



## The Birth of Ethics: Reconstructing the Role and Nature of Morality

Philip Pettit and Kinch Hoekstra

Print publication date: 2018

Print ISBN-13: 9780190904913

Published to Oxford Scholarship Online: October 2018

DOI: 10.1093/oso/9780190904913.001.0001

## Commentary on Philip Pettit's *The Birth of Ethics*

Michael Tomasello

DOI:10.1093/oso/9780190904913.003.0011

### Abstract and Keywords

Philip Pettit begins his account of the evolution of morality with early human individuals reporting their experiences to others linguistically, and he maintains that language is a crucial and necessary part of the process throughout. The key is that there were social pressures on early humans to be honest in their reporting in order to maintain a good reputation. But informing others of things truthfully and helpfully is a cooperative social action that may be effected non-linguistically, for example, by pointing to relevant referential situations. And the reputation one creates by engaging in such behavior is not as a skillful language user but as a cooperative one: telling the truth only matters if it either helps or hinders the recipient in her behavioral decision-making. And so, Tomasello's claim is that what is actually doing the work in Pettit's account is not language per se, but rather the cooperative intentions and social actions that underlie certain kinds of speech acts. This suggests the possibility of an account of the evolution of morality based not on language but on cooperation more generally.

**Keywords:** cooperation, evolution of ethics, evolution of morality, language

Michael Tomasello<sup>1</sup>

Many philosophers on many issues—including Philip Pettit in these thoughtful and very powerful lectures on human morality—subscribe to something like “In the beginning was the word.” On most of these issues—including human morality—I myself subscribe to something more like “In the beginning was the deed.” The basic point is that language gains its communicative significance

from the contexts, the forms of life, within which it is used. From a naturalistic point of view, the forms of life that led to human morality are all and only about cooperation. Cooperation is social action—action whose goal is to influence what others think, feel, or do—and linguistic communication is only one form of social action. And so my claim is that to the degree that Pettit's story works, it is not because of language per se, but rather because of the cooperative social action involved.

Pettit begins his story with an initial step in which individuals go around reporting their experiences to others, for example, "The berries on the hill are ripe." There are social pressures on them to be honest, to always tell the truth, in order to maintain a good reputation. But informing others of things for their benefit is a cooperative social action that may be effected non-linguistically, for example, by pointing to **(p.334)** relevant referential situations (Tomasello, 2008). And the reputation one creates by engaging in such behavior is not as a skillful language user but as a cooperative one: telling the truth only matters if it either helps or hinders the recipient in her behavioral decision-making. And so, what is actually doing the work in Pettit's account, I would argue, is not language per se, but rather the cooperative intentions and social actions that underlie certain kinds of speech acts.

Of course, Pettit recognizes the important role of cooperation as background for the evolution of human morality. But this recognition receives only unsystematic treatment in his account, and so it stays in the background and its role is never made explicit. A more systematic treatment puts cooperation in the foreground and highlights it as the foundational infrastructure of human moral psychology that imbues speech acts such as reporting with their moral significance.

#### The primal scene of uniquely human cooperation

In his evolutionary account, Pettit proposes a pre-moral starting state of self-interested creatures, but with some cooperative characteristics (Section 1.2, on "the input condition"). Thus (33), his initial Erewhonians work toward the satisfaction of their desires according to their beliefs with the ultimate goal of promoting their own (and their kin's) welfare. But these are not self-maximizing chimpanzees, because in addition "they have the capacity... to exercise joint attention" and the capacity "to act jointly with one another in pursuit of shared goals." They also, unlike chimpanzees, interact within an essentially egalitarian social system, and "they are able to rely on others, and able to get others to rely on them." And finally, "they are able to build on those capacities and use words in the communicative fashion of natural human language," which usage entails, I would argue, many cooperative structures and motives. If what Pettit is trying to picture here is a step in human evolution that is cooperatively beyond other primates—given **(p.335)** that other primates are not egalitarian and have only

weak or no capacities for acting jointly, attending jointly, and getting others to rely on them within joint activities—I have a more systematic proposal.

What we need first is an interactive context structured by certain morally relevant features. The most fundamental such feature, as Hume already recognized, is that individuals are dependent on one another for their basic necessities. Although the individuals of all social species depend on group mates in some ways, what is needed is something especially immediate and urgent. The obvious candidate is foraging, because nothing is more immediate and urgent than procuring food. After consuming food, primate individuals have only a few hours respite before they must find and consume more. Foraging thus structures almost all of their waking hours. The key observation in the current context is that, whereas other primates mostly forage on their own (they may travel in groups but still obtain and consume food individually),<sup>2</sup> humans, from at least several hundred thousand years ago, forage together with others collaboratively, procuring resources that single individuals could never procure on their own. Based on observations of contemporary human hunter-gatherers, we may also speculate that in their collaborative foraging, early humans had a large measure of partner choice, with a bias, of course, for choosing good cooperative partners. And so, if collaboration for obtaining one's daily sustenance was obligate (i.e., there were no satisfactory solo alternatives), we have a situation in which choosing a good cooperative partner, and being chosen by others as a good cooperative partner oneself, was a matter of life and death.

Pettit's scenario in which individuals report their observations to others can fit quite well into this scenario, and indeed he points out that mutual reliance is an important feature of Erewhonian life. But when **(p.336)** I report to group mates that the berries on the hill are ripe when in fact they are not, the reason my mates become upset with me is because this communicative act leads them to waste their time and efforts. What upsets them is not that I am an inaccurate reporter—they would not care if I report some completely irrelevant fact inaccurately—but rather that I am being uncooperative. If their chasing fruitlessly after ripe berries results in me getting more honey in the opposite direction, then they will become even more incensed that I misled them for selfish reasons. So the social pressure is not for accurate reporting in and of itself, but rather for cooperative, as against selfish, social action. The communicative import of linguistic acts of reporting must involve in some way cooperation, or lack thereof, if they are to be morally relevant.

Tomasello (2016) argues that early human populations adapted to obligate collaborative foraging (with partner choice) by evolving a new social psychology comprising species-unique skills and motivations of joint intentionality. Specifically, individuals became capable and motivated to form with a partner a joint agent “we” that could act together and also know things together in joint attention and common ground with a partner. Early human individuals

understood this “we” as comprising “I” and “you,” perspectively defined, each of whom had its own role and perspective in the joint activity. They also understood that the two of them were interdependent in the joint activity, and that they were equal causal agents in the production of the resource. After repeated instances of a particular form of collaboration, such as hunting antelope, partners came to understand together in common ground how each role in this activity had to be performed if there was to be joint success. These role ideals, as we may call them, were agent-independent, or impersonal, in the sense that they were standards that anyone had to meet if the joint agent was to achieve its joint goal, and also in the sense that the roles of the partners could be reversed and the exact same standards would apply to each role.

There are debates about the exact nature of plural agency in general (cf. Bratman 2014 and Gilbert 2015, e.g., as well as List and Pettit 2011), but for current purposes these are not critical. What is critical is that there is something going on that is not just individual agency. This can (p.337) be clearly seen in comparative experiments in which young children but not great apes do such things as attempt to re-engage a recalcitrant partner (who *should* be doing his part), take leave when they must abandon a collaboration (which they know they *should* not do), and wait to consume the spoils of a collaborative effort until both partners can do so (see Tomasello and Vaish 2013 for a review). Joint agency of this type—individuals sharing a goal while simultaneously recognizing their different roles (the so-called dual level structure)—is morally relevant in a most fundamental way: it creates the possibility of relating to others second personally (as described by Darwall 2006, and others). For current purposes, let us focus on three of its most important features.

First, Pettit claims (47) that a key early step in the evolution of human morality is that individuals discover “that they should embrace ethical standards of desirability and responsibility: they should recognize . . . that morality is the best policy.” But in his account, this has to happen by a kind of reciprocity; I cooperate (by speaking the truth) and hope for the same in return. But that is risky business for the first altruist, as the cooperation may never be returned. Much more stable is mutualism, in which we are simultaneously acting together for a common end, and non-cooperation by either of us means failure for both. In this alternative scenario, the standards that first arise are not general prescriptions for behavior, but rather arise only in the local role standards governing collaborative partners; we both come to understand together in common ground what is needed in each role for our success. These initial standards are not really ethical, only instrumental, but they are nevertheless socially shared standards that exclude not only incompetence but also uncooperativeness. The essential point is that in the concrete instrumental contexts of collaborative foraging (e.g., hunting for antelopes, gathering honey), ideal role standards that simultaneously benefit us both arise quite naturally for

individuals with a psychology built for collaboration. It is then not a long step to the full-blown ethical standards to come.

Second, Pettit talks very little about such things as equality, fairness, respect, and justice, especially as they manifest in the division of resources (a.k.a. distributive justice). He does mention all of these things, **(p.338)** but, as with cooperation in general, they are in kind of an assumed background. In my alternative account, in contrast, the complement to working together collaboratively—indeed a precondition for working together collaboratively—is a trust that in the end we will be able to divide the spoils in a mutually satisfactory way. To be ethical, this mutually satisfactory way of dividing of the spoils cannot be done based on some rule of reciprocity motivated by self-interest, but, rather, it must be done for the right reason. And the right reason is that I understand my partner to be equally deserving as myself. Tomasello (2016) argues that the basis for a genuine appreciation of my partner as equally deserving as myself is a cognitive insight, as first stressed by Nagel (1970): I see others as beings like myself. This is not a desire or preference, but an inescapable recognition of the facts of the matter. In the current account, this recognition comes from the basic structure of joint intentionality: our roles are reversible and the standards of each role apply equally and impersonally to us both (not to mention that we are equally important causal agents in producing the resource). The further argument is that this bloodless judgment of equality turns into a judgment of equal deservingness once collaborators feel the need to bond together to exclude do-nothing free riders from the spoils, which they obviously have done nothing to deserve. It is noteworthy in this context that in dividing resources, as a central ethical challenge in the evolution of human morality, language is not centrally involved.

Third, as Pettit and everyone else recognizes, all evolutionary accounts of human morality must provide some kind of reputational pressure on individuals: they must care what others think of them (in a way that chimpanzees and other apes do not seem to). Pettit's account is more or less the traditional one: I care what they think of me. This account is well-known to be unstable in the sense that it is in constant competition with my selfish motives: I will cheat if I can get away with it. But the shared intentionality account includes an interesting variation: I care what *we* think of me. That is to say, once I have formed a joint goal with someone, we together self-regulate the collaboration to make sure that each of us lives up to our role ideals. If either of us reneges, it is “we” who object, or, more precisely, one of us as **(p.339)** a representative of our “we.” This way of looking at things invokes a kind of Rousseauian mechanism in which I am not only the one being judged, but I am part of the judging collective, as well, which gives the judgment a special legitimacy and so a special force. I live up to the ideal role standards inherent in collaborative activities as a way of cultivating and maintaining my cooperative identity with the “we” in which I am participating. When I judge myself, and perhaps feel guilty, it is because the “I”

that is judging “me” (using socially shared standards) has the representative authority of our “we.”

Together these three structural features of joint intentionality comprise what I would call the cooperative infrastructure of human morality: socially shared role standards whose observance affects us both, a sense of equal deservingness of collaborating partners (to the exclusion of undeserving free riders), and a concern for how “we” are evaluating me. This way of operating is not yet fully moral, of course, but it nevertheless constitutes what we may call a kind of second-personal morality governing how collaborating individuals, or those considering collaboration, ought to relate to one another if they are to maintain their cooperative identities.

#### Joint commitment and legitimate protest

Pettit claims that the next step, beyond simple reports, is that individuals begin to avow their beliefs about the world and to pledge their commitment to the truth of their beliefs. Once again, the first point is that avowing and pledging are speech acts aimed, in the end, at cooperation with others and in maintaining a reputation for cooperation. They make defection or deception more costly as the avower or pledger puts her reputation publicly on the line.

In the shared intentionality way of viewing things, it is unlikely that individuals just went around avowing and pledging that they were telling the truth in their reports. More likely, they avowed and pledged their cooperation in the context of partner choice for collaborative activities: “Let’s hunt some antelope. I’ll be a great partner.” Since the **(p.340)** partner must accept the offer (either explicitly or by simply beginning the activity), and the acceptance itself in effect constitutes a pledge, what we have is a joint commitment (Gilbert 2015). It is true that a joint commitment requires some kind of intentional communicative act, and so in this case, it may be that language is poised to play a key role. But language itself is not necessary for a joint commitment, specifically in routine activities that are well known in the common ground of the participants. If we have previously net-fished together many times in the stream, a simple head nod at the appropriate time of day in the appropriate direction should suffice.

In any case, an additional key feature of moral relevance in the shared intentionality scenario of obligate collaborative foraging (with partner choice) is the mutual self-regulation created by a joint commitment. As Pettit emphasizes, in making a joint commitment, I put my reputation publicly on the line. But, as hinted earlier, the shared intentionality account also has a somewhat different take on these things: it is not about me giving “them” more reason to believe me, it is about me giving “us” the power to regulate me. A joint commitment thus sets the stage for what Darwall and others have called legitimate protest or moral protest or second-personal protest. If you do not live up to your role ideal, I am not just surprised or puzzled, I positively resent it (since Strawson (1962),

the quintessential reactive attitude). I call you on your breach: "Hey, what are you doing?" Because we know in our common ground what you should be doing in your role, I do not even need to tell you the problem. I know that you know the problem and that you will want to correct it, assuming that you want to keep your cooperative identity. If you do not care about your cooperative identity—if you not only cheat but ignore my protest—I will simply leave and choose another partner. So you are faced with the choice of affirming your current selfish identity (in the process of cheating) or reverting to the cooperative identity you expressed when making the joint commitment, in which case you say something like "Sorry, I had a lapse, but now I'm back on board." If you reaffirm the joint commitment in this way, you are in essence saying that my protest is legitimate. You reaffirm that we made a joint commitment to each play our role ideally, and so the **(p.341)** protest is warranted; we both agree you deserve it. This is not just cultivating a reputation with others, but, rather, it is preserving a cooperative identity with all of us in the pool of cooperating partners, including oneself.

Importantly, in this way of viewing things we can preserve two of Pettit's deepest insights into how all of this works. First, he notes that in avowing one's beliefs, one is intentionally forgoing the future possibility of making the excuse that I didn't know any better (I mistook my own mind). In the shared intentionality account, a similar function is performed by the fact that the role ideals of our collaborative activity are mutually known in our common ground, most commonly because we have performed the activity together before previously. If we have hunted antelopes successfully three times in the past, performing our roles in specific ways, it is not a valid excuse to say I did the wrong thing because I didn't know any better. Second, Pettit notes that in pledging one's beliefs, one is intentionally forgoing the future possibility of making the excuse that I changed my mind. In the current account, a similar function is performed by the joint commitment. If we both wander over to the stream and start net-fishing, it would be awkward, but I could change my mind. But if one of us has said "Let's go net fishing" and the other has said "Okay," then neither of us can simply change our mind without giving a legitimating excuse (such as "I hear a child calling in distress"). We are committed. And so again in this case we may translate Pettit's insights from his context of reporting beliefs to our context of acting together cooperatively, with the gain being that we now can firmly ground words in actions that matter to both partners.

A key outcome at which Pettit aims in his account is a sense of responsibility. In the shared intentionality account, making a joint commitment to jointly self-regulate the collaborative activity—and to accept as legitimate any criticism for deviation from our mutually known role ideals—is indeed accepting a responsibility, both to my partner and to our partnership. I will responsibly forgo a potentially beneficial selfish activity because my equally deserving partner does not deserve to be treated that way and, in addition, it would ruin our "we."

I act as an individual responsible to you and to our cooperative **(p.342)** relationship. Nevertheless, despite this nascent sense of responsibility, the early humans we are picturing still have not reached a fully human morality; they are still in the realm of a fairly concrete second-personal morality. The key point is that there is no universality; there is no generalized judgment of “objective” right and wrong. What we have here is “only” ways that individuals relate to one another second-personally.

#### Scaling up to an objective morality

To reiterate, the early humans who participated in obligate collaborative foraging with partner choice were not fully moral creatures. The complete story must therefore have a large second step in which this early second-personal morality got scaled up to something more “objective.” I will be brief.

Once again, at this step we must begin with a new social-interactive context structured by morally relevant features. One possibility (emerging, perhaps, with modern humans some 150,000 years ago or so) is that as human populations began to expand, they began to splinter, with the splintered sub-groups needing to stay together to compete successfully with other expanding groups. The result was so-called tribally structured societies, or cultures, in which different sub-groups were all held together by common ways of doing things; a common set of cultural practices and a common language. To fit in to this tribal structure including in-group strangers, and so to benefit from cultural life (including coordinating with strangers and enjoying protection from other groups), individuals had to conform. Conformity thus came to signal identification with one's cultural group—when the barbarians come, we're all on the same team. And so individuals in cultural groups created group-specific conventions, norms, and institutions to coordinate their activities with in-group strangers and indeed to govern all aspects of their social lives. One set of these could be called moral norms, as they urged conformity to ways of doing things in domains in which there were already second-personal moral practices involving judgments of deservingness. For example, among the most common moral norms **(p.343)** cross-culturally are those governing situations in which conflicts and so group disruption are most likely to occur, including such things as how to divide resources, how to settle conflicts, with whom to have sex, etc.

Unlike the earlier collaborative foragers, these modern human cultural beings did not fully control their social commitments. They could still make and dissolve second-personal joint commitments, but their largest commitments were to the norms and institutions into which they were born. Individuals did not view these as external constraints, however, but rather as legitimate guides to thought and action. The culture's norms and institutions were legitimate because individuals identified with the cultural group and its lifeways—we Waziris do things this way and not that way—in effect making themselves co-authors (with their cultural forebears) of these norms and institutions. As in the case of joint commitments,



norm-following individuals were self-regulating via a supra-individual social structure, it is just that, in this case, this structure was much more general and beyond the individual's control. Indeed, it was not the case that individuals viewed the norm as emanating from and governing just the members of the cultural group; they viewed conformity to the group's norms as identificational, and so the norms govern the behavior not of a finite set of individuals but rather of "anyone who would be one of 'us'," any rational and moral person (since only our culture is rational and moral).

Social norms of all kinds are transmitted to children in an authoritative and objectified voice: this is how we do things, this is how things are done, this is the right way to do things. And so the process of enculturation served to objectify the group's social and moral norms, thus giving them the kind of universality and absolute authority characteristic of moral rules. This is not to say that individuals lost control of their own individual decision-making in the process. One of the most distinctive features of human morality is the existence of moral dilemmas in which behavior in a certain specific context may be governed by multiple, conflicting norms, or else the prevailing norms may conflict with one's individual second-personal interactions and relationships. As just one example, I may promise my friend to do something, but it turns out that in doing it I would **(p.344)** harm the group. Thus, individuals construct for themselves moral principles that help them in navigating their way through myriad and sometimes conflicting moral demands.

Pettit attempts to capture the generality of moral norms by invoking co-avowing and co-pledging in which individuals make group commitments, or even agree that one individual can make commitments for them. This is completely consistent with the current account; it is just that, again, I would emphasize that the content and context of the co-avowals and co-pledges must, to be moral, be grounded in cooperative interactions and relationships.

### Conclusion

Philip Pettit's narrative is among the most insightful and original accounts we have of the evolution of human morality. The weak version of my critique is simply that he has not elaborated to the degree needed the cooperative infrastructure underlying any account based in linguistic interactions. The strong version of the critique is that one simply cannot base an account of human moral evolution in acts such as reporting whose normative dimension is aimed at truth. Evolutionarily, truth is a value in human life to the degree, and only to the degree, that it affects things that matter, such as collaborative success and second-personal relationships. And indeed in the shared intentionality story, cooperative success, sustained over time, is intimately related to how partners view and treat one another and how they expect to be viewed and treated in return.

And so I couch my critique in the form of a friendly suggestion: to provide a convincing account of human moral evolution, Pettit should work harder to ground his language-based account more systematically in the cooperative social interactions that give language its social and moral significance.

Notes:

(1.) Max Planck Institute of Evolutionary Anthropology and Duke University Department of Psychology & Neuroscience.

(2.) The best-known exception is chimpanzees' group hunting of monkeys. But the most plausible interpretation of this behavior is that each individual is attempting to get the monkey for itself—adjusting to the actions of others in the process—and the unintended effect is a kind of surrounding of the monkey. A telling fact is that once they have killed the prey, sharing the meat is difficult (see Tomasello 2014 for a fuller description and defense of this account).

Access brought to you by: