Philosophy of Science (2025), **00**, 1–18
doi:10.1017/psa.2025.18

PSA PHILOSOPHY OF SCIENCE ASSOCIATION
PHILOSOPHY OF SCIENCE

ARTICLE

Towards a Direct Role for Values in the Heart of Science

Tim Lewens

Department of History and Philosophy of Science, University of Cambridge, Cambridge, UK
Email: tml1000@cam.ac.uk

(Received 04 October 2024; revised 10 January 2025; accepted 23 March 2025)

Abstract

Douglas has argued that if values were accorded a "direct" role during the "internal" phases of science, this would amount to "wishful thinking" in place of evidence. This article draws on two claims that jointly threaten this position. First, building a cogent ethical case in favor of a claim about what ought to happen is not a simple matter of saying what one wishes were true; a good ethical case has a kind of discipline to it. Second, some theorists have argued that scientists do and should defend "mixed hypotheses," that is, hypotheses that implicate both factual and evaluative content.

1. The proper role for values

Douglas has made a series of powerful cases for the proper role of values of many kinds—including social, ethical, and political—within all phases of the sciences. Throughout this article, I will use the term *values* as a shorthand for social, ethical, and political values alone. Partly because of the pervasive role she accords to such values, Douglas also takes pains to explain the circumstances under which it is inappropriate to appeal to them. She has argued that values can sometimes play what she calls a "direct" role in science: As she puts it, values might sometimes "determine our decisions in and of themselves, acting as stand-alone reasons to motivate our choices" (Douglas 2009, 96). She argues that this "direct" role for values is illegitimate within "the heart of doing science—during the characterisation of data, the interpretation of evidence, and the acceptance of theories" (102). A "direct" role for values is, in her view, appropriate only in what she calls the sciences' "external" phases. Here scientists must grapple with matters of obvious ethical import, such as whether it is justifiable to research on live animals or whether to publish the results of research that might be misused by malevolent actors.

For Douglas (2009), giving a direct role to values in what she calls the "internal" phases of the scientific enterprise would be equivalent to using values in an evidential

capacity. That, in turn, would involve a pathological form of wishful thinking: "Values are not evidence; wishing does not make it so" (87). Values are needed in the core phases of scientific research, but their primary use is in determining whether enough evidence has been gathered to accept a hypothesis (Douglas 2018, 3). A decision about whether evidence is sufficient can be made responsibly only if the costs of error (i.e., the costs of accepting a hypothesis that is in fact false or the costs of rejecting a hypothesis that is in fact true) are taken into account. Those costs can be determined only via a consideration of values. This makes the proper role of values in the "heart of doing science" an indirect one: "The integrity of science cannot tolerate a direct role for values throughout that process" (Douglas 2009, 96).

This article explores the possibility of a direct role for values, even within the "heart" of the scientific enterprise. My way into these debates draws on a prima facie tension between the position defended by Douglas and the position defended by Alexandrova (2018). In section 3, I give a summary of Douglas's position. I suggest that the key claim on which she rests her overall view is that it is inappropriate for evaluative considerations to play an evidential role—even though they may play other roles—with respect to scientific hypotheses. In section 4, I endorse Anderson's (1995) insistence that disciplined ethical reflection need not involve any kind of "wishful thinking." In section 5, I summarize Alexandrova's position. She has argued that the sciences do, and should, often trade in what she calls "mixed" hypotheses. She argues that such "mixed" hypotheses simultaneously implicate factual and evaluative considerations. If one puts Anderson's and Alexandrova's views together, then on the face of things, one would expect to see a legitimate role for evaluative considerations in helping to support—in a "direct," albeit disciplined, way—the ethical aspects of these "mixed" scientific hypotheses. I am not the first to question the proscription on an evidential role for values with respect to scientific claims: Brown (2020, 110), too, has remarked, "This seems to me to depend on the context, on the nature of the claim in question, and on the nature of the relevant values." And he has specifically speculated that, "when value-laden thick concepts appear in the claim, values might stand as reasons for such claims" (110). But to my knowledge, no other work has examined the tension between Alexandrova and Douglas in detail.

In section 6, I explore a potential response that might dispel this tension and render Douglas's position compatible with Alexandrova's stance on "mixed" hypotheses. Alexandrova's account of what a "mixed" hypothesis is differs importantly (as does Anderson's account of "thick" concepts, on which Alexandrova draws) from the way the notion of "thick" moral concepts is used in meta-ethics. Put briefly, Alexandrova does not say that "mixed" hypotheses need have evaluative *content*; she says only that evaluative considerations play into decisions over how to define the terms that feature in mixed hypotheses. There is room, then, for a position congenial to both Douglas and Alexandrova. That position allows that evaluative considerations are put to work in deciding how best to define concepts, even when the resulting content of those concepts is purely descriptive. Evaluative reasoning would then play a "direct" role in fashioning these appropriate concepts, but that direct role would not fall within the internal phases of science. The internal phases would begin, on this view, only once already well-defined hypotheses are put to the test.

This potential reconciliation between Douglas and Alexandrova, put forward in section 6, preserves Douglas's approach to the proper role for values only if the determination of which concepts are the best ones to use in formulating hypotheses is not itself a job for the sciences' internal phases. However, in section 7, I use a case study from conservation biology to suggest that an evidential role for evaluative considerations reappears if one looks toward the right scientific disciplines. This section thus establishes one way in which evaluative considerations can play a "direct" role in the internal phases of at least some sciences. Section 8 concludes with some brief reflections on an implicit appeal to a criterion of demarcation between scientific and nonscientific research when discussing the proper role for values.

2. Values and evaluative claims

Before proceeding further, I clarify a terminological matter. The term *values* may bring to mind, for some readers, positively valorized concepts like "solidarity," "honesty," and "integrity." If this is how one thinks of values, the view that "values are not evidence" may seem particularly attractive. But this is not because of their evaluative aspect: It is no less troublesome to see how descriptive concepts—when untethered from the statements in which they are deployed—can serve as evidence for further descriptive statements. A claim like "all species are descended from a small number of common ancestors" is not justified by appealing to concepts like "homology" or "distribution" tout court. Descriptive concepts do not have the right grain to justify descriptive statements—but they can do this when deployed in further descriptive statements. So, instead of appealing to "homology" alone, one needs to appeal to the modified appearance of homologous structures under different conditions of use (the wing of the bat, the fin of the whale, the forelimb of the badger). Instead of appealing to "distribution" alone, one appeals to the pattern of distribution around the globe of anatomically similar species in geographically proximal conditions, while more dissimilar species are found more distantly, regardless of the local environmental conditions to which they are subjected. Likewise, consider a claim with obvious ethical content, perhaps a claim like "it is proportionate to use live animal models in assessing the efficacy of our experimental vaccine." It seems strained to me to say one can provide a cogent justification for such a claim by appeal to "values": Citing merely the value of human health, or the value of sentience in animals, typically does not have the right sort of grain to speak in any detailed way toward the focal claim. Instead, one might cite the urgency of the hoped-for intervention for large numbers of people; one might cite the minimal suffering one expects the animals to undergo.

In the discussion that unfolds throughout this article, it is therefore important to bear in mind that although I will often follow the standard practice of considering the role of "values," what I really have in mind is the role of claims, statements, or judgments with ethical evaluative content. These are at least worth discussing as potential candidates for evidential roles in the sciences. Douglas (2023, 57) has made a very similar point in recent work, where she addresses the potential justificatory roles of what she refers to as "normative claims," rather than values tout court. My stance is also consonant with earlier writing on this issue by Anderson (1995, 2002, 2004), Solomon (2012), Clough (2020), and especially, with far more detailed work,

Table 1. Douglas on the proper roles for values in science

	Direct role for values	Indirect role for values
Internal phases of science	**Illegitimate**	Legitimate
External phases of science	Legitimate	Legitimate

Brown (2020). He has likewise stressed that "there is no general problem of values, ordinarily so-called, standing as reasons for other claims. When we are engaged in ethical discourse, for example, values will justify other claims about values" (110). He, too, highlights the need to focus on considered value judgments when the question of values as evidence is under discussion.

3. Internal and external; direct and indirect

Douglas's overall position on the legitimate role for values draws on a pair of distinctions that she takes to be independent of each other (see table 1). The first concerns different phases of the scientific enterprise: These can be divided into internal and the external ones. The second concerns different roles for values: either direct or indirect. Her view is that three of the four possibilities to which these distinctions give rise for the use of values in science can be legitimate.

I will sketch the internal/external distinction only in a preliminary way here, in a manner intended to be true to Douglas's intentions. At the end of this section, I will suggest that there is no need to invoke it, even if one is seeking to articulate the role for values that Douglas wishes to defend. As I read her, the "external" phases of science involve a series of preparatory questions that need to be addressed prior to research getting under way. For example, a biomedical scientist may need to decide what topic to research, which colleagues to recruit to address it, whether it is important to engage in advance with individuals whom the research concerns (and if so, how), whether the research project justifies the use of live animals as experimental models, and so forth. The "external" phases also involve a series of questions about what to do with the results of research after it is complete. These might include whether to publish the results of a study at all, and if so, where; whether to share benefits of research in a particular way; how to present the work in ways that might engage policymakers; and so forth. The "internal" phases, to which Douglas also refers as the "heart" of science, consist of the gathering and analysis of evidence, the processes of experimentation, and so forth.

What Douglas means by the "direct" versus "indirect" distinction is also something that I will introduce in an intuitive way, before moving on to a more detailed analysis. When considerations play a "direct" role, they "determine our decisions in and of themselves, acting as stand-alone reasons to motivate our choices" (Douglas 2009, 96). A scientist's value-based commitment to aiding the Global South may be one reason for deciding to focus their research on the elimination of malaria, but it brings about that decision only in tandem with an empirical view about the prevalence of malaria in these regions. So I assume that Douglas does not mean by "stand-alone reasons," or by determination "in and of themselves," that values are sufficient to motivate choices alone and independent of additional empirical information. Instead, in the

"external" phases of science, values speak in a direct way to—that is, they constitute straightforward reasons for—decisions about (for example) whether to focus one's research on the elimination of malaria, whether to use live animals as study models, whether to seek consent from affected parties, how to share the benefits of research with stakeholders, and so forth. Many of the decisions with which scientists are presented in these external phases of science are of a plainly ethical nature, and so it is to be expected that values can be deployed directly when they are addressed.

Douglas claims that this "direct" role for values is inappropriate in the internal phases of science. The easiest way to understand why recalls the argument from inductive risk, which I put forward here in a simplified manner. (In this article, I neither endorse nor dispute the force of the inductive risk argument; the point is simply to see how it illuminates Douglas's intentions for the direct/indirect distinction.) Suppose a scientist is investigating coastal erosion and is charged with ascertaining whether a particular stretch of coastline will collapse into the sea in the next five years. The consequences of a false positive—accepting that the coastline will collapse, when in fact it will not—may be significant. Local residents may be needlessly relocated, tourism may be disrupted, and so forth. The consequences of a false negative—rejecting the hypothesis that the coastline will collapse, when in fact it will—may also be significant. Local residents, or visitors, may suffer considerable injury and loss of property. Assessing the weight of evidence gathered makes a difference to deciding whether to accept or reject P, but so does assessing the costs of errors. Hence decisions to accept or reject hypotheses turn, legitimately, on issues around the moral weight of (in the hypothetical case sketched) injury, damage to property, disruption to stable domestic life, and so forth. Douglas argues that reflection on the downstream consequences of errors should have far-reaching implications throughout all phases of science, including the internal ones. It should affect, among other things, a scientist's choice of methodology and of data sources. A choice of one method, or one data source, over another can be expected to lead to errors with different costs (Douglas 2016).

The morally adverse consequences of needless relocation say nothing at all about how *likely* it is that a stretch of coastline will collapse. As Douglas rightly insists, the role these values have with respect to this hypothesis is not to give evidential support. So, in her terminology, the inductive risk argument fails to give values a "direct" role. But values do have an indirect role with respect to the hypothesis under consideration, because of the specific way in which they influence acceptance of that hypothesis. The anticipated downstream ethical costs of accepting the hypothesis feed back into the decision about whether the hypothesis should be used as a basis for action.

For this reason (as Douglas [2023, 69] herself also mentions), it can also be legitimate to credit values with an indirect role when they are deployed in the "external" phases of scientific reasoning. If values play an indirect role because of the potential for downstream consequences of *empirical* error in the *internal* phases of science, they can also play the same indirect role because of the potential for downstream consequences of *normative* (and also empirical) error in the *external* phases of science. For example, a scientist might come to the view that it is most likely reasonable to use a particular model organism in the context of research on pain. Even so, they may be unsure of the ethical reasoning that has led them to this

position. They may therefore decide that the potentially grave consequences of making an error about this ethical matter justify taking a precautionary approach and forgoing the use of these model organisms just in case they are mistaken in their reasoning. Indeed, recent work has pointed to the possibility of taking a precautionary stance in our use of animals in case our moral reasoning is faulty (Birch 2017). Whether this is strictly *inductive* risk is an open question here, for it depends on whether moral epistemology proceeds via inductive argument. It may be better to think of this as an instance of more general "epistemic risk" (Biddle and Kukla 2017). The important point is that values can play an indirect role in our decision-making about the acceptance of evaluative claims, just as they can play an indirect role in our decision-making about the acceptance of empirical claims.

Kevin Elliott (2011, 2013) has pointed to several different possible readings of Douglas's detailed comments about the direct/indirect distinction. I suggest that progress can be made by focusing on the most significant thought underlying Douglas's claim that values should not play a "direct" role in the internal phases of science. This is her denial that values can legitimately serve in an evidential role with respect to scientific hypotheses. Recall her pithy claim that "values are not evidence; wishing does not make it so." In other words, says Douglas, the question of whether scientific hypotheses are likely to be true can never be legitimately informed by values.

This reading of Douglas is bolstered by much more recent work. In collaboration with Elliott, Douglas jointly writes, "Douglas has been very careful to distinguish the role of values from the role of evidence. The reason she is so emphatic about her distinction between direct and indirect roles for values is that she does *not* think values serve as evidence" (Douglas and Elliott 2022, 204). Elliott himself suggests that values could play "something like an evidential role" when the question being asked is whether a scientific hypothesis is adequate for some purpose; for example, he points out that if a risk-assessment method were to prioritize public health, this could act as a value-based reason for thinking that it is "the best method for achieving particular regulatory purposes" (205). Even so, his underlying stance remains close to Douglas's: If "scientists are considering whether or not to believe that the hypothesis is true, ethical values are logically irrelevant and have no legitimate role to play because belief is directed only toward achieving truth" (Elliott and Willmes 2013, 812).

Douglas justifies a blanket prohibition on using values as evidence by appealing to the pathological nature of wishful thinking: "As Hempel rightly pointed out, value judgements have no direct place in the argument for what should be taken as true" (Douglas 2000, 564; see also Brown 2013; Elliott 2013; Miller 2014). More recently, she has claimed that "values in a direct role during evidential assessment would be equivalent to allowing wishful thinking into the heart of science. If values could play a direct role in the assessment of evidence, a preference for a particular outcome could act as a reason for that outcome or for the rejection of a disliked outcome" (Douglas 2016, 619). So it is this principle—namely, that values cannot serve as evidence for scientific claims—on which I will focus for the bulk of this article.

This principle captures what I take to be the core of Douglas's view about the appropriate role for values, in a way that avoids having to rely on a clear-cut distinction between "internal" and "external" phases of science. This is just as well, because the internal/external distinction is slippery. As I read her, Douglas prohibits a

"direct" (i.e., evidential) role for values in the internal phases of the sciences because she conceives of "internal" phases (but not external ones) as those in which one tries to assess to what degree a hypothesis is supported by evidence. These phases are characterized by the analysis of newly gathered data, the conduct of experiments, perhaps the marshaling of existing pieces of work to provide an overall case for a new scientific claim. This suggests a correlative conception of the "external" phases of the sciences as those in which one asks very different sorts of questions about what topics to prioritize for research, what methods might be ethically acceptable to use, and so forth.

This means that, despite the convenient phrasing I reached for when first articulating Douglas's position, "internal" and "external" phases cannot be characterized in a simple, temporal way: The "internal" phases are not the elements of science that happen in the temporal period between "external" preparation and "external" dissemination. Even prior to getting into the lab, or going out to the field, the scientist needs to develop some set of views about what is worth investigating in the first place. This will be partially informed by their preliminary assessment of what is likely to be true. That evidential assessment, in turn, will be informed by a creative reading of cogent literature, by thought experiments, by casual observation around the lab and in the field. That informal evidential assessment helps to tell them what is so unlikely to be the case that it would be idle to spend time trying to test or establish it more formally. And even when deciding what research methods to use—whether, for example, to use live animals in a study—the scientist needs to take a decision on whether a hypothesis is likely enough to be worth testing in a way that might make demands on animals' well-being. In short, evidential considerations have a bearing even as research is being prepared.

Conversely, even when data are being analyzed, even when experiments are being conducted, and even when hypotheses are being accepted, explicitly evaluative questions remain in play. A scientist who has embarked on an experiment using rodents may face real-time decisions in the lab about whether a hitch in the procedure means the animals need to be destroyed on welfare grounds. And, of course, Douglas's own inductive risk argument draws attention to the ways in which ethical and evidential questions are also implicated in a series of methodological choices during the analysis of data and the interpretation of experimental results.

The claim that it is only in the "external" phases of science where one sees a legitimate "direct" role for values can thus be transformed into a claim that I take to be equivalent, but which eschews the troubling internal/external distinction. This is the claim that while sometimes science deals in explicitly ethically loaded claims—about, for example, which topics are the most urgent targets for study or about the need to share the benefits of research—the hypotheses scientists aim to establish are entirely descriptive in nature. As a result, it is only in the former context, and never in the latter, that the claims science deals with are apt for evidential support from evaluative claims. In the following two sections, I show a prima facie tension between this view and two other widely discussed claims within the science and values literature. The first concerns the disciplined nature of ethical cases; the second concerns the allegedly "mixed" character of at least some scientific hypotheses.

4. Wishful thinking about wishful thinking

In rejecting a direct role for values within the "heart" of science, Douglas (2000, 578) cites Hempel with approval: "As Hempel pointed out, the argument 'I want X to be true, therefore X is true' remains a bad argument, both within and without science." As she puts it elsewhere, and as I've already noted, she claims that "values in a direct role during evidential assessment would be equivalent to allowing wishful thinking into the heart of science" (Douglas 2016, 619). But even if wishful thinking is always epistemically inappropriate (for caveats, see Peters 2021), this would not mean that a direct role for evaluative claims is pathological. An ethical evaluative case would be defective if it had the simple structure "I want rats to be used in the lab; therefore the use of rats is justified." In practice, ethical evaluative cases are presented in much richer ways that obey principles of cogency. These points are all neutral regarding deeper matters in meta-ethics: A sophisticated expressivist like Blackburn (1998), for example, does not think that sensitive ethical case building has the simple structure of wishful thinking.

As Anderson (1995, 2002, 2004) has stressed in systematic detail, the means by which an evaluative conclusion is supported can have discipline (see also Sober 2007; Clough 2020; Brown 2020). Consider, once again, the claim that it is justifiable to use animals in a given pharmacological study. It is relevant to appeal to the use to which the research will eventually be put, to the suffering of the animals, and so forth. It is irrelevant to appeal to the cuteness of rodents or the need to give researchers something that will alleviate their boredom. Just as an X-ray crystallography image constitutes (inconclusive, partial) support for a claim about the structure of DNA, so a claim about the weighty benefits of a vaccine constitutes (inconclusive, partial) support for the legitimate use of rats in a laboratory safety study.

This section establishes merely that cogent ethical cases have a form of discipline; they are not instances of wishful thinking. By itself, this leaves Douglas's view intact, because even if "wishful thinking" is not itself a problem affecting all invocations of values in an evidential role, one might think it obvious that evaluative claims are never of relevance in providing support to the sorts of hypotheses entertained in the sciences. Douglas (2023, 57–58) herself has recently phrased her argument not in terms of the impermissibility of evaluative claims as evidence but in terms of the impermissibility of evaluative claims as evidence *for descriptive claims*: "When Carl Hempel argued back in the 1960s that normative claims can provide no confirmatory weight to a descriptive claim, he was right." This way of putting things rules out an evidential role for evaluative claims in the sciences only if scientific hypotheses are purely descriptive. As de Melo-Martín and Intemann (2016) have also suggested, in this respect Douglas's position rests on a specific version of the value-free hypothesis. The question for the next section is whether this form of value freedom can be sustained in the light of work by Alexandrova.

5. Mixed hypotheses

Alexandrova (2018) has highlighted the category of what she calls "mixed" claims in the human, social and medical sciences. Her case has not convinced all commentators (e.g., Peters 2023), but my own presentation here does not require that her position be accepted in full. One of her examples is "Early learning difficulties have a

disproportionate impact on life well-being" (Alexandrova 2018, 423, summarising Beddington et al. 2008). On the face of things, this claim implicates value judgments in the notions of "difficulties" (for one might think what is "difficult" is indexed to a class of suitably smooth learning interactions), of what is "disproportionate" (for that also presupposes a standard of proportionality), and of well-being itself (which involves fixing on a true, or appropriate, measure of how well a life goes).

Her claim that such statements are "mixed" is neutral regarding deeper matters in meta-ethics: It does not entail that all scientific hypotheses are inevitably mixed, nor does it involve a general denial of a distinction between fact and value. Alexandrova does not deny that one might, in principle, separate the factual from the evaluative. Her view, instead, is that nothing of practical value would be gained from attempting to strip the evaluative content from scientific hypotheses, even if such a thing were possible. In her view, it is better to preserve scientific fields that aim to deal in a rigorous way with providing support to value-laden claims.

Putting the thoughts of Anderson and Alexandrova together, if scientific hypotheses have mixed content, then (on the face of things) there is a legitimate evidential role for disciplined evaluative reasoning in establishing those mixed hypotheses. Consider one of Alexandrova's examples: "Happiness is not always conducive to well-being." This involves a mixture of the descriptive and the evaluative because (1) there is a choice over which conception of well-being is the most plausible that (2) is informed by the goal of understanding (for example) which emotions are good for us and which might harm us. This points to an evidential role for evaluative claims in the *internal* phases of those sciences that deal with mixed hypotheses: The evidential support given to the claim "Happiness is not always conducive to well-being" needs to draw on ethically loaded propositions about the right conception of well-being to use.

Let me recap: There is a prima facie tension between Alexandrova's argument for the legitimacy of mixed hypotheses and Douglas's prohibition on a direct role for values in the internal phases of the sciences. For if some sciences regularly deal in mixed hypotheses, and if mixed hypotheses require mixed support, then it should be easy to find examples when good evaluative reasoning contributes in an evidential way to the support of scientific conclusions. In the next section, I explore a potential line of response from Douglas that draws on the detail of how Alexandrova characterizes "mixed" hypotheses.

6. Reconciling Douglas and Alexandrova: Mixed hypotheses without mixed content

Alexandrova says this about what she takes a "mixed" hypothesis to be:

> A hypothesis is mixed if and only if:
>
> 1. It is an empirical hypothesis about a putative causal or statistical relation.
>
> 2. At least one of the variables in this hypothesis is defined in a way that presupposes a moral, prudential, political, or aesthetic value judgement about the nature of this variable. (Alexandrova 2018, 424)

As I will explain in a moment, value judgments of the relevant sort may be "presupposed" in the way that the variables within a hypothesis are defined, even when the variables themselves are not given value-laden content. The same can be said of Anderson's definition of a "thick" concept, on which Alexandrova draws in her work on science and values. Anderson (2002, 504–5) writes,

> A concept is thickly evaluative if (a) its application is guided by empirical facts; (b) it licenses normative inferences; and (c) interests and values guide the extension of the concept (that is, what unifies items falling under the concept is the relation they bear to some common or analogous interests or values).

Consider how these definitions of "mixed" hypotheses or "thick" concepts play out in a public health context. Public Health England (PHE) initially recorded any death following a positive COVID-19 test as a COVID-19 death. It later moved to daily reporting of those people who died within twenty-eight days of a positive COVID-19 test (Griffin 2020). One might think it a purely factual matter as to whether a death follows a positive COVID-19 test, or whether it comes within 28 days of a positive COVID-19 test. But the decision to define COVID-19 deaths in these various ways was informed by normative concerns: John Newton, PHE's director of health improvement, said at the time, "The way we count deaths in people with Covid-19 in England was originally chosen to avoid underestimating deaths caused by the virus in the early stages of the pandemic" (quoted in Griffin 2020, 1). In other words (and in ways that are reminiscent of the argument from inductive risk, where it is the anticipated downstream consequences of a scientific decision that feed back in an *indirect* way on the choice to make that decision), the definition of a COVID-19 death was intended to encourage overcounting, rather than undercounting. This was on the grounds that the former would be preferable in terms of its consequences for health policy.

This approach to defining a COVID-19 death appears to satisfy Alexandrova's and Anderson's definitions of "mixed hypotheses" and "thick concepts," respectively. In Alexandrova's (2018, 424) case, the way in which the concept is defined "presupposes a moral, prudential, political, or aesthetic value judgement about the nature of this variable": The choice to define the concept in a highly permissive manner reflects the prudential judgment that overcounting is better than undercounting. Likewise, in Anderson's case, it seems clear that the concept's application is guided by empirical facts (for example, by facts about COVID-19 diagnoses), that it licenses normative inferences (for example, about taking preventive measures), and that interests and values guide the extension of the concept (in the sense that the safeguarding of public health is what determines the decision to draw a permissive extension for the concept). Many other examples could be given: Take the example to define obesity in terms of body mass index (BMI) greater than 30. In one sense, the BMI concept has no normative content: It is nothing more than a calculation of weight in kilograms divided by height in meters squared. But the motivation for defining obesity in this (entirely descriptive) way is itself of a normative sort. As the WHO Expert Consultation Group (2004, 161) puts it, "The purpose of a BMI cut-off point is to identify, within each population, the proportion of people with a high risk of an undesirable health state that warrants a public health or clinical intervention."

This means that "thick" concepts in Anderson's sense or "mixed" hypotheses in Alexandrova's sense need not correspond to "thick" claims as that term is used in the bulk of the meta-ethics literature. There it is generally understood that thick concepts do carry normative content. Exactly how thick concepts should be understood in that latter domain is up for grabs, but paradigmatic examples tend to include concepts like "courageous" and "generous." They are similar to concepts like "good" and "right," in that they have positive evaluative content; however, they are dissimilar to "good" and "right" in that they can be properly applied only to people or actions meeting a relevant empirical description. As a recent overview puts it, "It is remarkably difficult to find a suitably neutral positive characterization of what thick concepts are, beyond the idea that they somehow combine evaluation and non-evaluative description" (Väyrynen 2021). But a hypothesis can be "mixed" in Alexandrova's sense even if its content is entirely nonevaluative, just so long as the background reasons for defining the terms within it in one nonevaluative way, rather than another, are themselves of an evaluative character. This was recognized, although not stressed, in a rather different context by John Dupré (2007, 28) when he noted that "sometimes we try to lay down rather precise criteria for applying interest-relative terminology to things . . . . In such cases, we might be tempted to say that the precision of the criteria converts an evaluative term to a descriptive one. It is important to notice, however, that the precision is given point by the interest in evaluation." It is the "interest in evaluation" associated with a concept that triggers Alexandrova's "mixed" hypotheses or Anderson's "thick" concepts, even if the term so defined is strictly free of evaluative content.

In this respect, Alexandrova's approach differs from a slightly earlier approach to similar issues by Brown (2021), who draws on an understanding of "thick" concepts that is much closer to the way that term is used in meta-ethics. As he puts it, "Concepts like 'gender,' 'equality,' 'poverty' and 'well-being' all have an inextricable mix of descriptive and evaluative content, and the attempt to regiment concepts into purely descriptive and normative ones distorts what such concepts mean" (55).

This all means that the case required to support a mixed hypothesis (sensu Alexandrova) need not itself be of an ethical evaluative character. The claim "There were thirty Covid deaths in a one-month period" does not itself require ethical support once the terms have been defined, even if the choice to define what a Covid death is in one way, rather than another, does require such support. And this, in turn, indicates a potential way for Douglas to hold on to the prohibition on an evidential role for evaluative claims in the internal phases of science. She might argue that, on one hand, evaluative claims do not give support to scientific hypotheses because scientific hypotheses do not have normative content. On the other hand, sciences do trade in "mixed" hypotheses (sensu Alexandrova) or hypotheses containing "thick" concepts (sensu Anderson), because the prior choice to provide descriptive content to hypotheses in one way, rather than another, is guided by ethical evaluative considerations.

7. Evaluative claims can play an evidential role within some sciences

The response sketched in the preceding section is not sufficient to defend the prohibition on values playing a direct role in the internal phases of the sciences. That

is because scientific effort itself is sometimes focused on building a case for defining a concept in one way rather than another. And though this scientific work may be "external" to the conduct of research in other domains that make use of these concepts, it has its own "internal" phases where the appeal to values needs to be (in Douglas's parlance) a "direct" one.

In this section, I make a case for this claim by focusing on the relationship between biological taxonomy and values (see also Ludwig 2015). It is uncontroversial that taxonomic decisions have downstream impacts on, for example, the allocation of resources for research and conservation efforts. A well-trodden instance of this phenomenon can be found in discussion surrounding the reclassification of the African elephants of the genus *Loxodonta*. For many years, taxonomists produced evidence indicating that two distinct extant species should be recognized within the genus: the forest elephant (*L. cyclotis*) and the savanna elephant (*L. africana*). In addition to data indicating genomic divergence between these groups of a similar magnitude to divergence between mammoths and Asian elephants (Rohland et al. 2010), there are differences in height, weight, time between pregnancies, and other traits, as well as limited hybridization between individuals of the two populations (Roca et al. 2001). However, it was only in 2021 that the International Union for the Conservation of Nature (IUCN) recognized these two distinct species for the purpose of its "Red List" assessments: The result was that, instead of African elephants being treated as a single species with the "vulnerable" listing, there were two species, listed as "critically endangered" (*L. cyclotis*) and "endangered" (*L. africana*). So splitting a single species into two created a sense of greater conservation urgency because of the concomitant change in Red List status: Amy Fraenkel from the Convention on the Conservation of Migratory Species of Wild Animals remarked that her organization welcomed "IUCN's recognition of two distinct African elephant species, and hope that it will lead to greater conservation actions for both species" (International Union for the Conservation of Nature 2021). Kathleen Gobush, who led the IUCN's African Elephant Specialist Group, also commented to *The Guardian* that the "reclassification allows detailed attention to each animal—the forest elephant and the savanna elephant—and then to tailor conservation plans according to each species' needs, which are different" (Greenfield 2021).

These remarks are all comparatively uncontroversial, because they simply point to the effects (which the scientists in question regarded as positive) of the new taxonomy. By themselves, they say nothing one way or another about whether the anticipation of such effects may also have influenced the IUCN's taxonomic decision in a manner that might indicate a role for evaluative considerations. Of course, the question of which species are recognized depends not just on facts about genomic divergence, morphology, hybridization, and so forth; it also depends on which species concept is deployed. This matter has caused well-documented controversy. I will not attempt to summarize in any detail here: I will note only that numerous options have been defended over the years. They include various "phenetic" approaches, which make species membership a matter of overall similarity; the "biological species concept" (BSC), which makes species membership a matter of which organisms can breed successfully with each other; and various "phylogenetic" species concepts, which consider species as twigs on the tree of life (Ereshefsky 2007). It is not important in this article to give a full discussion of these options or even of the nature

of the relations between them. But I do want to draw attention to a scientific paper in which Frankham et al. (2012) explicitly asked which species concept was best suited to the goals of conservation biology. Here I closely follow Conix's (2019) treatment of science and values, whereby he also examines this paper.

Frankham et al. (2012, 26) set themselves the task of evaluating "methods for defining species from the perspective of conservation biology, *advocating that definitions used in conservation biology should maximise conservation benefits*" (emphasis added). In other words, their goal was to build a case in favor of a specific species concept, and the case made in favor of that concept—that is, the evidential support that contributed to their overall conclusion—drew explicitly on evaluative considerations regarding the needs of conservation biology. They "recommend use of concepts that are most beneficial for conserving global biodiversity" (26). The type of case built is precisely the sort of "mixed" one to which Alexandrova draws attention, and it takes the sort of disciplined format that Anderson stresses.

Frankham et al. (2012) make a case against use of the diagnostic phylogenetic species concept in the context of conservation science. To understand what this species concept is, recall that a phylogenetic species concept is one that regards a species as a chunk of the evolutionary genealogical nexus, a "twig" on the tree of life. This leaves room for maneuver when one asks exactly which chunks will count as species: A family unit is a genealogical unit, and so is the entirety of the earth's biota, but the former is too narrow to be thought of as a species, and the latter is too broad. In putting forward the diagnostic phylogenetic species concept, Cracraft (1989, 34–35) proposed that "a phylogenetic species is an irreducible (basal) cluster of organisms, diagnosably distinct from other such clusters, and within which there is a parental pattern of ancestry and descent." Whether successful or not, the appeal to irreducibility is intended to exclude overly broad units from species status, and the appeals to (1) "diagnosable" distinctness and (2) patterns of ancestry and descent are intended to exclude overly narrow units.

Equipped with this rough understanding of background issues in taxonomy, it is now possible to understand the case made by Frankham et al. (2012). First, they note that human population growth, and its consequent effects on other species' habitats, leads to fragmentation of various populations. Fragmentation puts these populations under threat: Their small size in itself is a risk, and it also increases the chances of inbreeding depression as close genetic relatives become more likely to breed with each other. Fragmentation also makes it more likely, because these populations are small, that they will be affected by evolutionary processes in idiosyncratic ways. Frankham et al. note that genetic drift (which is more likely to occur in small populations) can lead to the accumulation of idiosyncratic differences from one population to the next. These differences may render the isolated populations diagnosable, even though individuals within one population would be able to breed with individuals in other isolated populations. So an isolated population may end up counting as a distinct species from the perspective of the diagnostic phylogenetic species concept, but not according to the BSC. That is because the BSC assigns two individuals to the same species whenever they can potentially breed with each other.

The question of which species concept is used becomes salient because of the twin threats of inbreeding depression and outbreeding depression. As already noted, inbreeding depression occurs when close genetic relatives mate. Outbreeding

depression occurs when very distant relatives that are adapted to different environments mate. Offspring can then be worse adapted than either parent. (Imagine, for example, that one parent is adapted to hunting deer, whereas the other is adapted to hunting mice; the poor offspring then end up adapted to hunting nonexistent prey of an intermediate size, and they starve as a result.) A species concept that encourages splitting will tend to exacerbate the problem of inbreeding depression. If an isolated population is regarded as a distinct species, then this taxonomic outcome will discourage the possibility of conservation biologists introducing measures to facilitate mating with less closely related organisms from different isolated groups, because those will be understood as members of a different species. On the other hand, species concepts that encourage lumping may also exacerbate the problem of outbreeding depression. Extreme lumping can end up encouraging the attempted rescue of isolated populations by introducing measures that enable mating with very distantly related members of what will then (given the lumping approach) be thought of as the same species, bringing concomitant problems of low hybrid fitness. In short, Frankham et al. make a case that the BSC is the better one for conservation biologists to use because of the ways in which it strikes a suitable middle way between the threats of inbreeding and outbreeding depression.

Their appeal to the interests of conservation science forms a direct part of their case—just as direct as the normative appeal one might make to the urgency of health problems in a case for the use of animals in research. That case is also informed by empirically and theoretically disciplined reflection on the ways in which habitat fragmentation can lead to both the diagnosability and the fragility of specific populations. Their case is potentially revisable in the face of countervailing concerns about the weaknesses for conservation science of the BSC itself, as well as in the face of counterevidence about the nature of diagnosability and the consequences of species delimitation.

It is not my aim here to give a full defense of the case Frankham et al. (2012) make. Indeed, it is important for my argument that their case is potentially vulnerable to cogent criticism. But it does seem to me that they cannot be accused of simple "wishful thinking," despite the appeal they make to evaluative considerations. Theirs is just the sort of disciplined, undogmatic, direct appeal to values to which the earlier discussion of Anderson's work drew attention. One might question whether the sort of work Frankham et al. are doing is really "internal" to the sciences. But it seems hard to deny this without introducing a series of potentially problematic assumptions about when research proper begins and when the mere validation of appropriate concepts ends. Frankham et al.'s work constitutes a case where evaluative claims form part of the evidential case for a scientific hypothesis—in this case the hypothesis that the diagnostic phylogenetic species concept is the wrong one for conservation biology—and in that sense, their work shows how values can play a direct role in the heart of this science, even if not in the heart of all sciences.

8. Evidence and demarcation

The preceding sections have shown one way in which evaluative considerations can legitimately act in an evidential role with respect to scientific hypotheses. I have tried to present this work in a manner that is neutral regarding questions in meta-ethics

about the general nature of moral judgment and putative moral facts. In the case I have been considering, descriptive and normative considerations combine to support a claim about the most appropriate definition of a concept for a scientific field. This case meets the conditions laid down by Solomon (2012); it does not show that worthy values in general always constitute relevant evidence for or against scientific hypotheses, nor does it show that all scientific hypotheses can be cogently supported by evaluative claims of some well-justified kind. But under some circumstances, evaluative claims can have evidential relevance when they are brought into contact with the right scientific hypotheses.

Can one draw any general lessons from all this for when evaluative considerations are admissible in the sciences? I have rejected one proposal: It is a mistake to say that values are never admissible when they are put forward as evidence for scientific hypotheses. The hypotheses entertained by the sciences are sometimes apt to receive support from evaluative claims. One might wonder whether the argument presented in this article licenses a slightly modified conclusion, namely, that evaluative claims can provide evidential support only when scientific hypotheses are of a "mixed" sort. In my view, this would also be too hasty. The only general claim that can be made with confidence is the trivial one that so long as evaluative claims are evidentially relevant to scientific hypotheses, they can be legitimately put forward in support of those hypotheses. Work from Longino (1990), among others, has stressed the highly contextual nature of evidential support: Whether one claim can provide evidence for another depends not only on the precise content of both claims but also on background propositions that render one relevant to the other. For these reasons, the articulation of anything like a domain-general set of guidelines telling scientists when evaluative claims can be used as evidence is likely to be exceptionally difficult: There is no reason to expect it to be any easier than giving domain-general guidelines for when one empirical claim can provide evidence for another. The argument of this article leaves open the issue of whether values might be able to provide evidential support to empirical claims that are not "mixed."

There is a line of response that alleges the triviality of the conclusions of this article. It is obvious that in some domains of knowledge production, the evidential case brought in support of key hypotheses draws, in part, on normative claims. This is clearly the case in domains such as empirical bioethics, where empirical research methods from the social sciences combine with explicitly normative reasoning to support propositions about, for example, the most appropriate ways to seek consent in a given medical situation, or the best way to empower patients through the presentation of information. In these domains, evaluative claims obviously have a "direct" role with respect to hypotheses: They provide evidence for them, and not (when the work is done well) in a manner that involves wishful thinking. One might then argue that on reflection, it should be no surprise that investigators working in some other domains of research—such as conservation biology—find themselves in positions where the crafting of suitable concepts can legitimately appeal to a mixture of empirical and normative considerations. One might then go on to say that empirical bioethics is no counterexample to Douglas's prohibition on a direct role for values in the sciences, because it is not a science. More generally, one might argue that the sciences properly construed are only ever engaged in supporting the sorts of hypotheses that require purely empirical evidence. Regardless of the merits of

Frankham et al.'s (2012) paper, this response says that their article is not a piece of science but instead an instance of some other kind of respectable epistemic endeavor.

From a certain perspective, the results of this article really are obvious: If one begins by thinking about domains of investigation where the evidential role of normative considerations is self-evident, then it becomes much easier to detect a similar role across a broad spectrum of cognate fields of research. But it is harder to build a plausible argument that denies these fields the status of science. Frankham et al.'s (2012) paper is written by salaried scientists, and published in a scientific journal, so there is at least a prima facie case that their work is a piece of scientific research. One cannot argue on merely definitional grounds that a field of knowledge counts as a science only if it deals in purely empirical forms of evidence, for that would trivialize the claim that there is no direct role for values in the internal phases of the sciences. It would not rule out the thought that there are plenty of ways for values to legitimately act as evidence in support of a whole variety of claims, but those claims would simply be denied the label of "science." Moreover, standards of cogency discipline the ways in which normative considerations can legitimately be marshaled in support of hypotheses. There is, then, no easy route from a claim about the epistemic authority of disciplines worthy of the name "science" to the nonscientific status of fields in whose "internal" stages values play a "direct" role. This article has aimed to show that there are such fields.

Acknowledgments. I am grateful to audiences in Bordeaux, Cambridge, and Paris for feedback on this work and especially to Anna Alexandrova, Ahmad Elabbar, Philippe Huneman, Stephen John, Maël Lemoine, Thomas Pradeu, Elliott Sober, Miriam Solomon, Marion Vorms, Joeri Witteveen, and two anonymous referees from this journal for comments. I am also grateful to Stijn Conix for teaching me about taxonomy and values.

References

Alexandrova, Anna. 2018. "Can the Science of Well-Being Be Objective?" *British Journal for the Philosophy of Science* 69 (2):421–45. https://doi.org/10.1093/bjps/axw027.

Anderson, Elizabeth. 1995. "Knowledge, Human Interests, and Objectivity in Feminist Epistemology." *Philosophical Topics* 23 (2):27–58. https://doi.org/10.5840/philtopics199523213.

Anderson, Elizabeth. 2002. "Situated Knowledge and the Interplay of Value Judgments and Evidence in Scientific Inquiry." In *In the Scope of Logic, Methodology and Philosophy of Science*, vol. 2, edited by Peter Gardenfors, Katarzyna Kijania-Placek, and Jan Wolenski, 497–517. Springer. https://doi.org/10.1007/978-94-017-0475-5_8.

Anderson, Elizabeth. 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. https://doi.org/10.1111/j.1527-2001.2004.tb01266.x.

Beddington, John, Cary L. Cooper, John Field, Usha Goswami, Felician Huppert, Rachel Jenkins, Hannah S. Jones, Tom B. L. Kirkwood, Barbara J. Sahakian, and Sandy M. Thomas. 2008. "The Mental Wealth of Nations." *Nature* 455:1057–60. https://doi.org/10.1038/4551057a.

Biddle, Justin, and Rebecca Kukla. 2017. "The Geography of Epistemic Risk." In *Exploring Inductive Risk*, edited by Kevin Elliott and Ted Richards. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190467715.003.0011.

Birch, Jonathan. 2017. "Animal Sentience and the Precautionary Principle." *Animal Sentience* 16 (1). https://doi.org/10.51291/2377-7478.1200.

Blackburn, Simon. 1998. *Ruling Passions*. Oxford University Press. https://doi.org/10.1093/oso/9780198247852.001.0001.

Brown, Matthew. 2013. "Values in Science Beyond Underdetermination and Inductive Risk." *Philosophy of Science* 80 (5):829–39. https://doi.org/10.1086/673720.

Brown, Matthew. 2020. *Science and Moral Imagination: A New Ideal for Values in Science*. University of Pittsburgh Press. https://doi.org/10.2307/j.ctv18b5d19.

Brown, Matthew. 2021. "The Descriptive, the Normative, and the Entanglement of Values in Science." In *The Rightful Place of Science: Science, Values and Democracy*, edited by Ted Richards. Consortium for Science, Policy, and Outcomes.

Clough, Sharon. 2020. "Using Values as Evidence When There's Evidence for Your Values." *Philosophy in the Contemporary World* 26 (1/2):5–37. https://doi.org/10.5840/pcw2020261/21.

Conix, Stijn. 2019. "Taxonomy and Conservation Science: Interdependent and Value-Laden." *History and Philosophy of the Life Sciences* 41:15. https://doi.org/10.1007/s40656-019-0252-3.

Cracraft, Joel. 1989. "Speciation and Its Ontology." In *Speciation and Its Consequences*, edited by Daniel Otte and John Endler. Sinauer.

de Melo-Martín, Immaculada, and Kristen Intemann. 2016. "The Risk of Using Inductive Risk to Challenge the Value-Free Ideal." *Philosophy of Science* 83 (4):500–520. https://doi.org/10.1086/687259.

Douglas, Heather. 2000. "Inductive Risk and Values in Science." *Philosophy of Science* 67 (4):559–79. https://doi.org/10.1086/392855.

Douglas, Heather. 2009. *Science, Policy and the Value-Free Ideal*. University of Pittsburgh Press. https://doi.org/10.2307/j.ctt6wrc78.

Douglas, Heather. 2016. "Values in Science." In *The Oxford Handbook of Philosophy of Science*, edited by Paul Humphreys. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199368815.013.28.

Douglas, Heather. 2018. "From Tapestry to Loom: Broadening the Perspective on Values in Science." *Philosophy, Theory, and Practice in Biology* 10:8. https://doi.org/10.3998/ptpbio.16039257.0010.008.

Douglas, Heather. 2023. "Science and Values: The Pervasive Entanglement." In *Democratizing Risk Governance*, edited by Monica Gattinger. Springer. https://doi.org/10.1007/978-3-031-24271-7_3.

Douglas, Heather, and Kevin Elliott. 2022. "Addressing the Reproducibility Crisis: A Response to Hudson." *Journal for General Philosophy of Science* 53:201–9. https://doi.org/10.1007/s10838-022-09606-5.

Dupré, John. 2007. "Fact and Value." In *Value-Free Science? Ideals and Illusions*, edited by Harold Kincaid, Alison Wylie, and John Dupré. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195308969.003.0003.

Elliott, Kevin. 2011. "Direct and Indirect Roles for Values in Science." *Philosophy of Science* 78 (2):303–24. https://doi.org/10.1086/659222.

Elliott, Kevin. 2013. "Douglas on Values: From Indirect Roles to Multiple Goals." *Studies in History and Philosophy of Science* 44 (3):375–83. https://doi.org/10.1016/j.shpsa.2013.06.003.

Elliott, Kevin, and David Willmes. 2013. "Cognitive Attitudes and Values in Science." *Philosophy of Science* 80 (5):807–17. https://doi.org/10.1086/673719.

Ereshefsky, Marc. 2007. "Systematics and taxonomy." In *The Blackwell Companion to the Philosophy of Biology*, edited by Sahorta Sarkar and Anya Plutynski, 99–118. Oxford: Blackwell.

Frankham, Richard, Jonathan D. Ballou, Michele R. Dudash, Mark D. B. Eldridge, Charles B. Fenster, Robert C. Lacy, Joseph R. Mendelson III, Ingrid J. Porton, Katherine Ralls, and Oliver A. Ryder. 2012. "Implication of Different Species Concepts for Conserving Biodiversity." *Biological Conservation* 153:25–31. https://doi.org/10.1016/j.biocon.2012.04.034.

Greenfield, Patrick. 2021. "African Elephant Recognised as Two Separate Species." *The Guardian*, March 25. https://www.theguardian.com/environment/2021/mar/25/african-elephants-now-red-list-two-species-both-nearer-extinction.

Griffin, Shaun. 2020. "COVID-19: England Comes into Line with Rest of the UK on Recording Deaths." *BMJ* 370:m3220. https://doi.org/10.1136/bmj.m3220.

International Union for the Conservation of Nature. 2021. "African Elephant Species Now Endangered and Critically Endangered—IUCN Red List." Press release, March 25. https://iucn.org/news/species/202103/african-elephant-species-now-endangered-and-critically-endangered-iucn-red-list.

Longino, Helen. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton University Press. https://doi.org/10.1515/9780691209753.

Ludwig, David. 2015. "Ontological Choices and the Value-Free Ideal." *Erkenntnis* 81:1253–72. https://doi.org/10.1007/s10670-015-9793-3.

Miller, Boaz. 2014. "Catching the WAVE: The Weight-Adjusting Account of Values and Evidence." *Studies in History and Philosophy of Science* 47:69–80. https://doi.org/10.1016/j.shpsa.2014.02.007.

Peters, Uwe. 2021. "Illegitimate Values, Confirmation Bias, and Mandevillian Cognition in Science." *British Journal for the Philosophy of Science* 72 (4):1061–81. https://doi.org/10.1093/bjps/axy079.

Peters, Uwe. 2023. "Values in Science: Assessing the Case for Mixed Claims." *Inquiry* 66 (6):965–76. https://doi.org/10.1080/0020174X.2020.1712235.

Roca, Alfred, Nicholas Georgiadis, Jill Pecon-Slattery, and Stephen O'Brien. 2001. "Genetic Evidence for Two Species of Elephant in Africa" *Science* 293 (5534):1473–77. https://doi.org/10.1126/science.1059936.

Rohland, Nadin, David Reich, Swapan Mallick, Matthias Meyer, Richard E. Green, Nicholas Georgiadis, Alfred L. Roca, and Michael Hofreiter. 2010. "Genomic DNA Sequences from Mastodon and Woolly Mammoth Reveal Deep Speciation of Forest and Savanna Elephants." *PLOS Biology* 8 (12):e1000564. https://doi.org/10.1371/journal.pbio.1000564.

Sober, Elliott. 2007. "Evidence and Value Freedom." In *Value-Free Science: Ideal or Illusion?*, edited by Harold Kincaid, Alison Wylie, and John Dupré. Oxford University Press.

Solomon, Miriam. 2012. "The Web of Valief." In *Out from the Shadows: Analytical Feminist Contributions to Traditional Philosophy*, edited by Sharon Crasnow and Anita Superson. Oxford University Press.

Väyrynen, Pekka. 2021. "Thick Ethical Concepts." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. https://plato.stanford.edu/archives/spr2021/entries/thick-ethical-concepts/.

WHO Expert Consultation. 2004. "Appropriate Body-Mass Index for Asian Populations and Its Implications for Policy and Intervention Strategies." *The Lancet* 363 (9403): 157–63. https://doi.org/10.1016/S0140-6736(03)15268-3.

Cite this article: Lewens, Tim. 2025. "Towards a Direct Role for Values in the Heart of Science." *Philosophy of Science*. https://doi.org/10.1017/psa.2025.18