

## *Maladaptive social norms, cultural progress, and the free energy principle*

Matteo Colombo

**ABSTRACT** Veissière and collaborators ground their account of culture and social norms in the free energy principle, which postulates that the utility (or adaptive value) of an outcome is equivalent to its probability. This equivalence would mean that their account entails that complying with social norms has always adaptive value. But this is false, since many social norms are obviously maladaptive.

Veissière et al.'s account of "the ability to perform inferences about the shared beliefs that underwrite social norms" (6) blurs the distinction between descriptive expectations (i.e., beliefs that enough people in a certain situation behave in a certain way), normative expectations (i.e., beliefs that enough people in a certain situation expect others to behave in a certain way), and preferences for conformity to social norms. This distinction grounds many existing accounts of social norms (e.g. Bicchieri 2006; Binmore 1994; Boyd & Richerson 2001; Colombo 2014; Elster 1989; Gintis 2007; Ullmann-Margalit 1977). In particular, according to one influential account, preferences for norm compliance are conditional on having the right descriptive and normative expectations (Bicchieri 2006). Thus, in order to identify and change social norms, the key is to find out and intervene on relevant expectations and their associated preferences for conformity.

Because the active inference model underlying Veissière et al.'s account reduces preferences to probabilities (more precisely, it identifies the utility of an observation with its log probability), it apparently rejects the separation between the probability of an outcome and its utility (Colombo 2017). This move raises the question of how those probabilities should be interpreted—are they subjective degrees of belief (or credences), or are they objective frequencies of outcomes within a given reference class, or are they objective propensities of entities in the actual world (cf., Colombo & Wright Section 3.1)? Clarifying this question is important to facilitate the comparability between Veissière et al.'s model and other accounts of social norms, but also—and especially—to evaluate the prospects of using Veissière et al.'s account for real-world interventions on *maladaptive social norms*.

Veissière and collaborators explain that their account "focuses on the conservative nature of human culture; its ability to ensure that certain well-bounded and highly valuable states are frequented" (38). This focus on conformity and conservation is potentially misleading, since it suggests that any social state brought about by social norm compliance is valuable, insofar norm compliance reduces one's sensory uncertainty. The problem with this suggestion is that there are many social norms that are maladaptive, harmful, morally abhorrent, or just lack any value or social function. Social norms such as female genital cutting, open defecation, binge drinking, and norms of revenge, are widespread in several present-day communities, but have no value. Although these norms may encapsulate statistical regularities in a community, and "act as a guide to what to expect from the future [...] putting uncertainty under control" (Douglas 1986, 48), they are maladaptive. Significant public policy efforts are in fact being made to intervene and change them, by targeting people's expectations and conditioned preferences for following those norms (Bicchieri 2016).

If Veissière et al.'s account is committed to the ideas that "[t]he action with the most affordance... is the one associated to the least *expected free energy*" (47), that expected free energy is equivalent to uncertainty, and that the adaptive value of an action is equivalent to the reduction of uncertainty brought about by that action, then their account cannot obviously explain the emergence and resilience of maladaptive social norms, which nonetheless reduce uncertainty in local cultural contexts. This would mean that Veissière et al.'s account of cultural practices is less unifying than what they claim. It could explain only adaptive social norms, and why they are resilient to cultural change.

Veissière et al. might respond by pointing out that the time scale at which uncertainty is reduced is the key factor for explaining the resilience of maladaptive social norms too. Considering longer time scales, it can make good sense for a community to sample social norms that have low utility or adaptive value—and, hence, high uncertainty. The local, short-term increase in uncertainty produced by sampling and complying with maladaptive norms would serve "the more general process of reducing free energy (either for the individual, because it prepares the organism for potential changes in adaptive contexts, and enlarges the repertoire of responses for the individual or the group)" (40). Compliance with maladaptive norms would thus be in the service of guiding the learning of social norms that are adaptive, and that constitute local or global minima in the larger free energy landscape of sensory samples of different communities.

One challenge, however, is that this response seems to assume that cultural dynamics must promote cultural progress at a longer time scale. That is, over the course of social and cultural history, adaptive social norms would replace maladaptive ones, as the human condition will continue to improve, and more and more (risk-prone) moral trendsetters deviate from accepted, socially harmful norms. The rise in income and wealth that the world has experienced in the past couple of centuries, the increase in average life expectancy worldwide, and the increasing resistance to sexist and racist norms in many cultural communities in the world could be cited as examples of cultural dynamics promoting

progress—although one may also offer several counterexamples, such as present global threats like human-induced climate change, which is sustained by an array of environmentally-harmful social norms.

One way to avoid this challenge is to deny that Veissière et al.'s account should be committed to the idea that cultural dynamics must promote cultural progress, and particularly that social states with high probability tend to have high adaptive value. Denying this commitment would mean, however, that at least in the social domain minimizing free-energy with respect to actions may not be equivalent to maximizing expected utility.

The upshot can be formulated as a dilemma. Either Veissière et al.'s account should be grounded in the free energy principle, or it need not. If it should, then its explanatory scope would be limited to the acquisition, production and stabilization of only adaptive social norms. If their account avoids the commitment to the free energy principle, its explanatory scope could obviously include social norms that are maladaptive, harmful or morally abhorrent, while it would still illuminate the important insight that social norms can helpfully be conceived of as uncertainty minimizing devices.

## REFERENCES

- Bicchieri, C. (2016). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. New York: Oxford University Press.
- Bicchieri, C. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.
- Binmore, K. (1994). *Game Theory and the Social Contract, Vol. I. Playing Fair*. Cambridge MA: MIT Press.
- Boyd, R., & Richerson, P. (2001). Norms and bounded rationality. In G. Gigerenzer & R. Selten (Eds.). *Bounded rationality: the adaptive toolbox*. Cambridge, MA: MIT Press, pp. 281–296.
- Colombo, M., & Wright, C. (2018). First principles in the life sciences: The free-energy principle, organicism, and mechanism. *Synthese*, 1-26. <https://doi.org/10.1007/s11229-018-01932-w>
- Colombo, M. (2017). Social motivation in computational neuroscience. Or if brains are prediction machines, then the Humean theory of motivation is false. In J. Kiverstein (Ed.) *Routledge Handbook of Philosophy of the Social Mind* (pp. 320-340). New York: Routledge.
- Colombo, M. (2014). Two neurocomputational building blocks of social norm compliance. *Biology & Philosophy*, 29(1), 71-88.
- Douglas, M. (1986). *How Institutions Think*. New York: Syracuse University Press.
- Elster, J. (1989). Social norms and economic theory. *Journal of Economic Perspectives*, 3(4), 99-117.
- Gintis, H. (2007). A framework for the unification of the behavioral sciences. *Behavioral and brain sciences*, 30(1), 1-16.
- Ullmann-Margalit, E. (1977). *The Emergence of Norms*. Oxford: Oxford University Press.