



**SERRC**  
Social Epistemology  
Review & Reply Collective

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

Nudging is Giving Testimony: A Response to Grundmann

Neil Levy, Macquarie University, [neil.levy@philosophy.ox.ac.uk](mailto:neil.levy@philosophy.ox.ac.uk)

---

Levy, Neil. 2021. "Nudging is Giving Testimony: A Response to Grundmann." *Social Epistemology Review and Reply Collective* 10 (8): 43-47. <https://wp.me/p1Bfg0-65w>.

Nudges (Thaler and Sunstein 2008) are ways of changing people's behavior by changing features of the context in which they choose, rather than by giving them explicit arguments and without removing or unduly burdening the options available to them. The nudge program draws on extensive work in behavioral economics and related disciplines, work that has apparently shown that trivial alterations in features of the context of choice can have significant effects on what is subsequently chosen. Thaler and Sunstein advocate deliberately nudging choice, to bring people to choose better. For instance, people tend to undersave for retirement, but they can be nudged to save more. We can take advantage of people's tendency to accept the default chosen for them (Smith et al. 2013), by writing employment contracts that specify a higher rate of savings into a retirement fund.

### **Nudging Belief**

Thomas Grundmann (2021) argues that we can nudge beliefs, as well as behavior. We can take advantage of the sorts of manipulations identified by psychologists and behavioral economists to lead people to change their minds about some topic. For example, we can take advantage of the affect heuristic—people's disposition to take their current emotional state into account when making evaluations—to change people's minds concerning a proposition. We might highlight an incident that is damaging to a politician to make his followers think ill of him and *thereby* make them more likely to accept that he lost the election. Grundmann argues that not only can we nudge belief, but we can also nudge people into acquiring *knowledge*. At least if we accept a particular safety condition on knowledge and a particular way of individuating methods of belief-acquisition, we should think it is possible to nudge people into knowledge.

Thaler and Sunstein call their program libertarian paternalism. It is libertarian because it (allegedly) doesn't unduly constrain options and it is paternalistic because it is in the interests of the person nudged. Grundmann is advocating a kind of epistemic paternalism. Nudging knowledge is paternalistic because (setting aside cases in which people might be worse off for knowing some fact—think of some medical diagnoses) such nudging is in the interests of the person nudged. It is important to note that both nudge programs are predicated on a particular, some would say pessimistic, picture of human beings. We must nudge better behavior or belief because people fall short of being rational animals. We will behave and believe in ways that reflect our biases and the many mental shortcuts we use, like it or not. It follows that rational argument has restricted power to improve us, and nudging might fill the gap.

Grundmann defends the following conditional:

If safe belief is identical to knowledge, and if the relevant method is externally individuated, then epistemic nudging that results in knowledge is possible (8).

I take no stand on the truth of that conditional. However, I reject the picture of human beings that motivates Grundmann and other advocates of nudging. I agree with Grundmann

that we can produce knowledge by nudging, but for different reasons to hm. Whatever knowledge is, it had better be the kind of thing that can be the result of nudging, because nudging is (typically, at any rate) giving reasons, and agents (typically again) respond to these reasons in ways that reflect their reason-giving force.

### **Nudging as Implicit Testimony**

On my view, nudging is usually a way of offering implicit testimony (Levy 2019, forthcoming). Broadly, to nudge behavior or belief (or rather belief, or belief and thereby behavior: we nudge behavior by changing people’s mental states, and that almost always entails changing their attitude toward some proposition), we make some consideration salient, and in making it salient we testify with regard to our attitude to that consideration. People respond in ways that reflect this implicit testimony: if they trust us (or if they don’t *distrust* us and they are indifferent or torn) they will tend to accept our testimony and it will influence their beliefs. If they distrust us, or if they have better sources of information, they will set it aside. In other words, though they don’t deliberate about it and may be entirely unaware of how they integrate nudges into their mental processes, they respond to nudges in precisely the same way they respond to testimony, and they are right to do so: nudging just *is* giving testimony (again, in typical cases).

As Grundmann recognizes, if his account is to apply to the real world, he must reject my take on how nudges work. He devotes several pages to this task. He argues that what he calls ‘brute nudges’—nudges that target Kahneman’s System 1, as he characterizes them—do not provide us with reasons in the way I envisage. Rather, they work by “triggering automatic non-rational mechanisms.” Grundmann suggests I over-intellectualize what is in fact typically a ‘brute’ process. More significantly, my view is simply false, at least of many canonical nudges. Before assessing his objections, let me note a couple of misunderstandings Grundmann has of my views.

Grundmann takes me to claim that nudges work by providing reasons to agents only in ‘good cases.’ I don’t make any such good case/bad case distinction, though I do make a distinction between good and bad reasons (of course): just as we can give someone misleading explicit testimony, so nudging can be used to give misleading implicit testimony. Perhaps Grundmann misreads me because he takes an earlier paper of mine (Levy 2017) to argue for the same view. That paper might be read as claiming that when (and only when) certain cues for trustworthiness line up, nudges function as testimony. In fact, I had not developed the testimonial account when I wrote the earlier paper, and I reject the view that only under certain conditions are nudges testimonial (in my considered view, cues for trustworthiness affect not whether a nudge provides testimony, but whether the testimony is given uptake).

Second, Grundmann assimilates my view to Gigerenzer’s account of ecological rationality. Gigerenzer has influentially argued that System 1 processes—the kinds that are supposed to be triggered by nudges—are broadly rational, in the sense that they are adapted to the environment in which they are (or were) usually triggered. In his view, though nudges don’t

*themselves* give us reasons, and in that sense aren't rational, they allow us to track reasons. We're rational to be irrational in the ways they entail (Goldstein and Gigerenzer 2002). That's not my view. I argue that nudges provide us with reasons in the ordinary way. My earlier paper might be assimilated to Gigerenzer's view, but that's not the view I defend in the later paper (or my new book), in which I explicitly make the case that nudges provide us with reasons.

These clarifications of my view should make it easier to refute. I argue that *at least* most of those nudges discussed in the literature work by providing agents with reasons. I can't wriggle out of counterexamples by arguing that it's in our epistemic interests to be nudged like that, or by suggesting that the example is a bad case of some special sort. Grundmann's objections are perhaps even more troubling for me than he thinks, if they are successful. But are they successful?

### **Default Effects**

Grundmann argued, first, that I over-intellectualize nudges. I claim that default effects work by providing implicit recommendations to agents, but that's treating a brute effect as though it had cognitive content. In fact, the selection of defaults is "often effective because of inertia": it simply takes too much effort to make a different choice. Now, I take default effects to be paradigmatic of nudges, both in the sense that they are central to debates over nudges, and in the sense that the mechanism that (I argue) underlies them is the same mechanism underlying many other canonical nudges: in one way or another, most nudges work by making some option salient to us and thereby recommending it—I claim). So if default effects don't provide testimony, then my entire project is shaky.

Do default effects work by taking advantage of our cognitive laziness? That's a common claim. It's important to recognize that we have *no* detailed—let alone validated—models of how System 1 processes work, not even at the functional level. Suggested mechanisms tend to be vague and often tautologous (the affect heuristic works by taking advantage of our disposition to rely on affect; that sort of thing). In suggesting that a whole range of these type 1 processes (I prefer this terminology to systems talk, because I doubt there are discrete systems in play) work by providing testimony, I fall well short of providing a mechanism myself, but I'm going a lot further than most in the field.

Why prefer my 'implicit testimony' account to more standard laziness talk? The best reason is that my view makes sense of and promises to integrate a whole range of such phenomena. The ballot order effect might plausibly be held to arise from cognitive laziness, but it's hard to see how framing effects could; effort is equalized across frames. Nor are ways of making options salient by placing them at eye level easily explained via laziness (it is worth noting that proponents of the laziness hypothesis can cite other evidence commonly interpreted as turning on laziness—in particular, I'm thinking of work on the Cognitive Reflection Test. But *that* interpretation is itself contestable). Further, the implicit testimony account neatly explains not only the power of nudges to influence choice and belief, but also why nudges

are often resisted: they're resisted when agents have reasons to discount the testimony offered. A cognitive laziness account has no neat explanation for the limits of nudges.

So I reject the claim that I over-intellectualize default effects. I think nudges and our response to them reflect adaptive and rational outsourcing of cognition to the environment and to other agents; cognition that is in fact notable for its efficiency and speed. Such cognition is largely automatic and effortless; it is, however, cognition. Let's turn now to Grundmann's second objection. Here he takes me head on: he suggests it's simply false that most nudges work by the provision of testimony.

### Considering Counterexamples

I must confess, I have no idea how many or what proportion of nudges work in the way I suggest. Again, we lack well-worked out models of the underlying processes. I claim that at least most, and a very high proportion of those that have featured in the nudge literature, work in this way. Grundmann offers counterexamples.

The first alleged counterexample is social referencing, whereby someone makes a decision or adopts a belief because they believe that a majority of their peers do the same. Grundmann takes this to be irrational (or arational): "The behavior of one's peer group is not a reason to adopt such behavior, and conforming to this behavior is certainly not reasons-responsive". I think it is: in fact, social referencing features heavily as example of rational offloading of cognition in my new book. Peer disagreement is after all widely held to provide us with higher-order evidence (Matheson 2015). But if peer disagreement provides us with higher-order evidence, that so should peer *agreement* (Levy 2021). If many people who are my epistemic peers (in some undemanding sense of 'peer') believe that  $p$ , then I should believe that  $p$ , other things being equal. Consensus is strong evidence in favor of a claim (again, other things being equal). In fact, that claim seems to follow from the widely accepted claim about peer disagreement: if I fail to adopt the belief that my peers adopt, I find myself in a peer disagreement case.

Similar sorts of things can be said about Grundmann's other two examples: sensitivity to frames and the affect heuristic. We know already that people frame options communicatively: if they prefer an option, they will present it in terms of successes or gains, rather than losses (Sher and McKenzie 2006). We also know that their communicative intent is given uptake (Fisher 2020; McKenzie et al. 2006). So framing is the provision of implicit testimony and responses to framing reflect the uptake of that testimony. The use of affect in guiding decision-making is also rational. Affective processes can integrate information more rapidly than conscious processes (Bechara et al. 1997), and relying on our gut feelings is typically a way of getting things right. Relying on the affect heuristic might be a way of relying on testimony from a surprising (but highly trusted) source: oneself. Deferring to oneself is in fact not limited to reliance on the affect heuristic; it is a reasonably common phenomenon (see Levy forthcoming for more examples).

There are important issues at stake in the debate between Grundmann and I. Most proponents of nudges—and, indeed, many of their opponents—take the evidence from behavioral economics and psychology to show that our much-vaunted rationality is in fact sadly limited. Proponents of nudges often argue that we should use them, because if we don't allow ourselves to be intelligently manipulated to our own benefit, we will be unintelligently or even malevolently manipulated nevertheless. I urge a very different interpretation of the same evidence. If I'm right, then nudging is not manipulative at all. It provides us with evidence, and we integrate and respond to it in just the same way as we respond to explicit evidence. If (and only if) we don't have countervailing evidence or mistrust the source, we use it to guide our decisions. If I'm right, we're rational animals after all.

## References

- Bechara, Antoine, Hanna Damasio, Daniel Tranel, Antonio R. Damasio. 1997. "Deciding Advantageously Before Knowing the Advantageous Strategy." *Science* 275 (5304): 1293–1295.
- Fisher, Sarah A. 2020. "Meaning and Framing: The Semantic Implications of Psychological Framing Effects." *Inquiry* 1-24. doi: 10.1080/0020174X.2020.1810115.
- Goldstein, Daniel G. and Gerd Gigerenzer. 2002. "Models of Ecological Rationality: The Recognition Heuristic." *Psychological Review* 109 (1): 75–90.
- Grundmann, Thomas. 2021. "The Possibility of Epistemic Nudging." *Social Epistemology* 1-11. doi: 10.1080/02691728.2021.1945160.
- Levy, Neil. forthcoming. *Bad Beliefs: Why They Happen to Good People*. Oxford: Oxford University Press.
- Levy, Neil. 2021. "The Surprising Truth About Disagreement." *Acta Analytica* 36 (2): 137–157.
- Levy, Neil. 2019. "Nudge, Nudge, Wink, Wink: Nudging is Giving Reasons." *Ergo: An Open Access Journal of Philosophy* 6 (10). doi.org/10.3998/ergo.12405314.0006.010
- Levy, Neil. 2017. "Nudges in a Post-Truth World." *Journal of Medical Ethics* 43 (8): 495–500.
- Matheson, Jonathan. 2015. *The Epistemic Significance of Disagreement*. Palgrave Macmillan.
- McKenzie, Craig R.M., Michael J. Liersch, and Stacey R. Finkelstein 2006. "Recommendations Implicit in Policy Defaults." *Psychological Science* 17 (5): 414–420.
- Sher, Shlomi and Craig R. M. McKenzie. 2006. "Information Leakage From Logically Equivalent Frames." *Cognition* 101 (3): 467–494.
- Smith, Craig N., Daniel G. Goldstein, and Eric J. Johnson. 2013. "Choice without Awareness: Ethical and Policy Implications of Defaults." *Journal of Public Policy & Marketing*, 32 (2): 159–172.
- Thaler, Richard H. and Cass R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth and Happiness*. New Haven, CT: Yale University Press.

*N. Lery*