

The Scientist, *Qua* Scientist, Is an Ethical Agent

PHILIP KITCHER, Department of Philosophy, Columbia University, New York, USA

KITCHER, F.: The Scientist, *Qua* Scientist, Is an Ethical Agent
FILOZOFIA, 79, 2024, No 3, pp. 231 – 243

I argue that the currently confused talk about “values in science” should be understood in terms of schemes of values. The new demarcation problem is then formulated as asking “Which scheme(s) of values should scientists adopt?” I defend the view that my title provides the answer.

Keywords: scientific progress – value-free science – epistemic values – ethical judgment – science and values

I. How to Talk about Values

The thesis that science ought to be value-free aimed to debar a particular form of scientific decision-making. Champions of the thesis condemned individual scientists for advocating supposed new findings, inadequately supported by the available evidence, simply because their widespread acceptance would advance their careers. They also castigated groups for campaigning on behalf of a theory just because it resonated with some political cause. In response, Thomas Kuhn (1977) argued for the legitimacy of appeals to some values (simplicity, scope, and similar “epistemic values”) in debates about large scientific changes. Heather Douglas and Torsten Wilholt (Douglas 2009, Wilholt 2009) broadened the critique of the value-free ideal by demonstrating how a larger class of value judgments can play unimpeachable roles in individual scientific decisions (and in the social practices that coordinate such decisions). To take one obvious and compelling example, researchers working against the clock to find ways of responding to a new eruption of infectious disease legitimately cut corners, moving through a sequence of investigations without exploring possibilities that a less urgent inquiry would reasonably consider.

So arises the “new demarcation problem” (Holman and Wilholt 2022). Allegedly some values belong in science, and others don’t. The philosophical challenge is to specify a criterion for separating the welcome candidates from

the interlopers. But, although my formulation follows the usual way of posing the issue, there is surely something peculiar about it. It invites a picture: there is a door to some desirable venue; outside stand people seeking entrance; what's the ticket for gaining admission?

Of course, values are not people. People have values, or so we tend to say. What are they, these things people have?

That's a bad question. We should recall some lessons about ontology taught long ago by Quine. People also sometimes do things for the sake of someone else; on other occasions, they do things for their own sake. What are these sakes, these things people have? Don't ask, Quine told us. Don't inflate ontology by pulling nouns out of syncategorematic expressions.

'Value' is both a noun and a verb. If we're tempted to ask what values are, we are treating 'value' as a noun. Perhaps, however, the verbal form is primary. People value. Their values are manifest in the things they value (manifest how, exactly?). If you start down this path, you might connect values with desires: to value a possible outcome is to want to bring it about; to value an object might be to want to acquire it, or to preserve it, or to restore it; to value a property is to want to see it realized more extensively; and so forth. Maybe valuing properties is the fundamental notion. You value an object or an outcome because of some property it exemplifies. So, you fight to preserve a natural landscape, because you value its beauty. You want it to be there for others to enjoy as you have done. To value is to desire (desire just what?). But values can't be linked to any old desire. Some objects of desire are just too unimportant to you to express your values: does anyone *value* chocolate? Maybe then the things you value correspond to *special* desires, the important ones, the desires that really matter. Or better, your values are expressed in the desires to which you give priority, the ones you really want to be satisfied.

I have staggered to a formulation that provides clues about how to go on. Our values, I suggest, are expressed in the *priority ordering* among our desires. Don't ask how highly ranked in this ordering a desire must be to count as expressing a value. Just pay attention to the ordering. People have *schemes of values*, standing psychological dispositions to give priority to some desires over others. Moreover, this ordering is affected by context. We give priority to one desire in the presence of other desires, but that may be reversed when a different set of background desires is in play. The citizens of the Weimar Republic who, under conditions of daily inflation, traded precious family heirlooms for food, only did so because of the threat to their continued

survival. Similarly, a scientist who gives priority to getting on with the follow-up experiment because the inquiry needs to go as quickly as possible to combat a wave of infection, would, under less urgent conditions, view that as an illegitimate exercise in corner-cutting.

We thus return to the phenomena that motivated the value-free ideal for Weber and his allies. What should move scientists is perceived conformity to standards of good evidence and objective criteria for pursuing particular questions. However high a consideration may rank in a scientist's scheme of values it should never outrank the wish to be true to those epistemic standards. Thus, a thesis:

Value-Free Ideal: In a scientist's scheme of values, epistemic values should outrank all other values.

Assume we understand what epistemic values are. Then, I propose, the demarcation problem is to identify the scientific contexts in which it is permissible for combinations of values, including at least some of which are not epistemic, to outrank any set of epistemic values, and so violate the Value-Free Ideal.

II. Schemes of Values

My principal aim here is to tackle the demarcation problem, reformulated as a question about permissible schemes of values for scientists. Before embarking on that venture, it's worth clarifying the concept, *scheme of values*, at the center of my approach. I think of this as a standing psychological disposition, operative over a period of time, that plays a causal role in a person's choices. Onlookers see the scheme of values expressed in her behavior. She has a self-conception that includes what takes priority over other considerations. She says, for example, "My family takes priority in my life. I always put them first." From time to time, though, others detect what they take to be a mismatch between her professions and something she does. Their verdict might be unjust; there may be no divergence between the scheme of values and her conduct, but simply a factual error on her part. "I thought," she says, "that working late would give me the means to provide important things for the children, but I now appreciate that spending the time with them would have been more beneficial to them."

Or the critic might be correct in seeing a mismatch, and she may, in response, modify her scheme of values to bring it into accord with her characterization of it. She isn't entirely self-deceived. Most of us have only a rough sense of our scheme of values, in that we can rarely (never?) anticipate

the full range of circumstances in which we shall face choices. We claim to give *X* priority over *Y*, not appreciating the extra pull *Y* would exert on us in a situation in which *Z* were also present. When such situations occur, others point out to us how we have diverged from what we professed, and we respond either by modifying our self-conception or the disposition that guided the choice of *Y* over *X*. We say “You’re right. It turns out that I don’t give *X* priority in all circumstances.” Or, instead, we alter our dispositions. Perhaps we now try to avoid situations in which *Z* shows up (keeping ourselves out of temptation.) Or, knowing of *Z*’s disruptive power, we are very careful to be on the lookout for its presence and inclined to stiffen resistance to *Y*’s charms. When someone pursues the latter course, there is something deeply correct about the initial self-description: the subject really *does* give priority to *X*, and that shows up in the strategy for responding to the apparent failure. People’s schemes of values aren’t only expressed in their choices, but in their regrets and their ways of *revising* when things appear to go wrong.

The structure of a scheme of values is that of a complex partial ordering. It’s complex because the features we pick out in assessing our options often don’t stand in determinate relations with one another. Perhaps in most, even almost all, combinations with other identifiable features, *X* dominates *Y*, but there are some combinations in which the relation is reversed. When we describe ourselves we do so by appealing to readily identifiable features, and, in consequence, our self-descriptions aren’t entirely accurate (although few people, I suspect, are massively self-deceived.) Moreover, as we have seen, more complexity enters because of our dispositions for responding to unforeseen situations. It’s partial because for some pairs of identifiable features, most prominently for those occurring at the top of the ordering, there may be no general presumption that one will outrank the other. Often, a person may struggle with a choice, seeing a “fundamental value conflict” or “incommensurable values.”

Finally, many people can be attributed several schemes of values. Elsewhere, I have credited an individual scientist with a “broad scheme of values,” an “epistemic scheme of values,” and a “probative scheme of values” (Kitcher 2011). The first is the overarching scheme within which the latter two have their place. The second covers the judgments made by “the scientist *qua* scientist” (Rudner 1953), and the third is expressed in a particular area of scientific work. Weber and his followers regard scientists as giving priority to epistemic standards once they enter the laboratory, taking on a different

scheme than the one guiding them in their non-scientific lives. In such cases, the overarching (broad) scheme of values must itself license the choice of steering by the subordinate scheme when the person takes on the appropriate role. The broad scheme allows taking on the role and conforming to its norms during the practice of that role.

Just as we can consider what scheme of values a scientist should adopt, we can reflect on what broad scheme a person should have. Of course, there are branches of inquiry devoted to aiding such reflections – and they have familiar names. Value theory. Moral theory. Ethics. If those address what broad schemes people should have, and if broad schemes issue licenses to take on particular roles, don't ethics and moral theory bear on the question of the proper scheme of values for scientists? And on the Value-free Ideal, and the New Demarcation Problem? And, if they do, isn't the Value-free Ideal clearly and absurdly false?

It is. The scientist *qua* scientist should make ethical judgments all the time. The decision to *join* a particular research enterprise is subject to moral norms: nobody ought to engage in work to make chemical weapons or to vindicate the harmlessness of tobacco. Choices of research questions ought to be guided by the wish to contribute to human welfare (Kitcher 2001). Experimental design, especially in the life sciences and the medical sciences, should not ignore the potential effects on the lives of the organisms the inquiry will study. No epistemic end justifies the practices of the Nazi doctors. Historically, vast numbers of laboratory animals have been appallingly mistreated. On a smaller scale, the everyday practices of many labs are rightly criticized for their lack of consideration of what is done in "sacrificing" animal lives (Kitcher 2015).

Many scientists today are inspired to follow Weber in viewing science as a value-free zone. They are moved by a small cluster of examples of scientific misconduct. Researchers who defend a claim because doing so offers career advantages are (almost always) lapsing from the standards that should govern their practice of science. Groups that campaign for a theory on the grounds of its resonance with their politics are often equally culpable. The faults, however, have nothing to do with the value-freedom of science. They are diagnosed by recognizing the deviation from a scheme of values that would agree with ethical norms. Moreover, diagnosing them in this way enables us to perceive how the practices can on occasion be condoned. To pretend a view, one you recognize as inadequately supported, is evidentially warranted is a form of dishonesty that would be defensible only under

extreme circumstances: you are subject to heavy-duty blackmail from the originator of the view, for example.

The political example leads us further. Critics of human sociobiology were frequently castigated for resisting truths about human nature simply because of their dedication to progressive politics; similarly, those who believe in socially relevant genetic differences among social groups (inequalities among sexes, races, castes etc.) have been accused of holding their views simply because they were politically attractive. Many of the people involved in these heated debates were innocent. They sincerely believed the claims they propounded were supported by the evidence.¹ But some of the accused have been guilty, and their ethical lapse was comparable to that committed by the ambitious investigators who tune their beliefs to their career prospects. Nevertheless, we should recognize the possibility of occasions on which scientists properly campaign for politically relevant conclusions *in advance* of the evidence.

In 1988, James Hansen warned the US Congress about the dangers of climate change. Tragically, his counsel has been disregarded. Like some of his fellow climate scientists, Hansen has since used public forums to urge large reductions in the use of fossil fuels. His reward for doing so has frequently been waves of criticism, some of it from other scientists, reviling him for abusing his position as a scientist (many of the accusations express the value-free ideal). Hansen's activities not only meet ethical standards; they appear heroic. His attempts to sway a reluctant public to respond to the greatest threat our species has ever faced are firmly based in evidence. Suppose, however, that were not the case. Would the charges against him then be justified? No. His behavior accords, both under the actual circumstances and in my hypothetical alternative, with an ethically excellent scheme of values. The Hansen who warns ahead of conclusive evidence is moved, like the real James Hansen, by deep concern for humanity. He places the preservation of our planet as a home for our species ahead of scientific restraint, *even at the cost of being vilified and of abandoning a career he loves and in which he has excelled.*

I draw a very simple moral. When the issue of science-and-values is reformulated in the terms I have adopted, the value-free ideal should appear

¹ I think it important to state this, since my own criticisms of human sociobiology have sometimes been interpreted as accusing practitioners of ethical lapses. This is particularly so with respect to the discussion of Lumsden and Wilson's theory of gene-culture co-evolution (see Kitcher 1985, Chapter 10). For the record: I think Lumsden and Wilson were mistaken; but it was an *honest* mistake.

completely indefensible. It is ludicrous to propose that an ethically appropriate role for scientists is to subordinate themselves to the search for truth, so that no consideration should make them deviate from it. Nietzsche saw the point long ago; so too did William James. Only someone deeply gripped by epistemological abstractions would come to think a commitment to epistemic values should “always and everywhere” take precedence over the other interests of sentient beings.

III. The Scientist’s Scheme of Values

The issue behind the demarcation problem can be formulated succinctly: What scheme of values would be ethically appropriate for the scientist’s role? First, some background clarification. Science is an institution, within which people take on the role of scientific investigator (the institution admits other roles as well.) A general thesis (one I think close enough to the truth for present purposes): Provided the progress of an institution contributes to human progress, the roles adopted within that institution should be defined so that they collectively contribute to the progress of the institution. Applying this to the role of scientific investigator, I conclude that the scheme of values appropriate to that role ought to be one guiding the investigator to contribute to the progress of science.

Weber ought to like this conclusion. He might think it gives epistemic values absolute priority in the scientist’s scheme of values, by directing researchers to contribute to scientific progress – through the accumulation of truth about nature. I reply: you have neglected my proviso. There are occasions on which your preferred scheme of values interferes with human progress. Science dovetails with other institutions to have an impact on human lives and human societies, and, in some situations, that means non-epistemic considerations take priority.

Yet there’s a deep mistake in the Weberian position I haven’t so far exposed. It lies in the conception of scientific progress. Is progress just amassing truth? That’s wrong for all sorts of reasons. First, since Kuhn, we’ve appreciated that there are many dimensions of scientific practice – the science of a time isn’t simply a list of propositions. Second, as Catherine Elgin has eloquently argued, the propositions we aim to accumulate need only be “true enough” (Elgin 2017). Third, and importantly for the new demarcation problem, the teleology implicit in the position has to be abandoned.

Many people talk of the aim (or aims) of science, and think of progress in teleological terms. Science makes progress by achieving more and more of

its aims. (Just as we make progress when we travel by decreasing the distance to our destination.) For people who emphasize the accumulation of knowledge, the teleological view is readily articulated as “learning more about the world.” A naïve version supposes us to come ever closer to the complete truth about nature. Trivially, however, in any finite spatio-temporal interval there are continuum many points, and thus at least continuum many truths about that interval. Far too many even to approximate. Most of them, of course, not worth knowing.

In characterizing the realism debate, Bas van Fraassen improves on the naïve version. “Science aims,” he explains, “to give us in its theories, a literally true story of what the world is like” (1980, 6). Van Fraassen is recapitulating Newton’s hopes in the *Principia*, when he envisages a complete set of fundamental laws governing the world, and hopes that science will discover them. Popular though it remains among many physicists, the idea is untenable. It’s vitiated by the impossibility of deriving the claims of many sciences from the principles of physics (even in principle) and by the rarity of the exceptionless generalizations that so fascinated Newton.

But, as Kuhn and Larry Laudan saw (Kuhn 1962 section XIII; Laudan 1977), there’s an alternative: progress *from*, rather than progress *to*. Pragmatic progress (as I shall call it) consists in overcoming problems and transcending limits. It’s evident in technological progress. The laptop on which I type has advanced on the clunky machine on which I hunted and pecked thirty-odd years ago, not in virtue of being closer to the Ideal Form of the Computer, but because clever people have learned how to solve problems and overcome limitations. Laudan’s own discussion was clearly based on this pragmatist idea.

Like Kuhn before him, though, Laudan focused on the internal development of the sciences. Both were insightful about progress, but the insight was limited by treating scientists as the ultimate authorities on identifying problems in their fields. Once we abandon the idea of science as being a search for “the whole truth and nothing but the truth,” we’re forced to ask what legitimately counts as a problem (or a limit to be pushed back), and what counts as an adequate solution. Kuhn and Laudan deferred those questions to the pertinent group of scientists. Was their deference justified? Could a scientific community be mistaken?

Of course it could. Sometimes a field spends years on an issue that eventually appears worthless (think of the nineteenth century search for models of the ether.) Yet that kind of trouble shouldn’t shake Kuhn and

Laudan in their deference. After all, they might reply, when scientific communities make mistakes, there's no group that would do better, nor is there any socially feasible way of improving on their judgment. That's correct for the ether-model case, and for others similar to it.

But by no means for all. When a community's chosen investigations are detrimental to the broader social good – or even when, despite the extensive support they receive from the ambient society, their selected inquiries generate only trivial social benefits – they don't discharge their assigned role. The proviso in my general thesis about the progress of institutions is violated. If things have gone wrong because of understandable factual errors (as in the ether-model example) they are excused. They have done their best to live up to their role, and have been unfortunate. Sometimes, however, the problem doesn't arise from some false belief about nature (that there's an ether, say) but from a value-judgment they ought not to have made.

Between 1900 and 1925, genetics made spectacular progress – or so it seems. After three researchers simultaneously rediscovered Mendel's laws, the community identified crossing-over and recombination, learned to make genetic maps, devised techniques for inducing mutations, and amassed a vast amount of information about inheritance in particular organisms (notably the fruit-fly, *Drosophila melanogaster*.) While that was occurring, however, nascent classical genetics became intertwined with eugenics, spawning an enterprise of seeking genes for "important" human traits, some of them dubious and amusing (thalassophilia), others selected for study because of morally flawed motives. Fears that human evolution piles up a "load of mutations" combined with concerns about the "decline of the West" to inspire efforts to uncover genes for crime and for feeble-mindedness, and to trace genetic differences among so-called racial groups (not only those listed on current census forms, but also subdivisions of the European population). The significance of these investigations was intensified by their relevance to social policy: potential criminals were to be identified in advance, the feeble-minded were to be sterilized, and various types of immigrants were refused asylum in countries where they might have survived.

The moral blunders of this movement are evident today. They also evoked protests at the time. The scientists who introduced inquiries into genes for "socially undesirable traits," even those who (like guards in Hitler's concentration camps) simply took the questions for granted and pursued answers to them, can't be absolved by pleading that sufficient moral

sensitivity wasn't available in their time. They should have known better. Their scheme of values was ethically unjustified.

Examples like this one show deference to the scientific community's view of what investigations are worth pursuing may be mistaken. Only if there were some reliable mechanism to ensure that scientific communities would always operate in accordance with some ethically defensible scheme of values should we be satisfied to grant them authority in selecting significant questions. Perhaps there are reasons for supposing some areas of science to be relatively invulnerable to moral mistakes – although *failure* to do the good you can is also a lapse. But, most importantly, because scientific research has such deep and broad impact on public policy, its choice of directions should not be a matter just for the scientists.

An answer, then, to the New Demarcation Problem. The scientist's role is to contribute to human progress. Scientists' schemes of values should be attuned to discharging that role. That requires ethical sensitivity to pervade their decision-making. The lab is no special place in which values can be ignored. Any consideration that can play a role in sensitive ethical decision-making can potentially figure in scientists' decisions. What's important is for scientists, like all of us, to reflect on what we do and on how it changes the world, for better or for worse.

IV. Too Vague?

Surely a disappointing answer! Lamentably inspecific. How should that reflection go? How often should it occur? And, above all, what do 'better' and 'worse' mean here?

It's best to begin with the last question. Again, I abandon teleology: the progress of the world doesn't occur through its increased proximity to any supposed ideal state. We make progress by overcoming the problematic features of its current condition.

What counts as a problem? I don't think there's any community of alleged experts who can be given authority to identify the problematic features of the current state of the world. Religious teachers, sages, gurus, even philosophers won't do. Instead, I see ethical progress as what emerges from collective ethical inquiry, in which all sentient beings are involved (with those unable to speak, infants, the demented, non-human animals represented by those who best understand their condition). A situation is justifiably regarded as problematic if an ideal deliberation would judge it to be problematic. A deliberation is ideal just in case the perspectives of all those

affected by the situation are represented in it, if all assertions are well-grounded in evidence, and if each participant is concerned to understand the perspective of each of the others and to arrive at an outcome all can tolerate. Similarly, a solution is justified if it emerges from the same type of deliberation (where the parties recommended now include representatives of all those affected by the proposed remedy.) Justified progress is change that would be endorsed by such deliberations. Progress (*tout court*) occurs when the justification continues through an indefinitely proceeding sequence of ideal deliberations – when justification sticks.

I've defended this approach to ethical progress and ethical inquiry elsewhere (Kitcher 2021). Plainly, the deliberations I envisage can't be realized in our world. But ideals are best seen as diagnostic tools, pointing to current flaws and ways of addressing them. My ideal exposes the faults of excluding points of view, of using unsupported claims, and (most important, I think) of failing to engage fully with others. It encourages us to decrease these kinds of faults.

A researcher's scheme of values ought not to be entirely restricted to the epistemic. Like many other scientists of his generation, Alan Turing's wasn't. He didn't say, "I'm in the process of exploring this research question, so I can't come to Bletchley just now," but abandoned his projects to try to crack the code the Germans were using. Once you recognize that scientific progress is a matter of addressing significant questions, you have the responsibility of asking what makes your preferred line of inquiry significant – and that requires thinking about the way it might affect the world's future. How should that thinking go?

Of course, you can't assemble a fully representative deliberation of well-informed mutually-engaged people to supply the answer. Instead, you have to do your best to simulate that kind of deliberation, and what its likely outcome would be. That demands considering how your research might bear on a variety of lives (most of them human), whether it would damage any of them, and, if so, whether others, to whom it might be beneficial, would be able to convince the potential victims of the greater value of going ahead. Simulation is far from easy. It requires developed capacities – "sensitivities" we might call them – of imagination, sympathy, and reasoning. Being responsibly reflective requires a more basic sensitivity, the ability to know when to stop and consider and when to plunge ahead. For ethical life is plagued by two kinds of mistakes, being too hasty in following

habit and dithering, Hamlet-like, when action is needed. (Kitcher 2021 Chapter 3)

Whether that makes any difference between what's legitimate in the lab and what's proper in everyday life is, I think, impossible to determine. Because the contexts in which scientists make decisions are so many and so various, there is no way to round up all the features that could legitimately figure in researchers' decisions. If that's what's wanted, the New Demarcation Problem is insoluble. If, on the other hand, what's required is to say how decision-making should go, then, I claim, it's now been solved. The scientist, *qua* scientist is an ethical agent.

A final moral: "Science and Values" has been on the agenda of the philosophy of science for well over a decade now. Yet people who write about it usually seem allergic to engaging in value inquiry. They sometimes write as if, when two parties dispute about what is valuable, there can be no reasonable resolution. Very few people believe that general thesis. It is high time for philosophers of science to take the next step, and start to consider how the value-laden disputes that sometimes arise in the sciences might be resolved. About the contextually varying details of what it means for a scientist to be an ethical agent.

Bibliography

- DOUGLAS, H. (2009): *Science, Policy, and the Value-Free Ideal*. Pittsburgh: University of Pittsburgh Press.
- ELGIN, C. (2017): *True Enough*. Cambridge, Massachusetts: MIT Press.
- HOLMAN, Bennett – WILHOLT, Torsten (2022): The New Demarcation Problem. *Studies in the History and Philosophy of Science A*, 91, 211 – 220. DOI: <https://doi.org/10.1016/j.shpsa.2021.11.011>
- KITCHER, P. (1985): *Vaulting Ambition: Sociobiology and the Quest for Human Nature*. Cambridge, Massachusetts: MIT Press.
- KITCHER, P. (2001): *Science, Truth and Democracy*. New York: Oxford University Press.
- KITCHER, P. (2011): *Science in a Democratic Society*. Amherst, New York: Prometheus Books.
- KITCHER, P. (2015): Experimental Animals. *Philosophy and Public Affairs*, 43, 287 – 311.
- KITCHER, P. (2021): *Moral Progress*. New York: Oxford University Press
- KUHN, T. S. (1962): *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- KUHN, T. S. (1977): Objectivity, Value Judgment and Theory Choice. In: Kuhn, T. S.: *The Essential Tension*. Chicago: University of Chicago Press, 320 – 339.
- LAUDAN, Larry (1977): *Progress and its Problems*. Berkeley and Los Angeles: University of California Press.
- RUDNER, R. (1953): The Scientist Qua Scientist Makes Value Judgments. *Philosophy of Science*, 20, 1 – 6. DOI: 10.1086/287231.

VAN FRAASSEN, B. (1980): *The Scientific Image*. Oxford: Oxford University Press.
WILHOLT, T. (2009): Bias and Values in Scientific Research. *Studies in the History and
Philosophy of Science A*, 40, 92 – 101. DOI: <https://doi.org/10.1016/j.shpsa.2008.12.005>

Philip Kitcher
Department of Philosophy
Columbia University
1150 Amsterdam Ave
708 Philosophy Hall
MC4971, New York, NY 10027
USA
e-mail: psk16@columbia.edu