SERRC
Social Epistemology
Review & Reply Collective

http://social-epistemology.com
ISSN: 2471-9560

Predictably Rational: A Further Response to Grundmann

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

---

Levy, Neil. 2021. "Predictably Rational: A Further Response to Grundmann." *Social Epistemology Review and Reply Collective* 10 (12): 75-79. https://wp.me/p1Bfg0-6px.

SERRC
Social Epistemology
Review & Reply Collective
10 (12): 75-79. 2021.
https://wp.me/p1Bfg0-6px

There's a certain pleasure in accounts that debunk some of our most highly prized traits, especially when they're seen to stem from science. Scientists and those who see themselves as scientifically minded often tell us that science shows that free will is an illusion (usually without a shadow of real knowledge of the philosophical free will debate, and often on the basis of the shakiest experimental evidence). We see the same thing with rationality. Websites and the press love to repeat the news. The phrase "not as rational as we think" returns more than 66,800 hits, ranging from ABC news to LinkedIn. We have a kind of irrational attraction to the idea we're irrational.

Nudging and Reasoning

Unlike debunking accounts of free will from the sciences, the near consensus from psychology is largely accepted by philosophers, and Thomas Grundmann is well within his epistemic rights in accepting it. It is on this consensus, or rather on a slice of the experimental evidence that supports it, that he builds his case. He accepts the standard story that nudging "bypasses reasoning altogether" (Grundmann 2021). I have offered an alternative explanation of nudges, on which (at least) most of them work by offering implicit evidence, to which agents respond rationally (Levy 2022; 2021; 2019). I continue to believe this account is correct. Here I offer a brief defence.

The arationality account of nudges might for two reasons be seen to be the default, with rivals having the burden of proof. First, the near consensus that we're not as rational as we'd like to think is based on a range of experimental evidence, not merely the evidence that underlies nudging. Second, we might think the psychologists are the experts here. After all, Daniel Kahneman won a Nobel Prize for the work that launched the heuristics and biases program! If the psychologists are the experts, then we should defer to their interpretation of their work (indeed, Grundmann criticizes my interpretation of default effects as implicit recommendations on the grounds that "the psychological literature does not privilege implicit recommendation as the exclusive explanation of the default effect" (30)).

I don't think either of these claims shift the burden of proof to me. The psychological literature is in need a rethink for two reasons. First, it is a heavily polluted literature. Much of the work prior to around 2010 is based on tiny samples and some of it is p-hacked. The replication crisis gripping multiple areas is not confined to social psychology (nor is it even especially bad in social psychology compared to, say, oncology) but it is very real. A lot of work that motivates nudges doesn't replicate. Unfortunately, that's probably true of one of the studies Grundmann cites: the famous experiment that purportedly shows that cues of watching eyes increases prosocial behavior (Northover et al. 2017).

Nevertheless, I have no doubt that a great deal of psychological research (especially, but not only, more recent research) is reliable. It is not because I don't think nudging occurs that I disagree with Grundmann, it is because I think it doesn't work by bypassing reasoning. That brings us to the second reason why I don't shoulder a burden of proof: the consensus that psychology shows us to be predictably irrational is cracking in multiple ways. Psychologists

and philosophers have become aware that a great deal of apparent evidence for motivated cognition failed to account for participants' background beliefs. Classic experiments that purport to show that people respond perversely to evidence fail to take into account their beliefs (about the reliability of the evidence, for instance), and once we take these beliefs into account we have good reason to see them as responding rationally (see, for example, Cook and Lewandowsky 2016; Tappin, Pennycook, and Rand 2020; Tappin and Gadsby 2019; Williams 2019). The case for pervasive irrationality is much weaker than once thought. I add to that case by emphasising the ways in which classic experiments often presented participants with higher-order evidence to which they respond rationally.

Regarding Grundmann's Examples

Let's turn to Grundmann's detailed examples. He accepts that some nudges may present participants with implicit recommendations, but maintains that others don't: they bypass reasoning. Even when an implicit recommendation is given, he suggests, "effort will play at least some role" (29). He gives the following example: suppose that shoppers are informed that products are distributed randomly across shelves. They then know that placement does not constitute implicit recommendation. Nevertheless, he predicts, products placed at eye-level will probably be chosen more often than those placed lower or higher. He suggests that "the simplest explanation" of this effect is that it will be easier "to select products at eye-level than picking up products from lower or higher shelves" (29).

I suspect Grundmann's prediction is correct: shoppers will prefer products at eye level. Does that show, however, that their reasoning has been "bypassed"? I don't think so. Here's one explanation of what might be going on in such cases: product placement is an implicit recommendation and the mechanisms that respond to it are distinct from the conscious reasoning mechanisms that are aware that placement isn't an implicit recommendation. They are encapsulated from domain-general information. That would give us disunified reasons-responsiveness. But I think a simpler and more unified story is more plausible: people simply don't have any strong preferences in these cases, so why not do what comes easiest? This isn't having one's reasoning bypassed: it's not using it when it doesn't add value.

Grundmann notes that on my account, default effects are evidence, but this evidence is weak. He notes, however, that default effects can be "quite substantial" (30). He gives the example of organ donation, and the dramatic effect that changing defaults can have on its rates. Here we need to be careful not to mix up the *strength* of evidence with the *size* of the effect it can have. Just as the proverbial straw can break the camel's back, so a very weak influence can decisively change behavior. Default effects have decisive effects only when people are indifferent between the options or know little about them—or so the evidence concerning the ballot order effect suggests (Pasek et al. 2014). Default effects can be decisive because testimony can be decisive in such circumstances: because it takes very little to tip the balance when things are evenly poised.

Let me turn now to social referencing. Grundmann accepts that peer consensus can be an excellent reason to believe something, but argues that social referencing persists in the face

10 (12): 75-79. 2021.
https://wp.me/p1Bfg0-6px

of defeating information, which indicates it's not rational. He cites the famous Asch conformity experiments, where many participants reported a belief that was plainly false, due to peer testimony supporting that false belief. There is good evidence that most of those who reported the false judgment did so without believing it (Mercier 2017). Nevertheless, a minority apparently accepted their peers' judgment. Is this evidence of irrationality? I doubt it. Different participants come to the experiment with different prior beliefs. Some are confident enough in their judgments to conform behaviorally without changing their judgments: they judge it wise to go along with group while keeping their opinions to themselves (we may think that's spineless, but it's not irrational). Some are less confident. They think it's more likely that their perception is unreliable than that well-placed peers are wrong. That, too, isn't irrational. Sometimes when it seems to one that p one *should* accept that not-p*,* on the basis of testimony.

On Affect and Testimony

I will conclude with a discussion of the affect heuristic. Affect is just information processing, in my view: it is not something that interacts with reasoning, rather it is a way of engaging in implicit reasoning. Here's one influential, albeit controversial, account of how the mind is set up: the mind consists of a set of discrete processing modules, which take information as an input, subject it to processing by the algorithms that together constitute reasoning, and then output a result that becomes an input for further processing. I think we should see affect as functioning in broadly the same kind of way: affect is (imprecise) information that becomes available for further processing. On this picture, being influenced by affect is like being influenced by any other reasoning mechanism.

In my initial response to Grundmann (Levy 2021), I called reliance on affect a kind of reliance on testimony from oneself. Perhaps that was unwise; Grundmann thinks that the idea can only be understood as metaphorical. I do think we can testify to ourselves (for example, when we consult our own diary we receive testimony from ourselves). I had in mind the kind of guidance by information about oneself we see in choice blindness experiments (Hall et al. 2013; Hall, Johansson, and Strandberg 2012). In these experiments, people report attitudes toward various items and their choices are subsequently altered. When they're asked to justify their choices (including choices they had not, in fact, made) they often fail to notice the switch and attribute to themselves a belief on the basis of evidence that they had previously expressed support for it. This is a rational mechanism, I suggest, and might be seen as a kind of reliance on testimony from oneself (perhaps it is the use of such a mechanism that occurs in Grundmann's example of framing outside a conversational context, though I'm not sure I understand the example). I don't want to fight over terminology, however. What matters to me is that such self-reliance (as it were) is a normal, adaptive, and rational way for agents with plans and projects that are temporally extended to behave.

I don't suggest that all nudges work by offering implicit testimony, or indeed by some other mechanism that is fully reasons-responsive. Humans are finite beings, and we cannot weigh all the information potentially available to us, nor even take all our own mental states into

account when making a decision. We are and must be boundedly rational. But the evidence that we respond perversely to evidence is much weaker than has been thought. Much of what looks like irrationality is ordinary (Bayesian) response. I agree with Grundmann we can (and do) acquire knowledge through nudging, but I think we do so in much the same way as we acquire knowledge via testimony.

References

Cook, John and Stephan Lewandowsky. 2016. "Rational Irrationality: Modeling Climate Change Belief Polarization Using Bayesian Networks." *Topics in Cognitive Science* 8 (1): 160–79. https://doi.org/10.1111/tops.12186.

Grundmann, Thomas. 2021. "The Possibility of Epistemic Nudging: Reply to My Critics." *Social Epistemology Review and Reply Collective* 10 (12): 28–35.

Hall, Lars, Petter Johansson, and Thomas Strandberg. 2012. "Lifting the Veil of Morality: Choice Blindness and Attitude Reversals on a Self-Transforming Survey." *PLOS ONE* 7 (9): e45457. https://doi.org/10.1371/journal.pone.0045457.

Hall, Lars, Thomas Strandberg, Philip Pärnamets, Andreas Lind, Betty Tärning, and Petter Johansson. 2013. "How the Polls Can Be Both Spot on and Dead Wrong: Using Choice Blindness to Shift Political Attitudes and Voter Intentions." *PloS One* 8 (4): e60554. https://doi.org/10.1371/journal.pone.0060554.

Levy, Neil. 2022. *Bad Beliefs: Why They Happen to Good People*. Oxford: Oxford University Press.

Levy, Neil. 2021. "Nudging Is Giving Testimony: A Response to Grundmann." *Social Epistemology Review and Reply Collective* 10 (8): 43–47.

Levy, Neil. 2019. "Nudge, Nudge, Wink, Wink: Nudging Is Giving Reasons." *Ergo: An Open Access Journal of Philosophy* 6. https://doi.org/10.3998/ergo.12405314.0006.010.

Mercier, Hugo. 2017. "How Gullible Are We? A Review of the Evidence from Psychology and Social Science." *Review of General Psychology* 21. https://journals.sagepub.com/doi/10.1037/gpr0000111.

Northover, Stefanie B., William C. Pedersen, Adam B. Cohen, and Paul W. Andrews. 2017. "Artificial Surveillance Cues Do Not Increase Generosity: Two Meta-Analyses." *Evolution and Human Behavior* 38 (1): 144–53. https://doi.org/10.1016/j.evolhumbehav.2016.07.001.

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.

Tappin, Ben M., Gordon Pennycook, and David G Rand. 2020. "Thinking Clearly about Causal Inferences of Politically Motivated Reasoning: Why Paradigmatic Study Designs Often Undermine Causal Inference." *Current Opinion in Behavioral Sciences* 34 (August): 81–87. https://doi.org/10.1016/j.cobeha.2020.01.003.

Tappin, Ben M. and Stephen Gadsby. 2019. "Biased Belief in the Bayesian Brain: A Deeper Look at the Evidence." *Consciousness and Cognition* 68: 107–14. https://doi.org/10.1016/j.concog.2019.01.006.

10 (12): 75-79. 2021.
https://wp.me/p1Bfg0-6px

Williams, Daniel. 2019. "Epistemic Irrationality in the Bayesian Brain." *British Journal for the Philosophy of Science*. 1-29. https://doi.org/10.1093/bjps/axz044.