For Values in Science: Assessing Recent Arguments for the Ideal of Value-Free Science

Matthew J. Brown*

Draft - January 28, 2024

Abstract

There is a fairly stable consensus among philosophers of science whose research focuses on science and values that the ideal of value-free science is untenable, and that science not only is, but normatively must be, value-laden in some respect. The consensus is far from complete; with some regularity, defenses of the value-free ideal (VFI) as well as critiques of major arguments against the VFI surface in the literature. I review and respond to many of the recent defenses of the VFI and show that they generally fail to meet the mark. In the process, I articulate what the current burden of argument for a defense of the VFI ought to be, given the state of the literature.

1 Introduction

There is a fairly stable consensus on the view that the ideal of value-free science is untenable and that science is and ought to be value-laden in some respect [Hicks, 2014, Lusk, 2021, Holman and Wilholt, 2022]. Considerations such as the endemic uncertainty of empirical science, the role of contingency in science, the nature of scientific practice, the pragmatic orientation of scientific inquiry, the way that the public relies on science for advice, and the normative weight of many scientific concepts have figured in a wide variety of arguments against the value-free ideal (VFI). Arguments against the VFI

*Center for Dewey Studies, Southern Illinois University, matt.brown@siu.edu, http://matthewjbrown.net

can be traced back at least to the late nineteenth century [James, 1896; see Magnus, 2013] and have grown increasing sophisticated since. The current priorities of the field have largely shifted from arguments about the VFI to questions about how to understand science and its role in society in the face of its value-ladenness.

However, the consensus is not without its challengers; indeed, defenses of the VFI appear with some regularity. The last time there was a concerted response to such challenges by philosophers of science working on values in science was in response to Betz [2013] [see Miller, 2014, John, 2015b, Steel, 2016, Douglas, 2017, Resnik, 2017, Frank, 2017, Lusk, 2021]. Betz's concerns have been answered in many ways, as have many of the arguments that came before. Many of the more recent challenges have not received the same response from the field. Thus, I will focus attention on defenses of the VFI published after 2013, surveying and evaluating them in light of the best arguments against the VFI in the recent literature. I will also consider critiques of arguments against the VFI or for values in science that do not go so far as to defend the VFI.

The literature on science and values at present is large and complex, but two kinds of questions are central. The first question is: should science be value-free, i.e., ought scientists to regulate their practice by aiming at value-freedom? There are various ways one might interpret the value-free ideal (VFI), but the most common way is the claim that only scientific or "epistemic" values can influence scientific reasoning or inference, while the only place for other values, including social and ethical values, should be in external aspects of science, such as choice of research projects or decisions about acceptable methods. I will call this question, about whether the VFI is the correct normative ideal for science, "the VFI question." The second question is, assuming that science is not value-free, how ought the role of values in science be managed, that is, when and how should values be permitted to operate in science? I will call this "the value-management question."[1]

The two questions are often conflated, giving rise to various confusions. One of these confusions is that there are a spectrum of positions on values in science, with the VFI at one extreme, and some sort of radical value-ladenness at the other. This characterization is a mistake, however, given that the VFI is an all-or-nothing affair—either social and ethical values should play a role in

[1]See Silk [2018] on this distinction. The "value-management question" has also recently been called "the new demarcation problem" [Holman and Wilholt, 2022].

the internal phases of scientific reasoning, or they should not. While different interpretations of key terms (e.g., "internal phases") can lead to somewhat different interpretations or versions of the VFI, any supposed "middle position" on the question *is* a rejection of the VFI. A related confusion is to treat various positions on the second question as if they were partial defenses of the VFI. I demonstrate this point below, when I discuss the supposed democratic defense of the VFI (which is more productively considered as a constraint on answers to the value-management question) and the various attempts to effect a partial rapprochement between defenses and criticisms of the VFI. Often, the best work that claims to be addressing the VFI question is better understood as a specific sort of answer to the value-management question. One might find the latter question more interesting, but it only makes sense to address it once the VFI has been rejected.

This paper takes up the VFI question. My argument is that the best recent criticisms of the VFI set a high burden of proof for defending the VFI, and that recent defenses of the VFI have by and large failed to meet the mark. Many of these defenses raise legitimate concerns against the older arguments that they target, but because they do not address recent moves, they are not successful in their goals. Other arguments raise important concerns about the role of values in science that must be addressed when we take on the value-management question, but they do not succeed as arguments in favor of the VFI.

I will begin with a survey of the strongest arguments against the VFI and the crucial moves and counter-moves that must be taken into account by anyone aiming to defend the VFI. Central emphasis here belongs on the *argument from inductive risk* as developed by Heather Douglas [2000]. I will synthesize this survey into a statement of the burdens of proof that must be met in order for a defense of the VFI to be considered a serious challenge to the consensus against the VFI. Next, I survey recent defenses of the VFI, grouping related defenses together, and showing that they all fail to meet one or more of the relevant burdens identified. Then, I consider various attempts to affect a partial rapprochement with the VFI, arguing that what is valuable in these arguments is better seen as addressing the value-management question. In conclusion, I argue that we either need better arguments in defense of the VFI, or we need to focus our attention on the value-management question and other issues that arise from the rejection of the VFI.

2 Arguments Against the VFI

The ideal of value-free science asserts that there is a part of science (sometimes called the "internal" or "inferential" part) where certain values ("non-epistemic values") ought not be permitted to have an influence [Douglas, 2009, Ch 3; Douglas, 2016]. According to the VFI, only epistemic values are allowed in the context of justification; for this reason, we could call it "the ideal of epistemic purity" instead [Biddle, 2013], though I will stick to the standard terminology. The VFI is typically understood by its defenders to be limited to certain parts of science; they will typically acknowledge that values have a legitimate role to play in determining which questions or problems scientists pursue or how they frame hypotheses for consideration. It is also widely acknowledged that scientific methods must be constrained by ethical considerations such as the wellbeing of human research subjects or environmental impacts. However, when it comes to the constitution of evidence or evaluating the way that evidence bears in support or against a hypothesis, only epistemic values may play a role. Scientists ought not consider non-epistemic values when making judgments in this part of science. Though there is room to debate which values must be proscribed from which phases of scientific inquiry to achieve the "epistemic purity" that the VFI demands, to defend the VFI is to insist that science ought to be absolutely value-free within those bounds.

The VFI is a normative ideal, not a descriptive claim. It is meant to govern the kinds of considerations scientists should weigh in making and justifying their decisions, actions, and claims. Defenders of the VFI can and do acknowledge that values sometimes cause scientists to act or reason in certain ways, or that values often motivate the way that they pursue their work. Defenders of the VFI readily admit that scientists are human and not perfect epistemic machines. What matters is that scientists explicitly strive for neutrality or impartiality in their explicit reasoning, as well as adopting individual and social practices that tend to minimize the influence of nonepistemic values.[2] As an *ideal*, the VFI is insulated from crass ought-implies-can arguments. It does not matter whether it is achievable; it may still be worth pursuing [see Menon and Stegenga, 2023]. The question that concerns us here is whether scientists *in practice* ought to be *guided* by the VFI, whether they ought to take it as a *regulative ideal* of their activity, or

[2] In terms of Ward [2021], it is values as "justifying reasons" that is primarily at issue in arguments for and against the VFI.

whether they ought to *strive* towards value-freedom in some sense.

An adequate critique of the VFI must be a normative argument that, all things considered, the VFI is not desirable to pursue or achieve *in principle*, even if it were possible in practice. Such critiques must show that it is normatively legitimate, even required, for scientists to consider what are typically termed "non-epistemic values" in the ordinary course of their work, in the *internal* part of science. Recent attacks on the VFI meet this burden.[3] In the rest of this section, I will review the strongest of these arguments, and some of the core moves that have been established in the back-and-forth over the VFI. In the following section, I will synthesize this work into an account of what the burdens of proof are for defending the VFI in such a way that shows genuine uptake of the criticisms from the recent literature.

2.1 The Argument from Inductive Risk (AIR)

The most influential and arguably the strongest argument against the VFI in the recent literature is the *argument from inductive risk* (AIR). The contemporary form of the AIR is due to the groundbreaking work of Heather Douglas [2000; 2009]. Havstad [2022] shows that, although Douglas adopts the term "inductive risk" from Hempel [1965], and despite the fact that Douglas's argument bears resemblance to earlier arguments from Churchman [1948], Rudner [1953], and others, Douglas's argument is distinct (296n6, 309n37). Importantly, as Havstad shows, Rudner's argument is weaker than Douglas's.

Brown and Stegenga [2023], building on Havstad [2022], have provided a reconstruction of Douglas's [2000] AIR that shows that the argument is valid and, arguably, sound:

> 1. If it is not the case that scientists ought to consider the predictable consequences of error (or inductive risk), then it is not the case that scientists are responsible for their actions as scientists.

[3] Earlier arguments against the VFI focused on the underdetermination of theory by evidence and the "gap" it shows between evidence and hypothesis. However, many of these earlier arguments are somewhat unclear on whether they take the VFI to be descriptively inadequate (in practice or in principle) or whether they take it to be untenable *even as an ideal*. The arguments reviewed below are clearer on the normative nature of the argument against the VFI.

2. If it is not the case that scientists are responsible for their actions as scientists, then it is not the case that scientists have the same moral responsibilities as the rest of us.
3. Scientists have the same moral responsibilities as the rest of us.
4. Therefore, it is not the case that scientists are not responsible for their actions as scientists.
5. Therefore, it is the case that scientists ought to consider the predictable consequences of error (or inductive risk).
6. Where scientists ought to consider inductive risks and the weighing of inductive risk requires the consideration of non-epistemic consequences, non-epistemic values have a legitimate role to play in the internal stages of science.
7. In the cases discussed by Douglas, the consequences of the choices include clear non-epistemic consequences.
8. So in these cases, scientists should weigh the inductive risks, and doing so requires consideration of clear non-epistemic consequences.

---

9. Therefore, in the discussed cases, non-epistemic values have a legitimate role to play in the internal stages of science.[4]

This argument is deductively valid. The first part of the argument (1-5) proceeds by two applications of *modus tollens*. The second part of the argument (5-9) is valid by *modus ponens* plus conjunction introduction. The conclusion (9) amounts to a denial of the VFI. It is important to point out that the argument applies in those cases in science in which "non-epistemic consequences of error can be foreseen" [Douglas, 2000, p. 578]. Where there are no non-epistemic consequences of error, or those consequences are unforeseeable, there may be no role for non-epistemic values on (this version of) the argument from inductive risk in those specific cases. On the other hand, when there are foreseeable non-epistemic consequences, weighing inductive risks means both considering the uncertainties involved in inference and evaluating the different possible consequences of error.[5]

[4]Premises have been renumbered; the rest is a direct quotation from Brown and Stegenga [2023].

[5]Different frameworks for doing the weighing might be considered (expected utility theory versus less formal and more qualitative approaches to value judgment), but this is a

Building on Douglas [2000; 2003; 2009] Havstad [2022] also ably defends the truth of each of the premises (305-309). What we have here is thus a putatively valid and sound argument to the conclusion that non-epistemic values legitimately act as justifying reasons in science, specifically in justifying how inductive risks are weighed. Because weighing inductive risks is part of the internal aspect of scientific inference, this is an argument against the VFI. As such, any defense of the VFI must show either that, appearances aside, the argument form is invalid, or that one of the premises is false.

As Havstad shows, it would seem that the only really plausible premise to attack is Premise (6). And indeed, historically, many defenses of the VFI can reasonably be reconstructed as attacks on Premise (6), attempting to show that scientists need not consider non-epistemic values when confronted with inductive risks that have non-epistemic consequences. Havstad makes two major points in defense of (6). First, it is part of the ordinary practice of science to weigh inductive risks and demand stricter evidence when the consequences of error seem more significant. She draws on Douglas [2021] in considering alterative responses to inductive risk, such as flipping a coin or using only epistemic values to decide the case, arguing that these approaches are self-undermining, because these non-standard procedures are motivated by nonepistemic values [Havstad, 2022, 308]. Second, Havstad builds on Havstad and Brown [2017] in arguing that responding to inductive risk by deferring or hedging rather than making nonepistemic value judgments is unworkable.

2.2 The Deferred Decision or Hedging Response

This second line of attack on Premise (6) is a classic response to both Rudner's argument and the AIR that has been called the "deferred decision response" or the "hedging response." Examples of this response can be seen in Jeffrey [1956], Mitchell [2004], Betz [2013], and (perhaps the best version) Edenhofer and Kowarsch [2015]. These responses argue that, though it is necessary to weigh inductive risks in order to reach conclusions about hypothesis acceptance or rejection, it is not an essential part of scientific work to reach such conclusions. Instead, scientists can defer decisions about acceptance to the context of application, providing instead all the relevant information (such as probabilities of the hypothesis given the evidence) that they would use to make such decisions. The conclusions that scientists do (or should)

value-management question, and so beyond our scope here.

reach (or, at least, assert in public) are suitably hedged in such a way as to putatively avoid inductive risks entirely. Thus, Premises (6) is rejected.

Several quite convincing lines of response have been pursued to this line of argument. The first style of response points towards a certain kind of regress of inductive risks, making it clear that scientists cannot in fact defer or hedge effectively and responsibly. This regress goes in two different directions, concerning whether the regress is *downstream* or *upstream* of hypothesis evaluation.

What I will call "the *downstream regress* of inductive risks" is pointed out already by Rudner [1953]. Rudner anticipated the deferred decision response and argues in response that inductive risks are also present in whatever information is presented to the decision-makers to whom value judgments are deferred. If instead of asserting H, the scientists assert $P(H) = p$, there are inductive risks here as well. Steel [2016] calls this "second-order uncertainty." The downstream regress has quantitative and qualitative aspects. Quantitatively, the assertion that $P(H) = p$ is not itself completely certain, but itself has inductive risks associated. Likewise, $P(P(H) = p) = p'$ and so on. Though the practical significance of these risks may decrease as one travels down the regress, certainty is never reached, and so inductive risk never goes away entirely. At a practical level, whether there are foreseeable non-epistemic consequences associated with decisions about N-order uncertainties cannot be determined in advance, *a priori*; it is itself a value judgment. Qualitatively, there are multiple options at play for models and methods of estimating probabilities that lead to different probability ascriptions and so different conclusions. In a related but more technical way, Steel [2015] shows that a Bayesian analysis of confirmation remains subject to the AIR.

There is also an *upstream regress* of inductive risks. The phenomenon of inductive risk does not only apply to the final judgment about whether the evidence supports a hypothesis or theory. There are many intermediate judgments made in scientific inquiry (or many premises and intermediate inferences made in scientific arguments) that themselves are uncertain, with potentially significant non-epistemic consequences for how they are made. Havstad and Brown [2017] follow Douglas [2009] and Winsberg [2012] in arguing that these upstream decisions are too many, and too complex, to be dealt with through the deferred decision response. What's more, the consequences of those upstream decisions may not align with those of the downstream decisions, but may bring in unique considerations. One might defer a small subset of value-laden decisions, but not all.

Frisch [2020] shows that the decisions cannot be legitimately deferred or hedged based on a principle articulated by Elliott (2011): the *no-passing-the-buck* principle. Elliott rightly points out that it can be harmful for scientists to withhold judgment on matters when their technical judgment can help inform decision-makers. As Frisch argues, "One way in which scientists may violate the no-passing-the-buck principle is to commit only to a strongly hedged variant of a hypothesis P while withholding judgment on P itself" [Frisch, 2020, p. 983]. There is a kind of trade-off between informativeness and certainty, which is itself a kind of epistemic risk different from but relevant to inductive risks.[6] Hedging gets you certainty at the cost of informativeness, but science must be *informative* and *policy-relevant* in order to be consequential to decision-makers and thus to fulfill the social role that gives it social and political weight [Elliott, 2011; Steele, 2012; Steel, 2016; Brown, 2018a; Frisch, 2020; see Menon and Stegenga, 2023]. It is not an either-or proposition (hedged or informative), but rather a trade-off that must be judged by relevant nonepistemic values. Hedging allows one to mitigate inductive risks only at the cost of limiting the social value of informative science. No one ought to deny that there are contexts where hedging or deferring to some extent turns out to be best; what opponents of the VFI rightly deny is the claim that it is *always* best in every context for scientists to present the *most strongly* hedged claims. Finding the right balance depends in part on nonepistemic values, thus undermining the argument against Premise (6).

These strategies for breaking the implication between non-epistemic consequences of inductive risks and the legitimate consideration of values in science thus fail; such critiques of the AIR are unworkable. The attempt to remove values from playing a justifying role in science has introduced a justifying role for values in science. Thus, it is reasonable to tentatively conclude that the AIR is a sound argument against the VFI.

The AIR is by no means the only important argument against the VFI in the contemporary literature. Another major argument against the VFI concerns the value-laden content of certain kinds of concepts and claims used in many of the sciences, especially in the biological, human, and social sciences [Putnam, 2002, Dupré, 2007, Alexandrova, 2017, 2018, Alexandrova and Fabian, 2022]. Biddle and Kukla [2017] and Brown [2020] present even more general arguments concerning the variety of contingencies or epistemic

[6] My gratitude to Jacob Stegenga for drawing my attention to this trade-off in a recent presentation.

risks beyond inductive risks that require value judgments. These arguments are, however, probably less strong than the AIR, lacking a deductively valid and sound presentation, and have received less attention from defenders of the VFI.

2.3 The VFI and the Actual Practice of Science

Answers to the question of whether science ought to be value-free, whether scientists ought to pursue the VFI, affirmative or negative, concern ideals. They are normative claims about what scientists ought to do or aim for, what science ought to be like. Defenders of the VFI acknowledge that scientific practice often fails to live up to the VFI, but this is no objection to the worth of the VFI. Contemporary opponents of the VFI argue instead that the VFI is unworthy as an ideal, that achieving or even pursuing it is undesirable.

I review the centrality of normative *ideals* in science to distinguish it from *idealized* images of science. Idealization of course may play an important role in philosophy of science as it does in science itself.[7] Perhaps it is useful, in constructing normative ideals, to abstract away from certain realities of scientific practice, and consider somewhat idealized accounts of what science is about. Some defenses of the VFI emphasize the conditions of rational *belief*. Others emphasize the logic of evidential support. Such defenses attempt to argue that the influence of nonepistemic values in science is irrational or misunderstands the logic of scientific inference. They present abstract accounts of inference or justification to do so. Are these idealized accounts legitimate ways of answering the VFI question?

Although we have already discussed Jeffrey's defense of the VFI in the context of the deferred decision response, another way to understand his argument is as a claim that scientific inference properly understood does not involve accepting or rejecting hypotheses, but appropriately apportioning credences to hypotheses in light of evidence (in the Bayesian way). Similarly, Lacey has long argued that the ultimate goal of science is to accumulate a stock of established scientific knowledge, which is the sort of knowledge contained in textbooks, and the decisions the scientific community makes about what counts about such knowledge ought to be value-free [Lacey, 1999]. According to Lacey, scientists "hold" a claim in this sense only after the

[7] I am grateful to an anonymous reviewer of an earlier draft of this paper for emphasizing this point.

elimination of all legitimate doubt, for all practical purposes, in the extreme long run [Lacey, 2015]. As such, there are no inductive risks involved, no contingencies remaining.

A serious problem with this family of argumentative strategies is that the operative ideas of belief, inference, or the stock of scientific knowledge are philosophical abstractions with little purchase on the socially relevant parts of actual scientific practice. The questions that guide this paper are not: Is there some context where it makes sense for philosophers to consider the abstract possibility of value-free science? Is some post hoc rational reconstruction of science possible where nonepistemic values are eliminated? Rather, the questions are: Should the VFI guide scientific practice? Should scientists strive to be value-free? Any normative work idealizes from current reality, and particularly when we are concerned with normative ideals. Otherwise, we would simply be left with descriptions of current practices. The question is, *can* the normative ideal we are defending be taken as a regulative ideal for the actual practice of science. In many cases, the philosophers' description of "science" is such an abstraction that the argument simply has no purchase on questions concerning what scientists can or should do.[8]

Elliott argues that, even if a "non-behavioral" account of belief which is value-free is defensible,[9] another central (arguably more central) cognitive attitude in science is *acceptance* rather than *belief*, where acceptance is understood as using a claim as a premise in reasoning or decisions about how to act [Elliott and Willmes, 2013, Elliott, 2013]. There's no doubt that many acceptance decisions require us to consider value judgments, as Lacey acknowledges [Lacey, 2015]. But at the cutting edge of science, and where it is pressing for scientists to deliver information for policy purposes, it is issues of *accepting* rather than *believing* or *holding* a claim that are relevant. One might argue that acceptance decisions should fall to policymakers or consumers of scientific information rather than scientists; but as we have already seen, acceptance decisions cannot universally be deferred in this way. These decisions are a proper part of science, and as they are subject to inductive risks, the AIR shows us that we must make value judgments in the process. (The role of non-scientist stakeholders in these processes should of course not be ignored, but this is a value-management question.)

[8]There may be other contexts in which it makes good sense to idealize from scientific practice and consider philosophical abstractions with little relevance to scientific practice.

[9]And there are good reasons to doubt that such accounts of belief and the belief-acceptance dichotomy are defensible, at least in the context of science [see Brown, 2015].

Science is a social practice; the results of particular scientific inquiries circulate in the community as public *assertions*, not personal *beliefs* nor even decisions to *accept* certain claims [Franco, 2017, Brown, 2021]. If the purpose of understanding whether or not science is value-free is to provide some form of normative guidance to scientists and not just to explore philosophical abstractions, then defenses of the VFI need to speak to what scientists do and how scientific knowledge actually manifests in practice. Assertions, both as items of public record and as things scientists *do* are necessarily subject to practical reasoning, i.e., value judgment. This aligns with a revised version of the AIR that Douglas [2021] has articulated recently. The virtue of this *pragmatic argument from inductive risk* is that it switches our view from more abstract and idealized (in the bad sense) discussions of inference relations towards the realities of scientific practice [Brown, 2020, 84].

Another aspect of the social nature of scientific practice is that scientific knowledge is not just a matter of individual cognition, just as it is not a matter of merely abstract inferential relations. In understanding what constitutes good scientific practice, or what is required for objective scientific knowledge, it is not sufficient to focus on individual cognition; indeed, what is good or bad for an individual reasoner may be the opposite for an epistemic community. The norms that ought to guide scientific reasoning apply at the level of community structure and process, on this view, not merely at the level of individual reasoning [Longino, 1990, Kitcher, 1990, Solomon, 2001, Peters, 2021], though the latter may also play a crucial role [Holman and Bruner, 2015, Brown, 2020, 17-18].

2.4 An Axiological Aside

Many defenders of the VFI, as well as many opponents, assume that (non-epistemic) values are inherently subjective (private, idiosyncratic) and misleading (biasing). For example, Stegenga and Menon [2023] assert that, "Evidence is truth-relevant, but values are not—values pertain to how one wishes the world were and clearly do not indicate how the world is." Here Stegenga and Menon clearly mean non-epistemic values, as epistemic values are often defined as truth-promoting values [Steel, 2010]. The assumptions about values that they make explicit are often left implicit; in either case, I have repeatedly criticized such a view of values [2013b; 2013a; 2018b; 2020, Ch. 3]. It is problematic in at least two respects: first, it conflates the important distinction between what one happens to prefer and warranted

judgments about what is prefer*able*. Second, it ignores the ways in which warranted value judgments incorporate empirical evidence, and so the ways in which highly warranted value judgments can be truth-relevant.

Some (apparently) nonepistemic values may be presupposed by science, such as honesty and what **?** called "communism" or "communalism" [see Bright and Heesen, 2023]. Others may be presupposed by democracy itself, such as liberty, equality or fairness, and solidarity. None of these values are subjective in the sense of being private and idiosyncratic. By contrast, many values typically classified as epistemic, cognitive, or pragmatic – often treated as unproblematic by defenders of the VFI – can be idiosyncratic and biasing preferences. As Douglas [2009] points out, "A simple theory, though elegant, may just be wishful thinking in a complex world" (107). Many such cognitive values and pragmatic considerations are better lumped with the so-called "non-epistemic values" for purposes of value management [Douglas, 2009, 108]. Even the value of empirical adequacy may mislead in contexts where there is good reason to doubt the available evidence [Brown, 2017, 70] or where it gets in the way of other legitimate epistemic aims [Bhakthavatsalam and Cartwright, 2017].

It is also worth recalling that statements of the VFI depend crucially on the distinction between epistemic and non-epistemic values; that distinction, and the ability of the distinction to support the VFI, has been strongly challenged in the literature [Rooney, 1992, 2017, Longino, 1996, Douglas, 2013a]. One might be inclined to accept Steel's [2010] influential version of the distinction, but that version comes with a problem for would-be defenders of the VFI. According to Steel, epistemic values are whatever values promote attainment of truths (intrinsically or extrinsically) in a specific context. That means that, if feminist political values have a tendency to uncover bias and promote better science, as feminist philosophers of science have shown they do, they count as extrinsic *epistemic* values [Clough, 2003; Anderson, 2004a; Hicks, 2014, p. 3284; Rolin, 2015, p. 159; Rooney, 2017, p. 41; Brown, 2020, p. 97]. What's more, it seems unlikely that we could know that feminist values were epistemic values *before* pursuing science according to such values. So the best account of epistemic values seems in tension if not directly in conflict with the VFI, and so itself may undercut any defense of the VFI.

3 The Burden of Proof for VFI

The previous section briefly reviewed the main normative arguments against the VFI in the science and values literature. From this review, we can see a few things that are clearly at stake in arguments for and against the VFI. First, the VFI is not a descriptive claim about what scientists do, nor what they can possibly do; nor are the arguments against the VFI claims about what scientists do or can do. It is a normative issue: critiques of VFI find it normatively inadequate or untenable, because, they argue, values are necessary considerations in justifying decisions made in the course of scientific inquiry or inference.

The AIR as articulated by Douglas [2000] and explicated by Havstad [2022] and Brown and Stegenga [2023] is the argument to beat; it seems clearly valid and likely sound. Many arguments in defense of the VFI can be reconstructed as attacks on Premise (6) of AIR, but those attacks have been carefully rebutted in a way that further discussions must take into account. Other arguments in addition to AIR raise the need to consider epistemic risks beyond the scope of inductive risk as well. It is also important not to treat the question as one of an abstract logical structure, but one of the social practices of acceptance, assertion, and knowledge-production in science. Finally, flat-footed claims about the nature of value and of the epistemic/non-epistemic value distinction must be avoided.

From the foregoing discussion, we can see that the existing literature on values in science sets a certain burden of proof for any would-be defenders of the VFI. One must do all of the following to successfully and *completely* defend the VFI:

(A) One must show that the AIR is either invalid, or one must identify which premise of AIR is false.
(B) In doing so, one cannot simply repeat the deferred decision / hedging response, at least, not without countering the main rebuttals of that response.
(C) One also must show that one's argument applies (as an ideal) to the social practices of science, and not merely some philosophical abstraction.
(D) A defense of the VFI cannot naively assert that values are subjective biases, nor rely on a naive version of the epistemic/non-epistemic value distinction without qualification.

This is what it takes to fully engage with the nuances of the current literature

on this topic in a way that is responsive to criticisms of previous defenses of the VFI. Insofar as we agree that responsiveness to criticism of the very claims one is defending and the type of arguments one is using to defend them is an important standard for philosophical argument, any defender of the VFI should be willing to accept these standards. That said, addressing all of (A-D) is a complex task. One need not expect that every author addresses all of these points. Perhaps an author wishes to focus exclusively on (A), and while this might require some consideration of (B) and (C), one could reasonably put aside (D) for future work. However, one *should* expect that defenses of the VFI not simply ignore these issues, and that they both acknowledge the problems and suitably qualify the scope of their argument in response. With respect to (D), one might note that one's argument relies on a fraught and disputed distinction, while putting off for future work providing an adequate account of the distinction that responds to the criticisms of it. However, at some point the defenders of the VFI must take up the issue. If it is continually put off, it becomes a serious concern about the defensibility of the VFI.

To be fair, some of the defenses I will canvass in the next section predate some of the key arguments canvassed above. In particular, one might find the insistence on (A), and particularly the criticism of arguments that fail to engage with Havstad [2022], unfair in at least those cases where the articles were written *after* the appearance of Havstad's article (in July 2021). But there are two things we can say here. First, Havstad only makes more explicit and clear the structure of arguments already made more than two decades ago by Douglas [2000]. (Indeed, as I will show, several authors ignore Douglas's text entirely, assimilating its arguments to the earlier and distinct argument of Rudner.) There has been plenty of time for authors to be responsive to those arguments. Second, anachronism is not really relevant, as the goal here is not to evaluate the authors but the current viability of their defenses of the VFI. I will do my best to reconstruct their arguments as responses to Havstad's best version of the AIR, but if they fall short, that means that their defense of the VFI cannot be considered adequate today.

On (D), there is significant dispute between those who *reject* the VFI on these questions about the nature of values and the epistemic/non-epistemic value distinction.[10] Why expect better from the defenders of the VFI? It would be too much to ask defenders of the VFI to give a definitive account of the nature of value. It is not too much to ask, though, that they not

[10]My thanks to anonymous reviewers for pushing on this point.

commit to claims that have been subjected to powerful criticisms, without acknowledgement or qualification. Simply asserting, without argument, that values are subjective biases, is to rely on a prima facie inadequate account of values. Likewise, to uncritically assume that there is a bright line between epistemic and non-epistemic values that can do heavy lifting runs up against strong arguments mentioned above. Doing so seriously weakens arguments for the VFI, and at minimum the defenders need to acknowledge that. Similarly, while (C) does not mean that defenders of the VFI should not entertain useful idealizations, it does mean that they need to defend the relevance of their results to actual practice.

4 Recent Defenses of VFI

In the previous section, I articulated four criteria for consideration as an adequate, complete defense of the VFI. In this section, I will review a variety of recent defenses of the VFI. Each of the arguments presents itself as an attempt to defend the VFI or refute an argument against the VFI, though in some cases these arguments are merely entertained rather than endorsed (as with the arguments by Bright and Lusk). Attempts to affect a partial *rapprochement* between the VFI and its critics, or to defend a thesis adjacent to the VFI, or to simply raise concerns about value-laden science will be discussed in the next section.

I group these defenses into a few types: there are those who would revive the deferred decision or hedging response, those that focus on the role of science in liberal democracy, those that retreat to a kind of ideal theory, and those that distinguish the value of pursuing versus achieving the VFI. I will show that each defense of the VFI fails to meet at least one of the criteria (A-D) relevant to the scope of its argument, and often more than one. There are several ways they fail to meet these burdens. In some cases, they fail to meet the criteria completely, because a severe lack of engagement with the literature. In other cases, they address the criteria inadequately, by making moves that have already been criticized thoroughly. Finally, even when novel responses to the criteria have been made, there are still serious problems with the argument that render them unconvincing. If I am right about these criticisms, it follows that there is a widespread failure in work defending the VFI to respond adequately to criticism of its arguments.

4.1 Deferring and Hedging Revisited

Several recent defenses of the VFI attempt to revive the deferred decision or hedging response [Henschen, 2021, Cassini, 2022, Carrier, 2022, MacGillivray, 2019]. Most of these attempts fall far short of the burden of proof, rarely even engaging with the major moves in the literature that have been made in the past in response to similar arguments. Henschen [2021] and Cassini [2022] only reconstruct Rudner's weaker precursor to the AIR, and do not even consider Douglas's argument in detail, nor the reconstruction by Havstad that seems to show the argument is sound.[11] Although many responses have been made to this style of argument, none even appear in the bibliography of Henschen's paper, except references to Douglas.

Although Henschen fails to consider the literature on the deferred decision response (B), he does make some arguments worth considering. The core of Henschen's argument is three-fold: (1) the regress arguments against the deferred decision response fail, because in scientific practice, scientists typically don't consider higher-order uncertainties (8-11); (2) if scientists accept or reject hypotheses, they need not do so categorically, but only hypothetically (11-12); (3) we can distinguish between using value judgments or pragmatic considerations in decisions to accept/reject or believe/disbelieve a hypothesis, and only pragmatic considerations are unavoidable (16). In that sense, science might not be entirely epistemically pure, but it can nevertheless remain value free. Although Henschen does not explicitly identify a fault with the form or premises of AIR (A), we can interpret him (and the other defenses in this category, as we saw in §2.2) as claiming that Premise (6) of AIR is false. (Recall that Premise (6) says, where weighing inductive risks involves non-epistemic consequences, it is legitimate to use non-epistemic values in the internal stages of science.)

With respect to Henschen's consideration the regress problem for the deferred decision response, he is aware of both the upstream and downstream versions as presented by Rudner and Douglas, respectively. His argument tries to reduce the upstream regress of inductive risks (IR)—which concerns decisions like the weighing of inductive risks in determining evidence prior to considering how the evidence bears on the hypothesis—to the downstream regress of IR—which concerns second-order uncertainty and the IR with

[11] As mentioned above, while it is not fair to blame Henschen for not citing an article that came out after his, it is entirely fair to say that his argument for the VFI fails by our current lights because it cannot refute Havstad's reconstructed AIR.

accepting probability assignments. This argument fails, because Henschen confuses IR with uncertainty. Henschen writes, "But the problem with Douglas's suggestion is that the inductive risk that is present at the three lower stages adds up to the total risk of the null hypothesis in accordance with the laws of probability" [2021, p. 7]. Uncertainties can be aggregated in the way Henschen describes; but IR involve not only probabilities but outcomes, which are typically qualitatively different things. The appropriate value judgments concerning aggregated uncertainties at the conclusion of inquiry cannot be the same as value judgments concerning decisions made in intermediate stages of the process, which may concern outcomes quite orthogonal to later decisions.[12]

Henschen's response to the downstream regress issue is also inadequate. Recall, there are both quantitative and qualitative aspects to the problem; he ignores the qualitative aspect of the problem entirely. With respect to the quantitative aspect, Henschen provides a description of common statistical procedure in the sciences, and then asserts without much argument that the last probability asserted by that procedure can legitimately be asserted with no further IR, without reference to Steel's [2016] argument that this cannot be done. One might defend Henschen's move here with reference to (C), since scientists do not in practice consider these higher-order uncertainties. But this won't do, as it is also not the case that scientists typically follow the hedge-and-defer strategy. In any case, this is not the right way to read (C): it does not bar deviations from ordinary scientific practice that can be normatively defended; it instead concerns abstract idealizations whose normative structure lacks clear application to scientific practice. What's more, Henschen and Cassini entirely fail to consider the trade-off of certainty and informativeness; their recommendations would have science advisors often passing the buck and providing deleteriously uninformative advice to policymakers.

This last point is relevant to Henschen's distinction between categorically and hypothetically accepting a hypothesis. Henschen argues that even if it is valuable for scientists to actually accept a hypothesis, they should not do so categorically (making bald claims); they should instead assert the hypothesis "only hypothetically" along with qualifications concerning the probability of the hypothesis and a cost-benefit analysis of relying on the hypothesis for a specific course of action being considered. It is not clear why Henschen

[12]It is strange that Henschen ignores this point, as it is actually quite central to the argument of Jeffrey [1956] that he relies on.

believes that this hypothetical hypothesis acceptance is somehow value-free. Although it is more hedged than bald, categorical assertion of the hypothesis, such assertions are still made under uncertainties and require decisions about how to trade-off uncertainty and informativeness. This approach could be a candidate answer to the value-management question, but cannot rehabilitate the VFI.

Henschen also revives a related objection from Levi [1960], that relies on a distinction between *belief* and *action* (or acceptance). In bringing on board Levi's objection, Henschen commits to an abstract "non-behavioral" account of belief that has no purchase on scientific practice (C). Henschen fails to clearly address what distinguishes epistemic values from non-epistemic, and this leads him astray insofar as he draws a poorly explicated distinction between value judgments and "pragmatic considerations." Henschen writes, "It is only in the case of value judgments that the antecedents refer to valuations of the utility of specific individuals or groups. In the case of conventional or pragmatic reasons, the antecedents make reference to technical goals" [Henschen, 2021, p. 16]. But these goals are not purely epistemic, and it seems clear that pragmatic considerations are a subset of value judgments, no less problematic than other nonepistemic values. Henschen also falsely attributes a concession that these considerations "can be regarded as epistemic values" to Staley (2017). Staley only acknowledges that these values can be considered "extrinsic epistemic values" [*sensu* Steel, 2010] in the right context; but so can moral and political values (see §2.4). Steel's account of epistemic values cannot support the VFI; Henschen has thus failed at (D). Henschen fails along all the criteria; even if we can reconstruct his argument as attack on Premise (6) of AIR, it is not a successful attack in light of the existing literature.

Cassini [2022] relies on a Bayesian account of model assessment to argue for the VFI in the case of simulation models.[13] Cassini acknowledges that

[13]Some have thought, incorrectly, that adopting Bayesian analysis instead of null-hypothesis significance testing (NHST) is sufficient to refute the AIR. MacGillivray [2019] argues that risk assessment should aspire to value-neutrality by assuming that the AIR depends on NHST. He gives good reasons to think that NHST is problematic as a risk assessment framework, but these reasons do not touch the argument in question. While NHST can provide an easy way to explain how the AIR can be applied, note that the actual argument does not refer to any specific features of NHST, only the possibility of error and the need to make decisions in the face of inductive risks. These and other problems with his argument have been thoroughly explored by Hicks et al. [2020].

decisions to accept are practical decisions, susceptible to practical reasoning (C). But like Jeffrey, Cassini denies that these decisions must be made by scientists, and so he falls into the deferred decision response, attempting to deny Premise (6) of AIR. Cassini is a little better at acknowledging the existing literature, but oddly, though he does cite the most relevant argument against his own [Steel, 2015], he rejects it without detailed argument as "too general to be examined here." He then asserts, without argument (and falsely), that the downstream regress problem can be solved on purely epistemic terms. The only arguments he addresses in detail are those of Rudner, Douglas, and Winsberg [2012; 2018b; 2018a]; in doing so, he really only considers the problem of the upstream regress.

In Cassini's analysis, all of the upstream IR issues can be aggregated into the prior probabilities associated with the hypothesis, the evidence, and the relevant background knowledge.[14] This conflates IR with uncertainty in ignoring the qualitative difference between valued or disvalued outcomes. Also, if values influence the priors in this way, we know that choice of priors influences the posterior probabilities that results. This is actually sufficient to reject the VFI [Steel, 2015]. Cassini's counter to this point depends on the fact that in the long-run, posterior probabilities should converge. This, like Lacey's account of value-free "holding," is an abstraction irrelevant to much of scientific practice (C). We rarely have time to wait for the long run to come. Cassini's strict separation between scientist and science advisor to policy is likewise an abstraction with little practical import, given the mixing of these roles in practice [see Douglas, 2009, p. 82]. Even if we consider (D) outside the scope of Cassini's argument, he clearly fails to meet the burdens of proof for defending the VFI.[15]

It seems clear that the deferred decision response has not advanced much. Some find the response compelling, but it is clear that those who do so have not really engaged with the arguments against it.

[14]These include the prior probabilities $P(H|B)$ and $P(\neg H|B)$ as well as the likelihoods $P(E|H\&B)$ and $P(E|\neg H\&B)$.

[15]Dressel [2022] appears at first to take the same approach as Henschen and Cassini. However, Dressel distinguishes a *descriptive* and *normative* sense of the VFI, and his argument only defends a *descriptive* version. This mistakes the stakes of the debate; we are concerned with whether values ought to, legitimately, play a role in science. Dressel acknowledges that AIR refutes the normative VFI, and defends a position close to Steel [2010], which is a rejection of VFI. As such, Dressel does not provide a defense of the VFI.

4.2 The Democratic Defense

Another argument made by Betz [2013] has continued to receive significant play in the literature: that value-laden science is incompatible with the requirements of liberal democracy. This has been called the *democratic defense* of the VFI or the *political legitimacy argument*. The argument is that we must uphold the VFI in order for science to play a legitimate role in democratic governance, which we cannot really do without [see Douglas, 2009, p. 8 & Ch. 2]. For instance, in one of the most interesting versions of this argument, Bright [2018] reconstructs W.E.B. Du Bois as arguing that, if scientists are not seen to be following the VFI, they will lose public support and trust, and this will undermine the attempt to use science to forward socially valuable goals.[16] Referring to the AIR and the older underdetermination arguments, Lusk [2021] says: "Despite the recent success these arguments have found, they fail to address one of the central historical motivations for adopting the value-free ideal: political legitimacy" (103).[17] Others make similar claims [Kappel, 2014-05, Kappel and Zahle, 2019, Carrier, 2022].

This is one of the most serious concerns about value-laden science, and in my view very much a live issue. It is useful to look at Lusk's version of the argument:[18]

> 2. [Legitimacy Premise] No set of non-epistemic values should have an undue influence in coercive democratic political decisions.
> 3. [Infiltration Premise] If non-epistemic values play a role in the empirical justification of political decisions, then those values have an undue influence[. . .]
>
> C. Therefore, it is not the case that non-epistemic values should play a role in empirical justification in democratic decision making. (104)

[16]This oft-repeated empirical claim about the public [see Menon and Stegenga, 2023] is doubtful in the face of the evidence presented by Hicks and Lobato [2022].

[17]Holman and Wilholt [2022] make a similar point concerning the significance of trustworthiness and social legitimacy of scientific knowledge.

[18]I have ommitted one unnecessary and one redundant premise from Lusk's original presentation. Note that Lusk articulates this argument, but ultimately rejects it. Nevertheless, he is right to urge us to take the argument seriously, and I think his exposition is helpful.

If this argument is correct, one of the most important background justifications of the AIR, that science has a clear and significant social impact because of its role in policymaking, becomes a reason instead in support of the VFI. However, each premise of this argument is shaky. Lusk [2021] concentrates on the third premise (infiltration), arguing that it is possible for non-epistemic values to have an influence that does not amount to "undue influence." Also, it seems likely that *some* non-epistemic values have a legitimate role to play in democratic decision making, e.g., the ethical values constitutive of science, the political values constitutive of liberal democracy, or the values of the majority (legitimacy premise).

The political legitimacy argument and the AIR cannot both be sound, but it seems clear that the former is on much shakier ground than the latter. As the political legitimacy argument is typically not accompanied by any direct refutation of premises of the AIR, one suspects something has gone wrong. In my view (following Lusk), although many of these arguments are presented as defenses of the VFI, they do not really amount to full defenses despite the intentions of the authors. Rather, they present a challenge that accounts of value-laden science need to address. In other words, they conflate the VFI question with the value-management question; their real value is in challenging us to think carefully about the later in the context of the political roles of science. And the literature has risen to the challenge, shifting a great deal of attention in recent discussions to matters of democratic values and the political legitimacy of value-laden science [e.g., Douglas, 2005, 2012, 2013b, Kitcher, 2011, Pinto and Hicks, 2019, Douglas, 2021, Boulicault and Schroeder, 2021, Schroeder, 2021, Lusk, 2021, Alexandrova and Fabian, 2022].

One could (by reading "legitimate" in the AIR as "politically legitimate") read this argument as a rejection of Premise (6). I do not think this is the right interpretation of AIR (where legitimacy should have both an epistemic and ethical aspect as well), not does it seem to be what is intended by most of these authors. In any case, this tactic would drive the debate back to the deferred decision response. As none of these arguments adds any new reasons to think that the deferred decision response works, this cannot vindicate the democratic defense of the VFI.

4.3 Ideal Theory Responses

Another approach to defending the VFI is to retreat from the messy world of scientific practice and the science-policy or science-society interface to

the realm of epistemic ideals. This sort of retreat automatically raises the specter of criterion (C), but we should not dismiss these arguments out of hand. We should ask, instead, can the ideal theory thus devised shed any light or provide any guidance to scientific practice, or our evaluation of it? Unfortunately, it seems that it cannot.

Hudson [2016] curiously separates two parts of the AIR into two separate arguments he calls "the uncertainty argument" (based on Levi's [1960] reconstruction of Rudner's argument) and "the moral argument" (based on Douglas [2009]). This *divide et impera* strategy is problematic, as it separates key premises that work together to make the argument sound.[19]

In his critique of "the uncertainty argument," Hudson deals with Rudner's weaker version of the argument, rather than the stronger version presented by Douglas [2000], presenting a version of the deferred decision response. He presents the argument in four steps [following Levi, 1960]:

(1) The scientist qua scientist accepts or rejects hypotheses.
(2) No amount of evidence ever completely confirms or disconfirms any (empirical) hypothesis but only renders it more or less probable.
(3) As a consequence of (1) and (2), the scientist must decide how high the probability of a hypothesis relative to the evidence must be before he is warranted in accepting it.
(4) The decision required in (3) is a function of how important it will be if a mistake is made in accepting or rejecting a hypothesis.

Hudson repeats Jeffrey's argument against premise (1), though he acknowledges the problems that have been raised with the deferred decision response (B). He dismisses these concerns with the claim that VFI is an "epistemic ideal" that sets out what is "epistemically preferable." This is an abstraction (C) that does not really address the terms of the argument about what is *legitimate* in the course of scientific reasoning. That is, rather than address those problems head-on and tell us what scientists could do to defer effectively and responsibly, he simply retreats to ideal theory. But we are not

[19]Although Havstad [2022] identifies two major argumentative moves in the AIR, the whole argument is needed to get to the rejection of the VFI. The reconstruction above §2.1 shows that AIR is one integral argument involving elements of the two arguments Hudson treats separately.

concerned with what is epistemically preferable, but what is preferable all things considered. We don't want an epistemic ideal, but a *scientific* ideal, that is, an ideal to guide scientists who have both epistemic and social duties.

Hudson's objection only works because the so-called "moral argument" has been separated out and treated as free-floating from the AIR, making (4) seem unmotivated when there are many ways one might resolve the decision about hypothesis acceptance. Hudson's analysis of the "uncertainty argument" also confuses the AIR with the older "background assumptions" model of the underdetermination argument by emphasizing his premise (3) as a kind of "gap"; for all of its historical importance, this "gap argument" is not as sharp or as strong as the normative AIR canvassed above. It is crucial to directly address the strongest version of the argument available, which Hudson failed to do by not addressing Douglas's argument directly. Today, that means the version presented by Havstad [2022] and tweaked by Brown and Stegenga [2023]. To his credit, Hudson recognizes the need to provide a compelling account of values and the epistemic/non-epistemic values distinction (D). However, while Hudson raises some potential concerns about older arguments about this by Longino [1990] and Rooney [1992], he doesn't actually provide an account himself, nor does he respond to the kinds of considerations raised by more recent arguments, such as Steel [2010], Douglas [2013a], and Rooney [2017].

What about Hudson's treatment of "the moral argument"? Here he refers directly to Douglas [2009]. He presents Douglas's argument as a hasty generalization from cases of reasonable ethical limitations on research methods (e.g., to protect research subjects) or research topics (e.g., those with potentially catastrophic applications) to the internal stages of science: "Inspired by these sorts of cases, Douglas' view is that every scientific decision is at the same time a moral choice" [Hudson, 2016, 178]. He sees her argument as "largely based on her intuitions" (179) rather than a sound argument. Hudson's rebuttal of the "moral argument" fails because it treats the argument as if it is totally disconnected from the issue of inductive risk. It makes Douglas's claim that evidential standards are value-laden seem unmotivated, whereas the inductive risks faced in the course of policy-relevant science are the precise motivators of the argument. These arguments do not stand separately. His main objections to the moral argument are that the rejection of the VFI will hurt the authority of science (i.e., its trustworthiness or political legitimacy), and further, that the objectivity of science will suffer. These are legitimate concerns about value-laden science, but they should be dealt

with as considerations for the value-management question, as Douglas [2009] has already done, and many following her have developed further. Many have argued that the value-ladenness of science is no threat to its objectivity [Harding, 1995, Hoyningen-Huene, 2023].[20]

Sheykh-Rezaee and Bikaraan-Behesht [2023] argue that the VFI is an "epistemic ideal," similar to Hudson [2016]. Their argument amounts to a kind of burden-shifting argument. Rather than attack the premises of the AIR, with which they identify no fault (A), they attempt to argue that, despite the conclusion that "non-epistemic values have a legitimate role to play in the internal stages of science," (even a *necessary* role, they concede), this is somehow not sufficient to defeat the VFI. Sheykh-Rezaee and Bikaraan-Behesht [2023] claim that any argument against the VFI should meet the following conditions:

> (1) There are *non-epistemic* values that should be shown to have some role in science.
> (2) The role the non-epistemic values play in science should be shown to be at the *internal* stages of science.
> (3) The role the non-epistemic values play at the internal stages of science should be shown to be *necessary* for scientific practice.
> (4) The role the non-epistemic values necessarily play at the internal stages of science should be shown to be *constitutive*. (146)

The AIR meets conditions (1-3), as they admit. What they introduce is this "constitutivity" condition; although non-epistemic values play a legitimate, even *necessary* role in scientific reasoning, they do not play a *constitutive* role, and so the VFI is unscathed. I find that this idea is not particularly clear nor well-motivated. More importantly, it relies on an abstraction that has no relevance to the kinds of decisions scientists make in practice (C). An ideal for scientists must be able to *guide* scientific practice. Labeling parts of

[20]Hudson [2021] pursues a unique strategy for defending the VFI, which in turn has received a thorough and convincing reply from Douglas and Elliott [2022]. According to Hudson, denying the VFI will exacerbate the replication crisis. Some of the same confusions from the earlier article persist, such as Hudson's confusion of the AIR with the background assumptions version of the underdetermination argument. Douglas and Elliott [2022] point out that Hudson conflates "value-laden" with "biased" (D), while Hudson's [2022] reply doubles down, dogmatically asserting that values inevitably lead to bias.

scientific practice "non-constitutive" does precisely nothing. What's more, when considering practically what scientists *should* do, Sheykh-Rezaee and Bikaraan-Behesht trot out the deferred decision response without responding to any of the criticisms of that response (B).

4.4 Pursuitworthiness of the VFI

Menon and Stegenga [2023] and Stegenga and Menon [2023] provide some of the best attempts at defending the VFI, in terms of responsiveness to the existing literature. They argue, correctly, that the VFI *might* still be worth pursuing, even if unattainable. That is, its unattainability does not imply that it is not pursuitworthy. This is widely acknowledged in the literature,[21] though opponents of the VFI obviously argue that it still does not turn out to be good to pursue the VFI. Menon and Stegenga also argue, intriguingly, that the VFI might be worth pursuing even if it would be undesirable to achieve as an end state in principle. They cast the VFI in a helpful new way: it aims at the elimination of "bifurcation points," i.e., decision points where difference in value judgments made in the course of inquiry would lead scientists to infer different conclusions from the same evidence.[22] According to their interpretation of the VFI, values should not make a "difference to inference" [Stegenga and Menon, 2023].

Understood this way, Menon and Stegenga still do not argue that science can or *should* be value-free. They acknowledge that the AIR shows that values should play a role in science in some cases. It is tempting to classify this as a kind of partial rapprochement with the VFI rather than a defense per se, but they insist that it is instead a legitimate alternative framing of the VFI, according to which: "Scientists should [typically] act *as if* science should be value-free" [Menon and Stegenga, 2023, *emphasis added*].[23] Between their reframing of their ideal and of the stakes of the debate (away from what ideal is worth achieving, towards what ideal is worth pursuing), they hope to obviate the key arguments against VFI, including most of those canvassed above. One tempting response here is that they have simply given up the game on the VFI as most philosophers of science understand it, and so no further consideration is necessary, if the VFI question is our target. In explicitly

[21] One simply has to think about what ideals are.

[22] They acknowledge that it is legitimate for values to play a role in many upstream decisions that determine both how conclusions are framed and what evidence is available.

[23] In the paper they call this version of the VFI, "VFI4."

denying the need to do (A), the thinking here goes, they concede the main substantive issue. This is too hasty, however, as there is a substantive issue when it comes action-guiding advice to scientists (C): should scientists make (appropriately warranted or constrained, non-epistemic) value judgments as a routine part of their inquiries, or should they behave as if science ought to be value free, and thus typically deploy strategies that obviate the need to make value judgments?

Menon and Stegenga provide three reasons to think that value-freedom is worth pursuing even if unattainable and undesirable as an end-state: (i) when values influence scientific reasoning, it becomes less truth-apt; (ii) value-free science advising is more democratic; (iii) public trust in science depends on value-freedom. (i) can be seen as cherry-picking examples of bad value-laden science while denying that positive examples are probative. (iii) is not supported by empirical studies on the relation of trust and value-ladenness [Hicks and Lobato, 2022]. (ii) & (iii) are also versions of the democratic defense of VFI, which, I have argued, is not really a defense so much as a desideratum for answers to the value-management question; in this case, the advice is to minimize their influence unless doing so runs afoul of more important constraints, such as those, "concerning resource use, research ethics and action-guidance" [Menon and Stegenga, 2023].

The cherry-picking issue is a common strategy in defenses of the VFI. Menon and Stegenga refer to Lysenko's critique of Mendelian genetics and pre-1970s androcentric primatology. These are unhelpful examples for their argument, though not uncommon ones. What makes the Lysenko case problematic is not the influence of values so much as the backing of the authoritarian Stalinist regime and the brutal repression of dissent; indeed, values have played a valuable role in the critique and revision of simplistic Mendelianism [Levins and Lewontin, 1985; cf. Graham, 2016]. The primatology case is of course a favorite in the context of feminist science studies, but many feminist philosophers of science have argued that one cannot simply see the move away from androcentrism in primatology as a move from value-laden to value-free, but rather as a replacement of androcentric with (better) feminist/egalitarian values [Harding, 1986, Longino, 1992]. It seems clear that the motivating examples they choose do not settle the case.

When it comes to practical advice for scientists (C), Menon and Stegenga suggest that scientists should typically adopt value-mitigating strategies, i.e., strategies that will help eliminate bifurcation points, except in cases where moral and practical constraints make adopting those strategies undesirable.

(In this way, they hope to accomodate Douglas's argument that value-free science is *irresponsible* science.) While there may be contexts where value-mitigation may be a good idea, in other cases it is crucial that scientists use (non-epistemic) values to weigh inductive risks, as they admit. More importantly, there is no way to tell ahead of time which case we are in; so in fact, scientists will have to continue weighing non-epistemic values throughout inquiry in order to determine whether value-mitigation is permissible or superior to explicit value judgment. That they miss this point suggests that they are thinking too abstractly about scientific inference (C). In any case, whether to pursue VFI must be judged in each case according to non-epistemic values, effectively undermining the idea that this approach is value-free (even in their *as-if* sense). This does not hamper the potential value of their approach as an answer to the value-management question instead of the VFI question, which would be a more productive way to frame their work.

Their recommendation of value-mitigation is also susceptible to another self-undermining worry, raised by Havstad: because value mitigation is not a typical canon of scientific procedure, and is justified in part on the basis of non-epistemic values, they introduce just that which they are attempting to remove. They could argue that many of the strategies for value mitigation that they promote (bias reduction, evidence strengthening, deferral, hedging) are indeed common methodological canons in science. There is a danger here of confusing values and biases [neither implies the other, as argued by Douglas and Elliott, 2022]. What's more, these strategies are typically deployed for reasons *other* than value mitigation. Whether in this case to prefer value mitigating strategies will require a complex non-epistemic value judgment concerning the values that value mitigation promotes (such as democratic legitimacy and public trust) as well as the values that pull in the opposite direction (including those they acknowledge, such as action guidance, research ethics, responsible use of resources). So the self-undermining concern remains, as far as I can see.

Stegenga and Menon [2023] pursue a similar strategy. The novel move in this paper is a focus on *scientific consensus*. Values in science (understood as bifurcation points that make a difference to what conclusions are inferred from evidence), they argue, impede the achievement of consensus. They defend what they call *strong consensus*—not only must scientists agree on the conclusion, but they must endorse the arguments that lead to the conclusion. It is not clear that this argument meets the mark; as Stegenga and Menon themselves point out, there is no necessary connection between values and

consensus. Values might play exactly the role the AIR specifies, but not interfere with strong consensus:

> For some scientific hypothesis, values could modulate the acceptable false positive and false negative probabilities for all relevant scientists, decision-makers, or people in general in the same way, such that all people either believe or disbelieve that hypothesis. That, in turn, would entail that values would not threaten the potential of a strong consensus about the hypothesis and thus the status of that hypothesis as scientific knowledge. [Stegenga and Menon, 2023, p. 435]

Indeed, on views like those of Kourany [2010], consensus about specific value judgments is at least as desirable as consensus about factual judgments. So whether or not strong consensus is desirable, it does not tell either for or against the VFI.

This points to an ambiguity in the "difference to inference" characterization of the influence of values in science. Values might make a kind of *counterfactual* difference to inference if adopting different values (or failing to consider values) would lead to different conclusions, while it nevertheless being the case that only one set of values is in fact considered or even defensible for consideration. This won't do; their argument instead requires the existence of actual (rather than counterfactual) irresolvable value disagreements influencing science. While it is easy to be pessimistic in the face of certain disputes of this kind, hope on this point is at least as attractive a regulative ideal as the VFI. And the relevance of this hope has been acknowledged at least since Rudner [1953], who argued that, "[A] science of ethics is a necessary requirement if science's progress toward objectivity is to be continuous" (6). In other words, one might address their concern by adopting an ideal of resolving value disputes rather than an ideal that would have us ignore the moral obligations of scientists.

What's more, strong consensus is a highly controversial claim; it is both unfeasible and undesirable in the views of many. Stegenga and Menon briefly discuss the views of Paul Feyerabend on this point, as the main proponent of the value of dissent, but arguably they miss the force of his argument. Not only is dissent productive for science, on his view, consensus is positively detrimental for scientific knowledge, for broadly Millian reasons (we lose our understanding of the ground and meaning of our beliefs without dissent).

Solomon [2001] argues persuasively, though on very different grounds, that there is not much epistemic significance to consensus.[24] If one's premises are more controversial than the controversial claim one wishes to argue, the argument is unlikely to convince. The advice to pursue "consensus-forming activity in science" instead of using values to weigh inductive risks likewise falls prey to all the same worries as "value-mitigating strategies."

It is best to think about what they are doing as answering the value-mitigation question. Although they present their work as a defense of the VFI, the changed terms of debate, and the nature of their argument, really puts them in the camp of the accounts discussed below.

5 Partial Rapprochements with the VFI

There is a set of arguments that, while not exactly defending the VFI, seek to accommodate what they take to be right in both defenses and critiques of the VFI by specifying value-laden and value-free moments of scientific inquiry, while acknowledging that the final result is thus value-laden in a sense. Defenders of such views imply that outright rejection of the VFI goes too far, while acknowledging that critics of the VFI have a point. However, this rhetoric presupposes a kind of error, the conflation of the VFI question with the value-management question. One can take a quite conservative view on the value-management question [e.g., Steel, 2010], but this does not uphold the VFI.[25]

Carrier [2022] attempts a qualified defense of the VFI, arguing that while nonepistemic values are essential for certain parts of the scientific process, scientists can nonetheless withhold commitment from those values "by elaborating a plurality of policy packages" that hypothetically involve different values. This appears at first to be a variety of the deferred decision response, very close to Mitchell [2004] and Edenhofer and Kowarsch [2015], one that also fails to account for much of the subsequent debate on that response (B). Carrier limits his defense of VFI to hypothesis assessment, while he acknowledges role of values in determining questions, concepts, relevance of

[24] My gratitude to Joyce Havstad for pointing out the connection to Solomon's argument.

[25] The error works both ways; de Melo-Martín and Intemann [2016] make this mistake when they argue that the AIR does not go far enough and so Douglas's positive account vindicates the VFI. Their disagreement is actually with Douglas's response to the value-management question, which presupposes a negative answer to the VFI question.

evidence, which already concedes the case against the VFI, as these elements cannot reasonably be classified as "external" to scientific reasoning proper. In other words, in response to the upstream regress problem, Carrier freely admits that these decisions are value-laden and must be made by the scientists.

What Carrier perhaps fails to acknowledge here is the point from Okruhlik [1994], which is that any value-free decision-procedure, operating on inputs that are value-laden, will reproduce rather than eliminate that value-ladenness. If value judgments determine in part which hypotheses are proposed, how concepts are operationalized, which evidence is considered relevant, etc., then the final results will be different on the basis of those value judgments, even if the final-stage decision were value-neutral or deferred to others. In Okruhlik's example, if patriarchal values inform all of the hypotheses under consideration, and if sexist biases determine which evidence is collected and how it is framed, then no matter how value-neutral the inference procedure, it will not erase the influence of those values. Nor, if we turn the final decision over to policymakers, even non-sexist, anti-patriarchal ones, will they be able to eliminate that bias. As Okruhlik puts it, "*even if we grant that the standards of theory assessment are free of contamination by non-cognitive factors*, nonetheless, non-cognitive values may permeate the very content of science... Even *granting* the transcendence of method, in other words, the scientific product could itself be radically culture-bound [i.e., value-laden]" (38-39, emphasis in original). To put it more succinctly: values in, values out.

Carrier does acknowledge that Douglas's argument is stronger than Rudner's, and like Mitchell, challenges Premise (6). He claims that the relevant information can be presented as a menu of policy options with background values specified, in line with Edenhofer and Kowarsch [2015]. Carrier provides no response to the concerns raised about this strategy by e.g. Havstad and Brown [2017]. However, this is not to say that, under certain conditions, presenting partially hedged information and deferring certain limited decisions to policymakers is not a fruitful approach to science advising; this could be a good answer to the value-management question. It just does nothing to vindicate the VFI, even in part.

Certain attempted rapprochements focus on sequestering values influences into a certain phase of scientific inquiry that seems less problematic. Recent work by Kareem Khalifa and collaborators, for instance, focuses on the role of questions in inquiry, and explicitly aligns itself with the value-free ideal in the sense of, "epistemic considerations being the only rational determinants for accepting or rejecting hypotheses" [DiMarco and Khalifa, 2019, p. 1022;

cf. Khalifa et al., draft]. A related approach focuses on the concept of adequacy-for-purpose in model assessment, which supposedly "opens the possibility of effecting a partial rapprochement between critics and proponents of the value-free ideal" [Lusk and Elliott, 2022]. Applying the adequacy-for-purpose approach from model assessment more widely, to all forms of scientific assessment, Lusk and Elliott argue that we may be able to recast what looks like value-laden scientific reasoning into value-free assessment of hypotheses about whether something is adequate for some purpose.[26] Yet another approach of this type is Wendy Parker's [2024] recent "epistemic projection approach."

There are a number of problems with the proposed rapprochements with VFI. First, it is not always possible to specify in advance the range of value considerations, so that they can be loaded into the question or purpose that inquiry begins with. The research question that inquiry is trying to solve may be only vague and inchoate at the start of inquiry and not properly settled until inquiry concludes [Brown, 2020, Ch. 1]. The relevant range of options and relevant values may likewise only be discovered in the course of inquiry, not settled ahead of time. Now, the defenders of this view might acknowledge this point, but insist that they can retain their view by packing the values into the posing of new questions throughout the course of inquiry or into the revision of the purposes for the sake of which adequacy judgments are made. This raises several concerns. One is Okruhlik's "values in, values out" problem. Perhaps most pressingly, it raises the worry about whether these accounts are merely a notational variant of the picture presented by the opponents of the VFI, describing a different way that value judgments figure in the internal processes of scientific inquiry, rather than denying that they do so. It is not clear to me that these views differ significantly from Douglas's [2009] use of the distinction between direct and indirect roles for values to answer the values management question. This is not to say that such alternative analyses of have no value, only that they do not in fact partially vindicate the VFI; rather, they address the value-management question in a particular way.

What's more, not all legitimate influences of values concern the purposes that guide the asking of questions or the proposal of hypotheses or models. Some values act as *side constraints*, limiting how we pursue inquiry quite

[26] By contrast, Harvard and Winsberg [2022] correctly note that the adequacy-for-purpose view means that model assessment is ineliminably value-laden.

independent of our aims and purposes [Brown, 2017, pp. 72-73]. As these sort of value judgments operate no matter what purposes the scientists pursue or what questions they seek to ask, they seem to escape the analysis in question. One can sweep all of this under the rug (that is, into the questions or purposes) only by doing violence to actual scientific practice via a problematic form of rational reconstruction (violating C).

It is unsurprising, given that these approaches do not try to fully defend the VFI, that they do not identify specific flaws with the AIR (A). That is, even if one does adopt the kind of semi-impartial view they aim at, it remains the case that the AIR shows non-epistemic values have a legitimate role to play in scientific reasoning, and is taken to do so in a limited way by the defenders of these views.[27] That is because they are really addressing the value-management question, not the VFI question. Furthermore, assessment of answers to questions or of the adequacy-for-purpose of models (hypotheses, etc) themselves are liable to the same kinds of concerns about the regress of upstream and downstream inductive risks as are the assessments of probabilities assigned to hypotheses given evidence as advocated by the hedging response. It is not clear on this basis that the rather conservative response to the value-management question is workable. In any case, what is clear is that the aspiration for rapprochement with the VFI, to to retain the rhetoric of epistemic purity, is misplaced.

Some of the most interesting work problematizing these issues is by John [John, 2015a, 2019]. John explicitly focuses on contexts of communication in science, especially public communication, solidly addressing the practical context in which these questions matter (C). But John's argument is not without its problems. John [2015a] "helpfully" attempts to improve AIR by restricting the consideration of IR to one's *intended* audience, with the aim of matching their values. He motivates this on the intuition that scientists cannot be held responsible for how bad actors interpret their work. This is a mistake; we should reject this amendment because the response of bad actors often forms part of the foreseeable consequences of assertion, i.e., foreseeable perlocutionary effects for which the speaker may be responsible [Franco, 2017]. Likewise, we should reject the claim that by virtue of the open-endedness of the audience for scientific publications, we cannot anticipate the reaction of the audience. AIR only requires that we consider the foreseeable consequences

[27] Khalifa in particular has said that he does not intend this work to side with the VFI against the AIR (personal communication).

of our assertions; and some reactions are clearly foreseeable.

John anticipates this objection, but his response is unconvincing. It amounts to two points: balancing all the foreseeable consequences of our actions is difficult, which is often true but irrelevant to the normative issues at hand. Second, the costs of error are complex, because it may have many downstream effects that are difficult to weigh. True, but again, irrelevant. Normatively, the consideration of values is needed, and the difficulty of the task concerns *how* and not *whether* inductive risks are managed via value judgments. He also argues that, while AIR applies, scientists should resolve IR by adopting universally high epistemic standards. But the reasons he adduces in favor are defeasible, and one can imagine them being outweighed in specific cases, thus making the role of value judgments unavoidable in principle, even if the best way to manage that role is to use those universally high epistemic standards as often as possible. It also raises the question: how high? It is doubtful that there is any one-size-fits-all answer to this question, especially in light of concerns about second-order uncertainty [Steel, 2016].[28]

These attempts at partial rapprochement with the VFI are generally valuable contributions to the literature, but framed in a pretty misleading and unproductive way. It is not that we need to find a middle ground in the debate for and against the VFI, because the VFI is rightly understood as an "ideal of epistemic purity" [Biddle, 2013], and so any middle ground is a rejection of the VFI. Rather, these raise important considerations when addressing the value-management question, and in many cases suggest serious frameworks for answering that question (if not entirely unproblematic ones). The urge towards rapprochement should be rejected in favor of more careful analysis of the issue of managing the roles of values in science.

6 Conclusions

One might object: isn't this all a bit too fussy? There are some arguments for the VFI and against it, and maybe some middle positions, and a lively debate that we should expect to continue on, as all deep philosophical debates do. Have I not put too much emphasis on one particular reconstruction of AIR?

[28]John [2019] revises his approach; in this essay, he makes clear that his view amounts to a rejection of the VFI in favor of the "value-apt ideal." This is similar in certain respects to the democratic ideal favored by Schroeder [Schroeder, 2021, Boulicault and Schroeder, 2021]. So it seems John now sees the error of affiliating with the VFI.

Is that not unfair, especially given the recency of that reconstruction? This is to mistake my point. We have, now, a strong contender for an argument against VFI that is not only valid, but sound. It is the state of the art in the field, drawing on the work by Douglas that gave the issue the renewed energy and has made it one of the most vibrant areas of research in philosophy of science. To defeat this argument, a definite flaw in terms of either invalid form or false premise should be identified. Sound deductive arguments are the gold standard for good arguments [Cartwright, 2013], and this one is decisive point against the VFI.

Perhaps you doubt, with Peirce, the wisdom in relying on a single deductive argument for a substantive conclusion:

> Philosophy ought to imitate the successful sciences in its methods, so far as to proceed only from tangible premises which can be subjected to careful scrutiny, and to trust rather to the *multitude* and *variety* of its arguments than to the conclusiveness of any *one*. Its reasoning should not form a *chain* which is no stronger than its weakest link, but a *cable* whose fibres may be ever so slender, provided they are sufficiently *numerous* and intimately connected. [Peirce, 1868, p. 141, emphasis added]

But the AIR is only the tip of the iceberg in terms of a variety of arguments pointing in the same direction. It is the strongest thread in a cable of arguments pulling against the VFI, which also includes the conceptual argument [Dupré, 2007, Alexandrova, 2018], the contingency and epistemic risk arguments [Brown, 2020, Biddle and Kukla, 2017], underdetermination arguments [Longino, 2004, 2008], and various others.

I have articulated four criteria (A-D) for defenders of the VFI to meet, if they wish to adequately and fully defend the VFI in a way that meets the concerns of the current literature. To this, based on the novel arguments of this paper, we might also add a fifth criterion: (E) Do not confuse the VFI question and the value-management question. Also, considering the other arguments in addition to the AIR, one might add one further criterion: (F) One must show that other epistemic risks/contingencies, including choices concerning thick concepts or mixed claims, can also be dealt with appropriately without reference to values. I have not defended or applied this last criterion here, but some opponents of the VFI would surely insist on it.

As philosophers of science, we may doubt the power even of a variety of philosophical arguments to win the day. One might consider the literature on

the VFI to be a worthy cause despite the problems canvassed above, because of the valuable diversity of viewpoint it brings to the discussion, and because continuing in this vein may produce more value in the future. A review like this one may have an unwelcoming chilling effect on a research program that should be allowed to flourish.

Perhaps we could zoom out and look at the question in terms of the VFI and value-laden science as Lakatosian research programs [Lakatos, 1970, 1974, Musgrave and Pigden, 2023, Plutynski, draft]. After all, it is not such a stretch, as Lakatos himself discussed not only scientific but also *philosophical* research programs [Lakatos, 1968, Harman, 2019, Musgrave and Pigden, 2023]. However, in my view this perspective would not rescue the VFI, for a few reasons. One is the severe failure of uptake of criticism[29] I have identified above. Another is the relative paucity of novel case studies compared to the many post-hoc reinterpretations of case studies used against the VFI. Finally, the retreat to ideal theory seen in Hudson and Sheykh-Rezaee and Bikaraan-Behesht, or the move to pursuitworthiness by Menon and Stegenga, could be considered degenerating problem-shifts. All this suggests that defense of the VFI should be classed as a *degenerating* rather than a *progressive* research program, in the Lakatosian sense. But a serious argument to this effect beyond this hasty sketch would be the difficult work of another sort of paper. I only mention it here to throw doubt on a kind of Lakatosian defense of the VFI.

So far, the legacy of the value-free ideal is a poor one, in that its defenders regularly fail to respond adequately to criticism. This is a serious concern; as Helen Longino wrote concerning the importance of uptake of criticism:

> ...the advocates of a point of view, and through them the point of view itself, may lose or even forfeit intellectual authority if their discursive interactions do not satisfy the... condition of uptake. That is, reiterating the same old complaint no matter

[29]This failure of uptake of criticism is not limited to the recent literature, but has been a feature of it for a very long time. Lloyd [1997] and Anderson [2004b] catalogue misrepresentations, bad arguments, and failed uptake from earlier defenders of the VFI and critics of feminist science studies. And, as Dan Hicks points out, "Contemporary defenders of VFI often focus almost exclusively on either Rudner or a narrow reading of Douglas (that treats her as just reiterating Rudner), as though nothing on science and values was written between 1953 and 2009, and often seem to work under an apparent complete ignorance of the contributions of feminists, Marxists, and Deweyan pragmatists throughout the twentieth century" (personal communication).

> what response is offered eventually disqualifies one as a member of a discursive community of equals. [Longino, 2002, p. 133]

Rather than this essay having a chilling effect, my hope is that by making explicit the current burdens of proof that the value-free ideal must meet and the failure of recent work to meet them, future work on this topic will be improved, and defenders of the value-free ideal can retain their authority in our discursive community. Alternatively, some work in this area might profit more from moving on to the value-management question than in revisiting the debate about the value-free ideal.

Conflict of Interest Statement

The author has no conflicts of interest to declare. This work has received no funding from public or private entities with a vested interest in the Value-Free Ideal.

Acknowledgements

My gratitude to the many folks who have given feedback or encouragement on drafts of this article, including especially Heather Douglas, Joyce Havstad, Dan Hicks, Jacob Stegenga, and Kareem Khalifa; to Jacob Stegenga, Tarun Menon, Kareem Khalifa, Roberta Millstein, and Anya Plutynski for sharing drafts of their work; and to Paul Franco, Kevin Elliott, Roberta Millstein, Gabriele Contessa, Boaz Miller, Greg Lusk, and Justin Biddle for other helpful input that contributed to the development of this paper.

References

Anna Alexandrova. *A Philosophy for the Science of Well-being*. Oxford University Press, 2017.

Anna Alexandrova. Can the science of well-being be objective? *The British Journal for the Philosophy of Science*, 69(2):421–445, 2018. doi: 10.1093/bjps/axw027. URL https://doi.org/10.1093/bjps/axw027.

Anna Alexandrova and Mark Fabian. Democratising measurement: or why thick concepts call for coproduction. *European Journal for Philosophy of Science*, 12(1):7, 2022. doi: 10.1007/s13194-021-00437-7. URL https://doi.org/10.1007/s13194-021-00437-7.

Elizabeth Anderson. Uses of value judgments in science: A general argument, with lessons from a case study of feminist research on divorce. *Hypatia*, 19 (1):1–24, 2004a.

Elizabeth Anderson. How not to criticize feminist epistemology: a review of *Scrutinizing Feminist Epistemology*, 2004b. URL https://web.archive.org/web/20230328013903/http://www-personal.umich.edu/~eandersn/hownotreview.html.

Gregor Betz. In defence of the value free ideal. *European Journal for Philosophy of Science*, 3(2):207–220, 2013.

Sindhuja Bhakthavatsalam and Nancy Cartwright. What's so special about empirical adequacy? *European Journal for Philosophy of Science*, 7(3): 445–465, Oct 2017. ISSN 1879-4920. doi: 10.1007/s13194-017-0171-7. URL https://doi.org/10.1007/s13194-017-0171-7.

Justin B. Biddle. State of the field: Transient underdetermination and values in science. *Studies in History and Philosophy of Science*, 44(1):124–133, 2013. URL http://dx.doi.org/10.1016/j.shpsa.2012.09.003.

Justin B. Biddle and Rebecca Kukla. The geography of epistemic risk. In Kevin C. Elliott and Ted Richards, editors, *Exploring Inductive Risk: Case Studies of Values in Science*, pages 215–238. Oxford University Press, 2017.

Marion Boulicault and S Andrew Schroeder. Public trust in science: Exploring the idiosyncrasy-free ideal. In *Social trust*, pages 102–121. Routledge, 2021.

Liam Kofi Bright. Du bois' democratic defence of the value free ideal. *Synthese*, 195(5):2227–2245, May 2018. ISSN 1573-0964. doi: 10.1007/s11229-017-1333-z. URL https://doi.org/10.1007/s11229-017-1333-z.

Liam Kofi Bright and Remco Heesen. To be scientific is to be communist. *Social Epistemology*, 37(3):249–258, 2023. doi: 10.1080/02691728.2022.2156308.

Matthew J. Brown. The source and status of values in socially responsible science. *Philosophical Studies*, 163(1):67–76, 2013a. doi: 10.1007/s11098-012-0070-x. URL http://dx.doi.org/10.1007/s11098-012-0070-x.

Matthew J. Brown. Values in science beyond underdetermination and inductive risk. *Philosophy of Science*, 80(5):829–839, 2013b.

Matthew J. Brown. John dewey's pragmatist alternative to the belief-acceptance dichotomy. *Studies in History and Philosophy of Science Part A*, 53:62 – 70, 2015. ISSN 0039-3681. doi: http://dx.doi.org/10.1016/j.shpsa.2015.05.012. URL http://www.sciencedirect.com/science/article/pii/S0039368115000722. Special Section: Cognitive Attitudes and Values in Science.

Matthew J. Brown. Against epistemic priority. In Kevin C Elliott and Daniel Steel, editors, *Current Controversies in Values and Science*. Routledge, 2017.

Matthew J. Brown. Is science really value free and objective? from objectivity to scientific integrity. In *What Is Scientific Knowledge? An Introduction to Contemporary Epistemology of Science.* Routledge, New York, 2018a.

Matthew J. Brown. Weaving value judgment into the tapestry of science. *Philosophy, Theory, and Practice in Biology*, 10(10), 2018b. doi: 10.3998/ptpbio.16039257.0010.010.

Matthew J. Brown. *Science and Moral Imagination: A New Ideal for Values in Science*. University of Pittsburgh Press, Pittsburgh, 2020.

Matthew J. Brown. The descriptive, the normative, and the entanglement of values in science. In Ted Richards, editor, *Science, Values, and Democracy: The 2016 Descartes Lectures by Heather Douglas*, pages 51–65. Consortium for Science, Policy & Outcomes, Arizona State University, 2021.

Matthew J. Brown and Jacob Stegenga. The validity of the argument from inductive risk. *Canadian Journal of Philosophy*, 53(2):187–190, 2023. doi: 10.1017/can.2023.37.

Martin Carrier. What does good science-based advice to politics look like? *Journal for General Philosophy of Science*, 53(1):5–21, 2022. doi: 10.1007/s10838-021-09574-2. URL https://doi.org/10.1007/s10838-021-09574-2.

Nancy Cartwright. Evidence, argument and prediction. In Vassilios Karakostas and Dennis Dieks, editors, *EPSA11 Perspectives and Foundational Problems in Philosophy of Science*, pages 3–17. Springer International Publishing, 2013. ISBN 978-3-319-01306-0.

Alejandro Cassini. Simulation models and probabilities: a bayesian defense of the value-free ideal. *SIMULATION*, 98(2):113–125, 2022. doi: 10.1177/00375497211028815. URL https://doi.org/10.1177/00375497211028815.

C West Churchman. *Theory of Experimental Inference*. Macmillan, New York, 1948.

Sharyn Clough. *Beyond epistemology: A pragmatist approach to feminist science studies*. Rowman & Littlefield Publishers, Lanham, Md., 2003. ISBN 0742514641 (alk. paper).

Inmaculada de Melo-Martín and Kristen Intemann. The risk of using inductive risk to challenge the value-free ideal. *Philosophy of Science*, 83(4):500–520, 2016. doi: 10.1086/687259. URL http://dx.doi.org/10.1086/687259.

Marina DiMarco and Kareem Khalifa. Inquiry tickets: Values, pursuit, and underdetermination. *Philosophy of Science*, 86(5):1016–1028, 2019. ISSN 0031-8248. doi: 10.1086/705446.

Heather Douglas. Inductive risk and values in science. *Philosophy of Science*, 67(4):559–579, 2000.

Heather Douglas. The moral responsibilities of scientists: Tensions between autonomy and responsibility. *American Philosophical Quarterly*, 40(1): 59–68, 2003.

Heather Douglas. Inserting the public into science. In Sabine Maasen and Peter Weingart, editors, *Democratization of Expertise? Exploring Novel Forms of Scientific Advice in Political Decision-Making*, volume 24 of *Sociology of the Sciences Yearbook*, pages 153–169. Springer, Dordrecht, 2005.

Heather Douglas. *Science, Policy, and the Value-Free Ideal*. University of Pittsburgh Press, Pittsburgh, 2009.

Heather Douglas. Weighing complex evidence in a democratic society. *Kennedy Institute of Ethics Journal*, 22(2):139–162, 2012.

Heather Douglas. The value of cognitive values. *Philosophy of Science*, 80(5): 796–806, 2013a.

Heather Douglas. Philip Kitcher, *Science in a Democratic Society*. *The British Journal for the Philosophy of Science*, 64(4):901–905, 05 2013b. ISSN 0007-0882. doi: 10.1093/bjps/axt006. URL https://doi.org/10.1093/bjps/axt006.

Heather Douglas. Values in science. In Paul Humphreys, editor, *The Oxford Handbook of Philosophy of Science*. Oxford University Press, 2016.

Heather Douglas. Why inductive risk requires values in science. In Kevin Elliot and Daniel Steel, editors, *Current Controversies in Values in Science*. Routledge, New York, 2017.

Heather Douglas. *Science, Values, and Democracy: The Rene Descartes Lectures*. Consortium for Science, Policy & Outcomes, Arizona State University, 2021.

Heather Douglas and Kevin C. Elliott. Addressing the reproducibility crisis: A response to hudson. *Journal for General Philosophy of Science*, 53(2): 201–209, 2022. ISSN 1572-8587. doi: 10.1007/s10838-022-09606-5. URL https://doi.org/10.1007/s10838-022-09606-5.

Markus Dressel. Inductive risk - riesgo inductivo: does it really refute value-freedom? *Theoria: An International Journal for Theory, History and Foundations of Science*, 37(2):181–208, 2022. ISSN 04954548, 2171679X. URL https://www.jstor.org/stable/27140154.

John Dupré. Fact and value. In Harold Kincaid, John Dupré, and Alison Wylie, editors, *Value-free science?: Ideals and illusions*, pages 27–41. Oxford University Press, Oxford, 2007. ISBN 9780195308969 (acid-free paper).

Ottmar Edenhofer and Martin Kowarsch. Cartography of pathways: A new model for environmental policy assessments. *Environmental Science & Policy*, 51:56–64, 2015.

Kevin C Elliott. *Is a Little Pollution Good for You?: Incorporating Societal Values in Environmental Research*. Environmental Ethics and Science Policy

Series. Oxford University Press, New York, 2011. ISBN 9780199755622 (alk. paper).

Kevin C Elliott. Douglas on values: From indirect roles to multiple goals. *Studies in History and Philosophy of Science Part A*, 44(3):375–383, 2013.

Kevin C Elliott and David Willmes. Cognitive attitudes and values in science. *Philosophy of Science*, 80(5):807–817, 2013.

Paul L. Franco. Assertion, nonepistemic values, and scientific practice. *Philosophy of Science*, 84(1):160–180, 2017. doi: 10.1086/688939.

David M. Frank. Making Uncertainties Explicit: The Jeffreyan Value-Free Ideal and Its Limits. In Kevin C Elliott and Ted Richards, editors, *Exploring Inductive Risk: Case Studies of Values in Science*. Oxford University Press, 08 2017. ISBN 9780190467715. doi: 10.1093/acprof:oso/9780190467715.003.0005. URL https://doi.org/10.1093/acprof:oso/9780190467715.003.0005.

Mathias Frisch. Uncertainties, values, and climate targets. *Philosophy of Science*, 87(5):979–990, 2020. doi: 10.1086/710538.

Loren R. Graham. *Lysenko's ghost: epigenetics and Russia*. Harvard University Press, 2016. ISBN 9780674089051 (alk. paper).

Sandra Harding. *The Science Question in Feminism*. Cornell University Press, 1986.

Sandra Harding. "strong objectivity": A response to the new objectivity question. *Synthese*, 104(3):331–349, 1995.

Graham Harman. On progressive and degenerating research programs with respect to philosophy. *Revista Portuguesa de Filosofia*, 75(4):2067–2102, 2019. ISSN 08705283, 2183461X. URL https://www.jstor.org/stable/26869262.

Stephanie Harvard and Eric Winsberg. The epistemic risk in representation. *Kennedy Institute of Ethics Journal*, 32(1):1–31, 2022. doi: 10.1353/ken.2022.0001.

Joyce C. Havstad. Sensational science, archaic hominin genetics, and amplified inductive risk. *Canadian Journal of Philosophy*, 52(3):295–320, 2022. doi: 10.1017/can.2021.15.

Joyce C. Havstad and Matthew J. Brown. Inductive risk, deferred decisions, and climate science advising. In Kevin C Elliott and Ted Richards, editors, *Exploring Inductive Risk*, pages 101–123. Oxford University Press, 2017.

Carl G. Hempel. Science and human values. In *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*, pages 81–96. The Free Press, New York, 1965.

Tobias Henschen. How strong is the argument from inductive risk? *European Journal for Philosophy of Science*, 11(3):92, 2021. ISSN 1879-4920. doi: 10.1007/s13194-021-00409-x. URL https://doi.org/10.1007/s13194-021-00409-x.

Daniel J. Hicks. A new direction for science and values. *Synthese*, 191(14): 3271–3295, 2014.

Daniel J. Hicks and Emilio Jon Christopher Lobato. Values disclosures and trust in science: A replication study. *Frontiers in Communication*, 7, 2022. ISSN 2297-900X. doi: 10.3389/fcomm.2022.1017362. URL https://www.frontiersin.org/articles/10.3389/fcomm.2022.1017362.

Daniel J. Hicks, P. D. Magnus, and Jessey Wright. Inductive risk, science, and values: A reply to macgillivray. *Risk Analysis*, 40(4):667–673, 2020. doi: https://doi.org/10.1111/risa.13434. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/risa.13434.

Bennett Holman and Justin P. Bruner. The problem of intransigently biased agents. *Philosophy of Science*, 82(5):956–968, 2015. ISSN 00318248, 1539767X. URL http://www.jstor.org/stable/10.1086/683344.

Bennett Holman and Torsten Wilholt. The new demarcation problem. *Studies in History and Philosophy of Science*, 91:211–220, 2022. ISSN 0039-3681. doi: https://doi.org/10.1016/j.shpsa.2021.11.011. URL https://www.sciencedirect.com/science/article/pii/S0039368121001886.

Paul Hoyningen-Huene. Objectivity, value-free science, and inductive risk. *European Journal for Philosophy of Science*, 13(14), 2023. doi: 10.1007/s13194-023-00518-9.

Robert Hudson. Why we should not reject the value-free ideal of science. *Perspectives on Science*, 24(2):167–191, 2016. ISSN 1063-6145. doi: 10.1162/posc_a_00199.

Robert Hudson. Should we strive to make science bias-free? a philosophical assessment of the reproducibility crisis. *Journal for General Philosophy of Science*, 52(3):389–405, 2021. ISSN 1572-8587. doi: 10.1007/s10838-020-09548-w. URL https://doi.org/10.1007/s10838-020-09548-w.

Robert Hudson. Rebuttal to douglas and elliott. *Journal for General Philosophy of Science*, 53(2):211–216, 2022. ISSN 1572-8587. doi: 10.1007/s10838-022-09616-3. URL https://doi.org/10.1007/s10838-022-09616-3.

William James. The will to believe. *The New World*, 5:327–347, 1896.

Richard C Jeffrey. Valuation and acceptance of scientific hypotheses. *Philosophy of Science*, 23(3):237–246, 1956.

Stephen John. Inductive risk and the contexts of communication. *Synthese*, 192(1):79–96, 2015a. ISSN 1573-0964. doi: 10.1007/s11229-014-0554-7. URL https://doi.org/10.1007/s11229-014-0554-7.

Stephen John. The example of the ipcc does not vindicate the value free ideal: a reply to gregor betz. *European Journal for Philosophy of Science*, 5(1):1–13, 2015b. doi: 10.1007/s13194-014-0095-4. URL https://doi.org/10.1007/s13194-014-0095-4.

Stephen John. Science, truth and dictatorship: Wishful thinking or wishful speaking? *Studies in History and Philosophy of Science Part A*, 78:64–72, 2019. ISSN 0039-3681. doi: https://doi.org/10.1016/j.shpsa.2018.12.003. URL https://www.sciencedirect.com/science/article/pii/S003936811830181X.

Klemens Kappel. The proper role of science in liberal democracy. *Unpublished draft*, 2014-05. URL https://www.academia.edu/7017103/.

Klemens Kappel and Julie Zahle. The epistemic role of science and expertise in liberal democracy. In *The Routledge Handbook of Social Epistemology*, pages 397–405. Routledge, 2019.

Kareem Khalifa, Jared Millson, and Mark Risjord. Inquiry and epistemic priority. draft.

Philip Kitcher. The division of cognitive labor. *The journal of philosophy*, 87 (1):5–22, 1990.

Philip Kitcher. *Science in a Democratic Society*. Prometheus Books, Amherst, N.Y., 2011. ISBN 9781616144074 (cloth : alk. paper).

Janet A Kourany. *Philosophy of Science after Feminism*. Oxford University Press, 2010. ISBN 0199732620.

Hugh Lacey. *Is science value free?: Values and scientific understanding*. Routledge, London, 1999. ISBN 0415208203 (hb).

Hugh Lacey. 'holding' and 'endorsing' claims in the course of scientific activities. *Studies in History and Philosophy of Science Part A*, 53:89–95, 2015. ISSN 0039-3681. doi: https://doi.org/10.1016/j.shpsa.2015.05.009. URL https://www.sciencedirect.com/science/article/pii/S0039368115000692. Special Section: Formal Epistemology and the Legacy of Logical Empiricism Special Section: Cognitive Attitudes and Values in Science.

Imre Lakatos. Changes in the problem of inductive logic. In *Studies in Logic and the Foundations of Mathematics*, volume 51, pages 315–417. Elsevier, 1968.

Imre Lakatos. Falsification and the methodology of scientific research programmes. *Criticism and the Growth of Knowledge*, pages 91–195, 1970.

Imre Lakatos. Science and pseudoscience. In G. Vesey, editor, *Philosophy in the Open*. Open University Press, 1974.

Isaac Levi. Must the scientist make value judgments? *The Journal of Philosophy*, 57(11):345–357, 1960. ISSN 0022362X.

Richard Levins and Richard C. Lewontin. *The dialectical biologist*. Harvard University Press, Cambridge, Mass., 1985. ISBN 0674202813 (alk. paper).

Elisabeth Lloyd. *Science and Anti-Science: Objectivity and Its Enemies*. Springer, 1997.

Helen E Longino. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton University Press, Princeton, N.J., 1990. ISBN 0691073422 (alk. paper).

Helen E Longino. Subjects, power, knowledge: Prescriptivism and descriptivism in feminist philosophy of science. In Linda Alcoff and Elizabeth Potter, editors, *Feminist epistemologies*, chapter 5, pages 101–120. Routledge, New York, 1992.

Helen E Longino. Cognitive and non-cognitive values in science: Rethinking the dichotomy. In Lynn Hankinson Nelson and Jack Nelson, editors, *Feminism, science, and the philosophy of science*, pages 39–58. Kluwer Academic Publishers, Dordrecht, 1996.

Helen E Longino. *The Fate of Knowledge*. Princeton University Press, 2002.

Helen E Longino. How values can be good for science. In Peter Machamer and Gereon Wolters, editors, *Science, Values, and Objectivity*, pages 127–142. University of Pittsburgh Press, 2004.

Helen E Longino. Values, heuristics, and the politics of knowledge. In Martin Carrier, Don Howard, and Janet A Kourany, editors, *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited*, pages 68–86. University of Pittsburgh Press, 2008.

Greg Lusk. Does democracy require value-neutral science? analyzing the legitimacy of scientific information in the political sphere. *Studies in History and Philosophy of Science Part A*, 90:102–110, 2021. ISSN 0039-3681. doi: https://doi.org/10.1016/j.shpsa.2021.08.009. URL https://www.sciencedirect.com/science/article/pii/S0039368121001254.

Greg Lusk and Kevin C Elliott. Non-epistemic values and scientific assessment: an adequacy-for-purpose view. *European Journal for Philosophy of Science*, 12(2):35, 2022. ISSN 1879-4920. doi: 10.1007/s13194-022-00458-w. URL https://doi.org/10.1007/s13194-022-00458-w.

Brian H. MacGillivray. Null hypothesis testing $\neq$ scientific inference: A critique of the shaky premise at the heart of the science and values debate, and a defense of value-neutral risk assessment. *Risk Analysis*, 39(7):1520–1532, 2019. doi: https://doi.org/10.1111/risa.13284. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/risa.13284.

P.D. Magnus. What scientists know is not a function of what scientists know. *Philosophy of Science*, 80(5):840–849, 2013.

Tarun Menon and Jacob Stegenga. Sisyphean science: Why value freedom is worth pursuing. *European Journal for Philosophy of Science*, 13(4):48, 2023. doi: 10.1007/s13194-023-00552-7.

973)]Merton:1942 Robert K. Merton. The normative structure of science. In *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press, Chicago, [1942] 1973.

Boaz Miller. Science, values, and pragmatic encroachment on knowledge. *European Journal for Philosophy of Science*, 4(2):253–270, 2014.

Sandra Mitchell. The prescribed and proscribed values in science policy. In Peter Machamer and Gereon Wolters, editors, *Science, Values, and Objectivity*, pages 245–255. University of Pittsburgh Press, Pittsburgh, 2004.

Alan Musgrave and Charles Pigden. Imre Lakatos. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition, 2023.

Kathleen Okruhlik. Gender and the biological sciences. *Canadian Journal of Philosophy*, 24(S1):21–42, 1994.

Wendy S. Parker. The epistemic projection approach to values in science. *Philosophy of Science*, 91(1):18–36, 2024. doi: 10.1017/psa.2023.107.

Charles S. Peirce. Some consequences of four incapacities. *The Journal of Speculative Philosophy*, 2(3):140–157, 1868. ISSN 0891625X, 15279383. URL http://www.jstor.org/stable/25665649.

Uwe Peters. Illegitimate values, confirmation bias, and mandevillian cognition in science. *The British Journal for the Philosophy of Science*, 72(4):1061–1081, 2021. doi: 10.1093/bjps/axy079. URL https://doi.org/10.1093/bjps/axy079.

Manuela Fernández Pinto and Daniel J. Hicks. Legitimizing values in regulatory science. *Environmental Health Perspectives*, 127(3):035001, 2019. doi: 10.1289/EHP3317. URL https://ehp.niehs.nih.gov/doi/abs/10.1289/EHP3317.

Anya Plutynski. Progress and contingency in cancer science. In *Lakatos @ 100*. draft.

Hilary Putnam. *The Collapse of the Fact/Value Dichotomy and other essays*. Harvard University Press, Cambridge, MA, 2002.

David B. Resnik. Dual-use research and inductive risk. In Kevin C Elliott and Ted Richards, editors, *Exploring Inductive Risk: Case Studies of Values in Science*, pages 59–77. Oxford University Press, 2017.

Kristina Rolin. Values in science: The case of scientific collaboration. *Philosophy of Science*, 82(2):157–177, 2015. doi: 10.1086/680522.

Phyllis Rooney. On values in science: Is the epistemic/non-epistemic distinction useful? In *PSA: Proceedings of the biennial meeting of the philosophy of science association*, number 1, pages 13–22. Philosophy of Science Association, 1992.

Phyllis Rooney. The borderlands between epistemic and non-epistemic values. In Kevin Elliot and Daniel Steel, editors, *Current Controversies in Values in Science*. Routledge, New York, 2017.

Richard Rudner. The scientist qua scientist makes value judgments. *Philosophy of Science*, 20(1):1–6, 1953.

S. Andrew Schroeder. Democratic values: A better foundation for public trust in science. *The British Journal for the Philosophy of Science*, 72(2): 545–562, 2021. doi: 10.1093/bjps/axz023. URL https://doi.org/10.1093/bjps/axz023.

Hossein Sheykh-Rezaee and Hamed Bikaraan-Behesht. Value-free ideal is an epistemic ideal: An objection to the argument from inductive risk. *Principia: An International Journal of Epistemology*, 27(1):137–163, 2023. doi: 10.5007/1808-1711.2023.e80487.

Matthew Silk. *Evaluation and Value Management in Science*. PhD thesis, University of Waterloo, 2018.

Miriam Solomon. Social empiricism, 2001.

Daniel Steel. Epistemic values and the argument from inductive risk. *Philosophy of Science*, 77(1):14–34, 2010.

Daniel Steel. Acceptance, values, and probability. *Studies in History and Philosophy of Science Part A*, 53:81–88, 2015. ISSN 0039-3681. doi: https://doi.org/10.1016/j.shpsa.2015.05.010. URL https://www.sciencedirect.com/science/article/pii/S0039368115000709. Special Section: Formal Epistemology and the Legacy of Logical Empiricism Special Section: Cognitive Attitudes and Values in Science.

Daniel Steel. Climate change and second-order uncertainty: Defending a generalized, normative, and structural argument from inductive risk. *Perspectives on Science*, 24(6):696–721, 2016.

Katie Steele. The scientist qua policy advisor makes value judgments. *Philosophy of Science*, 79(5):893–904, 2012. ISSN 00318248, 1539767X. URL http://www.jstor.org/stable/10.1086/667842.

Jacob Stegenga and Tarun Menon. The difference-to-inference model for values in science. *Res Philosophica*, 100(4):423–447, 2023. doi: 10.5840/resphilosophica2023928102.

Zina B. Ward. On value-laden science. *Studies in History and Philosophy of Science Part A*, 85:54–62, 2021. ISSN 0039-3681. doi: https://doi.org/10.1016/j.shpsa.2020.09.006. URL https://www.sciencedirect.com/science/article/pii/S0039368120301783.

Eric Winsberg. Values and uncertainties in the predictions of global climate models. *Kennedy Institute of Ethics Journal*, 22(2):111–137, 2012.

Eric Winsberg. Communicating uncertainty to policymakers: the ineliminable role of values. In Elisabeth A Lloyd and Eric Winsberg, editors, *Climate modelling: philosophical and conceptual issues*, pages 381–412. Palgrave Macmillan, Cham, 2018a.

Eric Winsberg. *Philosophy and climate science*. Cambridge University Press, Cambridge, 2018b.