Journal for General Philosophy of Science
https://doi.org/10.1007/s10838-022-09606-5

DISCUSSION

Addressing the Reproducibility Crisis: A Response to Hudson

Heather Douglas[1] · **Kevin C Elliott**[2]

Accepted: 10 February 2022

Abstract
In this response to Robert Hudson's article, "Should We Strive to Make Science Bias-Free? A Philosophical Assessment of the Reproducibility Crisis," we identify three ways in which he misrepresents our work: (1) he conflates value-ladenness with bias; (2) he describes our view as one in which values are the same as evidential factors; and (3) he creates a false dichotomy between two ways that values could be considered in science for policy. We share Hudson's concerns about promoting scientific reproducibility and reducing bias in science, but we reject his view that the value-free ideal provides helpful guidance for addressing these issues.

1 Introduction

Although we share Robert Hudson's concerns about the importance of promoting scientific reproducibility and reducing bias in science, we reject his claim that our views about the role of values in science would exacerbate reproducibility problems or lead to a neglect of bias. To the contrary, we consider concerns about reproducibility and bias in science to be central issues even while we reject the value-free ideal. In this response, we argue that Hudson misrepresents our work in three major ways: (1) he conflates value-ladenness with bias; (2) he describes our view as one in which values are the same as evidential factors; and (3) he creates a false dichotomy between two ways that values could be considered in science for policy. After addressing these sources of confusion, we clarify how we would recommend responding to the reproducibility crisis.

Heather Douglas
dougl239@msu.edu

Kevin C Elliott
kce@msu.edu

1 Department of Philosophy, Michigan State University, Michigan, United States

2 Lyman Briggs College, Department of Fisheries and Wildlife, and Department of Philosophy, Michigan State University, Michigan, United States

Published online: 11 May 2022

2 Value-Ladenness vs. Bias

One of Hudson's most important confusions is his conflation of bias and value-ladenness. After describing our view that non-epistemic values can legitimately influence scientific reasoning, he states, "It follows [from this view] that the scientists we have been describing who strive to remove the influence of bias on statistical thinking are somewhat naive in their aspirations: bias is not only inevitable in statistical analysis but should be embraced" (p. 22). We reject Hudson's inference that if we accept roles for non-epistemic values in scientific reasoning, we must embrace bias in science. On our view, one should reject the value-free ideal while insisting that scientists should do all they can to eliminate biases from their work. This is a coherent position because bias and value-ladenness are not the same thing.

Hudson does not clearly define bias in his paper, and this makes it easier for him to conflate value-ladenness with bias. We embrace a standard and precise definition of bias, as a systematic deviation from truth (or from accuracy). One can see this definition as underlying John Ioannidis's classic paper on bias and error in science (Ioannidis 2005), although he does not provide a precise definition of bias in that paper. This is also the definition used in the Cochrane Collective (2008, 8.2.1), which describes bias as "a systematic deviation from the truth, in results or inferences of studies."

With this definition, it would be incoherent to say one is biased towards the truth (contra May 2020). Bias is a tendency to lean away from the truth, and in a particular direction. If bias is defined more loosely, as a tendency either towards or away from the truth, it is no longer a term that should invoke epistemic worry or that marks "an epistemic shortcoming of the research in question" (Wilholt 2009, 92). If bias is defined as any movement away from truth, including nonsystematic error, it is not distinguishable from noise. Such looseness would muddy rather than clarify the concerns about bias.

With this more precise definition of bias, it is rather straightforward to show that bias and value-ladenness are conceptually distinct. While values can bias scientific research in some cases, many biases have nothing to do with values and much value-ladenness has directed science towards truth rather than away from it. Let us consider each of these possibilities.

2.1 Bias without Value-Ladenness

In the classic work on cognitive biases by Kahneman and Tversky (Tversky and Kahneman 1974), cognitive biases are detected that have no clear relationship to values. Cognitive biases such as availability heuristics and anchoring biases systematically deviate judgment away from more accurate judgments. These heuristics are cognitive shortcuts that happen without deeper reflection on the information presented, systematically driving judgment away from a more accurate assessment. As such, it is difficult to consider such biases the results of values affecting judgments.

2.2 Value-Ladenness Without Bias

Values influencing science can and have debiased science. The value of anti-sexism, the value of taking the perspective of women seriously, has worked against biases in science, for example. In contexts as diverse as primatology (Haraway 1989), medical research (Rosser 1994), and social science (Anderson 2004; Wylie 2002), having an explicit commitment to work against sexism in science and society has helped to create more accurate science and to fight against pervasive biases. We had more accurate understandings of primate behavior once primatologists began tracking the behavior of female primates as carefully as male primates (Haraway 1989). We developed more accurate understandings of how to detect heart disease in women once anti-sexist values drove attention to women's different experiences (Rosser 1994). We gained a more accurate understanding of the impact of divorce on families once feminist researchers asked about both positive and negative impacts of divorce (Anderson 2004). We now have a richer and more accurate understanding of the archaeological record (Wylie 2002). In these cases (as in many others, see e.g., Keller and Longino 1996), values helped to debias science.

Hudson might agree that value-ladenness can debias science in the early stages of science (i.e., deciding which studies to do and how to do them) while insisting that value-ladenness causes bias when it affects inference in science. We contend that even when values influence scientific inference, it does not necessarily cause bias. Consider the role of values in addressing inductive risk. It is important to keep in mind that when values influence science through inductive risk, they do so only by playing a role in addressing uncertainty. On Douglas's view, values can only tell us about the acceptability of uncertainty, or the sufficiency of evidence, if values are to be acting properly through inductive risk considerations (Douglas 2009; 2017). This means we do not know where the truth yet lies, or what is more accurate, when weighing the impacts of errors and their acceptability. By rejecting a claim or withholding judgment, one might be missing an opportunity to affirm a conclusion that is true, or one might be saving oneself from adopting a falsehood. It is thus only with hindsight, once we know more about what the correct assessment of evidence is, that we could say that values biased science through the consideration of inductive risk. In the middle of a dispute, whether scientists should find the existing evidence sufficient or demand stronger evidence is not determinable. At the very least, inductive risk-based disputes do not systematically drive scientists away from the truth. (Admittedly, Hudson might object that considerations of inductive risk can still drive scientists away from the truth in the limited sense that it allows them to lower their standards of evidence and thus potentially accept false claims, but this would presume that failing to adopt a potentially true claim is not a problem. We address this objection in more depth in the final section of our response.)

So what can we say about values and their potential for debiasing science when inductive risk is at issue? Divergent values can drive expert disagreement and foster more robust debate in science as experts consider how to respond to inductive risk. This process of debate and the further gathering of evidence it produces should narrow uncertainties and help scientists gain a more accurate understanding of the phenomena at issue. This occurs as long as scientists are acting with integrity (in the sense of Douglas 2014) or are not intransigently biased (Holman and Bruner 2015). If there are intransigently biased scientists distorting the debate, values can bias science, as we discuss next.

2.3 Value-Ladenness Producing Bias

In cases where scientists decide ahead of time what they want the outcome of a study to be, make methodological choices to ensure that outcome, and the outcome is wrong, biased science results. Value concerns are often central to such bias. In the cases discussed by Wilholt (2009, 93), for example, scientists looking to protect industry interests (and profits) designed and conducted studies to ensure no worrisome toxicological outcomes were found. Other scientists without such strong interests did find positive, worrisome results. Similar accounts can be found in Oreskes and Conway (2010). There are plenty of clear cases where value commitments have produced biased work in science, in the sense of systematically deviating from the truth. Value judgments can bias science.

Considering all three relationships between values and science discussed in this section, however, it is clear that value-ladenness is conceptually distinct from bias. Values in science can bias science, but they can also debias science or foster debate that debiases science. We hope now to have clarified why it is a conceptual mistake to conflate bias with value-ladenness.

Rejecting the value-free ideal neither embraces bias nor displays a lack of concern for bias. We are among those who reject the value-free ideal as a good ideal for science but also have a concern that science not be biased. Douglas, for example, has argued against bullshit that biases science and for integrity which guards against bias in science (Douglas 2006; 2014). Elliott has also proposed strategies for preventing values from biasing science (Holman and Elliott 2018). Indeed, given the epistemically positive role social and ethical values have played (and likely will continue to play) in the development of science, paying more attention to values in science, and not just to eliminate them, will help fight bias. It is not the only way to fight bias in science, however, as we discuss further below.

3 Values and Evidence

In addition to Hudson's fundamental confusion about how we view the relationship between biases and value-ladenness, a second major problem is his description of our views about the relationship between values and evidence. When presenting Douglas's view, he says that values play the role of "filling the evidential gap" (p. 22) and that they act as "indirect evidential factors" (p. 23). Later, he states in his conclusion that on our view, "moral or social values are needed to 'fill the relevant evidential gap'" (p. 27) when uncertainty is present. He also says we think that "the moral, social or other non-epistemic consequences of scientific decisions are themselves evidential in deciding what conclusions to draw" (p. 28).

This is not an accurate description of our views. Douglas has been very careful to distinguish the role of values from the role of evidence. The reason she is so emphatic about her distinction between direct and indirect roles for values is that she does *not* think values serve as evidence; instead, they establish the *amount* of evidence needed in order to accept a scientific conclusion (see e.g., Douglas 2009; Douglas 2017). Thus, it is incorrect to say that she sees values as "evidential" or that they could fill an evidential gap. She has consistently avoided describing values as playing this role in scientific inference. She draws a sharp distinction between the role of values and the role of evidence; evidence provides reasons to think that a conclusion is true or reliable, whereas values determine how much evidence should be demanded in order to accept a claim. Hudson could perhaps insist that values "fill evidential gaps" on Douglas's view in the limited sense that they establish which claims can be accepted on the basis of the evidence at a given time, but Douglas would resist this description because it could generate the confused impression that she takes values to be a form of evidence. It is ironic that Hudson describes Douglas's view as one in which values play an evidential role because she has been challenged by others in the values-and-science literature for maintaining such a sharp distinction between values and evidence (see e.g., Bluhm 2017; Brown 2013).

As Hudson notes, Elliott's views on the evidential status of values are somewhat more complex. He agrees with Douglas that it is better in situations of inductive risk not to regard values as playing an evidential role, and he would agree that it could generate confusion to say that they "fill evidential gaps" in such cases. However, Elliott is also interested in cases where values help to achieve the goals or aims of scientific inquiry (Elliott and McKaughan 2014). In some of these situations, values could play something like an evidential role in the sense that they provide reasons for thinking that a hypothesis or model is adequate for a particular purpose (see e.g., Parker 2020). For example, if scientists were developing a particular risk assessment method, the fact that the method prioritized public health could serve as part of the reason for thinking that it is the best method for achieving particular regulatory purposes (Fernández Pinto and Hicks 2019). This is somewhat beside the point of Hudson's article, though, because Hudson is focused on cases in which values are playing an indirect role in addressing inductive risk. Moreover, as we discuss below, values are still playing a pragmatic role in these cases (e.g., providing a reason for thinking that a method or hypothesis is *adequate for a purpose*) rather than serving as evidence that a hypothesis is *true*.

4 A False Dichotomy

This unfortunate way of describing our views about evidence is closely related to a third source of confusion, which involves the ways we understand science-policy relationships. According to Hudson, there are two ways in which one might import ethical or social values into science. First, one might "vouch" for hypotheses for pragmatic reasons, "without the further claim that these pragmatic reasons provide *evidence* for the hypotheses" (p. 25). Second, one might think that pragmatic reasons actually do provide evidence for hypotheses, such that they "provide a reason to believe, or justify, an [sic] hypothesis" (p. 25). He attributes the second view to us: "the view of Douglas and Elliot [sic] is *not* that when evidence is uncertain scientists leave it to policy makers to make decisions on a pragmatic basis – that is the viewpoint of those who uphold the value-free ideal. Rather, their view is the controversial one, that the moral, social or other non-epistemic consequences of scientific decisions are themselves evidential in deciding what conclusions to draw" (p. 28). There are two ways to critique this passage. One is already articulated above—that Hudson has misread our work on the relationship between values and evidence in science. The other is to notice that Hudson has created a false dichotomy, with policy makers making pragmatic decisions to accept hypotheses because of values as one option and scientists making evidential decisions to believe hypotheses because of values as the other option. We take a middle position: we think that scientists do need to draw conclusions on the basis of ethical and social values—they cannot leave these decisions solely to policy makers (Douglas 2009; Elliott 2011)—but we think scientists should treat these decisions as pragmatic rather than as evidential. Further, we argue that scientists should be clear about the value judgments used to come to their conclusions, so that decision-makers can more effectively use their scientific advice (Douglas 2009; 2021; Elliott and Resnik 2014).

Elliott is particularly explicit about the fact that scientists should not regard values as providing a reason to think that a theory is true. With David Willmes, he draws a distinction between two cognitive attitudes: belief and acceptance (Elliott and Willmes 2013). They argue that different considerations are relevant when taking these two different cognitive attitudes toward hypotheses. If scientists are considering whether to accept a hypothesis as a basis for policy making in a particular domain, then various ethical or social values are relevant. In contrast, "if … scientists are considering whether or not to believe that the hypothesis is true, ethical values are logically irrelevant and have no legitimate role to play because belief is directed only toward achieving truth" (Elliott and Willmes 2013). Thus, Elliott and Hudson differ not in their views about the evidential role of values but rather in their views about whether responsible scientists need to make pragmatic decisions about which hypotheses to accept (Elliott 2011, 67–68). Elliott thinks it is important to engage in a careful exploration of the range of cognitive attitudes appropriate to various areas of science because these sorts of pragmatic decisions permeate scientific work (Elliott and McKaughan 2015). Although Douglas has not written as much about cognitive attitudes, she emphasizes that her analysis of inductive risk is focused on situations in which scientists are "putting forward" or propounding claims for policy makers; she is interested in the ethical responsibilities of scientists who are considering "what to say" or "which claims to make" (Douglas 2009, 66). Thus, both Elliott and Douglas are focused on pragmatic decisions about what to *do* or *say* or *accept* rather than on epistemic decisions about what to *believe* (see also Franco 2017).

5 Addressing Concerns About Reproducibility

Hudson might still insist that all these clarifications fail to save us from his fundamental objection, which is that our willingness to adjust statistical standards for ethical or social reasons contributes to the reproducibility crisis. For example, he describes a psychological experiment involving the phenomenon of social priming. In the experiment, the research subjects were found to leave the laboratory slower after being primed with words associated with elderly stereotypes. Hudson notes that this experiment was not replicated, but he worries that we could justify lowering our standards of evidence and accepting the results of the original experiment if it served important social values.

We find this to be a strange critique of our views. As Douglas has argued, there is no "magic" statistical significance level that scientists are obliged to use in order to maintain their objectivity (Douglas 2017). No matter what level of evidence we were to recommend that scientists demand, Hudson could potentially criticize us by insisting that if one demanded a higher level of evidence it would help to prevent problems with reproducibility. To take this to the extreme, one could insist that scientists should never accept any claim because they never have absolute proof that they are correct. This would solve the reproducibility crisis because scientists would never make claims that later turn out to be false, but it would render science useless as a source of guidance for decision makers. Any level of evidence below this extreme could potentially be criticized. Based solely on epistemic considerations, we do not think there are compelling reasons to specify any particular standard of evidence as "correct." Thus, we think it is more important for scientists to be open and explicit about their methods, including statistical significance levels, rather than setting universal standards of evidential sufficiency (Douglas 2008; 2009; Elliott 2017).

But perhaps there is a more charitable way of putting Hudson's critique. Perhaps the real problem is that we are willing to let scientists depart from the conventional standards of evidence typically demanded in their fields. On this view, the worry is not that scientists are picking any particular standard but rather that they could confuse people by varying the level of evidence in support of their claims. This worry is fueled by the concern that it could be difficult to devise systems of openness or transparency that completely prevent confusion when scientists vary their standards of evidence. This is a reasonable worry that has been explored by others (e.g., John 2015). Elliott (2017, 97–98) has previously acknowledged that this worry could potentially provide reasons for conventionally maintaining fixed standards of evidence in some scientific contexts (see also Wilholt 2009). It is important to note, however, that both Elliott and John (2015) emphasize that these conventional standards of evidence would be justified based on social values, including Hudson's worries about reproducibility and the concern that scientific conclusions be usable and understandable for members of the public. Thus, far from abandoning the scientific community to the reproducibility crisis, we would consider concerns about reproducibility to be among the important social values that need to be considered once the value-free ideal for science is abandoned.

At the same time, we remain unconvinced that the most promising approach for addressing publication bias is to maintain fixed standards of evidence. We think a range of other approaches are more important to implement. Hudson discusses registered reports as one strategy that the scientific community has been discussing. This approach accords very well with our own emphases on maintaining openness about the methods and values that inform research (Douglas 2008; 2009; Elliott 2020; Elliott and Resnik 2014). Along these lines, Elliott has discussed the benefits of trial registries in previous work (Holman and Elliott 2018). Registered reports also have the advantage of providing careful peer review and feedback on methodologies before a study has been performed and resources have been committed. Review at the stage of methodology could also bring more attention to the power of studies, rather than the focus on the statistical significance of results. Finally, registered reports could help alleviate some of the publication biases that result from the tendency of scientific journals to want to publish statistically significant results. Whether or not registered reports will successfully disincentivize p-hacking or alleviate the placement of negative results in the dustbin rather than in a journal, we both commend the pursuit of this potentially valuable change in publication culture. Nothing in our work implies a rejection of registered reports.

6 Conclusions

We take bias (precisely defined) as a serious challenge for the social practices of science. We think it can be best addressed not by adhering to the value-free ideal, but rather by accepting the importance of social and ethical values for scientific practice, by acknowledging how those values can both reduce bias and produce bias in science, and by diagnosing the causes of bias more carefully. Some of those causes will indeed be traceable to values in science, and when they do, we need to assess carefully why the values drove scientists away from the truth. Debates over how best to define and manage conflicts of interest, how to mitigate career pressures to publish, or how to decide which research to fund (so that conservatism in science does not bias scientific results and so that resources are not wasted) are much more appropriate debates to have to address bias in science. The value-free ideal does not help clarify these issues and does not provide appropriate normative guidance for thinking about them. With bias and value-ladenness conceptually distinguished, we can more effectively address bias in science.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10838-022-09606-5.

References

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: 1–24.
Bluhm, R. 2017. Inductive Risk and the Role of Values in Clinical Trials. In *Exploring Inductive Risk: Case Studies of Values in Science*, edited by K. Elliott and T. Richards, 193–212. New York: Oxford University Press.
Brown, M. 2013. Values in Science beyond Underdetermination and Inductive Risk. *Philosophy of Science* 80: 829–839.
Cochrane Collective. 2008. *Cochrane Handbook: Version 5.0.2*. New York: Wiley and Sons.
Douglas, H. 2006. Bullshit at the Interface of Science and Policy: Global Warming, Toxic Substances, and Other Pesky Problems. In *Bullshit and Philosophy*, edited by G. L. Hardcastle and G. A. Reisch, 215–228. Chicago: Open Court Press.
Douglas, H. 2008. The Role of Values in Expert Reasoning. *Public Affairs Quarterly* 22: 1–18.
Douglas, H. 2009. *Science, Policy, and the Value-Free Ideal*. Pittsburgh: University of Pittsburgh Press.
Douglas, H. 2014. Scientific Integrity in a Politicized World. In *Logic, Methodology, and Philosophy of Science: Proceedings of the Fourteenth International Congress*, edited by Peter Schroeder-Heister, Gerhard Heinzmann, Wilfrid Hodges, and Pierre Edouard Bour, 253–268. London: College Publications.
Douglas, H. 2017. Why Inductive Risk Requires Values in Science. In *Current Controversies in Values and Science*, eds. K. Elliott, and D. Steel, 81–93. New York: Routledge.
Douglas, H. 2021. The role of scientific expertise in democracy. In *Routledge Handbook of Political Epistemology*, edited by Jereon de Ridder and Michael Hannon, pp. 435–445, New York: Routledge.
Elliott, K. 2011. *Is a Little Pollution Good for You? Incorporating Societal Values in Environmental Research*. New York: Oxford University Press.
Elliott, K. 2017. *A Tapestry of Values: An Introduction to Values in Science*. New York: Oxford University Press.
Elliott, K. 2020. A Taxonomy of Transparency in Science. *Canadian Journal of Philosophy*. Available at: https://doi.org/10.1017/can.2020.21.
Elliott, K., and D. McKaughan. 2014. Non-Epistemic Values and the Multiple Goals of Science. *Philosophy of Science* 81: 1–21.
Elliott, K., and D. McKaughan. 2015. Introduction: Cognitive Attitudes and Values in Science. *Studies in History and Philosophy of Science* 53: 57–61.
Elliott, K., and D. Resnik. 2014. Science, Policy, and the Transparency of Values. *Environmental Health Perspectives* 122: 647–650.
Elliott, K., and D. Willmes. 2013. Cognitive Attitudes and Values in Science. *Philosophy of Science* 80 (2013 Proceedings): 807–817.
Fernández Pinto, M., and D. J. Hicks. 2019. Legitimizing Values in Regulatory Science. *Environmental Health Perspectives* 127: 035001.
Franco, P. 2017. Assertion, Nonepistemic Values, and Scientific Practice. *Philosophy of Science* 84: 160–180.
Haraway, D. 1989. *Primate Visions: Gender, Race, and Nature in the World of Modern Science*. New York: Routledge.
Holman, B., and J. Bruner. 2015. The Problem of Intransigently Biased Agents. *Philosophy of Science* 82: 956–968.
Holman, B., and K. Elliott. 2018. The Promise and Perils of Industry-Funded Science. *Philosophy Compass* 13: e12544.
Ioannidis, J. P. A. 2005. Why Most Published Research Findings Are False. *PLoS Med* 2(8): e124. https://doi.org/10.1371/journal.pmed.0020124.
John, S. 2015. Inductive Risk and the Contexts of Communication. *Synthese* 192: 79–96.

Keller, E. F., and H. Longino, eds. 1996. *Feminism and Science*. New York: Oxford University Press.
May, J. 2020. Bias in science: Natural and social. Synthese. https://doi.org/10.1007/s11229-020-02937-0.
Oreskes, N., and E. Conway. 2010. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. New York: Bloomsbury.
Parker, W. 2020. Model Evaluation: An Adequacy-for-Purpose View. *Philosophy of Science* 87: 457–477.
Rosser, S. V. 1994. *Women's health–Missing from US medicine*. Bloomington: Indiana University Press.
Tversky, A., and D. Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science* 185: 1124–1131.
Wilholt, T. 2009. Bias and Values in Scientific Research. *Studies in History and Philosophy of Science* 40: 92–101.
Wylie, A. 2002. *Thinking from Things: Essays in the Philosophy of Archaeology*. Berkeley: University of California Press.