# A NATURAL HISTORY OF HUMAN MORALITY
## MICHAEL TOMASELLO

### Reviewed by Jonathan Birch

<u>*A Natural History of Human Morality*</u>

Michael Tomasello

Cambridge, MA: Harvard University Press, 2016, £27.95 (hardback)

ISBN 9780674088641

Working together creates mutual obligations. For example, the members of a football team owe it to each other to work hard for the good of the team. A player who doesn't try hard enough, or who makes a costly mistake, lets the side down. When the team loses, feelings of responsibility, guilt, and shame ensue. Players ought to feel committed to the team and responsible for its failures. If they don't, they deserve to be dropped.

The idea at the heart of Michael Tomasello's 'natural history' of human morality is that these oughts—the oughts of teamwork—were the first oughts. The lives of our Palaeolithic ancestors revolved around collaboration: initially in pairs, then in teams, and, finally, in large, 'tribal' groups. Effective social foraging, be it hunting or gathering, relied on agents forming and executing joint intentions. Tomasello hypothesizes that the basic architecture of human moral psychology—holding ourselves responsible for our actions, feeling that we owe things to others, and that others owe things to us—evolved to make us better collaborators.

This hypothesis implies a close relationship between the origin of morality and the origin of joint and collective intentionality, the main focus of Tomasello's research for over twenty years and the topic of his previous book, *A Natural History of Human Thinking* ([2014]). Tomasello makes a powerful case that these phenomena are indeed

related. If this is correct, then a great deal of previous work on the evolution of morality has been subtly misguided. The focus should never have been on acts of altruism, but on acts of mutualistic cooperation. Moreover, the focus should never have been on explicitly linguistic expressions of moral judgement, hypothesized here to be an evolutionary latecomer, but rather on the way normative judgement, construed more broadly, enters into in the deeper, older cognitive structures implicit in feats of cooperation as apparently simple as two people carrying a log together.

Although primarily a work of evolutionary anthropology, *A Natural History of Human Morality* is notable for its close engagement with moral philosophy. For example, in developing an evolutionary account of our obligations to others, Tomasello draws on Darwall's ([2006]) theory of the 'second person standpoint'. The upshot is that moral philosophers will probably find this book more rewarding than they might expect. Unlike many authors in this area (such as Joyce [2006]), Tomasello is not in the business of debunking morality or defending anti-realist views. His aim is to explain how humans became responsive to non-instrumental normative reasons, which in the first instance were 'second-personal' reasons to fulfil one's commitments to cooperative partners.

The book as a whole is rich, original, and insightful, and it makes a major contribution to the literature on the evolution of normative judgement. It is sure to stimulate a great deal of discussion and debate. With that in mind, my aim in this review is to press on some of the details of Tomasello's account. Rather than summarizing the whole argument, or listing the many points with which I agree, I will highlight two main points of disagreement.

*Morality and Trust*

There are two stages to Tomasello's evolutionary story. At the core of the first stage is the idea that a capacity for normative judgement would have made our ancestors better collaborators in the context of small-scale, two-agent cooperative activities, such as two people hunting an animal together. What, then, makes an agent with a capacity for normative judgement a better collaborator?

Tomasello's answer is that normative judgement strengthened the bonds of trust between cooperating agents. When two agents feel a non-instrumental obligation to perform their part in a collaborative task, and when this is a matter of common knowledge, they will trust each other more deeply than they otherwise would. Agents who trust each other more deeply will undertake riskier but more profitable tasks than agents who trust each other less:

> The kinds of collaborative activities so far described were inherently risky, because they were based only on strategic trust. I think I know your motives, but could easily be wrong [...] What we need to take the risk, then, is for each of us to trust one another more deeply, in a more committed way. What we need is for each of us to feel that we truly *ought* to follow through on our collaboration, that we truly *owe* it to one another. (p. 64)

I am not convinced by this picture of normative judgement as an adaptation for boosting trust. If the relevant design problem was a functional need for deeper bonds of trust between agents, why did natural selection respond by creating a suite of *de novo* cognitive adaptations in agents already capable of trust? Why not simply tweak the intensity of trust by, for example, boosting the release of oxytocin during cooperative activities (cf. Churchland [2011])? To make a convincing case for this adaptive hypothesis, one would need to explain why simpler, affect-based solutions to the putative design problem were not available.

It would, however, be unfair to single out Tomasello here, because virtually all evolutionary accounts of normative judgement face a problem of this form. A functional role for normative judgement is carved out, and it is hypothesized that normative judgement evolved to perform that function. Yet it is a real challenge to find any plausible functional role in the ecology of our Palaeolithic ancestors that could not have been fulfilled more simply by small alterations to existing cognitive or affective structures. For example, the same problem arises for accounts that attribute to normative judgement the function of making us more cooperative in prisoner's dilemma-type scenarios (why normative judgement, rather than more intense feelings of sympathy?), or the function of making us

better at evading punishment (why normative judgement, rather than better strategic reasoning?). Whatever the putative design problem, normative judgement always seems a sub-optimal, over-complicated solution. Ultimately, this is why normative judgement continues to provide such a tantalizing puzzle for adaptationists.

*Tribes as Group Agents*

For Tomasello, the evolution of a 'second-person morality', guiding our behaviour in the context of small-scale joint action, is only half the story of human normative evolution. The second stage is a story of scaling-up. Second-person morality was the foundation for a richer, more expansive collective morality: the enforcement across a tribe of shared social norms, regarded as 'objective' and applicable to all. As Tomasello puts it, a 'view from here and there' (a second-person standpoint) was joined by a 'view from nowhere', a 'fully agent-independent perspective on things' (p. 95).

In explaining this scaling-up, Tomasello follows Richerson, Boyd, Henrich, and colleagues in citing cultural group selection among tribes as the driving force (Richerson and Boyd [2005]; Henrich [2016]). But Tomasello adds his own twist, describing these tribes as full-blown group agents, in the sense of List and Pettit ([2011]):

> If an intentional agent is a living being who not only acts intentionally toward a goal but also knows what she is doing and so self-regulates the process as unexpected contingencies arise, then a cultural group acting towards a collective group goal, with collective commitments to self-regulate their progress toward that goal, may be seen as a collective agent. (p. 119)

Here too It to apply some critical pressure. This passage strikes me as an extreme example of what I see as a regrettable tendency in cultural evolution research: a tendency to reify tribal social groups. That is, a tendency to regard them as more stable, more bounded, and more cohesive than they would actually have been, and to credit them with a greater degree of agency and evolutionary individuality than they actually would have possessed.

This tendency can be criticized on both anthropological and evolutionary grounds. Social anthropologists, for various reasons, have long been sceptical of the 'tribe' concept. Contemporary small-scale societies vary greatly in their demography; and while seasonal aggregations of multiple residential groups are common, these aggregations do not always lead to cohesive, well-bounded tribes (Binford [2001]). Mid-twentieth century anthropologists, especially those working in Australia and Africa, increasingly came to see earlier classifications by 'tribe' or 'ethnolinguistic unit' as imposing sharp boundaries that did not exist on the ground (Berndt [1959]; Southall [1970]; Fried [1975]).

Meanwhile, recent work in evolutionary theory has indicated that the conditions for group adaptation are quite restrictive (Gardner and Grafen [2009]; Okasha and Paternotte [2012]). For group adaptations to evolve, either all within-group variation must be suppressed, or all within-group competition must be suppressed. Would these conditions have been met in Palaeolithic societies? We cannot rule this out, but it is a stretch. In modern small-scale societies, there is substantial variation between cultures, but also substantial variation within cultures (Lamba and Mace [2011]). Food-sharing practices partially suppress competition, but they do not eliminate it completely.

Gamble ([1999], Chapter 2) has argued that, in the context of understanding Palaeolithic societies, it is more helpful to focus on effective and extended social networks, centred on individuals, without assuming that these networks would have carved up neatly into discrete, stable tribes. This individual-centred, network-based approach avoids reifying tribal groups, while remaining fully compatible with the idea that humans have cognitive adaptations for managing effective and extended networks. It simply means we should think of such traits as adaptations of individual agents, not of group agents.

*Conclusion*

These points of disagreement notwithstanding, *A Natural History of Human Morality* is an excellent book.

Speculative histories of the origins of morality are sometimes criticized for being too far removed from empirical evidence. This criticism cannot be levelled at Tomasello. Although his 'natural history' is certainly speculative, the book is rich with detail of experiments he and his collaborators have carried out on chimpanzees and humans, children and adults, to probe the architecture of social cognition in both species, and these experiments are invariably fascinating.

Indeed, the book has a characteristic argument pattern: an ambitious claim about the nature of human morality is made, and just at the moment when one, as a philosopher, expects one's intuitions to be massaged by elaborate thought experiments, the discussion suddenly pivots towards actual data from real lab experiments. Inevitably, the relationship between theory and data is less clean when one does not simply imagine the data, and experiments on primates, human or otherwise, tend to be amenable to multiple interpretations. But Tomasello's determination to bring his conjectures about the evolution of morality into contact with serious empirical evidence is commendable.

The result is a book that, in less than 200 pages, manages to be incredibly wide-ranging, genuinely multidisciplinary, scientifically detailed, and philosophically stimulating. In these respects, it sets a new benchmark for evolutionary accounts of human morality.

*Jonathan Birch*
*Department of Philosophy, Logic and Scientific Method*
*London School of Economics and Political Science*
*j.birch2@lse.ac.uk*

**References**

Berndt, R. M. [1959]: 'The Concept of "The Tribe" in the Western Desert of Australia', *Oceania*, **30**, pp. 81–107.

Binford, L. R. [2001]: *Constructing Frames of Reference: An Analytical Method for Archaeological Theory Building Using Ethnographic and Environmental Data Sets*, Oakland, CA: University of California Press.

Churchland, P. S. [2011]: *Braintrust: What Neuroscience Tells Us about Morality*, Princeton, NJ: Princeton University Press.

Darwall, S. [2006]: *The Second Person Standpoint: Morality, Respect, and Accountability*, Cambridge, MA: Harvard University Press.

Fried, M. H. [1975]: *The Notion of Tribe*, Menlo Park, CA: Cummings Publishing Company.

Gamble, C. [1999]: *The Palaeolithic Societies of Europe*, Cambridge: Cambridge University Press.

Gardner, A. and Grafen, A. [2009]: 'Capturing the Superorganism: A Formal Theory of Group Adaptation', *Journal of Evolutionary Biology*, **22**, pp. 659-71.

Henrich, J. [2016]: *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter,* Princeton, NJ: Princeton University Press.

Joyce, R. [2006]: *The Evolution of Morality,* Cambridge, MA: MIT Press.

Lamba, S. and Mace, R. [2011]: 'Demography and Ecology Drive Variation in Cooperation across Human Populations', *Proceedings of the National Academy of Sciences of the United States of America*, **108**, pp. 14426–30.

List, C. and Pettit, P. [2011]: *Group Agency: The Possibility, Design, and Status of Corporate Agents,* Oxford: Oxford University Press.

Okasha, S. and Paternotte, C. [2012]: 'Group Adaptation, Formal Darwinism, and Contextual Analysis', *Journal of Evolutionary Biology*, **25**, pp. 1127–39.

Richerson, P. J. and Boyd, R. [2005]: *Not by Genes Alone: How Culture Transformed Human Evolution*, Chicago, IL: University of Chicago Press.

Southall, A. W. [1970]: 'The Illusion of Tribe', *Journal of Asian and African Studies,* **5**, pp. 28–50.

Tomasello, M. [2014]: *A Natural History of Human Thinking*, Cambridge, MA: Harvard University Press.

**Share this review**