The Bias that Unites Us: A Reply to Keith Stanovich

Neil Levy, Macquarie University, neil.levy@philosophy.ox.ac.uk

––––––––––––––––––––

Keith Stanovich (Stanovich 2021a) accuses me of misreading his book at multiple points. I think he's misread my review, so I guess we're even. Perhaps neither of us were as clear as we should have been. His main point is that I am mistaken in maintaining that the (supposed) irrationality of myside bias is a major theme in *The Bias That Divides Us* (Stanovich 2021b). Here is what I take Stanovich to be up to in the book. A principal aim is to assess the widespread view that myside bias is epistemically irrational (from now, I set aside ecological or instrumental conceptions of rationality, and focus on epistemic rationality alone). Following other Bayesians, Stanovich shows that this widespread idea is ill-founded: in fact, it is perfectly rational to assess the strength of evidence and argument in ways that are a function of one's priors. But he is unwilling to conclude that myside bias is in fact rational.

Rather, he concludes that myside bias is often only *locally* rational. In many cases, it is not *globally* rational. It is globally rational to use one's prior beliefs to evaluate evidence and arguments only "when the present prior belief is not itself determined by myside bias" (35; all page numbers refer to Stanovich 2021b). A good reasoner updates her beliefs on her priors and thereby achieves local rationality, but whether she is globally rational depends on how she acquired those priors. I take it that, on this view, when we are not globally rational we are globally *irrational* (though of course there's conceptual room for cognitive processing that it is neither rational nor irrational). Stanovich seems to suggest as much, writing that "if the prior probability is irrational …. then projecting the prior probability will lead to a globally irrational posterior probability" (43). This is no mere side issue for Stanovich. This is the central contention of the book: globally irrational (albeit locally rational) myside bias explains political polarization. *That*'s the bias that divides us.

So I am puzzled that Stanovich finds my contention that the irrationality of myside bias is the major theme of his book puzzling. I don't think I've gotten Stanovich wrong. I'm well aware that he defends the local rationality of myside bias (and the individual instrumental rationality of globally irrational myside bias). If we ignore the contention that myside bias is globally irrational, we ignore the heart of the book. This isn't supposed to be a gotcha: I take the argument that myside bias is globally irrational to be the major contribution of the book (along with Stanovich's demonstration of bias in some of the scales used in social psychology). After all, the Bayesian pushback against myside bias isn't original to Stanovich (as his careful citations and discussion acknowledge). What is original, and important, is the distinction between its local rationality and its (supposed) global rationality. I think he's wrong, but sometimes significant contributions are importantly wrong and yet still significant (so, yes, Kahneman and Thaler can each keep their Nobel Prizes).

Priors, Irrationality, and Rationality

Let's set aside these issues of exegesis to move on to the heart of our disagreement. Stanovich thinks that we are globally irrational when the priors that we (rationally) project in updating our beliefs are determined by our 'worldview'. Instead, our priors should be determined by evidence. Global rationality requires that H (the focal hypothesis) "must be what I am calling the 'evidence-favored hypothesis' rather than the 'personally favored

hypothesis" (36). The difference between worldviews (elsewhere, Stanovich calls them "peripheral beliefs" (176)) and evidence is that the latter, but not the former, are testable. Rather, we acquire our worldview "largely as a function of our social learning within the valued groups to which we belong and our innate propensities to be attracted by certain types of ideas" (94). My principal disagreement with Stanovich concerns the global rationality of myside bias. I maintain, and he denies, that projection of one's worldview is rational.

I won't repeat the argument in any detail; the reader can find it in the original review (Levy 2021) and much more fully in Levy (2022). Brutally condensed, the central idea is that Stanovich is wrong in maintaining that worldviews are not evidence-based. They are based on *higher-order* evidence, where higher-order consists in facts like the attitudes that those we trust take to propositions. Higher-order evidence is evidence about evidence. The attitudes those I trust take on issues like climate change and COVID-19 provides me with evidence about what evidence they have. Of course, higher-order evidence is often multiply removed from first-order evidence. I might trust the news anchors on Fox or MSNBC, despite knowing that they have no expertise with regard to vaccines, because I take them to be appropriately plugged into chains of testimony that trace back to genuine experts. There's a case for thinking that such chains of testimony – which are not linear but embedded in thick testimonial networks – usually help filter expert testimony and add epistemic value, though of course there are plenty of bad cases in which information degrades across a chain of transmission.

Our assessment of the evidential value of testimony is appropriately responsive to many features of the speaker, as well as the content of the testimony. Is the testimony plausible? Is the speaker well-placed to know what they assert? Do they represent a consensus? Are they trustworthy? (Mascaro and Sperber 2009; Sperber et al. 2010; Harris 2012). These are all *rational* cues to the evidential value of testimony. Trustworthiness, in particular, directly implicates ideology: I weigh the testimony of those whose ideology I share more heavily than that of political opponents (other things equal, of course) because sharing an ideology with me is evidence that they are trustworthy: they are less likely to seek to deceive me and more likely to take care to ensure that the information they impart is clear, accurate and relevant to me (Levy 2019). Sharing my values also makes it more likely that they get things *right*, by my lights: they are less likely to suffer pernicious biases and more likely to weigh the relative risks of false positives and false negatives appropriately. In short, though I agree with Stanovich that many of our priors stem from what he calls our worldviews, I deny that projecting these priors is globally irrational. Indeed, I don't believe there's any real alternative to doing so: any departure from doing so would be epistemically irrational.

Points of Agreement

Let me now emphasise where Stanovich and I agree. I agree that if we update on our worldviews, we will sometimes fail to converge on the truth, even given a substantial amount of (identical) evidence. Belief formation is path-dependent: because we update our beliefs sequentially, there's no guarantee of convergence (Hahn, Merdes, and von Sydow 2018). That's the epistemic condition, not evidence that something has gone wrong with regard to epistemic rationality. The solution to the problem is to harness myside bias, rather than to seek an alternative to it. Stanovich calls for greater ideological diversity in psychology and I

agree: we can't cure myside bias (nor should we seek to) but ensuring greater diversity will ensure that a wider variety of hypotheses are given credence, that more voices are trusted and the logical space explored much more thoroughly (Mercier and Sperber 2017). Though I don't accept his claim that projection of worldviews is globally irrational, we converge (despite *our* ideological differences) on a solution. We're all globally rational, but because rationality can diverge from accuracy, we need to harness myside bias to promote truth.

Points of Difference

Let me very briefly mention two other, relatively minor, points of difference. First, Stanovich seems to accuse me of confusion in my discussion of memetics. He notes that I use intentional language in describing my worries about his use of the idea. I said that I was "unable to imagine a mechanism whereby memeplexes with the capacity to recognize and fight off undesirable credences could develop." As he notes, memeplexes are thought by him to arise through evolutionary processes, and such processes do not involve entities to which intentionality can be literally ascribed. I make no apologies for the use of intentional language, which as he well knows is very common in discussions of evolution. It is common because it is useful, and it is very rare for it to give rise to any confusion on the part of those who use it. The point I was making is that I can't see how evolutionary processes could lead to the development of memeplexes capable of playing the roles he attributes to them. He does not offer a mechanism and I am unable to imagine one.

Second, I don't accept that there is much correlation between reflective processing and a capacity successfully to navigate hostile environments. Reflective processing has its strengths (Levy 2014), but it has its weaknesses too. It is possible to design epistemic traps that work by appealing to reflective processing, and there are many situations in which type-one processing outperforms type-two. We best respond to hostile environments by making them friendlier, rather than seeking solutions in our thinking dispositions

These two points are side issues. The heart and unique contribution of *The Bias that Divides Us*, as I read it, is the claim that myside bias is often globally irrational, and it is this claim that I am principally concerned to deny. It is that issue that principally divides Stanovich and I. Because I hold that myside bias is globally rational, even when we project our worldviews, I maintain that we promote truth principally by managing the environment in which thinkers are embedded, not by addressing their thinking dispositions; one way (far from the only way) we can do this is by ensuring more ideological diversity in psychology and other disciplines. On this, at least, nothing divides us.

References

Hahn, Ulrike, Christoph Merdes, and Momme von Sydow. 2018. "How Good Is Your Evidence and How Would You Know?" *Topics in Cognitive Science* 10 (4): 660–78.
Harris, Paul. 2012. *Trusting What You're Told.* Harvard University Press.
Levy, Neil 2022. *Bad Beliefs: Why They Happen to Good People.* Oxford: Oxford University Press.
Levy, Neil. 2021. "Is Myside Bias Irrational? A Biased Review of The Bias That Divides Us." *Social Epistemology Review and Reply Collective* 10 (10): 31–38.

Levy, Neil. 2019. "Due Deference to Denialism: Explaining Ordinary People's Rejection of Established Scientific Findings | SpringerLink." *Synthese* 196 (1): 313–27.

Levy, Neil. 2014. *Consciousness and Moral Responsibility*. Oxford: Oxford University Press.

Mascaro, Olivier, and Dan Sperber. 2009. "The Moral, Epistemic, and Mindreading Components of Children's Vigilance towards Deception." *Cognition* 112 (3): 367–80. https://doi.org/10.1016/j.cognition.2009.05.012.

Mercier, Hugo, and Dan Sperber. 2017. *The Enigma of Reason: A New Theory of Human Understanding*. London etc.: Allen Lane.

Sperber, Dan, Fabrice Clément, Christophe Heintz, Olivier Mascaro, Hugo Mercier, Gloria Origgi, and Deirdre Wilson. 2010. "Epistemic Vigilance." *Mind & Language* 25 (4): 359–93. https://doi.org/10.1111/j.1468-0017.2010.01394.x.

Stanovich, Keith E. 2021a. "A Rational Disagreement about Myside Bias." *Social Epistemology Review and Reply Collective* 10 (12): 48–57.

Stanovich, Keith. 2021b. *The Bias That Divides Us: The Science and Politics of Myside Thinking*. MIT Press.