

<CT>Rationalization enables cooperation and cultural evolution

<CA>Neil Levy

<CAA>Department of Philosophy, Macquarie University, Sydney, 2109 Australia;

University of Oxford.

neil.levy@mq.edu.au

<https://mq.academia.edu/NLevy>

<C-AB>**Abstract:** Cushman argues that the function of rationalization is to attribute mental representations to ourselves, thereby making these representations available for future planning. I argue that such attribution is often not necessary and sometimes maladaptive. I suggest a different explanation of rationalization: making representations available to other agents, to facilitate cooperation, transmission, and the ratchet effect that underlies cumulative cultural evolution.

<C-Text begins>

Rationalization is usually thought of as a way in which we deceive ourselves or enable ourselves to perform wrongful or unwise actions while preserving a flattering self-image. There are, however, many instances in which we apparently engage in rationalization when neither moral nor prudential goods are at stake. Take cognitive dissonance (Cooper 2007). In the classic essay-writing paradigm, participants are either paid or requested (in a way that is difficult to refuse but leaves individuals feeling that they complied freely) to write counterattitudinal essays. Those participants who were requested to write the essays, but not those who were paid to do so, altered their reported beliefs in the direction of the claims defended in the essays. A natural explanation of these data is that we infer our own beliefs, in part, on the basis of our own behavior. While those who were paid could explain their

behavior by reference to the financial inducement, those who were not paid inferred that they were committed to the view they defended.

Carruthers (2013) argues on the basis of evidence like this that belief attribution in the first-person case uses the same (observational) mechanisms as belief attribution in the third-person case. But if we are able to behave adaptively without attributing beliefs to ourselves, why do it at all?

Fiery Cushman suggests that such belief attribution is a useful fiction. When we rationalize our behavior, we extract information from it, taking ourselves to believe or desire whatever mental states best explain the behavior. This is a fiction when the behavior was caused by subpersonal mechanisms that lack such mental states. But it is a useful fiction, because these mechanisms encode responses that are adaptive. While our behavior is caused by mechanisms that are non-rational, beliefs and desires that *would* cause such behavior are themselves adaptive, and forming them allows us to rationalize – that is, make rational – future behavior. Rationalizing behavior allows us to transform their causes into explicit representations that then come to be available for planning.

The claim that explicitly represented information enables domain generality and thereby planning and flexibility of response is surely correct (Levy 2014). However, we should be wary of concluding that these gains explain the existence or yield the function of this kind of representational exchange. First, it is far from obvious that we need such beliefs and desires to engage in rational behavior. Cushman over-intellectualizes “reasoning,” understanding it as inference over explicit representations. But there is a compelling case for understanding reasoning in a much less intellectualized way, as a flexible response to environmental and internal information (Levy 2019b). Cushman also over-intellectualizes “belief” and “desire,” overlooking the fact that the great majority of these states are

dispositional and may never be explicitly represented (see, e.g., Schwitzgebel 2002). Of course, we may use words however we want, but the substantive point worth emphasizing is that subpersonal mechanisms often drive flexible response in the absence of explicit representations. That being the case, it is unclear what we gain from extracting explicit information from them.

Second, the extraction of information may change future behavior for the worse. Consider cognitive dissonance again. When apocalyptic prophecies fail, cult members often become more committed to the cult, apparently to explain why they have devoted their resources to it (Dawson 1999; Festinger et al. 1956). We are quite easily manipulated into attributing beliefs to ourselves that may not serve our interests, as both cognitive dissonance and choice blindness experiments (Hall et al. 2012; Johansson et al. 2014) show. We confabulate differences between consumer products when the true explanation of our choices is mere position (Nisbett & Wilson 1977), thereby becoming willing to spend more on them in the future. In all these cases, we would do better to refrain from forming explicit representations and allowing subpersonal mechanisms to drive our behavior.

Perhaps the gains in planning and flexibility Cushman points to are among the functions of rationalization. At least equally important, however, is the capacity to make extracted representations available to *other* agents. We are deeply social animals, and our ecological success is very significantly due to our capacity to exchange and aggregate information (Henrich 2015; Richerson & Boyd 2008). To allow for such exchange, we often make ourselves transparent to one another (Funkhouser 2017). We evince a variety of signals that allow others to know what we're thinking. Language is, of course, our most powerful and flexible system of signals. An important part of the reason why we rationalize our behavior is not to make it rational – it already is – but to allow us to communicate its causes

to others, thereby allowing them to make predictions about us and allowing them access to its contents for further epistemic work.

By making ourselves predictable, we enable more efficient cooperation, which is essential for the flourishing of social animals like us (Tomasello 2014). If each of us can predict how the others will behave, we can more efficiently play our part in joint actions, without interfering with one another or introducing redundancies. By communicating which aspects of our behavior are intentional, we indicate what should be copied and what may safely be ignored, thereby facilitating the transmission of cultural knowledge (Levy & Alfano, 2019). Perhaps most importantly, making beliefs explicit allows them to be displayed to others. This both facilitates the “ratchet effect” that makes cultural evolution cumulative (Tennie et al. 2009), whereby previous innovations come to be a platform for further development, and also allows others to critically assess our representations to our epistemic benefit as well as theirs (Levy 2019a). The mechanisms of cultural evolution may work very much more powerfully when our beliefs are made explicit; given the importance of cumulative cultural evolution to our adaptive success, these gains are likely an important part of the explanation of the existence and function of rationalization.

<C-Text ends>

<RFT>References [Neil Levy] [NL]

<refs>

Carruthers, P. (2013) *The opacity of mind: An integrative theory of self-knowledge*. Oxford University Press. [NL]

Cooper, J. (2007) *Cognitive dissonance: 50 years of a classic theory*. SAGE. [NL]

- Dawson, L. L. (1999) When prophecy fails and faith persists: A theoretical overview. *Nova Religio: The Journal of Alternative and Emergent Religions* 3(1):60–82. Available at: <https://doi.org/10.1525/nr.1999.3.1.60>. [NL]
- Festinger, L., Riecken, H. & Schacter, S. (1956) *When prophecy fails*. University of Minnesota Press. [NL]
- Funkhouser, E. (2017) Beliefs as signals: A new function for belief. *Philosophical Psychology* 30(6):809–31. Available at: <https://doi.org/10.1080/09515089.2017.1291929> [NL]
- Hall, L., Johansson, P. & Strandberg, T. (2012) Lifting the veil of morality: Choice blindness and attitude reversals on a self-transforming survey. *PLOS ONE* 7(9):e45457. Available at: <https://doi.org/10.1371/journal.pone.0045457>. [NL]
- Henrich, J. (2015) *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press. [NL]
- Johansson, P., Hall, L., Tärning, B., Sikström, S. & Chater, N. (2014) Choice blindness and preference change: You will like this paper better if you (believe you) chose to read it! *Journal of Behavioral Decision Making* 27(3):281–89. [NL]
- Levy, N. (2014) *Consciousness and moral responsibility*. Oxford University Press. [NL]
- Levy, N. (2019a) Due deference to denialism: Explaining ordinary people’s rejection of established scientific findings. *Synthese* 196(1):313–27. [NL]
- Levy, N. (2019b) Nudge, nudge, wink, wink: Nudging is giving reasons. *Ergo: An Open Access Journal of Philosophy* 6(10). (Online publication). Available at: <https://doi.org/10.3998/ergo.12405314.0006.010>. [NL]
- Levy, N. & Alfano, M. (2019) Knowledge from vice: Deeply social epistemology. *Mind*. Available at: <https://ora.ox.ac.uk/objects/uuid:0c055497-9903-4c7a-b17e-63bd35d25c1d>. <https://doi.org/10.1093/mind/fzz017>. [NL]

Nisbett, R. E. & Wilson, T. D. (1977) Telling more than we can know: Verbal reports on mental processes. *Psychological Review* 84(3):231–59. Available at:

<https://doi.org/10.1037/0033-295X.84.3.231>. [NL]

Richerson, P. J. & Boyd, R. (2008) *Not by genes alone: How culture transformed human evolution*. University of Chicago Press. [NL]

Schwitzgebel, E. (2002) A phenomenal, dispositional account of belief. *Noûs* 36(2):249–75.

Available at: <https://doi.org/10.1111/1468-0068.00370>. [NL]

Tennie, C., Call, J. & Tomasello, M. (2009) Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1528):2405–15. Available at: <https://doi.org/10.1098/rstb.2009.0052>.

[NL]

Tomasello, M. (2014) *A natural history of human thinking*. Harvard University Press. [NL]

<Refs end>