Cooperative Epistemic Trustworthiness

Marina DiMarco

Department of History and Philosophy of Science, University of Pittsburgh, Pittsburgh, PA, USA
Email: marina.dimarco@pitt.edu

(Received 11 January 2023; revised 17 March 2023; accepted 22 March 2023; first published online 27 March 2023)

Abstract

Extant accounts of trust in science focus on reconciling scientific and public value judgments, but neglect the challenge of learning audience values. I argue that for scientific experts to be epistemically trustworthy, they should adopt a cooperative approach to learning about the values of their audience. A cooperative approach, in which expert and nonexpert inquirers iteratively refine value judgments, better achieves important second-order epistemic dimensions of trustworthiness. I suggest that strong objectivity in the feminist standpoint theoretic sense is sometimes a prerequisite for trustworthiness.

1. Introduction

Scientists are often beholden to the aims of nonscientific audiences, whether these are generic democratic publics, specific communities that are uniquely affected by scientific research, or policymakers who depend on evidence-based recommendations. Most of these cases of epistemic dependence are asymmetric, or cases in which one set of knowers is dependent on another for information. Ideally, this asymmetric epistemic dependence is grounded not in unthinking deference but in considerations of epistemic trust (Hardwig 1991).

On most accounts, epistemic trustworthiness has two important features: It depends upon both an expert’s competence and orientation toward their audience (Baier 1986; Scheman 2001). By contrast, the role of audience values in the trustworthiness of scientific inquiry has been largely limited to decisions about responsibly placing or withholding trust, or to reconciling audience values with those of scientists (Wilholt 2013; Alexandrova 2018; Grasswick 2018). I argue that for scientific experts to be epistemically trustworthy, they should adopt a cooperative approach to learning about the values of their audience, in which expert and nonexpert inquirers iteratively refine value judgments about the aims of inquiry.

This argument has two parts. First, I argue that trustworthiness requires that experts have good reasons to think they know audience values and aims for inquiry. This follows from a second-order epistemic dimension of trustworthiness: experts must consider the limits of their own knowledge and whether it is well suited to the
task at hand (Daukas 2006, 2011). Second, I argue that these reasons are best achieved using a cooperative approach to characterizing values. A cooperative approach, in which expert and nonexpert inquirers iteratively refine value judgments, provides concrete methodological guidance as to how to achieve important second-order epistemic dimensions of trustworthiness. It also complicates the relationship between objectivity and trustworthiness. Whereas Scheman (2001) and Douglas (2009) seem to take trustworthiness as a precondition for objectivity, I argue that strong objectivity in the feminist standpoint theoretic sense is sometimes a prerequisite for trustworthiness.

2. Values-matching accounts
Philosophers of science have interpreted trustworthiness in terms of the legitimacy of scientific practices (John 2018); the honesty and reliability of scientific communication (Irzik and Kurtulmus 2019); the social conditions of science (Irzik and Kurtulmus 2019; Rolin 2020); scientists’ value judgments about the inductive risk (Wilholt 2013); and broader value judgments about scientific measurements and methodologies (Alexandrova 2018; Schroeder 2019). After Schroeder (2019) I’ll call these values-matching accounts, because they appeal to varying degrees to the idea that either the value judgments of trustworthy science should match those of an audience, or that the mismatch should be accounted for in some politically legitimate manner.

I want to highlight two important features of values-matching accounts. First, they often secure the trustworthiness of science by characterizing the role of values in scientific projects that are taken as given. For Wilholt, public values enter into methodological judgments about how to value the consequences of inquiry and the reliability of its results; for Alexandrova, they enter into decisions about how to operationalize concepts in “mixed claims” with normative and empirical content; and for Schroeder, they are explicitly confined to “the scientific process” and forbidden from influencing the choice of research project (2019, 554). These roles for values in securing trustworthiness are indexed to specific methodological questions. Second, values-matching accounts focus on aligning values that are sometimes uninterrogated. In section 3, I’ll show that trustworthy experts cannot afford to take the values of their audience for granted. Instead, second-order epistemic considerations inherent to trustworthiness demand that they actively negotiate the aims of inquiry: an important role for social values in science. Insofar as values-matching matters for trustworthiness, it is downstream of these considerations.

3. Trustworthiness and the aims of science
To anchor this argument, I will lean on Nancy Daukas’s (2006, 2011) feminist account of epistemic trustworthiness. The most important feature of this account is its requirement that trustworthy scientists attend to second-order epistemic considerations, that is, that they assess the limits of their knowledge and its relative suitability in a particular context.

1 I think we could draw similar lessons from Frost-Arnold (2014) or from Grasswick’s (2018) sincerity care condition for trust.
3.1 Second-order epistemic trustworthiness

For Daukas, one’s trustworthiness as a testifier depends not only on sincerity and accuracy, but also on one’s second-order assessment of one’s own knowledge: judging the salience and limits of what we might know in a particular context. If I hold forth on a topic I know relatively little about, I behave in an untrustworthy way even if I am accurate and sincere. To avoid this, Daukas recommends a self-reflexive approach to epistemic trustworthiness, which asks experts to responsibly characterize and steward their own epistemic resources. On this account, being trustworthy requires not only that experts make some reliable assessment about the limits of their own knowledge, but also that they have reason to think they are relatively well positioned to complete the epistemic task in question. This is a second-order dimension of epistemic trustworthiness: experts should attend not only to what they know but also to whether they have good reasons to think that they know it.

I endorse this account, and interpret it as demanding much more with respect to the knowledge we need to be trustworthy than some values-matching approaches seem to require. This is because in order to know whether our knowledge is well suited to a task, we need to have reasons to think we know what the task is. This is both prior to, and broader than, matching the values of scientists and the public(s) with respect to individual methodological choices within the scientific process. It requires that scientists learn about the values of their audience in ways largely neglected by values-matching approaches. Specifically, trustworthy experts may need to play an active role in negotiating value judgments about the aims of inquiry and in helping to characterize the values and aims of audiences for science. This is partly because some audiences may not be well positioned to know what these values or aims could be, and partly because scientists have not earned a monopoly on possible aims.

3.2 Asymmetric epistemic dependence

Epistemic asymmetries may impact agents’ abilities to characterize their values with respect to the aims of science. This is especially likely when expertise is relevant to knowing what our aims could be. To see how this might work, consider two examples.

In the first case, a cyclist enters a bicycle store and asks to buy the fastest bicycle in stock. The proprietor asks the cyclist what sort of riding he hopes to do, and he explains that he is new to cycling but hoping to do some long rides on dirt roads, bike packing, and so forth. Naturally, the proprietor would be happy to sell him a racing bicycle, but she suggests he consider a couple of alternative bikes that might better suit his purposes: perhaps one with larger tires and a more comfortable geometry for long days on the trail. She would have been perfectly within her rights to sell him the racing bicycle, but to do so would have exploited his relative epistemic disadvantage. Taking advantage of this epistemic asymmetry would be untrustworthy: Here I agree with Pielke (2007) and Almassi (2012).

---

2 For Daukas, this is reason enough to reject a solipsistic picture of trustworthiness because assessing how well we know in relation to those around us requires “other-attunement” (2011, 52).

3 The normative force of “require” scales with the strength of our (defeasible) commitment to trustworthiness. Scientists need not be equally trustworthy to all audiences (e.g., racist audiences) or all aims all the time (see Daukas 2006; Frost-Arnold 2014).
In the second case, a city resident asks an epidemiologist to explain why her neighborhood has higher rates of COVID-19 mortality than another neighborhood. Specifically, she asks what people in the other neighborhood are doing differently that protects them from severe disease. Like the bike shop proprietor, the epidemiologist could simply give her a true answer to her original question: Perhaps they have higher compliance with masking recommendations. But to do so would obfuscate other explanations that the epidemiologist could offer instead, such as the fact that the differences in the distribution of relevant comorbidities among neighborhoods is partly the product of environmental racism. Simply answering the resident’s question would align with the resident’s aims, but it would withhold information about other possible aims that she might prefer, such as assigning responsibility to industry or ineffectual regulators. This kind of obfuscatory move is a common feature of neoliberal discourse about health (Valdez 2021).

Although it might seem like scientists are less crassly incentivized than bicycle dealers to “sell” their expertise, feminist science studies scholars have shown that opportunistic and naive advertisements of the social salience of scientific pursuits are all too common (e.g., Roberts 2018). Along these lines, it is untrustworthy for the epidemiologist to take the resident’s original question for granted for the same reasons that it is untrustworthy for the bike shop proprietor to sell the cyclist the racing bike: It exploits her epistemic disadvantage with respect to characterizing her aims. Nothing about this means that, at the end of the day, the cyclist cannot choose a bicycle or that the resident cannot reasonably prefer to change her own behavior and leave the other issues to environmental activists. It simply reveals how insensitivity to aims characterization might render an expert intuitively untrustworthy, despite values-matching (see Haybron and Alexandrova 2013). Epistemic trustworthiness demands that we have good justification for our moral epistemology as well as our scientific knowledge, and this means that experts should attend to whether their audience is in a position to characterize their aims and values along these lines.

### 3.3 Values-matching: Reprise

This has two consequences for our understanding of trustworthiness. First, it shows that experts cannot take nonexpert aims and values for granted. Second, it means that values-matching efforts to align audience values with those of scientists cannot fully discharge the demands of trustworthiness. To see why this is the case, consider two strategies for dealing with audience values: appeals to best interests and deliberative polling.

Best interest approaches appeal to epistemic paternalism (Kelsall 2021). These approaches discount what agents want to know or do in favor of what some allegedly well-positioned expert thinks they should want to know or do, or what is allegedly in their “best interests” (see, e.g., John 2018). Deliberative polling approaches seek to manage diffuse and conflicting audience aims by using the process of deliberation to confer political legitimacy on the outcome of a value characterization process, often a choice among preset options (Fishkin 2021). Alexandrova (2018) and Schroeder (2019) both endorse something like this.

Both of these approaches to learning audience values can be criticized independently: Best interest approaches beg the question of audience aims and
are politically and epistemically illegitimate (Kelsall 2021); deliberative polling approaches can be compromised by nonideal epistemic situations such as testimonial smothering and silencing (Rolin 2020 aptly cites Dotson 2011 on this point). I mean to press a distinct, unified worry about these approaches, which is that they may neglect the challenge of negotiating audience values, either by assuming audiences are (or are not) well positioned to characterize values, or by preemptively restricting values to a set of options curated by experts. In contrast to paternalists, who would disregard audience values, I am arguing that experts should pay more attention to audience values, that they should cooperate in articulating these values, and that this cooperation may improve both experts’ and nonexperts’ epistemic situations regarding value judgments.

What I hope to have shown here is the following. An expert’s epistemic trustworthiness depends on her ability to assess the relative fit of her knowledge to the task at hand. To do this, she needs to understand the task. Sometimes she can learn this from her audience. But sometimes they are not well positioned to explore the values relevant to characterizing the aims of inquiry. In such cases, it would be easy for her to present them with a set of options. But to do so could inadvertently (in the best case) or maliciously (in the worst case) “sneak controversial values in through inattention” as Alexandrova (2018, 433) puts it, or simply fail to meet their needs. Their values may call for some bespoke epistemic goods. And maybe our expert can provide them! But to think her knowledge is well suited to the task, she may not merely assume that this is the case. She must earn the warrant for this inference, and this can be tricky business.

4. Cooperative epistemic trustworthiness

What can warrant a trustworthy scientist’s belief that both she and her audience have satisfactorily characterized their aims and values? Often, this will require some back and forth: The kind of negotiation that is common to conversations across different epistemic situations (Nagel 2019). Furthermore, our expert needs to expect her interlocutors to be able to participate unreservedly in such conversations. In other words, she needs to cultivate a cooperative approach to characterizing the aims of inquiry.

4.1 The cooperative account

Cooperative experts should prioritize audience voice over merely “consulting” an audience or offering them a fixed set of candidates. By this I mean the sense in which Anderson (1993) distinguishes voice, as a way of capturing the heterogeneity of ways in which we might value, from choice, a limited form of cardinal value ordering: we should not expect to meaningfully characterize what an audience might need or value simply by asking them to choose from a set of prefabricated options. Second, an expert should attend to the social conditions shaping a conversation or a process of

4 “Epistemic situation” in Currie’s (2018, 22) sense: “the set of challenges and advantages we have in generating knowledge in a given context.”
5 On “bespoke” communication, see Kelsall (2021, 298).
6 Daukas (2006, 113) says as much: “Confidence [with respect to others’ epistemic character and abilities] must be warranted.” I am concerned with this warrant in a particular domain: the aims of inquiry.
negotiating aims. As Rolin (2020) argues, contexts of asymmetric epistemic dependence are often characterized by power relations that can prevent agents from exercising their voice. Warrant for believing that aims are well characterized depends on the extent to which audience voices can be expressed. Third, experts should attend to the epistemic situation of an audience with respect to their values. Sometimes audiences may be far better positioned to characterize relevant values than scientific experts might be: they may be more familiar with a problem or better equipped to formulate a research question (Harding 1993; McHugh 2015). In other cases, experts may be able to inform values characterization by offering more information about possible aims and methods. Either way, this aspect of the cooperative approach will allow experts to improve their knowledge about audience values.

A cooperative account goes beyond making a process politically legitimate. If an expert wants good reason to think she knows the values of her audience, then she must account for the fact that both the values of her audience and her own understanding of those values are likely to change over the course of a conversation. The real decision to be made here is when to stop characterizing values and pursue a line of inquiry. Here I am indebted to Brown (2020, 152), who argues that because facts are relevant to values and vice versa, “value judgments are, or ought to be, a form of inquiry, sharing a common structure with scientific inquiry,” and that we must decide when these inquiries are adequate.

Part of the motivation for our expert to engage in value negotiation, rather than mere polling or consultation, is that it is not clear at the outset whether their audience is well positioned to characterize their values, socially or epistemically. If this is the case, experts should expect audience values to change as their epistemic situation changes. Sophisticated values-matching advocates (e.g., Alexandrova, Rolin, and Schroeder), and more general accounts of the aims of science that emphasize deliberation, are well aware of this. But importantly for the second-order considerations of trustworthiness, the expert should also expect her own epistemic situation to change in conversation with her audience. She knows more about what they know and want after each exchange, and so her own understanding of the task at hand should evolve as well.

Presenting an audience with options, as many values-matching accounts would do, allows scientists to retain control of, and responsibility for, the possibility space of scientific inquiry. But scientists are notorious for their failures of imagination, both moral and epistemological, in this regard (Stanford 2006; Brown 2020). A cooperative account avoids this by inviting the audience to expand or reject the possibility space as characterized by scientists and to participate in what Brown (2020) calls the “moral imagination” of science, or to refuse research altogether (Tuck and Yang 2018; Liboiron 2021). This is concordant with the view of value articulation in Irzik and Kurtulmus’s (2021) amendments to Kitcher’s (2011) well-ordered science, though not restricted to it.

Although the cooperative approach is demanding, it is already reflected in a number of existing models for inquiry. For instance, scientific/intellectual movements (SIMs), community-based participatory research (CBPR), and participatory action research (PAR) can all involve iterative refinement of the aims of inquiry in collaboration between scientific experts and communities affected by their research (Brydon-Miller et al. 2003; Rolin 2016, 2020; Brown 2020). These approaches have been
endorsed by philosophers of science on broadly moral and political grounds, and by feminist philosophers on epistemic grounds (Wylie 2015; Tekin 2022). Here, I am arguing that they are also valuable for moral epistemological reasons, and that these reasons are relevant to epistemic trustworthiness. To illustrate this, consider an example from health disparities research.

### 4.2 The preconception stress and resiliency pathways model

The preconception stress and resiliency pathways model (PSRP) is a framework for research designed by the NIH Community Child Health Network (CCHN). The problems at stake in the PSRP model are racial disparities in pregnancy outcomes, such as prematurity, infant mortality, and birthweight (Ramey et al. 2015). CCHN researchers used “community-academic partnerships” to develop a framework for studying these disparities. This involved sustaining long-term collaborative relationships between academics from “social, behavioral, and biomedical sciences” and “community representatives” from several research sites (ibid., 709). The academic researchers claim that community members were treated as coinvestigators (ibid.).

Many aspects of this project seem to embody a cooperative approach. The project began with a two-year “getting acquainted phase,” in which coinvestigators shared expertise and interests with one another, followed by iterative refinement of the conceptual framework for inquiry (ibid., 709–10). When disagreement arose about the direction of inquiry, each research site deliberated and received a single vote; this deliberation extended beyond the original framing of the aims of inquiry into choices that arose within the research process (ibid.). Ramey et al. report that the outcomes of this process were a framework for guiding inquiry, specific causal hypotheses, and a prospective study that tested two of these hypotheses. This framework stands out because it emphasizes paternal, as well as maternal, contributions to health; acknowledges community resources rather than focusing exclusively on stress; and deemphasizes specific mechanisms and biochemical pathways (ibid.).

The development of the PSRP model is imperfect. (We might ask: How do academic researchers ensure accountability to community co-investigators? Can community members refuse research if they are out voted? How does the model increase surveillance during the “preconception” period?) But it is an active, iterative approach to characterizing the aims of inquiry. It is at least nominally sensitive to “community” needs and interests, and is responsive to these interests not just at discrete stages of inquiry but also before, while, and after the model is implemented. All else being equal, this means that academic coinvestigators should have better reason to believe that the aims of community coinvestigators are well characterized. As Scheman (2001, 44) reminds us, trustworthiness is not an all-or-nothing attribute, but rather “a rolling horizon we move toward” and that admits of degrees. The iterative, collaborative elements of the PSRP process improve the degree of trustworthiness of the academic coparticipants.

Negotiating audience aims often calls for a process of collaboratively characterizing values and tailoring inquiry to these goals. To be trustworthy, experts need reasons to believe they understand the aims of their audience, and that their audience is well positioned to articulate these aims. “Well positioned” here can involve access to information about what science might have to offer, but it can also
invoke the social conditions that make it possible for an audience to express or negotiate their aims. In many cases, this will require scientists to take an active role in characterizing and negotiating aims, and this is best captured by a collaborative approach. My point is that SIMS, CBPR, PAR, and other such approaches are motivated not only by our intrinsic desire to institute collaborative research projects but also because they provide the epistemic warrant for scientists to know they are engaging with nonexpert audiences in a trustworthy way.

5. Objectivity and trustworthiness

So far, my primary aim has been to articulate how a collaborative approach to characterizing values might improve the epistemic trustworthiness of science. If I have succeeded, we are in a position in which critical moral epistemological reflexivity plays an important role in the trustworthiness of science. Now I want to explore some implications of this argument for our understanding of the relationship between trustworthiness and the objectivity of science.

Naomi Scheman (2001) and Heather Douglas (2009) have each argued that trustworthiness might ground the objectivity of scientific inquiry and, in doing so, serve as a viable alternative to the value-free ideal. For Scheman and Douglas, what warrants calling scientific inquiry “objective” is that it deserves our trust; this, in turn, depends on it being guided by the right values (for Scheman) or on several “bases for trust” (for Douglas, 116).

But I have argued that trustworthiness depends on values themselves being well characterized, and so I have a different way of thinking about the relationship between trustworthiness and objectivity. Like these authors, I endorse a value-laden approach to the objectivity of science. However, with respect to trustworthiness and objectivity, my argument runs in the opposite direction. This is because the cooperative approach to trustworthiness that I advocate bears a striking resemblance to the feminist standpoint theoretic notion of strong objectivity, and this makes objectivity prior to trustworthiness on my account. Strong objectivity, for standpoint theorists, is achieved in part by cooperative critical reflexivity about value judgments that justifies a specific normative commitment (Harding 1993; Intemann 2010). Cooperative epistemic trustworthiness demands very much this sort of starting point for inquiry, and just this sort of self-reflexive approach to values in science.7 Indeed, Daukas builds this self-reflexivity into her original account of trustworthiness and puts this in conversation with standpoint theory herself, albeit with different emphasis.

This complicates the relationship between objectivity and trustworthiness that we inherit from Douglas and Scheman. Whereas for these authors, objectivity depends on

---

7 In basing this affinity exclusively on the critical reflexivity dimension of strong objectivity, I am unhitching it from other important features of standpoint theory. I am cautiously optimistic about this, because I think the force of strong objectivity lies in its articulation of critical reflection about the warrant for value judgments as an account of objectivity, and this is important even in contexts where standpoints are not obviously relevant. Skeptical readers are invited to restrict the scope of the priority argument to cases in which standpoints are relevant and/or to disregard the priority argument and simply take this section as a grateful acknowledgment of the work that standpoint theorists have done to advance this line of reasoning.
trustworthiness, I am advocating for an understanding of trustworthiness that depends on objectivity. To be sure, this may be complementary to values-matching approaches (Intemann 2010; Daukas 2011; Frost-Arnold 2014). For instance, values-matching may often be appropriate once audience values are well characterized. Furthermore, I think Scheman would endorse this, as she cites strong objectivity in her argument that specific normative commitments are crucial for trustworthiness and therefore, objectivity. What I want to emphasize here is a role for strong objectivity in articulating more specific value judgments about the aims of inquiry. Insofar as we want to distinguish these roles for objectivity and trustworthiness, strong objectivity is prior to this broader sense of objectivity, because it is necessary—if insufficient—to ground the epistemic trustworthiness of science. If this amounts to an argument that thoughtful, cooperative deliberation is overdetermined by feminist reasoning about objectivity and trustworthiness, so much the better.

Acknowledgments. Thank you to Mike Dietrich, Sarah Richardson, Helen Zhao, Aja Watkins, Vivian Feldblyum, Will Conner, Kyra Hoerr, Katie Creel, audience members at the 2022 meeting of the PSA, and a kind reviewer for suggestions and solidarity in writing this paper.

References

8 Here I am closer to Rolin (2020).


