**A Rawlsian Solution to the New Demarcation Problem**

Frank Cabrera

University of Massachusetts, Lowell

**Abstract:** In the last two decades, a robust consensus has emerged among philosophers of science, according to which political, ethical, or social values *must* play some role in scientific inquiry, and that the “value-free ideal” is thus a misguided conception of science. However, the question of how to distinguish, in a principled way *which* values may legitimately influence science remains. This question, which has been dubbed the “new demarcation problem”, has until recently received comparatively less attention from philosophers of science. In this paper, I appeal to Rawls’ theory of justice (1971) on the basis of which I defend a Rawlsian solution to the new demarcation problem. As I argue, the Rawlsian solution places plausible constraints on which values ought to influence scientific inquiry, and moreover, can be fruitfully applied to concrete cases to determine how the conflicting interests of stakeholders should be balanced. After considering and responding to the objection that Rawls’ theory of justice applies only to the “basic structure” of society, I compare the Rawlsian solution to some other approaches to the new demarcation problem, especially those that emphasize democratic criteria.

**Acknowledgements:** I am grateful to Megan Fritts and Marcos Picchio for valuable discussion and comments on previous drafts. Many thanks also to the anonymous reviewers at *Canadian Journal of Philosophy.*

**1. Introduction**

 In the last two decades, a robust consensus has emerged among philosophers of science regarding the role that political, ethical, or social values ought to play in scientific inquiry. For a wide variety of reasons, most philosophers concerned with the question of values and science now accept that such values *do* or *must* play some role in scientific inquiry, and that the “value-free ideal” is thus a misguided conception of science (Hicks 2014, 327). However, the question of how to distinguish, in a principled way, *which* values may legitimately influence scientific inquiry remains. This question, which has been dubbed the “new demarcation problem” (Holman & Wilholt, 2022), has until recently received comparatively less attention from philosophers of science.

 In what follows, I will first review some of the main arguments for the claim that social, ethical, and political values are required in scientific inquiry. Then, I will spell out more explicitly why solving the new demarcation problem is necessary, in particular why we must pay attention to the *content* of the values that influence scientific inquiry. As I point out, in order to avoid objectionable results, it is indispensable that the values doing the influencing be the *right values*. To address this issue, I appeal to Rawls’ theory of justice (1971), on the basis of which I defend a Rawlsian solution to the new demarcation problem. As I will show, there is much to motivate an appeal to Rawls’ theory of justice in the context of the new demarcation problem. Furthermore, Rawls’ theory of justice places plausible constraints on which values ought to influence scientific inquiry and can be fruitfully applied to concrete cases—such as the Vioxx drug scandal (Biddle 2007)—to determine how the conflicting interests of stakeholders affected by scientific inquiry should be balanced. After considering and responding to the objection that Rawls’ theory of justice applies only to the “basic structure” of society, I compare the Rawlsian account to other answers to the new demarcation problem that focus on democratic criteria. I conclude by briefly revisiting the value-free ideal.

**2. The Demise of the Value-Free Ideal and the Rise of the New Demarcation Problem**

 According to the “value-free ideal”, scientists ought not to let any ethical, social, or political values influence their scientific practice (Douglas 2009, 1). Rather, scientists ought to strive to be objective and be guided only by the evidence. Ethical, social, and political values—i.e., “non-epistemic values”— might legitimately motivate us to commenceor finance some scientific research program. But, according to the value-free ideal, such non-epistemic values should not influence the scientist’s evaluation of hypotheses. Of course, scientists are human beings too, and so will still sometimes be influenced by their ethical and political views. For example, the biologist Stephen Jay Gould (1981) famously argued that many 19th century scientific investigations of intelligence were tainted by racist values. So too, feminist scientists, philosophers, and historians have extensively documented the illicit influence of sexist and androcentric biases on scientific inquiry (Fausto-Sterling 1985; Martin 1991; Schiebinger 2004; Longino 2013). According to proponents of the value-free ideal, when scientists are influenced in these ways, they fail in their duty as scientists. Even if the value-free ideal has never been realized by any practicing scientist, according to its proponents, that ideal is still worth persevering.

 While there are still contemporary defenders of the value-free ideal (e.g., Betz 2013; Lacey 2013), most philosophers are science working on these questions have argued that the value-free ideal is not worth preserving. There are several influential arguments for this conclusion. One of those arguments stems from a point that is familiar to most philosophers of science, the underdetermination of theory by data (Stanford 2017). As proponents of the value-free ideal claim, scientists in their choice of theories ought to be influenced only by epistemic considerations, sometimes called “cognitive values” (e.g., Douglas 2013). What counts as a cognitive value is somewhat debatable, but they are typically thought to consist in such criteria as predictive accuracy, empirical adequacy, logical consistency, unification, explanatory power, ontological parsimony, fruitfulness, etc. (Kuhn 1977, 320-1). As proponents of the underdetermination argument point out though, these criteria are not sufficient to uniquely pick out a specific scientific theory in most realistic cases. Not only will there often be cases in which two theories T1 and T2 score equally well, overall, with respect to the above criteria of theory choice, but there is also underdetermination in the choice of how to weigh these criteria, especially given that some of them can conflict. So, in order to move past the under-determination and select a theory, scientists must resort to “non-cognitive” (i.e., non-epistemic) values, including ethical, social, and political values (Longino 1990; Kourany 2003; Anderson 2004).

 Another argument against the value-free ideal is often referred to as the “argument from inductive risk” (Douglas 2000; 2009).[[1]](#footnote-1) The argument from inductive risk proceeds from the commonplace assumption that scientific inquiry is fallible. Ideally, scientists would assert a hypothesis if and only if there is sufficient evidence for the hypothesis (Havstad 2021, 3). Of course, since inductive reasoning is fallible, it is always possible that a scientist might assert a hypothesis without sufficient evidence or fail to assert a hypothesis for which there is sufficient evidence. Given the prominent role that science and scientists play in our culture, these mistaken decisions can potentially lead to bad consequences for the public. But as Douglas (2000, 563) points out, “scientists have the same moral responsibilities as the rest us”, and so they too have the moral responsibility to consider the “predictable consequences of error.” Thus, scientists too must make value judgments like the rest of us. Additionally, values must play a role in science even before inquiry takes place, specifically when deciding on evidential standards. As Douglas (2000, 566) observes, “the deliberate choice of a level of statistical significance requires that one consider which kind of errors one is willing to tolerate.” So, for instance, if making a Type I error, i.e., mistakenly rejecting a null hypothesis, would lead to very bad consequences for the public, then perhaps the scientist should adopt a higher threshold for statistical significance, e.g., p-value <.005, rather than the conventional standard of p-value <.05. The upshot is that scientists must consider and act on non-epistemic values in their capacity as scientists.

 There is much more that can be said about these[[2]](#footnote-2) and other arguments[[3]](#footnote-3) against the value-free ideal; however, one notable feature of the values-and-science literature is that most philosophers of science have primarily attempted to establish the thesis that non-epistemic values “are at least sometimes necessary in decisions at the core of scientific reasoning” (de Melo-Martín & Intemann 2016, 502). By and large, philosophers concerned with the role of values in science have spent comparatively less time, at least until recently, on the question of *which* values ought to influence scientific inquiry.[[4]](#footnote-4) For example, presumably there is something wrong with a scientist who leaves out data inconsistent with the hypothesis that some drug is both safe and causally efficacious merely in order to rush the drug to market and generate a profit. Such a decision might be driven by *values*, namely economic value, but this seems like a corruption of the scientific process. Even those who reject the value-free ideal express concerns about the dangers of dogmatism, wishful thinking, and biasing scientific inquiry toward pre-determined conclusions (Anderson 2004, 11; Elliott 2017, 13; Douglas 2009, 102). The question of how to distinguish, in a principled way, *which* values may legitimately influence scientific inquiry is what has recently been called “the new demarcation problem” (Holman & Wilhot, 2022). Now that the debate about the value-free ideal has “cooled significantly”, the “new direction” (Hicks 2014, 3271) for investigations of values in science should consist in tackling the new demarcation problem directly.

**3. Why the Content of the Values Matters**

Before defending my proposed solution to the new demarcation problem, it would be worthwhile to further motivate the present investigation. As Brown (2020, 63) points out in his recent synthesis of the many arguments against the value-free ideal, what the traditional arguments against the value-free ideal show is that, in general, epistemic considerations do not fully eliminate the “contingencies” or “unforced choices” that scientists often face in the course of inquiry. There are many junctures at which scientists must make important decisions, e.g.., what research project to work on, how much funding ought to be allocated to which projects, how to conceptualize the object of inquiry, which methods to employ, which background assumptions to accept, which experiment to set up, how much evidence to gather, how and where to disseminate findings, etc. (Brown 2020, 70). In the face of so many live options, traditional epistemic criteria cannot completely settle how these decisions ought to be made. Non-epistemic values will have to play a role in the deliberative process if scientists are going to make decisions at all. Since unforced choices are common, it is extremely important that the values that structure, influence, or determine these decisions be the *right values* (Brown 2018, 6). That is, the *content* of the values matters. When influenced by non-epistemic values in the making of these decisions, scientists ought to ensure that they are being influenced by goods that are actually worth respecting, preserving, or promoting. To be sure, having the right values may not suffice for a legitimate instance of values influencing science. Suppose the scientist knowingly falsifies, fabricates, and distorts data to promote the cause of gender equality. This would be a relatively clear case of an illegitimate influence of values on scientific inquiry, despite involving the right values.

 On the other hand, if the *wrong* *values* influence scientific inquiry, then no matter how those mistaken values manifest themselves, surely this is an instance of values having an illegitimate impact on science. For example, consider a scientist who, after considering all the evidence, must decide whether to reject H, which we can assume is the null hypothesis. Suppose the scientist decides on standards of acceptance that lead him to reject H because a Type I error will only have negative consequences for members of a minority racial group *A*, whereas a Type II error will have negative consequences for racial group *B*, the dominant group, of which the scientist is also a member. Needless to say, if any influence of values on science is illegitimate, surely this is one of those cases.

The problem then is that without substantive constraints on what sorts of values may influence scientific practice, we will be left with counterintuitive or objectionable results. If we wish to avoid this problem and incorporate substantive constraints about the values that may influence scientific practice, then we need to first decide which values are the right ones. A satisfactory answer to the new demarcation problem will thus require philosophers of science to draw from other areas of inquiry, including social, moral, and political philosophy. As Anderson remarks, “to make progress on these problems, we need to integrate moral philosophy and the philosophy of science” (2004, 2). Similar sentiments are expressed by Hicks, who claims that “philosophers of science should undertake a deeper engagement with ethics” (2014, 3273). Now that the debate over *whether* values ought to play a role in scientific inquiry has cooled, some consideration of *which* ethical, social, and political values should influence science must enter into the picture.

**4. Towards a Rawlsian Solution to the New Demarcation Problem**

 Recently, Schroeder (2020) has suggested distinguishing between ethics-based approaches and political-philosophy-based approaches to the new demarcation problem. As Schroeder (2020, 3-5) observes, most previous work on science and values owes more to normative ethics, although some approaches adopt concepts, methods, and principles more at home in political philosophy, (e.g., Kitcher 2011; Intemann 2015; Lusk 2021). Because of the social and political significance of scientific activity, I agree with Lusk’s (2021, 109) proposal—following Schroeder (2020, 9)—for philosophers to embrace a specifically “political philosophy of science”. In this section, I aim to demonstrate how Rawls’ theory of justice, as laid out in his *A Theory of Justice* (1971) and in subsequent work, can help to determine which non-epistemic values may legitimately influence science.[[5]](#footnote-5) But before explaining the key features of Rawls’ theory and then applying it to our central question, I will first consider a few background facts that serve to justify the choice of framework.

*4.1 Motivating the Rawlsian Framework*

 First, people value a diverse array of goods that are not obviously reducible to one common currency, e.g., autonomy, friendship, happiness, security, liberty, equality, health, stability, etc. Perhaps, in the final analysis, this rich set of values can be reduced to some single good, as for instance the hedonist argues. But as is also well-known, reductionist projects of this sort raise various difficulties.[[6]](#footnote-6) Consequently, some sort of pluralism about values seems to be the commonsense position.

 Second, although at a very abstract level there may be widespread agreement about the importance of certain values—perhaps because such agreement is a necessary precondition for the existence of any society at all (Rachels 2003, 26)—still there are many substantive disagreements about how to put commonly accepted values into practice and about how to navigate trade-offs. It may be a widely accepted “moral fixed point” (Cuneo & Shafer-Landau 2014) that liberty is valuable, but people who value liberty very often disagree about whether an infringement on liberty is justified. To make matters concrete, although American conservatives and American liberals both regard liberty as a core value, fierce disagreements persist over how liberty ought to be balanced against other commonly endorsed values, such as happiness or equality (Pew Research Center 2019).

 Third, some moral disagreements will not be resolvable by means of rational debate. Of course, some moral debates arise because of different beliefs about empirical matters of fact, e.g., whether genetic engineering is safe, in which case such debates are more easily rationally resolvable. However, it is implausible that all moral disagreements are of this kind. If *A* endorses some value that *B* does not, perhaps owing to certain metaphysical or theological commitments, then their disagreement might not be rationally resolvable. At some point, a disagreement between *A* and *B* might bottom out at a foundational difference in what *seems plausible* based on the totality of *A*’s experience, such that it is impossible for *A* to convey or give reasons to *B* that would be rationally convincing. Importantly, *A*’s inability to articulate reasons to *B* in this way does not necessarily mean that *A*’s original position was unreasonable. Cases of intractable moral disagreement seem to be a relatively common, albeit unfortunate, feature of our moral lives. Such disagreements should not be ignored in a discussion of the new demarcation problem.

 Finally, we should recognize that scientific inquiry is a *public* phenomenon, a complex social practice that requires for its success the *cooperation* of a diverse group of individuals (Longino 1990). The “division of cognitive labor” (Kitcher 1990) among different groups of researchers, all working on related questions, is one of the hallmarks of modern science and is likely a crucial reason that the sciences have progressed so much over the last few centuries. Not only are scientific researchers themselves involved in these complex interpersonal interactions, but so are other related parties, such as the students who are educated by scientists at colleges and universities, as well as government grant agencies, which provide funding for research and development.

 Clearly, if a significant amount of scientific research exists only because of public funding, then the ethical, social, and political values that structure or guide scientific research must be ones that are publicly justifiable. Once we jettison the value-free ideal, we must consider the question of the political legitimacy of scientific research that is guided, structured, or influenced by scientists’ non-epistemic values. This concern is especially pressing in light of the other three facts mentioned above, i.e., persistent value pluralism, the existence of substantive moral disagreement, and the irresolvability of some of these disagreements. Given the social embeddedness of modern science, our solution to the new demarcation problem should promote cooperation between scientists, even though they might possess diverse interests and different conceptions of the good. Any attempt to construct a replacement for the value-free ideal that does not acknowledge and undertake to deal with these backgrounds facts is, in my estimation, doomed to failure.

*4.2 Rawls’ Theory of Justice*

 In light of these background facts about moral disagreement and the public nature of the scientific enterprise, one promising approach to tackling the new demarcation problem is to apply the social contract theory outlined and defended by John Rawls in his *A Theory of Justice* (1971) and subsequent work. One reason that Rawls’ social contract theory appears promising for our purposes is that he is motivated by similar concerns, namely, to come up with a conception of justice given the fact that “free and equal citizens are deeply divided by conflicting and even incommensurable religious, philosophical, and moral doctrines” (Rawls 1993, 143).

 As Rawls states in the beginning of *A Theory of Justice*, what he seeks to uncover are “the principles of justice for the basic structure of society” (1971, 11). According to Rawls, the correct principles of justice are those that rational, self-interested contractors would unanimously choose in a one-time selection process from a fair and equal initial situation. This fair and equal initial situation is what Rawls dubs the “original position” and is characterized by the rational contractors being behind a “veil of ignorance”—possessing *general* knowledge of the world, e.g., facts about human nature, economics, psychology, etc. but possessing no *particular* knowledge about themselves, e.g., their race, gender, religion, economic status, conception of the good, natural talents, etc. The veil of ignorance is for Rawls a “useful analytic device” (1971, 189) intended to solve the problems that had plagued the traditional social contract theory. Clearly, if rational contractors know several key facts about what will maximize their own self-interest, unanimous agreement on any particular set of principles would be impossible. By contrast, from behind the veil of ignorance, “no one is advantaged or disadvantaged in the choice of principles by the outcome of natural chance or the contingency of social circumstances”, and since the original position is a fair initial situation, the “fundamental agreements reached within it are fair” (1971, 12).

 Famously, Rawls proposes that rational contractors in the original position would select two principles of justice. Rawls’ mature formulation of these principles appears in his *Political Liberalism* (1993, 271). The First Principle states: “Each person has an equal right to a fully adequate scheme of equal basic liberties which is compatible with a similar scheme of liberties for all.” The Second Principle, which has two parts, states that “Social and economic inequalities are to satisfy two conditions. First, they must be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they must be to the greatest benefit of the least advantaged members of society.” The First Principle, which concerns basic liberties, has “lexical priority” over the Second Principle, which concerns social and economic inequalities. Moreover, the first part of the Second Principle, which concerns fair equality of opportunity, takes precedence over the second part of the Second Principle, i.e., “the difference principle”, which determines the circumstances under which social and economic inequalities are justifiable.

 In *A Theory of Justice*, Rawls argues at length that rational, self-interested contractors would choose these principles to govern the basic institutions of society rather than some alternative principle of justice, such as the principle of utility. This is because behind the veil of ignorance it would be rational to employ the decision procedure known as *maximin.* Given an array of alternatives with different possible outcomes, the maximin rule counsels us to choose the alternative whose worst possible outcome is superior to the worst possible outcomes of all the other alternatives. So, for instance, given action A1 with possible outcomes of either −1000 utility or +1000 utility and action A2 with possible outcomes of either −10 utility or +5 utility, according to maximin, we ought to choose A2. We ought to choose A2 even though one possible outcome of A1 would be quite good, were it to obtain. Maximin is thus, as Rawls notes, a quite “conservative attitude” (1971, 153). It is to be recommended not in every circumstance, but only when three conditions hold: i) knowledge of the probabilities of the outcomes is uncertain or insecure, such that an expected utility calculation cannot be made, ii) it is not worthwhile for the person to gamble for some further advantage, and iii) some alternatives might lead to outcomes that are “intolerable” (Rawls 1971, 154-6).

 According to Rawls, all these conditions obtain in the original position, and so rational, self-interested persons would rely on the maximin rule. Since the rational contractors would rely on maximin, they would not select the principle of utility to govern the basic institutions of society. Once the veil of ignorance is lifted, individuals living in a utilitarian society might find themselves in the minority, whose rights and interests are consistently sacrificed to benefit the majority. Likewise, rational, self-interested persons would not enshrine in the basic constitution principles that overtly privilege one race, religion, gender, etc. over another, for when the veil of ignorance is lifted, those individuals might find themselves in the disadvantaged group. Instead, according to Rawls, rational contractors in the original position would select the two principles of justice discussed earlier. These principles of justice ensure the rational contractors a reasonable guarantee of securing the “primary goods.” These are the goods that “every rational [person] is presumed to want…whatever a person’s rational plan of life”, where such goods include: “rights and liberties, powers and opportunities, income and wealth” (Rawls 1971, 62). Importantly for Rawls, this contractarian method allows us to come up with a set of substantive principles that place “limits on fair terms of social cooperation” (1971, 21), while presupposing only a relatively thin conception of the good.[[7]](#footnote-7)

*4.3 Applying Rawls’ Framework to the New Demarcation Problem*

 Now that we have detailed the most significant aspects of Rawls’ theory of justice, we can begin to apply the Rawlsian framework to shed light on the new demarcation problem. There are two ways in which the Rawlsian framework can aid in the project of distinguishing the legitimate from the illegitimate influences of values on science. The first way, which I will call the “traditional method”, simply applies Rawls’ two principles of justice to generate value judgements to be used in scientific contexts. Crucially, Rawls’ two principles provide us with an avenue for ruling out pernicious values, e.g., sexist and racist values, in a principled and elegant way. This issue was one of the key considerations motivating inquiry into the new demarcation problem in the first place. Since science is a public phenomenon that requires for its success social cooperation, e.g., in the form of public funding, mutual trust and honesty, collaborative research, etc., the values that influence scientific inquiry must be ones that are publicly justifiable. This is so even in the case of scientific research that is not publicly funded because of the special role that science plays in a liberal democratic society: one may appeal to scientific conclusions as a “public reason” for politically justifying public policy (Rawls 1993, 224), provided these conclusions are not “controversial.” We can consider the benefits and innovations of scientific research to be a primary good, which is not an implausible addition to Rawls’ original list. Once we have made this addition to the list of primary goods, it is easy to condemn, as a matter of public or professional policy, the influence of pernicious values on science. Both principles of justice guarantee equal rights and liberties regardless of race, gender, religion, etc. Thus, Rawls’ theory of justice would prohibit the actions of the scientist who manages inductive risk by privileging members of his own racial group over those of minority groups. From behind the veil of ignorance, rational contractors in the original position would not permit scientific inquiry to be structured or influenced by racist or sexist values, owing to the sort of maximin reasoning described in section 4.2. Once the veil of ignorance is lifted, one might find oneself in the group that is systematically disadvantaged by those pernicious values. We would expect rational, self-interested contractors behind a veil of ignorance to support egalitarian values in science, instead.[[8]](#footnote-8)

 The second way in which the Rawlsian framework can help us to solve the new demarcation problem is through what I will call the “extended method”. In concrete cases in which no result clearly follows from Rawls’ two principles of justice, we can employ the veil of ignorance directly as a neutral framework for adjudicating specific conflicts of interests. This proposal looks similar to the “stakeholder theory” of corporate obligations (Freeman 2014), oft-discussed in the business ethics literature. According to the stakeholder theory, the corporation has obligations to all stakeholders, e.g., employees, local communities, etc., not just shareholders, whether or not this maximizes profit. To help solve the problem of how to manage the conflicting interests of stakeholders, we are to imagine that different stakeholders are “behind a Rawls-like veil of ignorance” (Freeman 2014, 190), and then decide which principles rational, self-interested contractors would choose to adjudicate conflicts of interest.

To illustrate the extended method, consider a case study involving the drug “Vioxx”, a painkiller manufactured by Merck & Co. and used to treat arthritis, which was withdrawn from the market in 2004 once evidence surfaced that the drug increased the risk of heart attacks (Biddle 2007). As Hicks argues in a discussion of this case, “commercial values” strongly influenced scientists’ decision to conclude that Vioxx was safe (2014, 3279). In the Vioxx scandal, scientists were faced with an instance of local underdetermination. Some research subjects taking Vioxx died during the study, and while there was evidence that the deaths were caused by heart attacks, ultimately the evidence was ambiguous. Consequently, the scientists chose to categorize the deaths as being due to “unknown causes”, rather than due to a heart attack. The latter categorization might have prevented the drug from being deemed safe to market. The *prima facie* problem here is that the values that influenced the decision to categorize the data in this way need not have violated extant demarcation criteria, e.g., the lexical priority of the evidence principle (Hicks 2014, 3287). Once value-free science is rejected, it is less clear whether this case can be ruled out as an illegitimate influence of values on scientific inquiry.

 Now, according to Hicks, the scientists working for the pharmaceutical company were mistaken because their decision to interpret ambiguous evidence in a way that promoted the value of profit undermined other things that are valuable *by their own lights.* As Hicks argues, the primary reason that the pharmaceutical company acted wrongly is that “the health and well-being of the industry’s consumers/patients” is one of “the constitutive values of the pharmaceutical industry”, and so “from the perspective of the pharmaceutical industry…it is wrong to sacrifice patient health for the sake of profit” (2014, 3291-2). While one might argue that perhaps the pharmaceutical company does not care as much about the health of its consumers as it does about profit, according to Hicks, since profit is not intrinsically valuable, “profit is not actually a constitutive value of health care” (2014, 3291). The pharmaceutical company may not care about “scientific values” such as truth or empirical adequacy for their own sake but promoting those scientific values “would have been a more effective way to promote the constitutive value of patient health”, and so ultimately “*from the perspective of the pharmaceutical industry it was wrong to sacrifice scientific values for the sake of profit*” (Hicks 2014, 3292).In giving this response, Hicks aims to condemn Merck & Co., while avoiding commitment to the indirect/direct distinction (Douglas 2009, 96-7), or the lexical priority of the evidence principle, both of which give implausible verdicts in other cases, e.g., feminist interventions in science (2014, 3288).

 The problem with Hicks’ analysis of the Vioxx case, though, is that it depends on a contentious claim about the constitutive values of the pharmaceutical company. The argument thus begs the question against the spokesperson of the pharmaceutical company who, we might imagine, seeks to defend the value-laden decision driven by profits as a legitimate instance of the influence of values on scientific inquiry. Suppose that the spokesperson for the company rejects Hicks’ claim that the health of the consumers, and not profits, is a constitutive value of her company. To do so, she might appeal to Friedman’s shareholder theory of the firm, according to which the sole purpose toward which corporate executives ought to direct their efforts is satisfying the *wishes of the shareholders*, which “generally will be to make as much money as possible while conforming to the basic rules of the society (2014, 180). It is unclear how one might proceed to defend the claim that the health of the consumer is an intrinsic, constitutive value of the company to someone, such as our imagined pharmaceutical representative, who rejects this claim.[[9]](#footnote-9) The best candidate for what the constitutive values of the company consist in will be those values of which the relevant spokesperson for the company sincerely informs us. At the very least then, the approach that Hicks advocates is dialectically ineffective. It is insufficient to generate the desired result that the influence of corporate values on scientific inquiry in the Vioxx case was illegitimate.

 By contrast, the Rawlsian account that I have proposed can provide a stronger argument that the influence of corporate values in the Vioxx case was illegitimate, for Rawls’ framework can be employed to adjudicate and balance the competing interests of different stakeholders. To fix ideas, we can imagine several groups that might be affected by the decision being made in the Vioxx case: i) the scientists, ii) the research subjects, iii) the corporate executives and the shareholders of the company, iv) the employees of the company, and v) the consumers of the drug being tested. These different parties have different interests, some of which clearly pull against others. The scientists might endorse several values relevant to the case, such as the value of truth or the value of attaining an estimable reputation in the scientific community. The corporate executives and the shareholders, let’s suppose, have interest in making profits, ultimately in order to satisfy their other desires. So too, the employees of the company might value profits since the success of the company has downstream effects on their livelihood. The research subjects in the study and the would-be consumers of the drug clearly value a product that *in fact* issafe and causally efficacious, since they are the ones taking the drug. If scientists were deciding how social, political, and ethical values ought to influence them in a case such as this where the evidence is ambiguous, then they should appeal to Rawls’ veil of ignorance. We should try to imagine what sort of policy self-interested contractors would adopt behind a veil of ignorance.

 To be sure, if the rational contractors knew that their role in society was that of a shareholder or the corporate executive, then they might be tempted to endorse what the scientists in the Vioxx case actually did. But we are supposing, as Rawls does in the original thought experiment, that the rational contractors do not know their social identity. So, they do not know if, when the veil is lifted, they will be so advantaged by a policy that subordinates scientific standards to commercial values. It is possible that when the veil is lifted, the rational contractors will find themselves occupying the role of the consumers of the drug, in which case they will be severely disadvantaged by this proposed way of balancing the evidential standards for causal efficacy against the standards for safety. The implicit policies that are adopted by scientists working for pharmaceutical companies can have a dramatic impact on the life prospects of the consumers of prescription drugs, where some of these possible outcomes are intolerable. It would be quite a bad state of affairs for consumers if the kind of value judgements that were made in the Vioxx case were widespread. If evidential standards were regularly subordinated to commercial values as they were in the Vioxx case, then consumers might find themselves in a society replete with hazardous, inefficacious pharmaceuticals. Consequently, this situation seems to be precisely one in which the rational contractors would rely on the maximin rule, thereby giving more weight to scientific values of truth and other values such as public health.

 It is worth pointing out that the rational contractors might also find themselves in the position of the corporate executives or the shareholders of the pharmaceutical company. Thus, the rational contractors would *not* *prohibit* the influence of commercial values on scientific inquiry entirely. The economic inequalities that result from profit-driven inquiry can sometimes be to the advantage of the least well-off, in which case such inequalities would be justified by the Difference Principle. However, the fine details of how exactly the trade-off between epistemic values, commercial values, and the health and well-being of the public ought to be made is not something that can be determined in the absence of particular details about a given case. But this limitation is also true of Rawls’ own principles in their original context. We do not get much action-guidance from the two principles of justice without detailed empirical knowledge about the actual political situation in which we find ourselves.

Furthermore, the sorts of solutions to the new demarcation problem that have been put forward suffer from similar limitations. The strategy favored by Hicks, for example, does not tell us how exactly to balance the sometimes-conflicting values of profit, health, well-being, and truth. I suspect that this limitation will affect any proposed solution to the new demarcation problem. It is doubtful that we can come up with a principle that neatly settles, in the abstract, whether some influence of values in scientific inquiry is legitimate or not for every case. General principles are useful, but there is no substitute for experience, sound judgment, and “good sense” (Duhem 1991).[[10]](#footnote-10) In some cases, a plurality of approaches to risk-management may be permissible (Winsberg et al. 2020, 148). Clearly though, an application of Rawls’ framework to the case at hand strongly suggests that a policy that would allow for “very, very low standards for efficacy claims and very, very high standards for hazard claims” (Hicks 2014, 3287) would not be favored by rational, self-interested contractors behind a veil of ignorance. Thus, we do get *some* guidance on how to handle specific cases.

Notably, feminist philosophers of science have stressed one of the main points of this paper, namely that the *content* of the values that influence scientific inquiry matters (e.g., Longino 1990; Hicks 2011; Kourany 2010). This point has been emphasized especially in critical discussions of the “social value management” ideal defended by Longino (1990, 2002). Longino’s ideal prizes diversity of values, but according to critics, diversity of values in the scientific community is not desirable if the values in question are explicitly anti-feminist (Hicks 2011; Kourany 2010; Intemann 2017). The Rawlsian solution does not fall prey to this problem. Although here I have focused primarily on the issue of values in the commercial context, it is likely that feminist interventions in science would be legitimated by the Rawlsian account that I defend. This is primarily because the egalitarian values enshrined in the two principles of justice can be used to justify feminist political commitments (Watson & Hartley 2018). Indeed, as the first application that I considered in this section demonstrated, the Rawlsian solution succeeds in protecting the rights of marginalized groups and minority stakeholders, coinciding with recent proposals by feminist philosophers for a new ideal of values in science that takes into account social justice and the interests of the least well-off (Intemann 2017, 140-1).

**5. The Basic Structure Objection and Two Responses**

*5.1 The Basic Structure Objection*

It is important to recognize though that in applying Rawls’ framework to the new demarcation problem, we have departed not insignificantly from Rawls’ original concerns and context. This is especially true of the extended method, which involves applying the veil of ignorance to concrete cases. While making recourse to Rawls’ theory in the context of science and values is relatively new, attempts to extend Rawls’ theory to other domains are not. For instance, “the literature in business ethics makes free use of Rawlsian procedures and concepts” (Cohen 2010, 565) to settle important normative questions related to commercial practices and corporate governance. However, such attempts to apply Rawls’ theory of justice in the corporate context have been met with a strong objection, which applies, *mutatis mutandis*, to the Rawlsian solution to the new demarcation problem.

The objection is that Rawls’ theory of justice applies only to the “basic structure” of society, and not in smaller-scale contexts, such as the corporation (Singer 2015). As Rawls claims:

For us the primary subject of justice is the basic structure of society, or more exactly, the way in which the major social institutions distribute fundamental rights and duties and determine the division of advantages from social cooperation. By major institutions I understand the political constitution and the principal economic and social arrangements (1971, 7).

Moreover, Rawls seems to cast doubt on the prospect of applying in other domains those principles of justice that are intended to govern the basic structure:

There is no reason to suppose ahead of time that the principles satisfactory for the basic structure hold for all cases. These principles may not work for the rules and practices of private associations or for those of less comprehensive social groups. They may be irrelevant for the various informal conventions and customs of everyday life… (1971, 8).

Because of this restriction, Rawls would say that Nozick’s footnoted objection concerning the “inappropriateness” of the difference principle as a “governing principle within a family of individuals who love one another” (Nozick 1974, 167) misses the mark. The principles of justice that Rawls articulates are not necessarily intended to apply to every decision at the micro-level.

 Because of the way in which Rawls appears to limit the scope of his theory of justice, Singer has forcefully argued that “attempts to formulate a position on corporate governance using the normative resources offered by Rawls are bound to fail” (Singer 2015, 75). As the passages from Rawls above make clear, his theory of justice applies only to the basic structure of society. However, on Singer’s reading of Rawls’ political philosophy, “Rawls indicates that the corporate form is not part of the basic structure” (2015, 78). One reason Rawls does not want the principles of justice to apply to voluntary associations is out of respect for “liberty of conscience” and “freedom of association” (Rawls 2001, 163). In a well-ordered liberal democratic society, individuals ought to be free to enter into voluntary associations, such as a corporation or a religious community, even if the internal governing structure of those associations is undemocratic or illiberal. As Rawls points out, “the two principles of justice (as with other liberal principles) do not require ecclesiastical governance to be democratic” (Rawls 2001, 164). What Rawls says here about the principles of justice and religious organizations would seem to apply to all voluntary associations, the corporation included.

This debate over the possibility of extending Rawls’ theory of justice is very much relevant to the Rawlsian solution to the new demarcation problem. In the first instance, this is because a lot of scientific activity, especially medical research, takes place in a private, corporate context. Additionally, even if we put aside the corporate context, many other organizations, associations, or institutions within which scientific activity is conducted (and where questions of *which* values ought to influence science arise), do not seem to count as part of the basic structure of a society. A university research lab, a professional scientific society, a non-profit organization, etc. might also be domains, in addition to the corporation, in which scientific inquiry is sponsored, and yet none of these seem to be what Rawls is talking about when he delimits the scope of his theory of justice. Instead, these are what Singer helpfully terms, “meso-level institutions”, ones that “lie between large ‘macro-level’ state institutions on the one hand and the actions of individuals on the other” (2015, 68). One succinct way of articulating the problem that the Rawlsian solution faces is that the principles of justice and veil of ignorance apply only to macro-level institutions, but much scientific activity occurs within a governing structure that exists at the meso-level, and thus outside the scope of Rawls’ framework.

*5.2 Two Responses*

There are at least two responses that the proponent of a Rawlsian solution to the new demarcation problem can give. First, some philosophers who seek to apply Rawls’ views to answer questions about corporate governance have argued that the principles of justice apply directly to the corporation because in fact the corporation *does* form part of the basic structure. For one thing, Rawls himself admits that “the concept of the basic structure is somewhat vague” (1971, 9), something which Singer (2015, 76) also acknowledges in his critique of the business ethics applications of Rawls. According to Abizadeh (2007, 319), we can discern in Rawls (1971) three different candidate criteria for which institutions count as part of the basic structure: i) those that “determine and regulate the fundamental terms of social cooperation”, ii) those that “have profound and pervasive impact upon persons’ life chances”, and iii) those that “subject persons to coercion.” In his critique of applications of Rawls, Singer endorses the “coercion” criterion for those institutions that count as part of the basic structure (Singer 2015, 77-8), which he regards as the one most consistent with Rawls’ broader liberal commitments. Since corporations cannot legally coerce individuals, it follows immediately from this criterion that they don’t form part of the basic structure.[[11]](#footnote-11) Others, such as G.A. Cohen, have defended the “profound and pervasive impact” criterion for when principles of justice apply, criticizing a disproportionate focus on legally coercive institutions (Cohen 1997, 23). It is doubtless true that corporations can have a profound and pervasive influence on people’s life chances, especially those large firms that provide services that increasingly structure everyday life, e.g., technology companies; and so according to this criterion, at least some corporations would count as part of the basic structure. So too, in their response to Singer (2015), Welch and Ly (2017, 11) adopt the “social cooperation” criterion, which on their view entails that “corporations are part of the basic structure, because they are institutions for economic production in society’s system of social cooperation.”

The upshot for my purposes is that it is not at all obvious that the corporation, or those other voluntary associations in which scientific activity is conducted, would fail to count as part of the basic structure. Much depends, of course, on what the correct criterion for being part of the basic structure is, something which Rawls does not tell us explicitly. It is highly plausible that these associations satisfy the “profound and pervasive impact” and the “social cooperation” criteria, given the immense impact that scientific activity has on our lives, including the degree to which they determine the terms of social cooperation, e.g., how we organize society in response to a deadly pandemic. If one of these two criteria is correct, then the basic structure objection fails, and there is no obstacle to applying the Rawlsian framework to meso-level institutions, where questions of science and values arise.

The second response to the basic structure objection is that even if the corporation or other voluntary associations do not form part of the basic structure, it does not follow that these areas are untouched by Rawls’ theory of justice (Welch and Ly 2017, 10). Rawls clearly endorses this general point when he discusses the relationship between the principles of justice, voluntary associations, and the basic structure:

Even if the basic structure alone is the primary subject of justice, principles of justice still put essential restrictions on the family and all other associations… A domain so-called, or a sphere of life, is not, then, something already given apart from principles of justice…If the so-called private sphere is a space alleged to be exempt from justice, then there is no such thing (2001, 166).

For example, the first part of Rawls’ Second Principle entails that employment discrimination on the basis of gender, religion, race, etc. is unjust; therefore, even if corporations do not form part of the basic structure of society because they are voluntary associations, the principles of justice nonetheless place constraints on the hiring activities of the corporation (Cohen 2010, 565).

The foregoing considerations can thus be used to defend the Rawlsian solution to the new demarcation problem from the basic structure objection. In the same way that the principles of justice place “essential restrictions” on the employment practices of business firms, so too the proponent of the Rawlsian solution to the new demarcation problem would say that the principles of justice place essential restrictions on the ways in which values structure or influence scientific inquiry, regardless of whether the domain in which scientific activity takes place is part of the basic structure. Thus, the application of Rawlsian principles and procedures to rule out racist and sexist values from influencing scientific inquiry, e.g., in how the threshold for sufficient evidence is determined, can be endorsed without going against the spirit of Rawls’ theory. The theory still has implications for the internal dynamics of voluntary associations, even if such matters are not the primary locus of justice.

**6.** **Ideal Approaches, Democracy, and Political Legitimacy**

 It is worth mentioning that the Rawlsian solution to the new demarcation problem bears some similarities to the framework defended by Phillip Kitcher in his *Science in a Democratic Society* (2011). According to Kitcher, the value-judgements that ought to inform standards of evidence, determine which questions are scientifically significant, establish which research programs are worthy of public funding, etc. are those that “would be endorsed by an ideal conversation, embodying all human points of view, under conditions of mutual engagement” (114). These conditions of mutual engagement can be epistemic, e.g., we should suppose that participants in these ideal discussions have correct factual beliefs and also affective, e.g., we should suppose that participants have enlarged sympathies and consider a wide variety of points of view (51-3). Kitcher’s ideal deliberation account is rooted in a naturalistic picture of the “ethical project”, according to which the original function of ethics is to solve problems caused by our imperfect capacities for altruism. According to Kitcher’s account, such conversations should concern all members of the human species, as well as future generations, and should be oriented toward affording all people an equal opportunity for a worthwhile life. The ideal conversation that Kitcher imagines would, of course, be impossible to implement in practice. Actual agents will inevitably fail to satisfy the conditions of mutual engagement, and any attempt to hold a “panhuman conversation” (112) would likely devolve into a “vast cacophony” (51).

 Although some significant differences exist[[12]](#footnote-12), both the approach envisioned by Kitcher and the Rawlsian solution that I have defended here attempt to answer questions about science and values by reference to some idealizing procedure rather than appealing directly to actual democratic mechanisms. Thus, both accounts are subject to what Keren (2015, 1285) has called the “responsiveness-of-science problem”: we should be alert to the possibility that scientists “may not be properly responsive to the values, needs, and interests of different segments of society.” As Keren argues, “counterfactual, informed democratic deliberations”, such as those that would be the outcome of Kitcher’s ideal conversations or those that result from “deliberative polling” (Fishkin 2009), may have “epistemic significance”, but they do not have “legitimizing force” in the same way that actual democratic decision-making does. Considerations of autonomy and self-determination demand that the value-judgements that structure scientific inquiry be decided by actual democratic mechanisms, and not through hypothetical or counterfactual procedures (Keren 2015, 1292). This objection mirrors traditional concerns about Rawls’ contractarian method. As Dworkin (1973, 501) famously puts it, “A hypothetical contract is not simply a pale form of an actual contract; it is no contract at all.”[[13]](#footnote-13)

 Because of the problems that idealizing approaches raise, one might think that the more straightforward solution to the new demarcation problem is to simply appeal to actual democratic procedures, e.g., voting, rather than hypothetical or counterfactual deliberations. According to what I will call the “democracy criterion”, which has several defenders (e.g., Douglas 2005; Intemann 2015; Schroeder 2019; Lusk 2021), the values that influence science ought to be ones that are actually democratically endorsed. There is much to recommend this proposal, especially in light of the point discussed above, namely that scientific inquiry ought to be responsive to the public’s needs, values, and interests. Failure to do so insufficiently respects people’s rights to self-determination. Since the general public both funds and is significantly affected by the structure and outcomes of scientific inquiry, failure to be responsive to the public’s needs, values, and interests seems politically illegitimate.

 While the democratic criterion may better avoid the charge of political illegitimacy than hypothetical approaches such as the Rawlsian solution and Kitcher’s ideal deliberation account, it is important to recognize that the democracy criterion suffers from some difficulties as well (Havstad and Brown 2017; Brown 2020, 72-4). First, it would be inordinately impractical to put to a vote every possible value-laden decision that a scientist faces, especially when dealing with complex areas of science such as climate modeling, which “involves literally thousands of unforced methodological choices” (Winsberg 2012, 130), and moreover, which often requires significant technical expertise to understand (Havstad and Brown 2017, 102). So, it seems that scientists in their capacity as scientists will still need to make value-judgments that cannot, for practical reasons, be decided on by a democratic majority. Thus, scientists will still need some normative principles to guide their decisions. Second, the democracy criterion might prove objectionably permissive, potentially licensing the influence of pernicious values on the scientific process, if, say, those values were held by the majority. Relatedly, a majority of the public might hold some value-judgment because of wildly false beliefs about the empirical facts, perhaps owing to the undue influence of special interest groups or irresponsible journalistic practices. We certainly do not want the scientific enterprise to be held hostage to “a tyranny of ignorance” (Kitcher 2011). Now, perhaps it is possible to overcome these difficulties by, for example, improving science education; or, rather than voting on every value-laden micro-decision, constituents can vote on general policies or strategies for scientists to follow[[14]](#footnote-14). Still, these problems are *prima facie* ones that ideal approaches tend to evade.

 Ultimately though, I think the distance between the Rawlsian solution and the democratic criterion is not so large as it appears at first glance. First, it is crucial to recognize that some of the problems highlighted above apply only to what Kitcher (2011, 117) calls “vulgar democracy”, a conception of democracy that gives normative authority to the “untutored preferences” of the majority. By and large, though, sophisticated approaches to democratic theory do not identify democracy with simple majority rules, focusing instead on more abstract notions such as “popular control”, the “will of the people”, or “responsiveness to settled preferences” (Kitcher 2011, ch. 3). For example, Dewey criticizes the conflation of democracy with simple majority rules, writing instead that democracy is “a mode of associated living, of conjoint communicated experience”, a notion closely connected to further ideals of freedom and equality (Festenstein 2018). Crucially, both of these ideals, freedom and equality, figure prominently in Rawls principles of justice.

Second, proponents of the democratic criterion themselves recognize some of the difficulties posed by the vulgar conception of democracy. For instance, Intemann (2015: 228) points out that simple majoritarian conceptions of democracy will tend to objectionably downplay the interests of “marginalized groups.” So too, Schroeder (2020, 11) points out that a solution to the new demarcation problem must be such that “politically illegitimate values”—e.g., racist, sexist, etc.— are “filtered” out, suggesting that such values cannot play a role in the scientific process, even if they are endorsed by a majority. Notably, the Rawlsian framework can deal with these two difficulties. As discussed in section 4.3, Rawls’ principles of justice succeed in protecting the rights of marginalized groups and minority stakeholders, and moreover, would filter out those politically illegitimate values that might be endorsed by a democratic majority. Thus, it seems proponents of the democracy criterion can strengthen their accounts by incorporating substantive normative principles provided by Rawls’ theory of justice.

Even for more sophisticated accounts of democracy, such as “deliberative democracy” (e.g., Lusk 2021), which build into the theory substantive democratic ideals, e.g., freedom, equality, etc., Rawlsian procedures still might have something valuable to offer. Deliberative democracy focuses primarily on the “conditions of debate”, where deliberators are engaged in the collective project to make a decision that is good for the whole (Lusk 2021, 107). Of course, one can easily imagine cases in which the norms and ideals of deliberative democratic decision-making are satisfied, i.e., “equality, reciprocity, absence of coercion, and fairness” (Lusk 2021, 107), and yet no viable consensus has been reached. Perhaps it is impossible to reach a consensus due to mutually incompatible conceptions of the good. In this case, we might think the fairest solution is to decide by a simple majority vote. An alternative possibility for the proponent of deliberative democracy would be to attempt to apply the veil of ignorance in the same way in which it was applied in the Vioxx case, i.e., the extended method, to see whether some fair compromise can be reached.

It would be interesting to further explore approaches to the new demarcation problem that incorporate both actual and idealized elements. Perhaps some combination of these two views could help to resolve indeterminacies in applying the veil of ignorance to concrete cases. What I hope to have shown here, however, is that the Rawlsian solution and democratic approaches to the new demarcation problem, need not be starkly opposed, but rather can be mutually reinforcing. Because the Rawlsian solution can be combined with actual democratic procedures, there is good reason to believe that the objection from political illegitimacy (e.g., Keren 2015) can be overcome.

**7. Concluding Remarks: The Value-Free Ideal Revisited**

 Having thus shown how Rawls’ framework can help to solve the new demarcation problem, I would like to conclude by briefly reconsidering the value-free ideal. While proponents of the value-free ideal will reject the presupposition of the present inquiry, namely that values may legitimately influence scientific inquiry, the Rawlsian solution to the new demarcation problem might prove persuasive to those who remain attracted to the idea of value-free science. This is because the Rawlsian account respects the intuition that motivates the traditional value-free ideal, namely that science ought to be impartial and unbiased. For instance, W.E.B. Du Bois writes in defense of value-free science that “The [American Negro] Academy should be *impartial* in conduct; while it aims to exalt the people, it should aim to do so by truth not by lies, by honesty not by flattery” (2008, 186).[[15]](#footnote-15) What we have learned from several decades of work on the role of values in science is that this ideal of impartiality cannot be achieved, *per impossibile*, by banishing ethical, social, and political values entirely from scientific practice.

Nevertheless, the ideal of impartiality can be divorced from the value-free ideal and maintained by replacing the value-free ideal with one that is inspired by Rawls’ theory of justice. Throughout *A Theory of Justice*, Rawls emphasizes that one of the goals of employing the veil of ignorance method is to eliminate bias, prejudice, and partiality from the decision procedure. As Rawls points out, “if knowledge of particulars is allowed, then the outcome is biased by arbitrary contingencies” (1971, 141). For this reason, the veil of ignorance deprives the rational contractors of “knowledge of those contingencies which sets men at odds and allows them to be guided by their *prejudices*” (Rawls 1971, 19; emphasis mine). Ultimately, in applying Rawls’ social contract theory, “we try to work out what rational legislators suitably constrained by the veil of ignorance, and in this sense *impartial*, would enact to realize the conception of justice” (1971, 284; emphasis mine). Crucially then, the Rawlsian account allows us to endorse an ideal of impartiality in science without being committed to the value-free ideal. We can safeguard an impartial and unbiased science by ensuring that the value judgments that influence scientific inquiry are ones that would be endorsed by rational, self-interested contractors from behind a veil of ignorance. For this reason, the Rawlsian solution may serve as a congenial compromise position for those who are attracted to the traditional value-free ideal.

As I have already indicated in my discussion of other answers to the new demarcation problem, the Rawlsian solution may not be the whole story. One might also, for instance, want to add to the account deontic constraints ruling out scientific fraud, regardless of whether Rawlsian principles are motivating such misconduct. Or perhaps we should, in general, demand that scientists be as transparent as possible about the way in which their value judgements influence their inquiry (Elliott 2020). Still, as I hope to have shown, appealing to Rawls’ theory of justice can shed a great deal of light on how to distinguish legitimate from illegitimate values in science.

References

Abizadeh, A. (2007). “Cooperation, Pervasive Impact, and Coercion: On the Scope (not Site) of Distributive Justice,” *Philosophy & Public Affairs*, 35(4): 318-358.

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.

Betz, G. (2013). “In Defence of the Value Free Ideal.” *European Journal for Philosophy of Science* 3(2): 207–20.

Biddle, J. (2007). “Lessons from the Vioxx Debacle: What the Privatization of Science Can Teach Us About Social Epistemology,” *Social Epistemology*, 21(1): 21-39

Biddle, J. (2013). “State of the Field: Transient Underdetermination and Values in Science.” *Studies in History and Philosophy of Science* 44(1): 124–33.

Blanc, S. (2016). “Are Rawlsian Considerations of Corporate Governance Illiberal? A Reply to Singer.” *Business Ethics Quarterly*, 26(3): 407–421

Bright, L.K. (2018). “Du Bois’ Democratic Defence of the Value Free Ideal.” *Synthese* 195(5): 2227–45.

Brown, M.J. (2018). “Weaving Value Judgment into the Tapestry of Science,” *Philosophy, Theory, and Practice in Biology*, 10(10): 1-8.

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

Cohen, G.A. (1997). “Where the Action Is: On the Site of Distributive Justice,” *Philosophy & Public Affairs*, 26(1): 3-30.

Cohen, M.A. (2010). “The Narrow Application of Rawls in Business Ethics: A Political Conception of Both Stakeholder Theory and the Morality of Market,” *Journal of Business Ethics*, 97: 563–579

Cuneo, T., & Shafer-Landau, R. (2014). “The moral fixed points: New directions for moral nonnaturalism.” *Philosophical Studies*, 171(3): 399–443.

de Melo-Martín, I. & Intemann, K. (2016). “The Risk of Using Inductive Risk to Challenge the Value-Free Ideal.” *Philosophy of Science*, 83(4): 500–520.

Doppelt, G., (1989). “Is Rawls Kantian Liberalism Coherent and Defensible?” *Ethics*, 99(4): 815–851.

Douglas, H. (2000). “Inductive Risk and Values in Science.” *Philosophy of Science*, 67(4): 559–79.

Douglas, H. (2005). “Inserting the Public into Science.” In S. Maasen and P. Weingart (eds.), *Democratization of Expertise? Exploring Novel Forms of Scientific Advice in Political Decision-Making*, Dordrecht: Springer, 153–169.

Douglas, H. (2009). *Science, Policy, and the Value-Free Ideal*. Pittsburgh: University of Pittsburgh.

Douglas, H. (2013). “The Value of Cognitive Values.” *Philosophy of Science* 80(5): 796–806.

Du Bois, W.E.B. (2008). *The Souls of Black Folks*. Oxford: Oxford University Press.

Duhem, P. (1991). *German Science: Some Reflections on German Science/German Science and German Virtues*. John Lyon, trans. La Salle, IL: Open Court.

Dupré, J. (2007). “Fact and Value.” In H. Kincaid, J. Dupré, and A. Wylie (eds.), *Value-Free Science? Ideals and Illusions*. Oxford: Oxford University Press, 27–41.

Dworkin, R. (1973). “The Original Position,” *The University of Chicago Law Review*, 40(3): 500-33

Elliott, K.C. & Richards, T. (2017). *Exploring Inductive Risk*: *Case Studies of Values in Science.* New York: Oxford University Press.

Elliott, K.C. (2017). *A Tapestry of Values: An Introduction to Values in Science*. Oxford: Oxford University Press.

Elliott, K.C. (2020). “A Taxonomy of Transparency in Science.” *Canadian Journal of Philosophy*, 1–14. doi:10.1017/can.2020.21

Fausto-Sterling, A. (1985). *Myths of Gender: Biological Theories about Women and Men.* New York: Basic Books.

Festenstein, M. (2018). “Dewey’s Political Philosophy”, *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), Edward N. Zalta (ed.), URL= <https://plato.stanford.edu /archives/win2019/entries/dewey-political/>.

Fishkin, J. (2009). *When the People Speak*. New York: Oxford University Press.

Freeman, R.E. (2014). “Stakeholder Theory of the Modern Corporation.” In W.M. Hoffman, R.E. Frederick, & M.S. Schwartz (eds.), *Business Ethics Readings and Cases in Corporate Morality, 5th ed*. Hoboken, NJ: Wiley-Blackwell, 184–91.

Friedman, M. (2014). “The Social Responsibility of Business is to Increase its Profits.” In W.M. Hoffman, R.E. Frederick, & M.S. Schwartz (eds.), *Business Ethics Readings and Cases in Corporate Morality, 5th ed*. Hoboken, NJ: Wiley-Blackwell, 180–83.

Gould, S. J. (1981). *The Mismeasure of Man*. New York: Norton & Company.

Hartley, C. (2009). “Justice for the Disabled: A Contractualist Approach,” *Journal of Social Philosophy*, 40(1): 17–36.

Havstad, J. & Brown, M.J. (2017). “Inductive Risk, Deferred Decisions, and Climate Science Advising,” in K.C. Elliott & T. Richards (eds.), *Exploring Inductive Risk: Case Studies of Values in Science*. Oxford University Press, pp. 101-123.

Havstad, J. (2021). “Sensational Science, Archaic Hominin Genetics, and Amplified Inductive Risk,” *Canadian Journal of Philosophy*, 1-26. DOI: https://doi.org/10.1017/can.2021.15

Hicks, D. (2011). “Is Longino’s Conception of Objectivity Feminist?” *Hypatia* 26(2): 333–51.

Hicks, D. (2014). “A New Direction for Science and Values.” *Synthese* 191(14): 3271–95.

Holman, B. & Wilholt, T. (2022). “The New Demarcation Problem,” *Studies in History and Philosophy of Science,* 91: 211-20.

Hursthouse, R. (1996). “Normative Virtue Ethics,” in R. Crisp, (ed.), *How Should One Live?* Oxford: Oxford University Press, 19–33.

Intemann, K. (2005). “Feminism, Underdetermination, and Values in Science.” *Philosophy of Science*, 72(5): 1001–12.

Intemann, K. (2015). “Distinguishing between Legitimate and Illegitimate Values in Climate
Modeling.” *European Journal for Philosophy of Science* 5(2): 217–32.

Intemann, K. (2017). “Feminism, Values, and the Bias Paradox: Why Value Management Is Not Sufficient,” in J. Turri, K.C. Elliott, & D. Steel (eds.), *Current Controversies in Values and Science*. Routledge.

Keren, A. (2015). “Science and Informed, Counterfactual, Democratic Consent”, *Philosophy of Science*, 82 (5): 1284-1295

Kitcher, P. (1990). “The Division of Cognitive Labor.” *Journal of Philosophy* 87(1): 5–22.

Kitcher, P. (2011). *Science in a Democratic Society*. Amherst, NY: Prometheus Books.

Kittay, E. (1999). *Love’s Labor: Essays on Women, Equality, and Dependency*. New York: Routledge.

Kourany, J. (2003). “A Philosophy of Science for the Twenty‐First Century.” *Philosophy of Science*, 70(1): 1–14.

Kourany, J. (2010). *Philosophy of Science after Feminism*. Oxford: Oxford University Press.

Kourany, J. (2018). “Adding to the Tapestry.” *Philosophy, Theory, and Practice in Biology*, 10(9): 1–6. doi: 10.3998/ptpbio.16039257.0010.009.

Kuhn, T.S. (1977). “Objectivity, Value Judgment, and Theory Choice.” In *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press, 320–39.

Lacey, H. (2013). “Rehabilitating Neutrality,” *Philosophical Studies*, 163(1): 77-83.

Longino, H. (1990). *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, NJ: Princeton University Press.

Longino, H. (1996). “Cognitive and Non-Cognitive Values in Science: Rethinking the Dichotomy.” In L. Nelson and J. Nelson (eds.), *Feminism, Science, and the Philosophy of Science*. Dordrecht: Kluwer Academic Publishers, 39–58.

Longino, H. (2002). *The Fate of Knowledge*. Princeton, NJ: Princeton University Press.

Longino, H. (2013). *Studying Human Behavior: How Scientists Investigate Aggression and Sexuality*. Chicago: Chicago University Press.

Lusk, G. (2020). “Political Legitimacy in the Democratic View: The Case of Climate Services”, *Philosophy of Science*, 87(5): 991-1002.

Lusk, G. (2021). “Does democracy require value-neutral science? Analyzing the legitimacy of scientific information in the political sphere”, *Studies in History and Philosophy of Science Part A* 90: 102-110.

Martin, E. (1991). “The egg and the sperm: How science has constructed a romance based on
stereotypical male-female roles.” *Signs* 16(3): 485–501.

Nozick, R. (1974). *Anarchy, State, and Utopia*. New York: Basic Books.

Pew Research Center. (2019). “In a Politically Polarized Era, Sharp Divides in Both Partisan Coalitions.” Available at: https://www.pewresearch.org/politics/2019/12/17/in-a-politically-polarized-era-sharp-divides-in-both-partisan-coalitions/

Putnam, H. (2002). *The Collapse of the Fact/Value Dichotomy and Other Essays*. Cambridge, MA: Harvard University Press.

Rachels, J. (2003). *The Elements of Moral Philosophy, 4th ed.* New York: McGraw-Hill.

Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Belknap Press.

Rawls, J. (1993). *Political Liberalism*. New York: Columbia University Press.

Rawls, J. (2001). *Justice as Fairness: A Restatement*. E. Kelly (ed.), Cambridge, MA: Harvard University Press.

Rooney, P. (2017). “The Borderlands between Epistemic and Non-Epistemic Values.” In K. Elliott & D. Steel (eds.), *Current Controversies in Values in Science*. New York: Routledge, 31–45.

Rudner, R. (1953). “The Scientist Qua Scientist Makes Value Judgments.” *Philosophy of Science,* 20(1): 1–6.

Sandel, M.J. (1984). “The Procedural Republic and the Unencumbered Self,” *Political Theory*, 12(1): 81-96.

Schiebinger, L. (2004). *Nature’s Body: Gender in the Making of Modern Science*. New Brunswick,
NJ: Rutgers University Press.

Schroeder, S.A. (2019). “Democratic Values: A Better Foundation for Public Trust in Science.” *The* *British Journal for Philosophy of Science*, 1–20. doi: 10.1093/bjps/axz023.

Schroeder, S.A. (2020). “Values in Science: Ethical vs. Political Approaches.” *Canadian Journal of Philosophy*, 1–10. doi:10.1017/can.2020.41

Singer, A. (2015). “There Is No Rawlsian Theory of Corporate Governance,” *Business Ethics Quarterly*, 25(1): 65-92.

Stanford, K. (2017). “Underdetermination of Scientific Theory.” *The Stanford Encyclopedia of Philosophy*(Winter 2017 Edition), E.N. Zalta (ed.), URL = <https://plato.stanford.edu/ archives/win2017/entries/scientific-underdetermination/>.

Stark, C. (2000). “Hypothetical Consent and Justification,” *Journal of Philosophy*, 97(6) :313-334

Watson, L. & Hartley, C. (2018). *Equal Citizenship and Public Reason: A Feminist Political Liberalism.* Oxford: Oxford University Press.

Welch, T. & Ly, M. (2017). “Rawls on the Justice of Corporate Governance*,” Business Ethics Journal Review*, 5(2): 7–14.

Winsberg, E. (2012). “Values and Uncertainties in the Predictions of Global Climate Models.” *Kennedy Institute of Ethics Journal*, 22(2): 111–37

Winsberg, E., Oreskes, N., & Lloyd, E. (2020). “Severe weather event attribution: Why values won't go away.” *Studies in History and Philosophy of Science Part A*, 84: 142–149.

1. It is a common misconception that Douglas’ argument from inductive risk is a “revival” of an argument found in Rudner (1953). However, see Havstad (2021, 15, fn. 37) for a lucid and convincing explanation of why Douglas’ argument is both significantly different from and stronger than Rudner’s argument. [↑](#footnote-ref-1)
2. For more discussion of the argument from inductive risk, see Elliott & Richards (2017). For more discussion of the underdetermination argument, see Intemann (2005) and Biddle (2013). [↑](#footnote-ref-2)
3. One of these arguments stems from the claim that there is no sharp distinction between “facts” and “values”, and that as a result many scientific hypotheses are unavoidably “value-laden” (Putnam 2002; Dupré 2007). Another argument challenges the distinction between epistemic and non-epistemic values (Longino 1996; Rooney 2017). [↑](#footnote-ref-3)
4. This is not to say that proponents of popular arguments against the value-free ideal have been silent on the new demarcation problem. In section 6, I will discuss some other ways of approaching the problem. [↑](#footnote-ref-4)
5. The idea of invoking Rawls to determine which values are legitimate in science is not unprecedented. Kourany (2018) briefly suggests this idea. In addition, as discussed in section 6, the approach outlined by Kitcher (2011) has some commonalities with key elements of Rawls’ framework. [↑](#footnote-ref-5)
6. Consider standard objections to a hedonic theory of value, e.g., the experience machine (Nozick 1974, 42-5). [↑](#footnote-ref-6)
7. Rawls’ theory has, of course, been the object of much criticism over the last few decades, which has generated responses from commentators in turn. For instance, communitarians, such as Sandel (1984), have criticized Rawls’ theory for its unsustainable vision of the rational agent as an “unencumbered self”, i.e., as “a self understood as prior to and independent of purposes and ends” (86). Similarly, some philosophers such as Kittay (1999) argue that Rawls’ theory, and social contract theories in general, cannot adequately account for the interests and rights of persons with disabilities, or those in similar states of dependency. Addressing foundational objections of this sort is beyond the scope of the present inquiry; however, see Hartley (2009) for a response to the objection from disability, and see Doppelt (1989) for a response to Sandel’s objection. [↑](#footnote-ref-7)
8. An anonymous reviewer wonders whether the Rawlsian solution would rule out controversial areas of the human sciences, e.g., research into a biological basis for gender differences. In this paper, I have been especially concerned with the equitable distribution of inductive risk and not the content of scientific theories *per se*. I lack the space to explore this issue in depth here, but I see no reason that research into these areas need be ruled out by Rawls’ framework, provided the concepts employed therein are not morally compromised. As the anonymous reviewer points out, such work may potentially yield “liberatory conclusions or policies”. However, insofar as the research itself often proves controversial on purely epistemic grounds, Rawls’ caveat about “controversial” science (1993, 224) suggests that this work may not be an adequate basis for public policy. [↑](#footnote-ref-8)
9. Perhaps the scientists *qua* physicians have therapeutic obligations, but they also have obligations to their employer. So, at the very least, we are left with a conflict of values with no method for adjudicating them. [↑](#footnote-ref-9)
10. This is a theme often emphasized by virtue ethicists (e.g., Hursthouse 1998) critical of principle-based approaches to ethics. [↑](#footnote-ref-10)
11. Perhaps this entailment is not so immediate. See Blanc (2016) for a defense of the claim that corporations count as coercive in the relevant sense. [↑](#footnote-ref-11)
12. While Kitcher connects his ideal conversation to Rawls’ notion of publicity (2011, ch. 6), a salient difference between the two views is that the ideal conversation does not rely on the veil of ignorance. Rather, Kitcher’s ideal deliberators are fully informed about their own perspectives and their society, unlike the rational contractors in Rawls’ framework, who owing to reasons of fairness, *lack* crucial information about themselves and society. [↑](#footnote-ref-12)
13. Despite the prominence of this objection, in my view, it misconstrues Rawls’ project. As Stark (2000, 334) convincingly argues, Rawls’ hypothetical contractarianism “is not designed to generate political obligation; rather, it is designed to justify political principles”. As discussed in section 4.2, the veil of ignorance is a device for generating principles of justice that are *fair*. The principles have normative force because they are fair, not because actual individuals “hypothetically consent” to them. [↑](#footnote-ref-13)
14. Thanks to an anonymous reviewer for this suggestion. [↑](#footnote-ref-14)
15. See Bright (2018) for an exposition of Du Bois’ arguments for the value-free ideal. [↑](#footnote-ref-15)