

Philosophical tradition has long defined humanity as the rational animal. We – alone, or nearly so among the animals – are capable of standing back and assessing the reasons bearing on our beliefs and on our decisions and carefully weighing them. Rational deliberation sets a norm for us: that’s how we *ought* to behave and to believe. Moreover (tradition has it), this is a realistic norm: while everyone acknowledges that most of us fall short of fully rationality on many occasions, we may realistically aspire to having our most important and most consequential beliefs governed by rational deliberation. Departures from rationality – akrasia, self-deception, and the like – are criticisable because they manifest our falling short of a standard we ought to abide by.

Today, this consensus sounds quaint to many philosophers and psychologists. The new consensus is that rationality is a scarce resource; most of our decisions, most of the time, are governed by intuitive processes. Moreover (though there’s more controversy over this point), many think that it’s unrealistic to expect us to do better. Our minds are structured such that Type (or system) one processes predominate, and Type one cognition is responsive only to a narrow set of cues, not the full range of considerations that a genuine rational processing system must be sensitive to (Evans and Stanovich 2013; Kahneman 2011). We *are* capable of genuine – Type two – rational deliberation, but we engage it only rarely. Worse, we seem *least* likely to engage in such deliberation when we might most need it. When we deliberate about politics or morality – that is, about how to behave in some of the most important spheres of life – we are least likely to be able to deliberate well.

*Bad Beliefs* (Levy 2022) challenges this account of human rationality. It challenges both the empirical and the normative claims that dominate contemporary views of reasoning. It challenges the view that most of us, most of the time, respond to cues *rather than* reasons in acquiring and updating our beliefs. And it challenges the view that we ought to step back and carefully deliberate about important issues.

The empirical claim –that as a matter of fact most of the time we make up our minds by responding to cues and not to evidence – is supported by a great deal of experimental work supposedly demonstrating how Type one cognition crowds out true reasoning. *Bad Beliefs* challenges the standard interpretation of this experimental evidence. I argue that these experiments present participants with more (and different) evidence than the experimenters intend.

Take the *ballot order effect* (Darcy and McAllister 1990): the disposition to prefer candidates higher up on a ballot to those lower down. In a recent article, John Doris cites the effect as the principle evidence showing that people respond to cues, not evidence (Doris 2018). After all, being listed higher in the ballot is not in fact a genuine reason in favor of a candidate – is it? Actually, it might be. We use order to implicate importance or choiceworthiness all the time. The most important items are mentioned earlier in a news broadcast, and if someone asks me for my top 10 restaurants in a town, I am likely to list those I think more highly of earlier than those I think less of. For those people who find the choice between candidates difficult (because they know little about them, or because they have weak preferences), it makes sense to defer to the implicit recommendations they see in ballot order. The ballot order effect is in fact only consequential for those who find the choice difficult (Pasek et al. 2014). They think and act rationally in deferring to a recommendation.

In general, I argue, psychologists have overlooked the extent to which ordinary people make decisions by taking onboard their evidence because they have focused on *first-order* evidence

alone. They have ignored *higher-order* evidence. First-order evidence refers to the evidence that bears directly on belief formation or action (is this candidate trustworthy? What are her policies? Does she have relevant experience?) Higher-order evidence is evidence *about* evidence; the principal source is *other people's opinions*. I may not know anything about the candidates and their policies, but the fact that my well-informed and compassionate friend is voting for the Green candidate might be enough for me: her opinion is higher-order evidence that the Green candidate has good policies (good policies *by my lights*, moreover, given plausible assumptions about my friend and our relationship).

Adducing evidence from experimental psychology and from cultural evolution, *Bad Beliefs* argues that this kind of rational deference to the opinions of others is how human beings always have acquired and updated their beliefs. It promotes a view of human knowledge as a deeply social product: the product of communities of inquirers, each of whom makes a very small contribution on their own. The dispositions that psychologists have tended to see as dispositions to ignore evidence in favor of non-evidential cues are in fact adaptations (genetic or cultural) for the acquisition and transmission of socially produced knowledge. For the most part, they're dispositions to defer, and to defer reliably; that is, to defer to others when they're more likely to be right than we are. We defer to the majority (the *conformity bias* (Henrich and Boyd 1998)), and we're rational to do so, because numbers make a genuine difference to the strength of higher-order evidence. We defer to those who are successful (the *prestige bias* (Henrich and Gil-White 2001)) and we're rational to do so, because their success indicates they're getting things right. We copy behavior when it seems to reflect a convention (we *overimitate* (Lyons, Young, and Keil 2007)) and we're rational to do so because conventions develop as a result of cultural evolution and typically embody knowledge. But we deploy these dispositions to defer and to copy flexibly and intelligently; for example, ignoring the majority if the question is a technical one and the experts take a different view, and placing less weight on testimony if the testifier has a motive to deceive.

So as a matter of fact, we don't ignore evidence in favor of non-evidential cues: instead we defer intelligently, guiding our behavior by reference to higher-order evidence. Of course, this is no guarantee we get things right. It's easy to think of examples where people get things badly wrong by (say) deferring to the prestigious. It might be true that we take (higher-order) evidence into account much more than the psychologists and many philosophers have thought, but the normative claim – that we would do much better by attending to the first-order evidence alone or at least attending to it very much more than we tend actually to do – might nevertheless be true.

*Bad Beliefs* takes issue with the normative claim as well. Of course, it's important to carefully consider first-order evidence. But we should consider the first-order evidence (or give it much weight, at any rate) only when we possess the expertise to assess it. If a question is at all technical (what rate of taxation maximizes government revenues without unacceptably reducing motivation to work? Is the climate changing and is the cause anthropocentric? How should we respond to the COVID-19 pandemic?) we do better to defer to expert opinion than to make up our own minds in the light of the first-order evidence. If we engage seriously with the first-order evidence, we risk losing knowledge. Deference is much more reliable, in general. Of course, expert opinion may be wrong, or it may not yet have had sufficient time to be reliable (the early days of the COVID-19 pandemic is a good example). Even under these conditions, we tend to do at least as well by deferring as by attempting to grapple with the first-order evidence.

Suppose it's true that we rely very heavily on higher-order evidence in acquiring our beliefs and that we're rational to do so. What follows? I argue that we ought to ensure that the higher-order evidence that is available is in fact reliable. To put it another way, given that people form their

beliefs by responding to their epistemic environment, we need to be good stewards of that environment. We live in a heavily polluted epistemic environment, in which apparent (but misleading) evidence is deliberately as well as inadvertently manufactured by others. One aim of policy should be to regulate the epistemic environment, so that bad evidence is more clearly marked as bad and good evidence promoted.

This policy prescription raises a familiar worry, however: isn't it paternalistic to control agents' epistemic environment so that they form better beliefs? I argue it is not. I work through the debate over nudges and paternalism to show that such environmental stewardship is respectful of agents' autonomy. Nudges are ways of altering behavior to get people to choose better by altering cues in their environments (Thaler and Sunstein 2008). For example, one can increase the proportion of their income which people set aside for their retirements by changing the default savings rate on employment contracts. Nudges are often said to be paternalistic because they bypass reasoning (Bovens 2008; Wilkinson 2013). Instead, they take advantage of our cognitive laziness. Of course, that's a view I reject. Defaults are implicit recommendations: in being guided by them, people respond rationally. They're higher-order evidence. How can it be paternalistic to ensure that people are presented with better (higher-order) evidence than with worse? In fact, we can reverse the accusation: *refusing* to nudge (to the extent to which nudges function in this kind of way) is disrespectfully refusing to ensure that people are given reliable evidence.

We began by describing the traditional idea that humanity is the rational animal. Where does *Bad Beliefs* stand in relation to that tradition, and to its Enlightenment apotheosis in Kant's injunction "Dare to know!" (Kant 1991)? Insofar as it insists that human beings respond more rationally than recent naturalistic philosophy has tended to think, it might be held to return to the more traditional view of ourselves as rational animals. But it recasts that rationality, rejecting the individualism that was implicit and often explicit in how tradition conceived it. Kant calls on us to emerge from "immaturity," which he characterizes as "the inability to use one's own understanding without the guidance of others." He thus calls on us to use our "*own* understanding" (Kant, 1991: 54; emphasis in original). Similarly, Locke (1689/1975) warned "we may as rationally hope to see with other Men's Eyes, as to know by other Men's Understandings." *Bad Beliefs* rejects these individualistic views: knowledge is a social product, and we produce it together, by relying on one another. Our understanding should be deployed largely in the service of deferring better, not in thinking for ourselves without the guidance of others.<sup>1</sup>

## References

- Bovens, Luc. 2008. 'The Ethics of Nudge'. In *Preference Change: Approaches from Philosophy, Economics and Psychology.*, edited by Mats J. Hansson and Till Grüne-Yanoff, 207-20. Berlin: Springer, Theory and Decision Library A.
- Darcy, R., and Ian McAllister. 1990. 'Ballot Position Effects'. *Electoral Studies* 9 (1): 5-17. [https://doi.org/10.1016/0261-3794\(90\)90038-A](https://doi.org/10.1016/0261-3794(90)90038-A).

---

<sup>1</sup> *Bad Beliefs* was supported by grants from the Wellcome Trust (WT104848/Z/14/Z) and the Australian Research Council (DP180102384). I am grateful for their support. I am also grateful to Catarina Dutilh Novaes for organizing (and speaking at) the symposium on the book where these contributions originated, and to Lisa Bortolotti for suggesting that *Philosophical Psychology* publish it.

- Doris, John M. 2018. 'Précis of Talking to Our Selves: Reflection, Ignorance, and Agency'. *Behavioral and Brain Sciences* 41: 1-12. <https://doi.org/10.1017/S0140525X16002016>.
- Evans, Jonathan St. B. T., and Keith E. Stanovich. 2013. 'Dual-Process Theories of Higher Cognition: Advancing the Debate'. *Perspectives on Psychological Science* 8 (3): 223-41. <https://doi.org/10.1177/1745691612460685>.
- Henrich, Joseph, and Robert Boyd. 1998. 'The Evolution of Conformist Transmission and the Emergence of Between-Group Differences'. *Evolution and Human Behavior* 19 (4): 215-41. [https://doi.org/10.1016/S1090-5138\(98\)00018-X](https://doi.org/10.1016/S1090-5138(98)00018-X).
- Henrich, Joseph, and Francisco J Gil-White. 2001. 'The Evolution of Prestige: Freely Conferred Deference as a Mechanism for Enhancing the Benefits of Cultural Transmission'. *Evolution and Human Behavior* 22 (3): 165-96. [https://doi.org/10.1016/S1090-5138\(00\)00071-4](https://doi.org/10.1016/S1090-5138(00)00071-4).
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Penguin UK.
- Kant, Immanuel. 1991. 'What Is Enlightenment'. In *Political Writings*, by Immanuel Kant. Cambridge: Cambridge University Press. <http://www.columbia.edu/acis/ets/CCREAD/etscc/kant.html>.
- Levy, Neil. 2022. *Bad Beliefs: Why They Happen to Good People*. Oxford: Oxford University Press.
- Locke, John. 1975. *An Essay Concerning Human Understanding*. Oxford: Oxford University Press.
- Lyons, Derek E., Andrew G. Young, and Frank C. Keil. 2007. 'The Hidden Structure of Overimitation'. *Proceedings of the National Academy of Sciences of the United States of America* 104 (50): 19751-56. <https://doi.org/10.1073/pnas.0704452104>.
- Pasek, Josh, Daniel Schneider, Jon A. Krosnick, Alexander Tahk, Eyal Ophir, and Claire Milligan. 2014. 'Prevalence and Moderators of the Candidate Name-Order Effect: Evidence from Statewide General Elections in California'. *Public Opinion Quarterly* 78 (2): 416-39. <https://doi.org/10.1093/poq/nfu013>.
- Thaler, R.H., and C.R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth and Happiness*. New Haven, CT.: Yale University Press.
- Wilkinson, T. M. 2013. 'Thinking Harder About Nudges'. *Journal of Medical Ethics* 39 (8): 486-486. <https://doi.org/10.1136/medethics-2013-101687>.