errors of type I or II when they make decisions in the production processes of data that lead to the acceptance or rejection of a hypothesis. The situation Powers envisages seems to be of a different kind. Paying attention to language choices is often important for political, moral, and social reasons, and it clearly implies epistemic and non-epistemic risks. But it is not a decision that implies *inductive* risk.

6 Summary

There are, of course, a number of quibbles and details that would deserve further discussion. Moreover, the volume also demonstrates that there is a need for further research on how to deal with the normative consequences of inductive risk. However, while the first point would be beyond the scope of this book review, it is only fair to say that the latter point would have been beyond the scope of this volume.

The editors provide a very elucidating summary and an outlook on future questions that largely fall, as they show, into three main categories: the nature of inductive risk, evading the inductive risk argument, and responsibly addressing inductive risk (p. 274). As Heather Douglas states in her foreword, “the idea of inductive risk […] has been regularly rediscovered by philosophers of science. I hope this time, it will stick” (p. x). The volume provides good reason to hope that this wish will be fulfilled.