Ergo AN OPEN ACCESS JOURNAL OF PHILOSOPHY

Science and Rationality for One and All

P.D. MAGNUS

A successful scientific community might require different scientists to form different beliefs even when faced with the same evidence. The standard line is that this would create a conflict between the demands of collective rationality which scientists face as members of the community and the demands of individual rationality which they face as epistemic agents. This is expressed both by philosophers of science (working on the distribution of cognitive labor) and by epistemologists (working on the epistemology of disagreement). The standard line fails to take into account the relation between rational belief and various epistemic risks, values of which are a matter of personal and social commitment. This introduces the possibility of conflicts the standard line does not recognize, because someone with extreme values might be individually rational but too far beyond the pale to have a place in the scientific community. More importantly, it introduces at least a possibility for good scientists to be rational individuals.

Science would probably go best if individual scientists at least sometimes form differing beliefs even when faced with the same evidence. The standard line, both among philosophers of science and epistemologists, is that this creates a conflict between the demands of collective rationality which scientists face as members of the community and the demands of individual rationality which they face as epistemic agents. I argue that this conflict presumes an inadequate conception of individual rationality. When we take into account the relation between rational belief and the values of various risks, it is at least possible for good scientists to be rational individuals.

1. The Standard Line in Philosophy of Science

It seems obvious that a community of one thousand scientists working together to make discoveries and solve puzzles should arrange itself differently than would

Contact: P.D. Magnus <pmagnus@fecundity.com>

http://dx.doi.org/10.3998/ergo.12405314.0001.005

one thousand scientist-hermits working independently. Because of limited time, resources, and attention, an independent scientist can explore only some of the possible approaches to a problem. Working alone, each hermit would explore the most promising approaches. They would needlessly duplicate the work of others and would be unlikely to develop approaches which look unpromising but really have tremendous potential. Contrariwise, a large community can more rigorously explore the space of possible approaches. Most scientists should work on the most promising approaches, but a smaller number can be committed to approaches that initially look less promising. Exploratory work can reveal if one of those initially unpromising approaches has unrealized potential, and more scientists can adopt it once its potential becomes more apparent.

This division of cognitive labor is an important feature of science. It is a central theme in the work of Thomas Kuhn, who emphasized both the importance of a shared approach and the need for some scientists to look beyond it. This is his famous distinction between *normal science* (governed by a paradigm) and *revolutionary science* (in which one paradigm is replaced by another). Kuhn (1977) calls this need to abide by shared commitments and also to innovate "the essential tension" in science. Importantly, he recognizes that this tension is not felt internally by every scientist. Instead, some explore wilder alternative approaches while most hew more closely to the dominant paradigm. He writes,

> Most anomalies are resolved by normal means; most proposals for new theories do prove to be wrong. If all members of a community responded to each anomaly as a source of crisis or embraced each new theory advanced by a colleague, science would cease. If, on the other hand, no one reacted to anomalies or to brand-new theories in high-risk ways, there would be few or no revolutions. (2012: 185–6)

What makes a novel approach appealing to a particular scientist cannot be its objective promise, because otherwise pursuing it would not seem to be a risk. Rather, a scientist might initially adopt an approach for personal or aesthetic reasons, and their work on it can reveal its potential. Kuhn writes, "If they had not quickly taken it up for highly individual reasons, the new candidate for paradigm might never have been sufficiently developed to attract the allegiance of the scientific community as a whole" (2012: 155).

Philip Kitcher develops this idea in a formal way, acknowledging "a large debt to the ideas and writings of Thomas Kuhn" (1990: 5). Subsequent philosophers have both extended Kitcher's formal model and developed others.[1] I am not concerned here with the technical details, but instead with the fact that many authors

1. See recent surveys by Ryan Muldoon (2013) and Michael Weisberg (2010).

in this tradition take it to involve an opposition between the epistemic demands of individual rationality which weigh on a scientist and the epistemic demands of collective rationality that weigh on the scientific community. For example, Kitcher claims that this work poses "a mismatch between the demands of individual rationality and those of collective (or community) rationality" (1990: 6). Kevin Zollman writes, "Looking at . . . scientists from the perspective of individualistic epistemology, one might be inclined to criticize the scientists' behavior. However, when viewed as a community, their behavior becomes optimal" (2010: 33).

One may wonder whether these philosophers of science *really* think that science puts the demands of responsible science in tension with the demands of individual rationality. Beyond brief passages like the ones that I have quoted, they do not whinge at length about it. This is not because they resolve the problem, however, but because they give up thinking about scientists as rational agents with full information.

Kitcher's solution is to "*dissolve* the notion of rationality" (1993: 194, emphasis in original). On his view, we should not expect or enquire after whether scientists are rational. Instead, we should ask only whether they employ strategies of belief formation which promote cognitive goals, recognizing that there are limits and conflicts involved in their doing so. In cases of science having functioned well, Kitcher writes, "Society, nature, and sound individual reasoning combined to drive the social learning machine to a new success" (1993: 218). He demands *sound reasoning* not in the technical sense (valid with true premises) but just in the minimal sense of employing reliable-enough strategies to promote collective progress. This allows for scientists to use different strategies, to take different risks, and to draw different conclusions even from the same evidence.

Zollman skirts the problem in a different way. Instead of arguing that scientists might responsibly draw different conclusions in the face of the same evidence, he models scientists as having limited information. In his models, the community structure of science limits which scientists share their findings with which others. Individuals react to different evidence and so may disagree, even though rationality would not allow disagreement in the face of the same evidence.

2. The Standard Line in Epistemology

Suppose we accept a roughly Kuhnian account of the distribution of cognitive labor. Given the same evidence, some scientists pursue research which has P as a presupposition. A few scientists pursue research which has incompatible presuppositions; to simplify the case, just suppose that the rival approach has $\sim P$ as a presupposition. One might advise them all to believe neither P nor $\sim P$, to remain agnostic while guiding their research by one presupposition or the other. There are

reasons to think this would be bad advice. First, scientists who cultivate agnosticism might not pursue their chosen research program with the necessary vigor. The community would then do better if those individuals fully embraced the presuppositions of their approach. Second, there is some reason to think that any declared agnosticism would just be ersatz. A scientist whose research presumes *P* might never have occasion to act on anything but that presupposition. Their lab is outfitted to pursue the *P*-presuming approach and they have filled it with *P*-presuming postdocs, after all, and one of the advantages of dividing cognitive labor is precisely that a scientist committed to that approach no longer needs to worry about the possibility of ~*P*. Arguably, this kind of practical commitment to *P*— acting as if *P* at every turn—is tantamount to belief that *P*.

Although these reasons might be contested, I will accept without further argument that the collective rationality of the community requires each individual scientist in this situation either to believe *P* or to believe ~*P*. If the scientists are individually rational, then rationality must allow agents in comparable circumstances to come to different beliefs; that is, epistemology must be *permissive*.

Many contemporary epistemologists deny this kind of permissiveness. For example, Roger White (2005) argues against any epistemological account which allows for agents in possession of the same evidence to reach different conclusions. Anthony Brueckner and Alex Bundy rebut White's arguments but nevertheless accept what they call "the plausible 'Pyrrhonian' principle whereby one always ought to withhold judgment when one's evidence equally favors a belief of *P* and a belief of ~*P* . . ." They add, "While Pyrrhonian skepticism may not be attractive, this principle certainly is" (2012: 166).[2] The principle condemns at least some scientists as irrational if, in the face of the same evidence, some come to believe *P* and others to believe ~*P*.

Similar positions are common in the literature on the epistemology of disagreement, where an epistemic peers' belief that ~*P* is taken to undercut one's belief in *P*. According to the *equal weight view*, advocated by Adam Elga (2007) and others, the peer's belief that ~*P* weighs against *P* just as much as one's own reasons for *P* weigh in favor of it.[3] David Christensen notes the tension between these responses to disagreement and the group-level strategies like the ones I discussed in the previous section. He writes,

> It's quite plausible that knowledge is best advanced by people exploring, and attempting to defend, a variety of answers to a given question. Perhaps, human psychology makes this easier to do when investigators actually have

2. The Pyrrhonian skeptic appeals to the possibility of acting for all the world as if *P* while not believing *P*, a possibility which I set aside here.

3. Ruth Weintraub (2013) surveys and rebuts arguments against the possibility of rational disagreement among epistemic peers.

> a lot of confidence in the hypotheses they're trying to defend. Certain sorts of inquiry might well work best when a variety of investigators have irrationally high levels of confidence in a variety of pet hypotheses. . . . . But I would argue that the patterns of belief are no more epistemically rational for all that. (2007: 215)

So Christensen insists that, from the standpoint of epistemology, scientists participating in the collective system would be irrational.

3. Putting It Together

The previous sections have pointed to an apparent tension between the demands of individual rationality and the demands of the scientific community. Although this tension is recognized both by philosophers of science and by epistemologists, I want to argue that both sides are mistaken. Scientists can responsibly participate in scientific communities.

The problem is that both sides presume that individual, epistemic rationality is just aimed at one univocal thing. The alternative is what I have elsewhere called the *James-Rudner-Douglas* or *JRD thesis*: "Anytime a scientist announces a judgement of fact, they are making a tradeoff between the risk of different kinds of error. This balancing act depends on the costs of each kind of error, so scientific judgement involves assessments of the value of different outcomes" (Magnus 2013).

The label is taken from the name of three exponents of it across the last century or so: William James, Richard Rudner, and Heather Douglas. James famously argues that we are subject to inconsistent epistemic demands. He writes, "Believe truth! Shun error!—these, we see, are two materially different laws; and by choosing between them we may end up coloring differently our whole intellectual life" (1948: 100). James argues that this opens up a space for belief in the absence of sufficient evidence. If my personal beliefs overextend in this way, I accept a greater risk of error but gain the chance of believing a truth which I otherwise would not have gotten to believe. He is primarily thinking of personal and religious matters, and argues simply that we should "respect one another's mental freedom" (1948: 109).[4] Rudner (1953) makes a parallel argument for scientific claims, but hopes that the appropriate value judgments might be provided by an objective science of ethics. Douglas (2009) reiterates Rudner's argument but without the dream of objective values. She recognizes that belief choice necessarily involves a balance between the risk of believing something false and the risk of failing to believe

4. Weintraub (2013) invokes James to argue for the possibility of rational disagreement. Her primary focus is philosophical disputes, so she (like James) does not extend the result to scientific disagreement.

something true—a balance which can only be struck by personal or social commitments. She writes, "Within the parameters of available resources and methods, some choices must be made, and that choice should weigh the costs of false positives versus false negatives. Weighing these costs legitimately involves social, ethical, and cognitive values" (2009: 104).

If the evidence favors P, then one might still suppose that the only rational options are to believe P or to believe nothing. For example, Richard Foley (1993) invokes James in arguing for the possibility of rational disagreement but argues that personal commitment only sets a threshold of rational confidence that will suffice for belief. On this view, the JRD thesis just allows for some individuals to rationally believe while others are rationally agnostic. This fails to recognize the full force of the thesis. Commitments do more than define a single threshold to balance the value one might gain by believing truth against the cost one might pay for believing in error. Rather, different possible truths may offer different value, and different possible falsities may impose different costs. This allows for some people to rationally believe P and others to rationally believe $\sim P$.

Imagine a community of scientists who are arranged so as to have the best collective distribution of their labor. In this community, the work of some scientists presupposes P, and the work of others presupposes Q which is incompatible with P. Moreover, presume that things will go best for science if the former scientists believe P (and $\sim Q$) while the latter believe Q (and $\sim P$). With the JRD thesis in mind, it becomes clear that agnosticism could make their personal epistemic situation worse too.

Consider a particular scientist—call her Jane—who believes P and structures her work around the assumption. The scientific evidence is suggestive but not entirely decisive, and a more cautious scientist might be agnostic in precisely the same circumstances. Jane values believing a truth about the matter more highly than avoiding a false belief. More specifically, she values the possibility of believing P if it is true more than she is averse to having a false belief if P is false. Another scientist—call her Sarah—structures her work around the alternative approach. She values the possibility of believing Q if it is true more highly than the cost of falsely believing Q if it is false. Since Q is incompatible with P, Sarah misses out on the opportunity to believe P and is willing to accept the risk involved in believing $\sim P$. It is rational for Jane to believe P and for Sarah to believe $\sim P$.[5] The demands that individual and collective rationality place on them coincide.

One might think that Jane and Sarah are illicitly cutting corners. If a scientist

5. A standard claim in the epistemology of disagreement is that finding rational, well-informed peers who have opposing beliefs should cause Jane to doubt her own belief. That could be made compatible with my claim here by denying that Jane and Sarah are epistemic peers because they assign different values to the risks and benefits involved, although 'epistemic peer' is not usually defined in this way.

has any business believing *P* or *Q*, then evidence and overwhelming reasons will eventually be uncovered. Yet no evidence is ever absolutely overwhelming, and evidence which would generally be taken as convincing might not be uncovered until after they are dead.[6] For each scientist, agnosticism would guard against the danger of her having a false belief. However, it increases the chance that she will *never* end up believing the truth. The JRD thesis means that both are legitimate epistemic considerations for her.

Note that the JRD thesis only provides what Douglas calls an *indirect role* for values, one which turns on "the sufficiency of evidence, the weighing of uncertainty, and the consequences of error" (2009: 103). The values that enter into Jane's believing *P* are the benefit of believing *P* if *P* is true, the cost of not believing *P* if *P* is true, the benefit of not believing *P* if *P* is false, and the cost of believing *P* if *P* is false. These are precisely the benefits of accuracy and the cost of error. There is no place for how much she would value unconditionally believing *P* (regardless of its truth) or how much she would value *P* being the case. Deciding belief based on the latter values would just be wishful thinking.[7]

One might still worry that the JRD thesis opens the door to madness. Since there is nothing in this argument which depends on the content of *P*, a hypothetical Jane can be posed for any belief whatsoever. Someone with a crazy array of value commitments can rationally come to have crazy beliefs. However, a parallel possibility arises for any theory of rationality, because someone who believes crazy premises can rationally derive crazy conclusions from them. The informal solution, the path for avoiding madness, is to eschew crazy beliefs and crazy value commitments. There is no formal test for craziness, of course, but this is true as much for beliefs as for commitments.

Moreover, there are some qualitative constraints which arise especially for the scientific case. As Kuhn points out, some commitments would be tantamount to giving up science altogether (Kuhn 2012: e.g. 79). So the socially-defined bounds of science can give us some implicit guidance in distinguishing legitimate disagreement from craziness. Although I do not claim this to be exhaustive, let's briefly consider two kinds of constraint.

First, few scientific examples pose choices between *P* and ~*P*. Instead they pose a choice between incompatible but non-exhaustive options like *P* and *Q*. There are indefinitely many other possibilities which have no scientific plausibility whatsoever. Any commitments which were sufficient to lead to Jane's believing unscientific *R* would be sufficient to make her not a scientist anymore. It may be that sometimes

6. They can even have good reason to think that they will not live to see the evidence, if they are old and part of a long-term research program.

7. Some authors argue for a broader or direct role for commitments in deciding belief. For example, Biddle (2013) suggests that arguments from underdetermination, if they make any place for values in scientific inference, would ultimately make a place for any contextual factors whatsoever.

the range of possible commitments a scientist might have and still be a scientist is narrow enough to rule out any freedom to believe. For Kuhn, the possibility of hewing to something besides the dominant paradigm only opens in time of crisis.

Second, a scientist must be honest about the evidence. As the research guided by *P* fails to progress and becomes weighed down by anomalies, there is no determinate point at which reason demands that Jane give up either the research or her belief in *P*. However, she ought not pretend there is progress when there is not or deny that there are anomalies. Her commitments give her freedom to dispute the *significance* of these things but not their *existence*. As Imre Lakatos insists, there is scope within rationality both for "pigheadedness" and "modesty" in pursuing rival research programs (1978: 113). He writes, "One may rationally stick to a degenerating research programme until it is overtaken by a rival *and even after*. What one must *not* do is to deny its poor public record" (1978: 117). Some value commitments would make it difficult or impossible for Jane to even recognize salient evidence, and those commitments would make her scientifically beyond the pale.

4. Taking It Apart

The previous section concluded with the idea that there are nonformal constraints on what value commitments scientists can have *qua* scientists. The constraints are fallible, defeasible, historical products, just like all the structure of science.

This makes for a possibility that seems not to have been noticed in the existing literature: At some point in the history of science, there could be premises or commitments which are scientifically out of bounds but which would not count as crazy by the standards of individual epistemology. So the commitments of some individuals might allow them to be individually rational while making them unable to count as scientists—*a fortiori* making them unable to meet the demands of collective rationality as part of the scientific community.

The standard line, discussed above, is that collective rationality might require more disagreement than individual rationality can allow. The point here is the reverse: Given particular constraints and commitments at a time, collective rationality might allow *less* disagreement than individual rationality requires.

To make this concrete, consider the search for extraterrestrial intelligence (SETI). Let *P* be the proposition that there is intelligent life on some other world which could be detected by a systematic SETI project. Direct evidence for *P* is lacking. One indirect argument relies on the Drake Equation, a formula which calculates the number of detectable civilizations from the total number of stars which could support life, the fraction of those which will have planets that could support life, the fraction of those which actually have life on them, and so on. Every term in the equation is subject to considerable uncertainty, making it reasonable to be agnostic about *P*. However, it is

plausible to think that SETI would be pursued with greater vigor and ingenuity by scientists who believe *P*. Since alleged alien signals should not be accepted too readily, it would be good to have some scientists who believe ~*P* and who would subject the signals to critical, sceptical scrutiny. So the community might maximize its chances of identifying alien civilizations (if there are any) by having some members who already think that there are (or probably are) such civilizations and others who already think that there are (or probably are) none. If individual rationality required scientists to be agnostic—or even if it just required them all to come to the same conclusion on the basis of available evidence—then there would be a conflict between the epistemic good of the scientific community and the individual rationality of scientists. The force of the JRD thesis is that such a conflict can be avoided. Enthusiast scientists would be rational if they saw more positive value in correctly believing *P* (if it is true) than they saw negative value in wrongly believing *P* (if it is false). Cynical scientists would be rational if their values were the reverse of that.[8]

Other cases of scientific controversy might have this same structure. For example, reasonable scientists disagreed about heliocentrism in the time after Copernicus but before Kepler. It is a historical question whether they did so because of disagreement about values and costs and so were rational due to considerations I have discussed here, or whether they disagreed for other reasons. One might worry that, because heliocentrism is logically contingent, the JRD thesis entails that there is a possible array of values such that if someone held them today then it would be rational for that person to reject heliocentrism—but surely no responsible scientist today could be a geocentrist! This just illustrates that some values would be crazily extreme. A present-day geocentrist might obey the demands of individual rationality but, because they could not be a responsible member of the astronomical community, would violate the demands of collective rationality.

Since the requirements of individual and collective rationality are logically independent, there can be no *a priori* guarantee that their demands will harmoniously align. Whether conflict or harmony is realized depends on the structure of the community, the state of science, and the value commitments of would-be scientists.

What I hope to have shown, contra the standard line, is that conflict is not inevitable. The space opened by the JRD thesis makes it possible for scientists to obey individual rationality while participating in collectively rational arrangements.

Acknowledgements

I would like to thank Jason D'Cruz, Gunnar Babcock, John Milanese, and Brad Armour-Garb for helpful feedback on earlier versions of this paper.

8. Since it would be bad for the whole community to obsess over SETI, perhaps it would be best for the community if many scientists were agnostic. The agnostics would be rational if they saw no significant cost in suspending judgement.

References

Biddle, Justin (2013). State of the Field: Transient Underdetermination and Values in Science. *Studies in History and Philosophy of Science*, 44(1), 124–133. http://dx.doi.org/10.1016/j.shpsa.2012.09.003

Brueckner, Anthony and Alex Bundy (2012) On "Epistemic Permissiveness". *Synthese*, 188(2), 165–177. http://dx.doi.org/10.1007/s11229-011-9921-9

Christensen, David (2007). Epistemology of Disagreement: The Good News. *The Philosophical Review*, 116(2), 187–217. http://dx.doi.org/10.1215/00318108-2006-035

Douglas, Heather E. (2009). *Science, Policy, and the Value-free Ideal*. University of Pittsburgh Press.

Elga, Adam (2007). Reflection and Disagreement. *Noûs*, 41(3), 478–502.

Foley, Richard (1993). *Working Without a Net: A Study of Egocentric Epistemology*. Oxford University Press.

James, William (1948). The Will to Believe. *Essays in Pragmatism*, 88–109. Alburey Castell (Ed.). Hafner Publishing Co.

Kitcher, Philip (1990). The Division of Cognitive Labor. *Philosophy of Science*, 87(1), 5–22.

Kitcher, Philip (1993). *The Advancement of Science*. Oxford University Press.

Kuhn, Thomas [S.] (1977). The Essential Tension: Tradition and Innovation in Scientific Research. *The Essential Tension: Selected Studies in Scientific Tradition and Change*, 225–239. University of Chicago Press.

Kuhn, Thomas [S.] (2012). *The Structure of Scientific Revolutions*, fourth edition. University of Chicago Press.

Lakatos, Imre (1978). History of Science and its Rational Reconstructions. *Falsification and the Methodology of Scientific Research Programmes: Philosophical Papers*, volume 1, 102–138. Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511621123.004

Magnus, P.D. (2013). What Scientists Know is not a Function of What Scientists Know. *Philosophy of Science*, 80(5), 840–849. http://dx.doi.org/10.1086/673718

Muldoon, Ryan (2013). Diversity and the Division of Cognitive Labor. *Philosophy Compass*, 8(2), 117–125. http://dx.doi.org/10.1111/phc3.12000

Rudner, Richard (1953). The Scientist qua Scientist Makes Value Judgments. *Philosophy of Science*, 20(1), 1–6. http://dx.doi.org/10.1086/287231

Weintraub, Ruth (2013) Can Steadfast Peer Disagreement Be Rational? *The Philosophical Quarterly*, 63(253), 740–759. http://dx.doi.org/10.1111/1467-9213.12065

Weisberg, Michael (2010) New Approaches to the Division of Cognitive Labor. *New Waves in Philosophy of Science*, P.D. Magnus and Jacob Busch (Eds.), 250–269. Palgrave MacMillan.

White, Roger (2005). Epistemic Permissiveness. *Philosophical Perspectives*, 19(1), 445–459. http://dx.doi.org/10.1111/j.1520-8583.2005.00069.x

Zollman, Kevin J. S. (2010). The Epistemic Benefit of Transient Diversity. *Erkenntnis*, 72(1), 17–35. http://dx.doi.org/10.1007/s10670-009-9194-6