Inductive Risk and Regulatory Toxicology: A Comment on

de Melo-Martín and Intemann

Daniel J. Hicks

Science and Technology Policy Fellow, American Association for the Advancement of Science, 1200 New York Ave NW, Washington, DC, 20005

hicks.daniel.j@gmail.com

Abstract
In a recent paper, Inmaculada de Melo-Martín and Kristen Intemann consider several ways in which, from the perspective of the argument from inductive risk, ethical and political values might "sometimes [be] necessary in decisions at the core of scientific reasoning." Specifically, they consider whether these kinds of values are logically, epistemically, pragmatically, or ethically necessary; and argue that there are significant conceptual problems in each case. In this comment, using regulatory uses of high-throughput toxicology at the US Environmental Protection Agency [EPA] as a case study, I suggest some clarifications and corrections to some of their claims about pragmatic necessity. I conclude that, while an inductive risk framework has some significant limitations, it is still conceptually and rhetorically valuable.

In a recent paper (de Melo-Martín and Intemann 2016), Inmaculada de Melo-Martín and Kristen Intemann consider several ways in which, from the perspective of the argument from inductive risk, ethical and political values might "sometimes [be] necessary in decisions at the core of scientific reasoning," that is, in the "gathering and characterization of evidence and/or appraisal and acceptance of hypotheses" (de Melo-Martín and Intemann 2016, 501. In what follows, unheaded citations are to this paper). Specifically, they consider whether these kinds of values are logically, epistemically, pragmatically, or ethically necessary; and argue that there are significant conceptual problems in each case. In this comment, I focus on their analysis of pragmatic necessity. Using regulatory uses of high-throughput toxicology at the US Environmental Protection Agency [EPA] as a case study, I suggest some clarifications, elaborations, and corrections to some of their claims about pragmatic necessity. I also briefly offer a general critique of a proposed response to the argument from inductive risk, made independently by Wendy Parker and Gregor Betz and discussed by de Melo-Martín and Intemann. I conclude that, while an inductive risk framework has some significant limitations, it is still conceptually and rhetorically valuable.

The argument from inductive risk is one of the major arguments discussed in the contemporary science and values literature. The argument is frequently associated with Heather Douglas (2000; 2009), but first received substantial discussion in the twentieth century by Richard Rudner (1953) and Carl Hempel (1965), and has been supported and expanded by several other contemporary philosophers of science (Wilholt 2012; Winsberg 2012; Parker 2014; Biddle 2016). In its simplest versions, inductive risk arguments claim, first, that when deciding whether or not to accept a hypothesis (data source, model, etc.), scientists should take into account the significance of making epistemic errors — the significance of accepting bad data or an unreliable model, or of rejecting good data or a reliable model.[1] Second, taking into account the significance of making epistemic errors requires taking into account the significance of the non-epistemic consequences of such errors. In the context of chemical safety, such consequences might include harming human health and the environment, or the economic consequences of regulation (Douglas 2000, 10–11). Third, the significance of non-epistemic consequences of epistemic error depends on social and political values. Thus, the argument concludes, scientists should take into account social and political values when deciding whether or not to accept a hypothesis. This might mean, for example, considering the non-epistemic consequences of making a type I vs. type

---

[1] An aphorism in statistics, attributed to George Box, is that "all models are wrong, but some are useful." That is, all statistical models depend on simplifications and assumptions, and it's highly plausible that at least some of these are strictly false in any given case; but, despite this, some models can still be useful for producing accurate estimates or drawing other correct inferences about the target population. So the simplest kind of inductive risk associated with accepting a model is not the risk of believing its false assumptions, but rather the risk of using a model that does not produce accurate estimates or that leads to other incorrect inferences. This is what I mean by "accepting an unreliable model."

II statistical error, and using these considerations to select acceptable error rates (compare Di Stefano 2001; Wahl and Ammann 2007, 40). This argument challenges the ideal of value-free science, which is generally defined as the view that social and political values have no legitimate role to play in the "core" stages of scientific research, such as gathering evidence and accepting or rejecting hypotheses.

In their analysis, de Melo-Martín and Intemann argue that challenges to the value-free ideal must satisfy what they call the *necessity condition*: Social and political values "are at least sometimes necessary in decisions at the core of scientific reasoning" (502). They consider, and challenge, several senses of "necessity" according to which an inductive risk perspective might be thought to satisfy the necessity condition. Due to limited space, here I focus on their discussion of "pragmatic necessity."

De Melo-Martín and Intemann characterize the "pragmatic necessity" of social and political values in the core of scientific research in terms of "advancing science and policy aims" (505). In regulatory agencies such as the EPA, science is done not (only) for its own sake — not (only) in order to satisfy researchers' curiosity or discover deep truths about the physical world, as might be the case with fields such as cosmology or particle physics — but primarily in order to promote the agency's mission of protecting human health and the environment. Philosophers of science generally agree that values, such as the aim of protecting human health and the environment, have a legitimate role to play in formulating research questions. At EPA, this is manifest in the mission of the Chemical Safety for Sustainability research program, which is developing methods to produce chemical safety data thousands of times faster than conventional animal-based methods at a fraction of the cost and without requiring the use of animals (US Environmental Protection Agency 2015). These methods are known as "high-throughput toxicology" [HTT]. While there are epistemic concerns about HTT — I'll discuss this more below — EPA believes that these methods have the potential to address the fact that we have no or limited data on the toxic effects of thousands of widely-used chemicals (R. Judson et al. 2009, 691, table 3). Without these data, the agency argues, we simply have no way to make empirically informed regulatory decisions that can protect human health and the environment. In other words, EPA's decision to develop HTT methods is guided explicitly by pragmatic considerations about how best to promote the agency's mission.

But a decision *to* develop HTT methods is not the same as decisions within the core scientific activity of *developing* these methods. Given that these methods will be developed, the goals of protecting human health and the environment might not be pragmatically necessary for making decisions about how to gather and characterize data, or how to evaluate and whether to accept a model. In other words, the ideal of value-free science can accept the use of social and political values to choose what research will be conducted; but rejects any role for these values within the research process itself. By contrast, inductive risk arguments are arguments that social and political values have a legitimate role to play even within the research process itself. For HTT, this would mean taking into account (at least) that some epistemic errors will lead to detrimental effects on human health and the environment (due to underregulation); others will lead to detrimental economic effects (due to overregulation); and then weighting the relative significance of these two types of error

in making decisions about (among other things) acceptable error rates. Thus, it seems that inductive risk considerations recommend incorporating social and political values into the core of HTT research in order to promote the EPA's mission; and in this sense these values seem to be pragmatically necessary.

De Melo-Martín and Intemann offer two arguments against this picture of the pragmatic necessity of social and political values in scientific research. First, they argue that "it is not obvious what would distinguish this position from the [value-free ideal]," because "it no longer appears that [social and political values] are playing a role in the core of scientific reasoning, but rather in decisions about which models, methodologies, or research strategies to pursue solely for the aim of continuing research or policy advising" (506). But this argument simply seems to miss the possibility sketched in the previous paragraph, on which social and political values play a role within the core scientific activity of developing HTT methods, and not only in the decision to develop these methods. Decisions about what error rates are acceptable are certainly within the core of scientific reasoning.

Following Mitchell (2004), it might be argued that there is a distinction between scientists acting "as policy advisors" and as "produc[ers] of knowledge" (506). But the explicit point of HTT research is to produce knowledge for the sake of advising policy. While Mitchell's distinction might be relevant to other cases (e.g., an astrophysicist who becomes Director of the National Science Foundation, and so whose research does not inform the policy decisions that she makes), it does not seem to apply to HTT researchers. Similarly, one might appeal to a distinction between belief and acceptance, according to which "whether a proposition is worthy of belief varies only on the amount of evidence available and is not based on other contextual factors because belief is always focused on arriving at truth," while a claim might be accepted (among other possible reasons) "because it serves as the best basis for formulating a particular regulatory policy" (Elliott and Willmes 2013, 811). But HTT researchers are focused on both arriving at truth (or, with respect to inductive risk, avoiding error) and informing regulatory policymaking. Indeed, truth is valuable because and insofar as it provides the best basis for regulatory policy. More generally, we cannot cleanly separate the epistemic and pragmatic aspects of HTT research — whether in terms of different research stages, different social roles, or different cognitive attitudes — not only because HTT research has both epistemic and pragmatic aims, but further because the pragmatic success of HTT research requires its epistemic success, and vice versa. Inductive risk arguments make this connection between pragmatic and epistemic concrete, in terms of the downstream social consequences of epistemic error.

De Melo-Martín and Intemann's second argument is that "when uncertainties exist, rather than using contextual value judgments to provide policy advising, scientists might instead utilize a plurality of alternative models, ranges of observational values, or parameters of statistical significance" (506-7). Here de Melo-Martín and Intemann borrow (without endorsing) proposals from Parker (2010) and Betz (2013). To develop, and ultimately respond to, this proposal, I will consider a particular technical-regulatory controversy involving EPA's HTT research.

In June 2015, EPA's Endocrine Disruptor Screening Program [EDSP] announced that they would begin to accept a particular HTT model — the "ER Model" — for regulatory purposes (Environmental Protection Agency 2015).[2] This decision was based on a series of scientific publications by a collaboration of researchers and regulators at EPA and the National Institute of Environmental Health Sciences [NIEHS], which argued that the ER Model could reproduce the results of conventional (i.e., non-HTT) "guideline-like" tests with a low false negative rate of 10.9%, i.e., a 10.9% chance of arriving at an erroneous conclusion that a chemical is not estrogenic when in fact it is (Kleinstreuer et al. 2015; Browne et al. 2015; R. Judson et al. 2015).

In public comments on the EDSP announcement, the Natural Resources Defense Council [NRDC], an environmental advocacy organization, argued that the ER Model had an unacceptably high false negative rate of 30.9% (Pullen 2015). This, of course, is much higher than the false negative rate calculated by the EPA-NIEHS collaboration. The disparity is due to different ways of handling 11 (out of 55 total) "inconclusive" chemicals that had low but non-zero scores in the ER Model. See table 1. The EPA-NIEHS collaboration counted these chemicals as "hits," that is, interpreted these chemicals as being estrogenic, and thereby counted them as true positives. NRDC, on the other hand, counted these same chemicals as "no-hits," that is, assumed that they would be interpreted as non-estrogenic, and thereby counted them as false negatives.

More precisely, the ER Model generates "scores" for chemicals, generally between 0 (no activity in any of the HTT assays) and 1 (the same level of activity as a reference estrogen). Browne et al. (2015) truncated scores below .001 to 0, on the grounds that a score of .001 corresponds to estrogenic activity only at a concentration "several orders of magnitude greater than the highest concentrations tested" (Browne et al. 2015, 8806); chemicals in this range are counted as negatives. (Note that *low* scores correspond to activity only at *high* concentrations, that is, *low* potency.) Chemicals with scores greater than or equal to .1 are counted as positives. The inconclusive chemicals are those with scores in the interval [.001, .1). Scores in this range indicate "limited activity and low potency" with

---

[2] Briefly, EDSP's mission is to screen chemicals as potential endocrine disruptors. EDSP uses a two-tier approach, where Tier 1 involves a relatively fast, inexpensive, but less reliable set of tests; chemicals that are judged to be potential endocrine disruptors based on Tier 1 are then sent on to Tier 2 for more rigorous scrutiny. Tier 1 involves several different subsets of tests, designed to check whether chemicals interact with distinct endocrine subsystems — whether they're estrogen mimics or testosterone inhibitors, for example. According to a chemical industry estimate, the established set of Tier 1 tests, despite being fast and inexpensive relative to Tier 2, still cost on the order of one million dollars and take three years (Becker and Tipaldo 2013). The July 2015 announcement was that the ER Model — a HTT model of the intracellular estrogen activity pathway — could be used in place of three Tier 1 tests. The tests required for this HTT model can be conducted much faster and less expensively than the conventional tests, and do not use animals.

"questionable relevance to in vivo bioactivity" (Browne et al. 2015, 8806). In other words, chemicals with ER Model scores in the inconclusive range appear to have a non-zero but low level of estrogenic activity. These scores are not inconclusive because of limited accuracy or precision, but instead because it is not clear whether these low levels of estrogenic activity are large enough to be concerning. Because their scores are non-zero, there is some reason to think that they are estrogenically active; but because their scores are low, there is some reason to think that this activity only takes place when someone is exposed to these chemicals in large quantities. Exposure estimates might determine that 95% of people will never be exposed to a given low-potency chemical in such large quantities (Wetmore et al. 2012); but in the US this still indicates that tens of millions of people might be exposed to large quantities of the chemical, and cumulative exposures to many low-potency chemicals might have effects similar to (or even greater than) exposures to high-potency chemicals.

Table 1. Three ways of calculating False Negative Rates [FNRs] in two sets of reference chemicals in Browne et al. (2015).

	in vitro	GL uterotrophic
# true positive	25	38
# inconclusive	1	11
# false negative	2	6
total w/o inconclusives	27	44
total w/ inconclusives	28	55
FNR w/o inconclusives	2/27 (7.4%)	6/44 (13.6%)
FNR (Browne et al. 2015)	2/28 (7.1%)	6/55 (10.9%)
FNR (NRDC)	3/28 (10.7%)	17/55 (30.9%)

"Inconclusive" chemicals had ER Model scores in the interval $[.001, .1)$*; scores below* $.001$ *were truncated to 0 and counted as negatives, and scores greater than or equal to* $.1$ *were counted as positives (Browne et al. 2015, 8806). The first FNR calculation ignores inconclusives completely. The second, used throughout Browne et al. (2015), interprets inconclusives as positives, thereby counting them as true positives. The third, used in the NRDC comments on the EDSP Pivot, interprets inconclusives as negatives, thereby counting them as false negatives. All chemicals used for these calculations were considered active based on a curated literature search (Kleinstreuer et al. 2015).*

I want to make six points about this disparity between EPA-NIEHS and NRDC. First, there is uncertainty in the calculation of the false negative rate for the ER Model. This uncertainty carries inductive risk, since the false negative rate estimates the reliability of the

ER Model for determining whether chemicals are estrogenic. But, second, this uncertainty is not due to limitations in the available evidence; instead, it is due to the need to make what Eric Winsberg calls "unforced methodological choices" (2012) about the range of activity and potency values that should be considered "relevant," that is, concerning. More accuracy or precision in the estimation of ER Model scores will not tell us what level of potency is sufficient to give us a reason to be concerned about a potential hazard. What's inconclusive about an inconclusive score is not the score itself, but rather its significance as a reason for regulatory action.

Third, decisions about how to interpret inconclusive chemicals — the decision about how to calculate the FNR — may have substantial downstream consequences for human health and the environment. NRDC's calculation is much more pessimistic about how the ER Model will be interpreted; and so it is highly plausible that their calculations are directly motivated by their concern to protect human health and the environment. But EPA-NIEHS's interpretation of inconclusive chemicals — which treats them all as positives — if adopted into policy, would be much more protective of these values than the interpretation assumed by NRDC. More generally, any way of calculating the false negative rate will have non-epistemic consequences for human health and the environment, whether or not these values play an explicit role in the decision to adopt one method or another.

Fourth, the second and third points together mean that this case provides an important counterexample or exception to Douglas' claim that "More evidence usually makes values less important ..., as uncertainty reduces" (Douglas 2009, 96). The uncertainty here is not a result of limited evidence; additional evidence (of the kind used by the ER Model) would simply improve the precision and accuracy of the measurements of the chemicals' low potency. But again, these scores are not inconclusive because of limited accuracy or precision, but instead because it's not obvious what range of potency values should trigger regulation. For instance, additional evidence that refines an ER Model score, say, from .01 to .00738 does not tell us whether scores in the interval [.001, .1) should be interpreted as potentially hazardous.[3] The values of protecting human health and the environment are needed to settle this indeterminacy; that is, these values are pragmatically

---

[3] Other kinds of studies might be used to narrow the range of inconclusive scores. For instance, exposure studies might be used to estimate the number of people who might be exposed to a given chemical beyond the level indicated by the ER Model score (Wetmore et al. 2012; Isaacs et al. 2016). However, such studies are themselves extremely complex research projects, which involve their own unforced methodological choices — about whom to survey and how, for instance — and raise their own inductive risk concerns — such as the possibility of over- or underestimating exposure. Thus, such studies have their own uncertainties, and combining them with the ER Model might, as it were, increase rather than decrease overall uncertainty (Sarewitz 2004).

necessary. (I do not consider here whether this one example is enough to challenge the "usually" in Douglas' claim.)

Fifth, in line with the proposal that de Melo-Martín and Intemann borrow from Parker and Betz, the EPA-NIEHS collaboration could have included multiple different false negative rate calculations in their analysis. Indeed, they did exactly this, considering several different sets of reference chemicals, several different accuracy statistics, and making calculations with and without the inconclusive chemicals (Browne et al. 2015, 8808, table 4). However, they did not include the particular calculation used by NRDC, primarily because it simply did not occur to them to do so; that is, it simply did not occur to them that inconclusive chemicals might be treated as false negatives. This interpretation of inconclusive chemicals was first suggested by NRDC, and there was no reason why the EPA-NIEHS collaboration should have considered the possibility beforehand.

Finally, in light of the preceding observations, I argue that the Parker/Betz proposal does not eliminate the pragmatic necessity for social and political values in the core of scientific reasoning. Taken in a very strong sense, the proposal would require that scientists anticipate and conduct numerous alternative analyses. But, from the perspective of both EPA and NRDC, such a requirement would indefinitely slow analysis and so delay regulation, which conflicts with the pragmatic aim of protecting human health and the environment. Even if the proposal were understood more modestly — say, as a requirement that scientists anticipate and conduct multiple analyses that cover a range of "reasonable" values and interests — it could still be quite demanding, leading to pragmatic conflict with the mission of protecting human health and the environment that both EPA and NRDC share. Simply put, the more analyses that EPA must conduct, the longer protective regulation is delayed. Following a recent suggestion from Steel (2016), this can be understood in terms of an epistemic error with potential harmful downstream consequences, that is, inductive risk; specifically, a stubborn withholding of acceptance/belief that potentially leads to preventable harm to human health or the environment.

Further, even the modest version of the Parker/Betz proposal depends on judgments about which values and interests are "reasonable." Inductive risk arguments can be applied to these judgments themselves, insofar as they are generalizations about the range of views and interests held by reasonable people. For consistency, the proposal would seem to require also working with a range of potential judgments about what counts as "reasonable," which in turn would depend on judgments about the range of reasonable criteria for judging reasonableness, and so on.

All together and in general, the Parker/Betz proposal requires at least two unforced methodological choices, regarding the number of alternative analyses to conduct and which alternative assumptions, methods, etc., to use in these analyses. These choices involve tradeoffs between breadth of coverage and other, social or political values (such as timely response to threats to human health and the environment), as well as judgments about which assumptions are "reasonable." Since both factors depend on social and political values, adopting the Parker/Betz proposal requires making decisions within the core of scientific reasoning that depend on social and political values. That is, social and political values are

pragmatically necessary for the Parker/Betz proposal. Thus, inductive risk considerations show that the Parker/Betz proposal is not a viable defense of the value-free ideal.

The preceding criticisms of de Melo-Martín and Intemann's discussion of "pragmatic necessity" do not mean that I think that inductive risk considerations exhaust the range of legitimate influences for social and political values in science. Indeed, my own full position is quite similar to their "aims approach," according to which "social, ethical, and political value judgments are legitimate in scientific decisions insofar as they promote democratically endorsed epistemological and social aims of the research" (516-7), and like them I am critical of Douglas' direct/indirect role distinction (509-10). In particular, I agree with Intemann and de Melo-Martín (personal communication) that, when the aims of protecting human health and the environment are used as a reason to adopt a certain method of calculating the false negative rate (accepting claims that this method is reliable and so on), this appears to be a direct use of values in the core of scientific inquiry.

However, while inductive risk considerations have their limitations, I find that they are still useful both analytically — as in my analysis of the disagreement between EPA and NRDC — and rhetorically — as a basic challenge to the value-free ideal for an audience of nonspecialists. Specifically, the notion of "pragmatic necessity" is useful for clarifying the relationship between the epistemic and social and political aims of scientific research: In many cases, such as at EPA, we pursue some epistemic aims for the sake of some other, social or political aims. In these cases, social and political aims appropriately regulate how and in what ways we pursue knowledge (compare Richardson 2002, 101): the EPA should use a false negative rate calculation that works to protect human health and the environment. Many portrayals of wishful thinking and politicized science portray epistemic and social and political aims as essentially antagonistic, but "pragmatic necessity" lets us recognize that often the pursuit of truth is properly informed by its utility for promoting further social and political ends. Pragmatic necessity does not seem to have the conceptual problems that de Melo-Martín and Intemann claim; so in this sense of necessity an inductive risk framework can hold that social and political values are necessary even in the core of scientific inquiry; and so an inductive risk framework can satisfy the necessity condition and reject the value-free ideal.

References

Becker, Richard, and Emily Tipaldo. 2013. "Comment Submitted by Emily Tipaldo, American Chemistry Council." *Regulations.gov* public comment. *https://www.regulations.gov/document?D=EPA-HQ-OPP-2013-0075-0009*.

Betz, Gregor. 2013. "In Defence of the Value Free Ideal." *European Journal of Philosophy of Science* 3 (2): 207–20. doi:*10.1007/s13194-012-0062-x*.

Biddle, Justin. 2016. "Inductive Risk, Epistemic Risk, and the Overdiagnosis of Disease." *Perspectives on Science* 24 (2): 192–205. doi:*10.1162/POSC_a_00200*.

Browne, Patience, Richard Judson, Warren Casey, Nicole C. Kleinstreuer, and Russell Thomas. 2015. "Screening Chemicals for Estrogen Receptor Bioactivity Using a Computational Model." *Environmental Science and Technology* 49 (14): 8804–14. doi:*10.1021/acs.est.5b02641*.

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

Di Stefano, Julian. 2001. "Power Analysis and Sustainable Forest Management." *Forest Ecology and Management* 154 (1-2): 141–53. doi:*10.1016/S0378-1127(00)00627-7*.

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

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

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

Environmental Protection Agency. 2015. "Use of High Throughput Assays and Computational Tools; Endocrine Disruptor Screening Program; Notice of Availability and Opportunity for Comment." *Federal Register* 80 (118): 35350–5.

Hempel, Carl. 1965. "Science and Human Values." In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, 81–96. New York: The Free Press.

Isaacs, Kristin K., Michael-Rock Goldsmith, Peter Egeghy, Katherine Phillips, Raina Brooks, Tao Hong, and John F. Wambaugh. 2016. "Characterization and Prediction of Chemical Functions and Weight Fractions in Consumer Products." *Toxicology Reports* 3: 723–32. doi:*10.1016/j.toxrep.2016.08.011*.

Judson, Richard, Felicia Maria Magpantay, Vijay Chickarmane, Cymra Haskell, Nessy Tania, Jean Taylor, Menghang Xia, et al. 2015. "Integrated Model of Chemical Perturbations of a Biological Pathway Using 18 in Vitro High-Throughput Screening Assays for the Estrogen Receptor." *Toxicological Sciences* 148 (1): 137–54. doi:*10.1093/toxsci/kfv168*.

Judson, Richard, Ann Richard, David J. Dix, Keith Houck, Matthew Martin, Robert Kavlock, Vicki Dellarco, et al. 2009. "The Toxicity Data Landscape for Environmental Chemicals." *Environmental Health Perspectives* 117 (5): 685–95. doi:*10.1289/ehp.0800168*.

Kleinstreuer, Nicole C., Patricia C. Ceger, David G. Allen, Judy Strickland, Xiaoqing Chang, Jonathan T. Hamm, and Warren M. Casey. 2015. "A Curated Database of Rodent Uterotrophic Bioactivity." *Environmental Health Perspectives* 124 (5): 556–62. doi:*10.1289/ehp.1510183*.

Mitchell, Sandra. 2004. "The Prescribed and Proscribed Values in Science Policy." In *Science, Values, and Objectivity*, edited by Peter Machamer and Gereon Wolters, 245–55. Pittsburgh: University of Pittsburgh Press.

Parker, Wendy. 2010. "Whose Probabilities? Predicting Climate Change with Ensembles of Models." *Philosophy of Science* 77 (5): 985–97. doi:*10.1086/656815*.

———. 2014. "Values and Uncertainties in Climate Prediction, Revisited." *Studies in the History and Philosophy of Science* 46 (June): 24–30. doi:*10.1016/j.shpsa.2013.11.003*.

Pullen, Kristi. 2015. "Comment Submitted by Kristi Pullen, Ph.D., Staff Scientist, Natural Resources Defense Council (NRDC)." *Regulations.gov* public comment. *https://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2015-0305-0024*.

Richardson, Henry. 2002. *Democratic Autonomy*. Oxford University Press.

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

Sarewitz, Daniel. 2004. "How Science Makes Environmental Controversies Worse." *Environmental Science and Policy* 7 (5): 385–403. doi:*10.1016/j.envsci.2004.06.001*.

Steel, Daniel. 2016. "Climate Change and Second-Order Uncertainty: Defending a Generalized, Normative, and Structural Argument from Inductive Risk." *Perspectives on Science* 24 (6): 696–721. doi:*10.1162/POSC_a_00229*.

US Environmental Protection Agency. 2015. "Chemical Safety for Sustainability: Strategic Action Plan 2016-2019." US Environmental Protection Agency. November. *http://www.epa.gov/sites/production/files/2016-02/documents/strap_2016_css_508_compliant_final_optimized_0.pdf*.

Wahl, Eugene, and Caspar Ammann. 2007. "Robustness of the Mann, Bradley, Hughes Reconstruction of Northern Hemisphere Surface Temperatures: Examination of Criticisms Based on the Nature and Processing of Proxy Climate Evidence." *Climatic Change* 85 (1-2): 33–69. doi:*10.1007/s10584-006-9105-7*.

Wetmore, Barbara A., John F. Wambaugh, Stephen S. Ferguson, Mark A. Sochaski, Daniel M. Rotroff, Kimberly Freeman, Harvey J. Clewell, et al. 2012. "Integration of Dosimetry, Exposure, and High-Throughput Screening Data in Chemical Toxicity Assessment." *Toxicological Sciences* 125 (1): 157–74. doi:*10.1093/toxsci/kfr254*.

Wilholt, Torsten. 2012. "Epistemic Trust in Science." *British Journal for the Philosophy of Science* 64 (2): 233–53. doi:*10.1093/bjps/axs007*.

Winsberg, Eric. 2012. "Values and Uncertainties in the Predictions of Global Climate Models." *Kennedy Institute of Ethics Journal* 22 (2): 111–37. doi:*10.1353/ken.2012.0008*.