European Journal for Philosophy of Science (2023) 13:48
https://doi.org/10.1007/s13194-023-00552-7

PAPER IN GENERAL PHILOSOPHY OF SCIENCE

Sisyphean science: why value freedom is worth pursuing

Tarun Menon[1] · **Jacob Stegenga**[2]

Received: 30 January 2023 / Accepted: 15 September 2023

Abstract
The value-free ideal in science has been criticised as both unattainable and undesirable. We argue that it can be defended as a practical principle guiding scientific research even if the unattainability and undesirability of a value-free end-state are granted. If a goal is unattainable, then one can separate the desirability of accomplishing the goal from the desirability of pursuing it. We articulate a novel value-free ideal, which holds that scientists should act as if science should be value-free, and we argue that even if a purely value-free science is undesirable, this value-free ideal is desirable to pursue.

Keywords Value freedom · Values in science · Value free ideal

1 Introduction

Ancient Indian philosophical texts from both the Vedic and Buddhist traditions describe the ultimate good for a person as the complete dissolution of all egotistical attachment, attaining a mental state free of all desire, hope, anxiety, pleasure and pain. If this state is meant to represent an ideal, then presumably it is a desirable state to attain. To many, however, an austere state of complete detachment does not seem very attractive. Our usual understanding of a desirable state is that it provides us with some subjectively accessible benefit, but if a subject has eliminated their desires, it is unclear if there can be such a subjective benefit. In what sense can a state be good for us if we are impervious to influence either benign or malign? This is not

✉ Tarun Menon
tarun.menon@apu.edu.in

[1] School of Arts and Sciences, Azim Premji University, Sarjapura, Bengaluru, Karnataka 562125, India

[2] Department of History and Philosophy of Science, University of Cambridge, Cambridge, UK

Published online: 06 October 2023

a straightforward incoherency—it is possible to articulate a notion of desirability entirely divorced from subjective desire—but it is a philosophical challenge.

The Indian philosophical tradition has developed a variety of responses to this puzzle. Jonardon Ganeri and Peter Adamson suggest that perhaps the desirability of this state of complete detachment comes not from the benefit one gains upon attaining the state, but from the benefit from actively pursuing it. They say,

> "Perhaps the very fact that the ideal states are described in such unappealing terms shows us that these are not really intended as descriptions of the good for human beings. We should instead ask: how does the idea of striving to achieve such a state help one make progress?… If I genuinely believe that the ideal state involves no pleasure at all, I am apt to allow myself to be nourished by the pleasures I do have without being distracted from my other goals by the need to seek out new pleasures. In other words, I may come to lead a life of restraint and self-control." (Adamson & Ganeri, 2020, 8).

What is interesting about this response is that it acknowledges that attaining the ideal state may be undesirable while still maintaining that actively pursuing that state is desirable. In this paper, we deploy this decoupling of pursuit desirability from end-state desirability in a very different philosophical context—the debate about the ideal of a value-free science—in order to develop a novel response to challenges to the value-free ideal.

The value-free ideal for science holds that scientific reasoning should not be influenced by non-epistemic (social, cultural, political, or ethical) values. It is on its face extremely plausible. One might even think that it hardly needs philosophical defence. Science is our best source of objective knowledge about the world, and at least on many accounts of scientific objectivity, the involvement of non-epistemic values threatens such objectivity. Despite its *prima facie* appeal, however, the value-free ideal faces a number of philosophical challenges, which have led to an emerging consensus that value-freedom is not attainable. Moreover, its status as an ideal is contested, with some arguing that a value-free science is not desirable. The ambition of this paper is to articulate a novel version of the value-free ideal which avoids the existing philosophical challenges.

Our core argument is that the feasibility and desirability of attaining an end can be decoupled from the feasibility and desirability of pursuing that end, as illustrated in the example from the Indian philosophical tradition. A particular end-state may be unfeasible or undesirable, yet pursuing that end may nevertheless be feasible or desirable. The ideal of world peace may be impossible to achieve, yet pursuit of that ideal is both possible and good; de-escalating conflicts, disarmament, and a more equitable distribution of resources are all potential means to approach world peace, and these means are, at least to some degree, feasible to enact. Things are a little trickier when we consider desirability. It may seem absurd to suggest that it is desirable to pursue an end that one judges undesirable, but as we have already seen, some philosophers

take this possibility seriously.[1] Challenges to the value-free ideal have focused on the feasibility and desirability of a value-free end-state. Yet, one can grant that the value-free ideal is neither end-state feasible nor end-state desirable, while maintaining that pursuit of the end-state is both feasible and desirable. That is our goal.

We start by articulating the primary arguments in favour of the value-free ideal, the main challenges to the ideal, and various versions of the ideal (§ 2). We then argue that pursuit of the ideal is feasible (§ 3). In § 4 we argue that pursuit of the value-free ideal is desirable even if the end-state is undesirable. The conclusion of our argument is a specific—and as far as we know, novel—version of the value-free ideal, which holds that scientists ought to act as if science should be value-free.

2 The value-free ideal

The value-free ideal is a somewhat nebulous notion, and has been interpreted in a number of ways (Proctor, 1991; Lacey, 1999). It is uncontroversial that values influence the choice of scientific research projects, constraints on research methods, and applications of scientific results. The issue under dispute in recent philosophical literature is whether values influence the internal structure of scientific reasoning, or the inference from evidence to conclusion. The value-free ideal, as we construe it, only calls for an elimination of values that play a role in this inferential stage.

Another standard caveat: Since science involves ampliative inference, if we are to distinguish between better and worse inferences, we must appeal to at least some normative considerations that go beyond mere deductive validity. So if we use the term "values" to encompass *all* normative considerations, then aiming for an entirely value-free science would be foolish, undermining the inferential basis of science. The values that further the goals of scientific understanding, prediction and knowledge-gathering are standardly dubbed *epistemic values*. These values usually specify some feature of scientific theories—archetypical examples include simplicity, empirical adequacy and fruitfulness—to be used as a basis for judging how a theory fares relative to its rivals on epistemic grounds. The value-free ideal does not target these sorts of normative considerations; it targets those values that do not serve an epistemic function.[2] Among the non-epistemic values, attention usually focuses on social, moral and political values.[3]

Taking these qualifications into account, we arrive at this preliminary characterization: The value-free ideal holds that non-epistemic values ought not influence the

[1] It is, of course, obviously true that some of the means to a particular end can be desirable even if the end itself is not. A state of complete starvation is undesirable, but cutting out junk food from one's diet may well be desirable, even though it is, *inter alia*, one of the means to the end of starvation. What we are talking about here is something stronger: not merely the desirability of taking certain actions that happen to be means to an undesirable end, but the desirability of actively pursuing the undesirable end itself.

[2] The sharpness of the distinction between epistemic and non-epistemic values has been challenged (Rooney, 1992). We believe this is a useful and important distinction, although the boundary may be vague. Constraints of space do not permit a defence of our claim here.

[3] From here on, when we use the word "values" without qualification, we are referring to non-epistemic values.

internal features of scientific reasoning. "Influence", however, is a broad notion. We use the term to refer to the role played by values in putatively justifying scientific claims.[4] Specifically, the version of the ideal we defend opposes values playing a crucial role in the inference from evidence to conclusions, in the following sense: if an appeal to value V_1 appears in the inferential chain connecting evidence E to conclusion C_1, then replacing V_1 with an alternate value V_2 should not lead to a conclusion C_2 that is incompatible with C_1. In other words, given a fixed evidential basis, whether a researcher regards a conclusion as true or false should not depend on the values they endorse. A scientific inference that is not value-free will have *bifurcation points*, places in the inferential chain where choosing between two value sets will lead to incompatible conclusions.[5] The value-free ideal, as we conceive it, holds that science should aim at the elimination of all bifurcation points, and progress towards the ideal can be made by reducing the number of bifurcation points.

If any conceivable value set defines a bifurcation point, then there is plausibly an infinity of bifurcation points involved in every scientific inference. Thus, if progress toward the value-free ideal is to be understood in terms of reduction in the number of bifurcation points, then there must be some restriction on the kinds of values that are relevant to defining bifurcation points. For our purposes, we restrict bifurcation points to those points of contention involving actually-existing value disagreement. Exotic value frameworks that nobody holds and that are mere conceptual possibilities do not interest us here, not least because eliminating the influence of such values does little (if anything) to further the benefits of value-freedom we discuss in the next section.

We emphasize that a bifurcation point must involve incompatible conclusions drawn from the same evidence, due to differing value sets. In other words, one set of values leads to a conclusion C, while the other set of values leads to a denial that C can be concluded.[6]It is not enough that the conclusions be merely different from one another, even if that difference is attributable to a difference in values. We exclude mere difference in conclusions because value-based differences in research focus may lead to different conclusions based on the same evidence, simply because the researchers care about different aspects of the evidence. But the value influence in such cases would appropriately be identified with external features of scientific reasoning – the research questions being asked – rather than internal features. So our conception of value-freedom allows non-epistemic values to play some role in determining the way the conclusion is framed—what terminology is used or what aspects of the evidence are highlighted, for instance—but does not permit them to impact the putative truth of the conclusion.

[4] In terms of the taxonomy developed in Ward (2021), we are considering values as justifying reasons for scientific choices, while remaining agnostic about whether they also function as causal effectors for those choices.

[5] There may well be three or more value-based considerations in conflict, all leading to mutually incompatible conclusions. We focus on pair-wise contestation (bifurcation) as the smallest unit of value influence in the inferential process.

[6] This does not mean that the second set of values must lead to the contradictory conclusion ~C. It could, for example, be the case that they lead to a denial that either C or ~C can be concluded, due to a lack of sufficient evidence.

While our conception of value-freedom may not capture every purported example of the influence of values on science, it does provide a precise account of the core intuition that the value-free ideal must focus on the internal features of scientific reasoning. And, as we argue further in § 2.2, it also captures many core cases of value influence discussed in the literature.

2.1 Arguments for the value free ideal

There are several considerations that offer prima facie support to the value-free ideal. First, value-ladenness threatens the reliability of science (we will see below that this is in fact controversial). A constitutive aim of science is the discovery of truths. Some epic failures in the history of science can be understood as resulting from violations of the value-free ideal. The disaster that befell genetics and agricultural biology in the Soviet Union was the result of Lysenko's rejection of Mendelian genetics and his support of Lamarckism, based on the belief that the latter was more consistent with Soviet principles. Primatology until the 1970s was dominated by men, and this resulted in an implicit androcentric value framework which biased the observation and interpretation of primate behaviour, leading to erroneous theories about female reproductive strategies and social hierarchies.[7] Such cases suggest that science's constitutive aim of truth can be hampered when values influence scientific reasoning. The influence of values on science may not always hinder the aim of truth; indeed Longino (2004) and many others argue that values can enhance the reliability of science —we address this argument below.

Second, value-free science can more justly inform public policy. Since science is often harnessed for policy, and democratic ideals entail that policy should reflect the values of the broad citizenry rather than the values of a handful of scientific experts, science (especially policy-oriented science) should be value-free (Betz, 2013; Bright, 2018), or, if value influence is inevitable, science should be influenced by representative democratic values rather than the potentially idiosyncratic values of particular scientists (Schroeder, 2021). If values influence the internal features of scientific reasoning in policy-relevant research, a narrow and non-representative set of values can opaquely influence policy.

A third and related argument for the value-free ideal is that public support and public trust in science is influenced by the extent to which the public views science as value-free (Bright, 2018; John, 2015).[8] The value-free ideal should be upheld as a real regulative constraint, so that the public supports and trusts science.

2.2 Challenges to the value free ideal

The putative merits of the value-free ideal have been contested, and both its feasibility and desirability have been challenged.

[7] The cases of Soviet genetics and mid-twentieth century primatology are staples in the literature on values in science; for a useful introduction, see Elliott (2017).

[8] This argument is, of course, empirical, and while it may more or less hold based on contingent facts about our contemporary society, it is not obvious that it is generally true.

One challenge is that certain concepts used in framing scientific hypotheses are irreducibly value-laden. Choosing to use a particular concept when describing, analysing or explaining data may indicate certain value commitments, and those who disagree with those commitments might therefore reject a conclusion framed in those terms. Alexandrova (2018), for example, discusses the indispensability of 'thick concepts' in some domains of science. Anderson (2004) talks about how values play a role in framing questions in research about divorce. Atkinson (1998) argues that when measuring poverty rates, value-influenced choices must be made when defining and operationalising poverty—for instance, the choice between relative and absolute standards of poverty may be influenced by whether one considers it more significant to be below a certain percentile of household income or to be unable to afford a particular bundle of goods. If making such value judgements is genuinely unavoidable in certain areas of research, then the value-free ideal is unattainable. We call this the *framing problem.*

Differences in framing will only count as bifurcation points (and, hence, on our view, genuine examples of value influence on the internal features of science) if they lead to contradictory conclusions about the same concept. Suppose researcher A, motivated by her set of values, chooses to measure absolute poverty, while researcher B chooses to measure relative poverty. There need not be a bifurcation point unless both A and B see themselves as talking about the same concept – poverty. It is only if they are drawing contradictory conclusions about this single concept due to their different conceptualizations that there will be a bifurcation point. If they see absolute and relative poverty as distinct concepts, then the claims they make about these measures will not be contradictory.

A second challenge is what Betz (2013) refers to as the *methodological critique*. One version of this challenge holds that evidence underdetermines support for hypotheses, and values must fill the gap between evidence and hypothesis (Longino, 1990, 2004; Elliott, 2011). Another version holds that decisions to accept or reject hypotheses are always uncertain, and since errors have practical consequences, our valuations of those consequences influence our decisions to accept or reject hypotheses (Rudner, 1953; Douglas, 2000).[9] This is known as the argument from inductive risk. Here is the argument in a little more detail: inferring the acceptance or rejection of a hypothesis from a given set of evidence depends on deciding whether the extent to which the evidence supports the hypothesis surpasses a threshold of sufficiency. There are two kinds of errors one could make – rejecting a true hypothesis because one's evidential threshold was too high, or accepting a false hypothesis because the threshold was too low. The evidence itself cannot tell us where to set the threshold; that must be determined by an evaluation of the relative consequences of these two types of error, and values play a role in this evaluation. Therefore, a completely value-free science is infeasible.[10]

[9] Miller (2014) makes a similar argument, framed in terms of the notion of pragmatic encroachment.

[10] The infeasibility of the value-free end state has been argued for in numerous contexts, including vaccine safety (Goldenberg, 2021), enzyme classification (Conix, 2020), the science of well-being (Alexandrova, 2018), epidemiological modelling during the Covid-19 pandemic (Winsberg et al., 2020), and especially climate science (Havstad & Brown, 2017; John, 2015; Frisch, 2020; Winsberg, 2012). Hoyningen-Huene (2023) argues that the inductive risk argument does not pose a serious threat to the objectivity of science.

Steel (2016) points out that this argument could be interpreted as making the descriptive point that values *do* as a matter of fact play an inevitable role in determining the threshold of evidential sufficiency, or as making the normative point that values *should* play this role, whether or not they actually do. The descriptive reading of the inductive risk argument challenges the feasibility of the value-free ideal, while the normative reading challenges the desirability of the ideal.

Douglas (2000) has argued that the argument from inductive risk applies not just to the conclusions drawn from evidence, but to the characterization of the evidence itself. It may appear that our definition of value influence cannot pertain to the characterization of evidence, because our definition is based on bifurcation points in the inference from evidence to conclusions, and this could appear insufficient to represent a value-based dispute about the evidence itself. However, what counts as "the evidence" may be contextually determined. What is a shared body of evidence in one context may be a disputed conclusion in another. In cases in which there is disagreement about the characterization of evidence, there could be a more basic or rawer description of the evidence about which the parties agree. For instance, Douglas (2000) provides an example in which different teams of pathologists disagreed about the extent of toxicity visible in rat liver slides. But it is plausible that there would be some other description of those slides, one that didn't mention toxicity, about which the teams would agree. Their disagreement, then, could be framed in terms of what *conclusions* about toxicity are appropriately inferred from these shared descriptions. So although this example could be described as one where there is disagreement about the characterization of evidence, it could also be described as one in which there are bifurcation points in the inference from evidence to conclusions about toxicity.

Another challenge to the desirability of the ideal, prominent in feminist philosophy of science, says that claims of value-freedom usually serve to disguise the unexamined dominance of a narrow range of hegemonic values. We call this the *value-enrichment critique*. We ought, goes this challenge, to enrich the range of values that influence science to avoid the biasing effect of this narrowness. Longino (2004), for example, argues that complex subjects of scientific study (such as human behaviour) are unlikely to be fully understood using a single approach, so in such domains the involvement of diverse value perspectives enhances our scientific understanding and mitigates the influence of extant biasing values.[11] Since scientific reasoning is anyway inevitably value-laden, those values should be sufficiently diverse and representative of a broad range of interests, and therefore rather than purifying scientific reasoning from values we ought to be enriching scientific reasoning with even more values. From this it follows that the value free ideal is not desirable. Longino (2004) states the value-enrichment critique explicitly, and Anderson (2004) offers a detailed case which is framed as highlighting the merits of value-enriched scientific reasoning. In § 5 we offer a reframing of the value-enrichment critique.

[11] For further discussion of this critique, see Intemann (2005), Kourany (2008) and Hicks (2018).

2.3 Reformulating the value-free ideal

Our aim in the remainder of this paper is not to deny the significance of the criticisms of the value-free ideal canvassed in the previous section. We believe they highlight important considerations that any clear-eyed view of scientific practice must take into account. Our defence of the value-free ideal does not rely on denying the conclusion they arrive at, namely, that an end-state of value freedom is neither feasible nor desirable. Our strategy will instead be to argue for a reformulation of the ideal that makes it clear why pursuing value-freedom is both feasible and desirable, even if the end-state is neither attainable nor desirable. The criticisms of the ideal should not lead us to abandon value-freedom completely, but instead to reconceptualize what it means to say that value-freedom is an ideal.

The full scope of our argument will be elaborated in subsequent sections, but we lay some preliminary groundwork here by distinguishing four conceptions of the value-free ideal. Three of these four conceptions are either straightforwardly refuted or rendered implausible by the arguments of the critics of the ideal, but the fourth conception—which is, as far as we can tell, a novel formulation of the value-free ideal—can be maintained even if one fully accepts the critics' conclusions.

First, however, let us dispense with the other three formulations. The first is:

(VFI1) Science is, as a matter of fact, value-free.

This version might seem so obviously wrong that it is not worth discussing. After all, philosophers of science have presented many case studies where values have played a role in scientific inference. It would be exceedingly naïve, perhaps even wilfully ignorant, to maintain that values play no role in science. However, as Lacey mentions (1999, 1), perhaps value-freedom could be understood as an idealization of science. No doubt the messy reality of research often involves value-based considerations playing a role in inference, but maybe our best science permits a rational reconstruction purged of all such considerations. Even if values are involved in the context of discovery, in the context of justification we may be able to arrive at the same conclusions without appeal to values. So while VFI1 may not be true strictly speaking, one might think that it could be true as an idealized reconstruction of science.

However, we do not think that even this interpretation of VFI1 is defensible. Worries about the ideal cannot be dispelled by moving to the context of justification, because these are worries about the inferential structure of science. The critics discussed in § 2.2 point to inferential gaps between evidence and conclusion that cannot be filled solely with epistemic values, so non-epistemic considerations must take up the slack. A rational reconstruction that defuses this criticism would have to show that one can in fact reason from the same evidence to the same conclusion without appeal to any non-epistemic considerations. This would only be possible either if the inferential gap does not actually exist, or if there are additional epistemic values that were not considered in the context of discovery. While there may be some cases that could be understood in one of these two ways, we do not think it plausible that this is a general diagnosis of all or even most cases where there is an inferential gap.

One may recognize that science, even in an ideal reconstructed form, is not value-free, but still argue that scientists should not concern themselves with any value-based considerations while engaged in scientific reasoning. Researchers, when considering scientific inferences, should restrict themselves to evaluating epistemic reasons, and ignore the intrusion of non-epistemic reasons. This is another potential formulation of the value-free ideal:

(VFI2) Scientists should act as if science is value-free.

We do not think there is much to be said for this conception of the ideal. If there are in fact values involved in scientific reasoning, then ignoring them does not serve the ends of value-freedom we discussed in § 2.1. Ignoring non-epistemic considerations will not make science more truth-apt, nor will it help fulfil the democratic ideal that policy-oriented science reflect the values of the citizenry. It is possible that VFI2 may serve the third end discussed in § 2.1—public trust in science—because scientists claiming value-freedom might convince the public even if the claim is false. However, we do not think that public trust in science predicated on a misrepresentation of scientific practice is worth having. If value-freedom contributes to public trust, it should be because science is genuinely worth trusting, not because of a false image of science.

VFI1 and VFI2 are the versions of the ideal that seem to be the primary target of criticisms raised by feminist philosophers of science. They point out that if scientists ignore the role of values in scientific inference, the consequence will be an entrenchment of the values of a socially dominant class, since dominant values usually serve as the invisible default in discourse that is not critically examined. To combat the bias inherent in prioritizing a single set of values, philosophers like Longino recommend explicit recognition of the influence of values. We agree—science is not, as a matter of fact, value-free; the purported value-freedom of science often hides the hegemony of an invisible value scheme that is taken for granted, and the appropriate response should be the acknowledgement and mitigation of the dominance of that value framework.

We have rejected the conceptions of the value-free ideal that present scientific practice as value-free, either as a matter of fact or as a pretence. A more plausible alternative is to think of the ideal as articulating a goal towards which science should be directed, rather than a description of current science:

(VFI3) Science should be value-free.

VFI3 does not deny the value-ladenness of scientific practice, nor does it argue that scientists should ignore the role of values. It expresses the desirability of a value-free end-state. One might assume that arguments against the feasibility of this end-state would immediately scuttle VFI3. After all, doesn't ought imply can? If it is impossible to attain value-freedom, how could the end-state be desirable? But this is too quick. "Ought implies can", if it is a sound maxim, applies to action—if it is impossible for an agent to act in a certain way, then the agent cannot be obliged to act thus. The maxim does not preclude the moral evaluation of unattainable states of affairs.

Even if complete world peace is unattainable, one can still argue for the desirability of the state. Along similar lines, one may project value-freedom as a regulative ideal that can guide our action even if it cannot be attained.[12]

However, as we noted in § 2.2, critics dispute the claim that the end-state of value-freedom is desirable, which would be a direct repudiation of VFI3. These critics present arguments for the undesirability of the ideal that are targeted at the ideal itself, because a fully value-free science would have characteristics that would undermine the practical significance of scientific inquiry. We will discuss these arguments at greater length below. For now, let us grant that the value-free end-state may well be undesirable, and if it is, VFI3 would be false.

Nevertheless, a fourth formulation of the value-free ideal is available:

(VFI4) Scientists should act as if science should be value-free.

Unlike VFI2, this formulation does not require pretence about the value-freedom of current science. VFI4 is compatible with acknowledging the value-ladenness of science. Unlike VFI3, it is not committed to the desirability of the value-free end-state. It merely says that scientists should act as if the end-state is desirable. It can thus sidestep the arguments critics raise against the desirability of value-freedom. It is based on the distinction between end-state desirability and pursuit desirability, and the idea that these two notions can come apart. Even if an end-state is not desirable, it is still possibly desirable to proceed with the goal of attaining that end-state.

The reader might object that while VFI4 does not call for the same sort of pretence that VFI2 demands, it still seems to call for something resembling pretence. Scientists must act as if a value-free end-state is desirable even if it is not. While this is true, the problem with VFI2 is the fact that the pretence is pointless and indeed very likely harmful—acting in the manner recommended by VFI2 will not help fulfil the purported ends of value-freedom. This is not the case with VFI4, as we argue in § 4. We will not only argue that VFI4 evades the arguments of critics of the value-free ideal, but also that it guides scientific practice in a manner that helps bring about the advantages of value-freedom mentioned in § 2.1. First, though, we argue that the value-free ideal as we conceive it is feasible to pursue.

3 Pursuit feasibility of the value-free ideal

The methodological critique and the framing problem challenge the feasibility of attaining a value-free end state. However, ideals can fail to be end-state feasible yet be pursuit feasible—even if the state described by the ideal is unattainable, there are still often strategies available to move closer to the ideal.

Values may play a direct biasing role in scientific reasoning, determining the degree of certainty assigned to hypotheses. For instance, the influence of values may lead researchers to consciously or unconsciously cherry-pick the data they regard as salient, or to collect data in a manner that favours a particular conclusion. But

[12] A similar point is made in de Melo-Martin and Intemann (2016).

there are ways to eliminate or reduce this biasing effect. Methodological strategies like random assignment and blinding have removed particular bias-prone decisions from the discretion of researchers (*bias reduction*), and the introduction of diverse value perspectives in a research discipline can expose and mitigate implicit and unacknowledged bias (*bias neutralization*).

Values may also play an indirect role in scientific inference, where they do not directly influence the degree of certainty assigned to hypotheses, but they help determine the threshold of certainty sufficient to accept or reject a given hypothesis (see Douglas 2009 on the direct-indirect distinction). Here, too, strategies for moving towards value freedom are available. If the degree of certainty of a hypothesis is placed between the evidential thresholds recommended by two conflicting value perspectives, then researchers could gather more evidence until the certainty falls either above or below both thresholds, so that partisans of both perspectives would agree on whether to accept the hypothesis (*evidence strengthening*). This would eliminate at least one bifurcation point. Alternatively, one could defer the dichotomized acceptance or rejection of the hypothesis to a later decision-making stage beyond the scope of scientific inference (*deferral*). This could be achieved by directly reporting degrees of certainty for various hypotheses, either precisely quantified or vaguely qualified, rather than committing to one specific hypothesis as the conclusion of scientific inference.

Strategies that mitigate (but don't necessarily eliminate) the influence of values on scientific reasoning—by reducing empirically- or theoretically-underdetermined methodological choices, by strengthening evidence, or by deferring decisions to accept or reject hypotheses—suffice to show that steps can be taken toward a state in which scientific reasoning is value-free. Such strategies exist and indeed are routine aspects of science. Therefore, the value-free ideal is, in general, pursuit feasible.

In the remainder of this section, we discuss the strategies of evidence strengthening and deferral in more detail, before defending our novel version of the value-free ideal in § 4.

3.1 Evidence-strengthening

John (2015) notes that a scientific conclusion can be regarded as beyond reasonable doubt when the evidence for that conclusion is extremely strong. In such a circumstance, even if different value perspectives have different evidential thresholds for accepting the conclusion, this will not result in bifurcation points, because the actual evidence would be strong enough to surpass all actually existing thresholds. So this would be a case in which values may be involved in determining the conclusions, but different values do not lead to different conclusions, making the inference value-free in our sense of the term.

The evidence-strengthening strategy can be illustrated by comparing observational and experimental trials. Evidence from an observational trial studying the association between A and B might not differentiate between the hypotheses that A causes B, that B causes A, or that a common cause C causes both A and B. Scientists with different values, and correspondingly different evidential thresholds, may differ in the conclusions they draw from the trial. The evidence-strengthening strategy would advise

performing an experimental trial, if possible: if A and B are correlated after A is administered to a randomly-assigned group and a comparison intervention is administered to a control group, then evidence supporting the hypothesis that A causes B is considerably strengthened. This makes it harder for values to influence the choice of hypotheses. Even if our values warn strongly against the risk of wrongly concluding that A causes B, leading us to set a high evidential bar for that conclusion, the results of a well-conducted experiment might cross that bar. It is, of course, not impossible for values to influence hypothesis choice even after evidence strengthening—no trial is perfect and the possibility of epistemic error remains, and thus the methodological critique continues to apply—but our conclusion is not that the value-free end-state is attainable, but rather that its pursuit is feasible. Steps can be taken toward eliminating the influence of values on scientific reasoning.[13]

3.2 Deferral

Betz (2013) notes that the methodological critique relies on the premise that generating policy-relevant scientific results requires making decisions regarding the acceptance or rejection of hypotheses which are not fully determined by empirical constraints, and it is this slack in the scientific process which affords the influence of values. Betz rearticulates a response due to Jeffrey (1956), which denies that scientists must accept or reject hypotheses and argues instead that they should characterise their uncertainties and report their findings accordingly, leaving questions of acceptance or rejection of findings to policy-makers. Rather than reporting their conclusions in terms of 'plain hypotheses' (that such-and-such is the case), scientists ought to report their conclusions in terms of 'hedged hypotheses' (given this evidence, the probability of such-and-such is so-and-so). This will shift assessments of the sufficiency of evidential support for acceptance or rejection of hypotheses outside scientific reasoning and into policy deliberations.

A challenge to the fruitfulness of this strategy is that the evidential sufficiency problem arises again when it comes to quantifying uncertainty (Rudner, 1953; John, 2015; Frisch, 2020). Different value sets may disagree about the threshold of evidence necessary to assign a particular degree of certainty to a given hypothesis. In response, Betz deploys the evidence strengthening strategy when it comes to determining degrees of certainty. Choose a representation of uncertainty imprecise enough that all contending value perspectives would agree that it crosses the threshold of sufficient evidence. If there is contention about point probabilities assigned to hypotheses, then perhaps use intervals instead; if intervals prove controversial, maybe qualitative ascriptions of uncertainty will not; if even that fails, perhaps a simple enumeration of serious possibilities will suffice; and so on. Betz seems to think that this combination of deferral and evidence strengthening would, in principle at least, allow for the complete elimination of values from scientific inquiry.

But this strategy may come at a cost. If we were able to attain broad agreement among all reasonable value perspectives with a fairly precise quantification of uncertainty, things would be looking up. But that is an ideal case that is unlikely to hold in

[13] See also Boulicault and Schroeder (2021) for discussion of this strategy.

many (maybe even most) policy-relevant scientific domains. We may need to make our description of uncertainty less precise in order to meet all reasonable evidential thresholds, but as the precision drops, so does the action-guiding potential of our results. After all, we want our scientific conclusions, especially those relevant to policy, to give a relatively clear basis for decision-making. If our climate scientists gave us precise probabilities for a variety of future outcomes, policy-makers could plug them into decision-making algorithms in order to strategize about climate action. But if all we get from climate science is a set of possible outcomes, with no further indication of their relative likelihoods, it is much less clear how one could translate that information into action. While choosing less informative representations of uncertainty helps with deferral and evidence strengthening, at some point the representation becomes too uninformative to be useful for guiding action.[14]

This worry challenges the end-state desirability of value-freedom. If attaining value-freedom would mean that our scientific conclusions sacrifice their action-guiding character, the state of value-freedom would come at a severe cost. This criticism has bite if the desirability of the end-state is a presupposition of one's defence of the value-free ideal, as it appears to be for Betz. However, as will become clear in § 4, we do not share this presupposition.

The deferral strategy can also help mitigate the framing problem. As discussed above, Atkinson (1998) argues that the study of poverty involves choosing between value-laden conceptualisations. In such cases, hedged conclusions might help mitigate the influence of values. An economist might be motivated by value-based considerations to convince others that poverty rates during the 1980s were greater in the United Kingdom than they were in France, and that would be supported by some empirical data, using a particular operationalisation and definition of poverty. However, with the development of more nuanced ways of conceptualising and operationalising the measurement of poverty, one could note that the French economist is relying on particular choices, while on other choices it appears that poverty in the United Kingdom was less than that of France. The articulation of alternative conceptualisations shows that the reliance on a single conceptualisation and measurement may not be robust. A reasonable thing for economists to do would be to hedge their inferences accordingly.

4 The value-free ideal is pursuit desirable

All other things being equal, reducing the influence of values on scientific reasoning is desirable, given the benefits of value-freedom discussed in § 2.1. But, of course, all other things are usually not equal. Strategies for eliminating bifurcation points can come with associated costs, and the costs trade off against the benefits of increased truth-aptness and democratic legitimacy. The costs may outweigh the benefits, such

[14] This objection has been pressed by several authors. Steele (2012) notes that in complex domains deferral would require complicated reports which would mitigate their policy-relevance, Elliott (2011) argues that the deferral strategy is harmful because it involves 'passing the buck' to decision-makers who must formulate policy, and Steel (2016) and Brown (2019) argue that the deferral strategy renders science practically inconsequential. Frisch (2020) argues for a 'no-buck-passing' principle for science.

that reduction of the role of values is, all things considered, undesirable. Nevertheless, we argue in this section that despite such undesirability of a value-free end state, that state is good to pursue.

Before delving into the implications of this possibility, let us first consider what these costs might look like. Some of the costs may simply be straightforward resource costs. When we adopt methodological strategies to reduce potential value-induced bias, the new methodology might be costlier in terms of the money, technology, effort or time required. There is very rarely such a thing as a free lunch; improving the quality of our evidence typically implies a greater resource burden. The same problem holds if we want to increase the *quantity* of our evidence in order to ensure that the degree of certainty of a hypothesis does not lie between two contested evidential thresholds. At some point, the marginal cost associated with gathering more or better evidence may outstrip the marginal benefit.

The cost may also be ethical. Conducting a randomized controlled trial in some context might help with bias reduction, but it may be ethically unacceptable to implement certain treatments on study participants. Gathering a lot of evidence on the efficacy of a new drug before publicly presenting the conclusion may strengthen our evidence, but the ethical urgency of a serious disease ravaging a community might mean we have to settle for a smaller quantity of evidence.

Another cost worth considering is the problem with the deferral strategy that we discussed in § 3.2—an overly-hedged hypothesis might not be able to guide action appropriately. If the aim of a scientific endeavour is to influence policy in certain ways, then we must ensure that the conclusions are not hedged to the point that it is unclear how they should inform policy. The purpose for which research is being conducted sets constraints on how the conclusion can be represented.

So strategies for reducing the influence of values have potential costs in terms of resource use, ethics and action guidance. When considering whether a particular strategy is desirable, we must balance the costs and benefits. But this cost-benefit analysis is itself an application of normative considerations. If, in attempting to reduce the influence of values, we need to make value-based decisions about whether the costs of pursuing value-freedom outweigh the benefits, then are we really moving closer to value-freedom at all? Eliminating a bifurcation point appears to introduce a new bifurcation point corresponding to disagreement about whether eliminating the first bifurcation point was worth it.

But recall that we restrict our defence of the value-free ideal to a very specific stage of scientific inquiry, the inferential stage, where we draw conclusions from the available evidence. The kinds of value considerations relevant to weighing the costs and benefits of value-minimization strategies are not considerations that apply at that stage. They are, rather, considerations that place prior restrictions on what we can accomplish at the inferential stage. This is most obvious when we consider the choice of methodology. Different methodological choices place different constraints on the quality and quantity of evidence from which we can draw conclusions. This does not mean that the ethical considerations that go into our methodological choices should be targeted by the value-free ideal. Similarly, a prior specification of the amount of action guidance required from a research endeavour will constrain the kinds of con-

clusions we draw at the inferential stage, but it will not be directly involved as a step in the inferential chain.

The values we seek to eliminate as we pursue the value-free ideal are those that make a difference to the conclusion that is drawn based on evidence. The values involved in performing the cost-benefit analysis described above play a different role. They help determine the kind and amount of evidence available to us (by guiding methodological choices), and also the particular form in which the conclusions are presented (i.e. how we choose to represent uncertainty), but they do not affect the content of the conclusions if the evidence is kept fixed, in the sense relevant to our analysis. Disagreements about what level of hedging in scientific results is best for policy purposes, or about which methodology makes the most ethical use of available resources, cannot lead us to conflicting conclusions from the same evidence base. In other words, these disagreements do not introduce new bifurcation points. These decisions are better thought of as setting side constraints on the process of scientific inference rather than as adding links to the inferential chain, so they do not threaten the value-free ideal as we conceive it. Value disagreements that would introduce new bifurcation points – disagreements about the appropriate threshold of evidence necessary to draw a particular conclusion, for instance – cannot be assimilated into the external constraints.

Returning to our example where there is a serious disease ravaging a community, ethical disagreement about the urgency of the situation might lead to disagreement about how much evidence should be collected. A decision about the permissible amount of evidence would then function as a constraint on our ability to eliminate bifurcation points, since certain evidence-strengthening strategies would be ruled out. But the disagreement about urgency is not a bifurcation point relative to this set of evidence; it is a disagreement about whether this is all the evidence we can collect. That disagreement may itself be due to divergent conclusions drawn based on some prior evidence about the virulence of the disease, and in that context it might constitute a bifurcation point. Ethical considerations that are involved in bifurcation points in one context of inquiry might function as constraints in another context.

It might be argued that disagreements about the level of hedging *do* lead to bifurcation points. If one group of scientists believes a more informative conclusion is warranted than the other group, then there is an incompatibility in the conclusions drawn. However, the value disagreement leading to this incompatibility is not part of the inferential chain from evidence to conclusion. The disagreement is at a meta level – once the two groups have determined their conclusions, they are disagreeing about how much they can hedge to eliminate bifurcation points while still fulfilling the action-guidance constraint. One way of seeing that the evidence is irrelevant to this disagreement is that we could hypothetically present the groups with the more and less informative conclusions before they even see the evidence, and they would still disagree about which conclusion would satisfy the demand for action-guidance. It is not a disagreement about what the evidence tells us, it is a disagreement about how the conclusion should be framed.

4.1 The minimal value set

Let us imagine an idealized research group, or research community, attempting to attain the value-free end-state by mitigating the role of values where possible using the strategies we have discussed. However, their activity is governed by constraints concerning resource use, research ethics and action-guidance. If a value-mitigation strategy would run afoul of these constraints, they do not proceed. As a consequence, the researchers might not be able to attain value-freedom even though they are aiming for it, because at some point there may be no available strategy that takes them closer to the value-free end-state without violating the constraints. In other words, they may reach a stage at which the value-free ideal is no longer pursuit feasible.

There may be some degree of path-dependence here, where how close one can get to the end-state without violating the constraints depends on what particular sequence of value-mitigation strategies one employs. In the idealized scenario we are considering, let us suppose that the researchers have enough time and persistence to get as close to the end-state as possible across all available paths. Assuming that complete value-freedom is not attainable, the state they end up at will still involve values playing a decisive role in the inference from evidence to conclusions. We label the set of all remaining value-based considerations – the set of values that is as small as one could possibly get subject to the constraints – the *minimal value set*. The minimal value state – the state of enquiry where only the value considerations in the minimal value set remain – is where a rational reconstruction of the inferential link between evidence and conclusion will have the fewest possible number of bifurcation points.

Even if attaining the value-free end-state is undesirable, arriving at the minimal value set is, we maintain, desirable. This is because the undesirability of the end-state is captured by the constraints on the optimization process we have been describing. All aspects of the end-state that make attaining it not worth the benefits of value-freedom discussed in Sect. 2.2 are encoded in the constraints. Since the minimal value state is, by definition, allowed by the constraints, it is a state in which the costs imposed by the constraints do not trump the benefits of reducing the influence of values. A properly motivated scientific community should want to arrive at the minimal value state.

It is worth noting that our conception of minimality does not call for minimising *reference* to values in scientific argumentation. It calls for minimising the extent to which values play a crucial role in inference, as measured by bifurcation points. This distinction gets to the heart of Longino's worries about the value-free ideal, which we discussed above. She correctly points out that even if scientific reasoning does not explicitly invoke value-based considerations, there is usually still an implicit value framework guiding arguments. Her call for an increased diversity of perspectives in order to expose this implicit value system is fundamentally a call to make explicit the already existent bifurcation points, and we are in full agreement with that. And once the role of values is made explicit, it might afford the possibility of a bias neutralization strategy, adjusting the results of research so that they don't encode preference for a particular set of contentious values. Such a strategy would refer to multiple value perspectives, but would do so in order to eliminate bifurcation points. As a simple example, one may move from the contentious claims "Based on evidence E, conclu-

sion C1" and "Based on evidence E, conclusion C2" to the potentially uncontentious claim "Based on evidence E, conclusion C1 if you hold value V1, and conclusion C2 if you hold value V2". This latter conclusion makes reference to two competing values, but may be agreed upon by partisans of both. On our account, this diversification of perspectives is therefore a move towards the minimal value set rather than an explosion of value-based considerations.

4.2 The indeterminacy of minimality

Attaining a minimal value state is desirable, and pursuing a state of value-freedom within the limits set by the side constraints will, ideally at least, lead us to a minimal value state. Does this suffice to establish the pursuit desirability of the value-free ideal?

It does not. The pursuit desirability of a goal does not simply amount to the desirability of taking steps towards the goal, it means that actively and consciously pursuing the goal is desirable. Consider an unattainable and possibly undesirable value-free end-state S, and suppose the closest we can get is the minimal value state M, which is both attainable and desirable. Perhaps by aiming at S we will end up at M, but that alone doesn't justify aiming for S. There might, after all, be an even better strategy for getting to M. And such a strategy suggests itself immediately—why not just aim for M directly, instead of S? If you are successful, you will end up in the same place, and you have the added advantage of not misrepresenting to yourself and others the true target of your endeavours. If one can get to the same destination without having to sustain a pretence along the way, is that not preferable? That would suggest abandoning the value-free ideal and replacing it with the minimal value ideal, which would state: don't aim for an elimination of all values, a target that is both unattainable and undesirable, but rather, aim for the state of minimal value involvement.

This challenge to the pursuit desirability of the value-free ideal does not work, however, because while it is true that the minimal value state is our actual desirable end-state, consciously aiming towards that state is not, in general, an action-guiding strategy. This is because we have no prior means to determine what the minimal value state is. Its status as the minimal value state only possibly becomes apparent once we actually get there and realise there are no further value-mitigation moves we can make without violating the constraints. If a state cannot be recognized as the minimal value state in advance, then we cannot adopt the strategy of aiming towards the minimal value state. The dictum "Attempt to attain a state where your scientific inference relies on a minimal value set" cannot guide action. On the other hand, the value-free state can be given a prior specification, so we can consciously aim towards it and use it as a guide to further action. We know what it means to try to get closer to that state. So even though our ultimate end is the minimal value state, the best action-guiding strategy available to us in order to get there is to attempt to attain value-freedom.

To illustrate this point, suppose that you are attempting to find your way in a largely featureless desert landscape. You want to get to a particular patch of land where, years ago, a friend buried some treasure. But of course, the treasure is underground, and the patch of ground looks to the naked eye just like any other boring

patch of desert. There are no road signs or addresses, so your friend could not directly tell you the location of the treasure. There is, however, one clearly visible feature in the desert—a giant red rock rising up in the distance. You do not want to get to the giant red rock—it is a sheer cliff and is teeming with snakes. You could not get to the red rock even if you wanted to—surrounding it is a deep crevasse that cannot be crossed. But you know that the treasure is buried at the reachable point in the desert that is closest to the red rock.

In these contrived circumstances, the best way to get to the treasure is to try to get to the red rock. Once you realize you can get no closer to your represented goal, the rock, you will in fact be at your actual goal, the location of the treasure. You know that the end-state of being at the red rock is neither attainable nor desirable, but the state you actually want to get to is not one you can meaningfully aim towards. Aiming towards the surrogate goal of the red rock is the best way to get there.

The relationship between the minimal value state and the value-free state is similar. Advances in scientific methods and concepts are often impossible to predict. Thus, it is impossible to predict in advance which values could be eliminated via the evidence-strengthening and conclusion-deferral strategies. As an example, it took Fisher's methodological innovation of randomization in experimental design to both expose the biasing impact of confounding factors and to design a method to mitigate that biasing impact. Before this development, it was difficult to even conceive of the biasing threat of what is sometimes referred to as selection bias (Hacking, 1988). After this development, experimentalists could block any influence of values which would have intruded via selection bias. Moreover, the imperative to act as if science should be value-free itself motivates scientific innovation. Consider for example the role of publication bias in pharmaceutical research—critics of this practice have argued that publication bias affords the influence of values (say, industrial executives' profit motive) on conclusions about the effectiveness of pharmaceuticals; in turn, clinical science has begun to employ tactics to block this threat, for instance by requiring the pre-registration of clinical trials.

Given that the minimal value state is determined by the strategies available for value mitigation and those strategies are often developed in response to immediate scientific challenges rather than pre-determined, one can't know that one is at the minimal value state until one gets there and exhausts the search for strategies to get any further towards value freedom. Like the red rock, the end-state of value freedom serves as our goal for the purpose of guiding action. Getting to the minimal value state requires researchers to constantly act as if they are trying to get to the value-free state, even if, in moments of reflection, they may admit that is not their actual goal. Moreover, the minimal value state can only be recognized based on an inability to get beyond it, and even then, there will be uncertainty about whether further value reduction is genuinely impossible or whether an appropriate strategy simply hasn't been contrived as yet. So even when they are at what appears to be the minimal value state, researchers should continue to attempt to go beyond it.

Our point is not simply that the minimal value state must be defined or identified with reference to the value-free state. That in itself would not preclude using it as a target to guide scientific action. After all, we often directly pursue targets that are only specifiable in terms of their relation to another putative target. We might aim to

order the second most expensive wine on a menu, or to park in the space to the left of the closest space to a house.[15] In these examples, however, the reference to another target is needed simply to identify the actual target, which can then be actively pursued. In the case of the minimal value state, there is no way to actively pursue it independent of the pursuit of value-freedom. As argued above, the minimal value set is not a static and easily identifiable state. Choosing the second most expensive wine on a menu is a well-defined and static goal, and knowing if one has achieved the goal is epistemically trivial. The minimal value set has none of those properties. Technological developments change what the minimal value set is, and knowing that one is at the minimal value set is only possible once one is already there. Thus, the minimal value state cannot be recognized except through the active pursuit of the value-free state, since its status as the minimal value state only becomes apparent as we make attempts to get beyond it and find we cannot.

Moreover, it is not just that the minimal value state cannot be identified without aiming at value freedom. If that was all, then one could plausibly characterize the process of bifurcation point elimination as a search for a minimum value state subject to constraints. After all, in many constrained minimization algorithms, the aim is explicitly to find a minimum, but this is accomplished by trying to reduce the relevant quantity as much as possible without violating the constraints. Procedurally, then, a strategy of minimization would look the same as a strategy of elimination.

But this assumes that the minimal value state is static. As argued above, technological and methodological developments in science can change the minimal value state: that is, such developments can permit the elimination of bifurcation points that were ineliminable prior to the development, all the while satisfying the relevant constraints. Thus, the strategy is not straightforward constrained minimization. Besides the exogenous constraints defined by resource limitations, ethical guidelines, and requirements of action guidance, there is also an endogenous constraint defined by the technological and methodological resources available to the scientist. We call this an endogenous constraint because it usually presents as a hard constraint at the level of the individual research project, but when we consider the development of the scientific discipline as a whole, the constraint is malleable. Scientific progress often involves the discovery of technology that permits elimination of bifurcation points that previously could not be eliminated. As a consequence, while constrained minimization may be an appropriate description of the way scientists should deal with value ladenness at the scale of the individual project, it misses the point that at a larger scale, scientists should be trying to change the minimal value state by altering the technological constraint.[16] Describing the process as one of elimination of bifurcation points, rather than constrained minimization, captures the important practical lesson that scientific progress does not just involve eliminating bifurcation

[15] Our thanks to an anonymous reviewer for suggesting these examples.

[16] It is, of course, true that one could still describe this larger scale process as constrained minimization. Trivially, any elimination strategy in the real world could also be described as constrained minimization, as there are always normative constraints restricting how we can pursue elimination. No one would suggest, for instance, that the goal of poverty elimination be fulfilled by bringing about the extinction of humanity. We do not take this to mean that we should never talk about a search strategy as aiming at elimination rather than minimization.

points that can be eliminated given the current state of technology and methodology, but also thinking about how we can advance technology and methodology to permit the elimination of further bifurcation points. There is, in principle, no pre-identifiable limit to the extent to which technological advancement might permit us to move towards value freedom. The process of bifurcation point elimination should not be modelled as constrained minimization, and we do not think it appropriate to describe it as aiming towards the minimum value state. It is not just that we cannot identify the minimum value state, it is that there is no fixed minimum value state to aim towards.

The indeterminacy of the minimal value state—the fact that we cannot recognize the state except through continuing failed attempts to get beyond it to value-freedom—is what justifies the claim that value-freedom is pursuit desirable. It is desirable to consciously aim towards that state, to act as if science should be value free, even if that state is end-state undesirable. And the pursuit desirability of value freedom is enough to ground a version of the value-free ideal.

A related point is made in Larroulet Philippi (2020). Referring to Kitcher's notion of well-ordered science, the author argues that the notion is an end-state ideal which offers little concrete guidance to how the scientific research agenda should be set. Channelling Sen (2009), Larroulet Philippi notes that end-state ideals are neither necessary nor sufficient for guiding improvements; Sen suggests more emphasis on 'transitional accounts', motivating Larroulet Philippi's distinction between 'ideal answers' and 'ideal procedures'. Ideal answers are normative ideals which answer a normative question (like "what does a just society look like?") while ideal procedures are normative ideals which specify how the normative question should be addressed (like Rawls' 'original position'). Just as Kitcher's well-ordered science is an ideal answer, the value-free ideal has tended to be characterized in terms of an ideal answer. Larroulet Philippi's conclusion is that the notion of well-ordered science cannot determine science's research agenda. The same can be said about the value-content of the ideal end-state of science. It is impossible to know *ex ante* what values will be biasing and what biases can be mitigated by future methodological developments. Thus, the minimal value set for science is indeterminable and therefore not action-guiding. But what science *can* do is to pursue a procedural ideal of value-freedom, and our version of the value-free ideal is just that.

5 Conclusion

We have argued for a principle which is stronger than merely advising scientists to minimise the influence of values. If our thesis was merely a value-minimization principle, then our view would be similar to that of some of the prominent critics of the value free ideal, such as Douglas (2009). Instead, we have argued for a version of the value-free ideal, advising scientists to aim at the elimination of all value-based influence on scientific reasoning, subject to the external constraints (financial, ethical, action-guidance) that are standard in scientific investigation.

Given our concession that there are value-based constraints on how much value-freedom is desirable, one might say that our account of values in science is not in fact a value-free ideal. Our putative ideal, goes this response, is in fact something like

"scientists should act as if science should be value-free, subject to value-laden constraints." What follows the comma in that statement of our ideal, goes this response, entails that our ideal is not a *value-free* ideal. However, all ideals and indeed all goals have this structure. Give an Olympic sprinter the goal of running as fast as she can, and what follows the comma is "… subject to value-laden constraints, such as not doping." Give a country the policy of achieving net-zero carbon emissions, and what follows the comma is "… subject to value-laden constraints, such as the maintenance of some minimum degree of quality of life." Give a judicial system an ideal of innocent-until-proven-guilty, and what follows the comma is "… subject to value-laden constraints, such as not deciding all legitimate suspects as innocent." There is nothing infelicitous in saying that the Olympic sprinter has the goal of running as fast as she can, that the country has the policy of achieving net-zero carbon, or that the judicial system has the ideal of innocent-until-proven-guilty.

We disagree with philosophers who maintain that there is a positive, productive role for values in scientific reasoning, particularly as instantiated in what we called the value-enrichment critique. For example, Douglas and Elliott (2022) argue that many well-known cases in the values in science literature show that "values helped to debias science." Douglas (2009) similarly suggests a productive role for values in science: "In place of the value-free ideal, we need a new ideal for science, one that accepts a pervasive role for social and ethical values in scientific reasoning, but one that still protects the integrity of science" (1). Longino writes that values "can be understood as a rich pool of varied resources, constraints and incentives to help close the gap left by logic" (2002, 128). She further suggests that this consideration "turns the value-free ideal upside down—values and interests must be addressed not by elimination or purification strategies, but by more and different values" (2004, 137). To illustrate, Longino refers to interventions by feminist anthropologists and primatologists since the 1970s as an example of "value-driven research that has improved quality of science in those areas" (2004, 137). Brown (2013) goes as far as claiming that evidence should not even have priority over values. In contrast, our view maintains precisely the opposite, namely, that scientists ought to do all they can to resist the influence of values in scientific reasoning.

While we agree that science is often 'value-driven', and that value-motivated scientists can improve the quality of research in a particular domain, we suggested in § 4.1 that the cases motivating Longino's position can be persuasively re-framed. Our view is that a more complete description of the cases appealed to by Longino (in various publications cited above), Anderson (2004), and many others would be to say that the infusion of new sets of values helped to debias the relevant science by exposing the impact of pre-existing biases, i.e., by revealing previously unacknowledged bifurcation points. Though the enrichment of values in a scientific domain can indeed improve the quality of science in that domain, the improvements can be completely characterized in epistemic terms.[17] The enriching values may motivate the pursuit of additional research questions and thereby the gathering of novel evidence, but that is value-influence on the external aspects of scientific reasoning. The enriching values may also motivate the hunt for biases in existing theories and methods, but the ulti-

[17] Ruphy (2006) defends a similar argument.

mate benefit of exposing bias is the eventual elimination of newly revealed bifurcation points. So the exposure of those biases is a crucial step towards eliminating the influence of values on scientific reasoning, rather like fighting fire with fire.

Nevertheless, one might be tempted to think that the view being defended here is not that different from the views of at least some of the prominent critics of the value-free ideal, since these critics agree value influence should be minimised, not eliminated, and we have also argued that the desirable end-state is a state of minimal value influence, not the value-free state. However, these critics do not explicitly recognize the difference between the pursuit desirability and end-state desirability of value freedom. If our argument that the minimal value state can only be attained through the explicit pursuit of value-freedom is sound, it might turn out that some prominent opponents of the value-free ideal are in fact committed to the version of the ideal we propose.

To summarize, we have argued that it is possible to pursue the value-free ideal, it is good to pursue the value-free ideal, and we have given the value-free ideal a novel articulation: scientists should act as if science should be value-free.

> *"The struggle itself towards the heights is enough to fill a man's heart. One must imagine Sisyphus happy."* (Camus, 1942).

Acknowledgements We are grateful to Cristian Larroulet Philippi, Adrian Erasmus, Hamed Tabatabaei Ghomi, Sophia Crüwell, Ina Jäntgen, Zinhle Mncube, Adrià Segarra, Oliver Holdsworth and Kit Patrick for helpful discussion.

Author contributions Both authors have contributed to the development of the argument in the work. The authors have not made discrete contributions that can be clearly distinguished.

Funding The authors did not receive support from any organization for the submitted work.

Declarations

Financial or non-financial interests The authors have no relevant financial or non-financial interests to disclose.

Ethical approval This work has no empirical component and does not raise any issues requiring ethical approval.

Informed consent This work has no empirical component and does not involve any participants other than the authors, so considerations of informed consent do not apply.

References

Adamson, P., & Ganeri J. (2020). *Classical Indian philosophy*. Oxford University Press.

Alexandrova, A. (2018). Can the science of well-being be objective? *British Journal for the Philosophy of Science*, *69*(2), 421–445.

Anderson, E. (2004). 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.

Atkinson, T. (1998). *Poverty in Europe*. Blackwell.

Betz, G. (2013). In defence of the value free ideal. *European Journal for Philosophy of Science*, *3*, 207–220.
Boulicault, M., & Schroeder, A. (2021). Public trust in science: Exploring the idiosyncrasy-free ideal. In K. Vallier & M. Weber (Eds.), *Social trust.* Routledge.
Bright, L. K. (2018). Du Bois' democratic defence of the value free ideal. *Synthese*, *195*, 2227–2245.
Brown, M. J. (2013). Values in science beyond underdetermination and inductive risk. *Philosophy of Science*, *80*, 829–839.
Brown, M. J. (2019). Is science really value free and objective? From objectivity to scientific integrity. In K. McCain, & K. Kampourakis (Eds.), *What is scientific knowledge? An introduction to contemporary epistemology of science* (pp. 226–242). Routledge.
Camus, A. (1942[1955]). *Le mythe de Sisyphe*. Gallimard.
Conix, S. (2020). Enzyme classification and the entanglement of values and epistemic standards. *Studies in History and Philosophy of Science Part A*, *84*, 37–45.
De Melo-Martin, I., & Intemann, K. (2016). The risk of using inductive risk to challenge the value-free ideal. *Philosophy of Science*, *83*(4), 500–520.
Douglas, H. (2000). Inductive risk and values in science. *Philosophy of Science*, *67*(4), 559–579.
Douglas, H. (2009). *Science, policy, and the value-free ideal*. University of Pittsburgh Press.
Douglas, H., & Elliott, K. (2022). Addressing the reproducibility crisis: A response to Hudson. *Journal for General Philosophy of Science*, *53*, 1–9.
Elliott, K. (2011). *Is a little pollution good for you? Incorporating societal values in environmental research*. Oxford University Press.
Elliott, K. (2017). *A tapestry of values: An introduction to values in science*. Oxford University Press.
Frisch, M. (2020). Uncertainties, values, and climate targets. *Philosophy of Science*, *87*(5), 979–990.
Goldenberg, M. (2021). *Vaccine hesitancy: Public trust, expertise, and the war on science*. University of Pittsburgh Press.
Hacking, I. (1988). Origins of randomization in experimental design. *Isis*, *79*(3), 427–451.
Havstad, J. C., & Brown, M. J. (2017). Inductive risk, deferred decisions, and climate science advising. In K. Elliott, & T. Richards (Eds.), *Exploring inductive risk: Case studies of values in science* (pp. 101–126). Oxford University Press.
Hicks, D. J. (2018). Inductive risk and regulatory toxicology: A comment on de Melo-Martín and Intemann. *Philosophy of Science*, *85*(1), 164–174.
Hoyningen-Huene, P. (2023). Objectivity, value-free science, and inductive risk. *European Journal for Philosophy of Science*, *13*, 14.
Intemann, K. (2005). Feminism, underdetermination, and values in science. *Philosophy of Science*, *72*, 1001–1012.
Jeffrey, R. C. (1956). Valuation and acceptance of scientific hypotheses. *Philosophy of Science*, *23*(3), 237–246.
John, S. (2015). The example of the IPCC does not vindicate the value free ideal: A response to Gregor Betz. *European Journal for Philosophy of Science*, *5*, 1–13.
Kourany, J. A. (2008). Replacing the ideal of value-free science. In M. Carrier, D. Howard, & J. A. Kourany (Eds.), *The challenge of the social and the pressure of practice: Science and values revisited* (pp. 87–111). University of Pittsburgh Press.
Lacey, H. (1999). *Is science value free? Values and scientific understanding*. Routledge.
Larroulet Philippi, C. (2020). Well-ordered science's basic problem. *Philosophy of Science*, *87*(2), 365–375.
Longino, H. (1990). *Science as social knowledge*. Princeton University Press.
Longino, H. (2002). *The fate of knowledge*. Princeton University Press.
Longino, H. (2004). How values can be good for science. In P. Machamer & G. Walters (Eds.), *Science, values, and objectivity* (pp. 127–142). University of Pittsburgh Press.
Miller, B. (2014). Science, values, and pragmatic encroachment on knowledge. *European Journal for Philosophy of Science*, *4*(2), 253–270.
Proctor, R. N. (1991). *Value free science: Purity and power in modern knowledge*. Harvard University Press.
Rooney, P. (1992). On values in science: Is the epistemic/non-epistemic distinction useful? In D. Hull, M. Forbes, & K. Okruhlik (Eds.), *Proceedings of the 1992 Biennial Meeting of the Philosophy of Science Association*. Philosophy of Science Association.
Rudner, R. (1953). The scientist qua scientist makes value judgements. *Philosophy of Science*, *20*(1), 1–6.
Ruphy, S. (2006). Empiricism all the way down: A defense of the value-neutrality of science in response to Helen Longino's contextual empiricism. *Perspectives on Science*, *14*(2), 189–214.

Schroeder, S.A. (2021). Democratic values: A better foundation for public trust in science. *The British Journal for the Philosophy of Science*, *72*(2), 545–562.
Sen, A. (2009). *The idea of justice*. Belknap.
Steel, D. (2016). Climate change and second-order uncertainty: Defending a generalized, normative, and structural argument from inductive risk. *Perspectives on Science*, *24*(6), 696–721.
Steele, K. (2012). The scientist qua policy advisor makes value judgments. *Philosophy of Science*, *79*(5), 893–904.
Ward, Z. B. (2021). On value-laden science. *Studies in History and Philosophy of Science Part A*, *85*, 54–62.
Winsberg, E. (2012). Values and uncertainties in the predictions of global climate models. *Kennedy Institute of Ethics Journal*, *22*(2), 111–137.
Winsberg, E., Brennan, J., & Surprenant, C. W. (2020). How government leaders violated their epistemic duties during the SARS-CoV-2 Crisis. *Kennedy Institute of Ethics Journal*, *30*(4), 215–242.