

Forthcoming in *Inquiry*.

Counterfactual Genealogy, Speculative Accuracy, & Predicative Drift

MANUEL VARGAS

UNIVERSITY OF CALIFORNIA SAN DIEGO

ABSTRACT: Explicitly fictional armchair reconstructions of the past are sometimes taken to be informative about philosophical issues. What appeal a counterfactual genealogy has depends on its speculative accuracy, that is, its accuracy in identifying relevant causal, functional, or explanatory particulars. Even when speculatively accurate, counterfactual genealogies rarely secure more than proofs of possibility. For more ambitious deployments of genealogy—for example, efforts to show what properties the target concept in fact predicates—genealogies are hamstrung by the possibility of predicative drift, or changes over time in what the concept predicates. Still, even when counterfactual genealogies fail to tell us about our current practices or concepts, they may identify appealing successors. Philip Pettit's (2018) innovative defense of moral realism and his employment of a counterfactual genealogy provides an illuminating instance of the promise and challenges facing accounts employing counterfactual genealogies.

1. The question

Can explicitly fictional armchair reconstructions of the past teach us about the nature of things?

Accounts of putatively actual and potential pasts have a long and varied history in philosophy, including Enlightenment-era contract theories, nineteenth century debunking genealogies, and contemporary evolutionary proposals. These accounts tend to get their power from the thought, sometimes explicit and sometimes implicit, that the proffered reconstructions are or could be true. On the face of it, a genealogy that is flatly implausible does little to make plausible the ensuing theory, be it a theory of contracts, government, morality, or gender roles.

Therein lies the puzzle. It is not obvious that strict truth is a requirement of these accounts, nor is it obvious that the theorists proposing these accounts have thought that the accounts were true. Indeed, intentional mythmaking has been a part of philosophical discourse at least since Aristophanes' proposal in Plato's *Symposium* of a race of humans who had combined male and female forms. How, then, does an explicitly counterfactual reconstruction—a history that never happened—teach us something about phenomena in the actual world?

In what follows, I take up the question of whether and in what way explicitly counterfactual reconstructions or genealogies are promising for illuminating the nature and role of their target notions. Despite the familiarity of methodologies that draw from reconstructed histories that might never have been, there has been relatively little philosophical discussion about the appeal of the approach, what exactly it hopes to secure, and whether the method is adequate to various theoretical ends. A welcome exception is Philip Pettit's (2018) impressive defense of moral realism, for it is framed by a richly developed account of a version of counterfactual genealogy. The sophistication of his approach provides an opportunity to consider the power and limits of an ambitious form of the methodology.

My view about counterfactual genealogy is one of attenuated skepticism. I argue that what appeal counterfactual genealogies have depends on their speculative accuracy, that is, their accuracy in identifying relevant causal, functional, or explanatory particulars. For example, a genealogy of romantic love that invokes being shot with arrows by a divine being is plausibly less speculatively accurate than a genealogy that begins with pair-bond mating in altricial species. (Altricial species are those where the young require care for a long duration, as opposed to precocial animals where the young are relatively mature from the outset.) As this example suggests, the effort to identify the relevant causal, functional, or explanatory elements will often benefit from engagement with the relevant sciences, where such matters are the stock and trade of the endeavor.

However, even when speculatively accurate, it is difficult for counterfactual genealogies to secure more than comparatively unambitious "proofs of possibility," i.e., demonstrations of the mere possibility that the target notion is as the genealogist claims. For more ambitious deployments of genealogy—for example, efforts to show what properties the target concept in fact predicates (e.g., in debates about moral realism, free will, social ontology, and so on)—genealogies are hamstrung by the possibility of predicative drift, or changes over time in what the concept predicates. Despite doubts about the special power of counterfactual genealogies, I go on to argue that there remains a reason to welcome rigorously constructed counterfactual genealogies: even when they fail to tell us about our current practices or concepts, they may identify appealing successors.

2. Counterfactual genealogy

In this section, I attempt to motivate a general strategy for making plausible the realist bona fides of some concept in a naturalistically respectable fashion. It is not a picture I myself am inclined to adopt, but it provides a useful place to begin exploring the appeal and limitations of counterfactual genealogies.

Suppose we wanted to provide an account of the origins of spatial concepts that would show that the best candidate for properties predicated by spatial concepts would be

those that would be recognized by natural science, broadly construed. One way we might try to produce such an account would be to study the anthropological record. Alternatively, we might study continuities in spatial concepts or practices with proximal species that share some of our cognitive architecture. Or we might try to study the physiology of humans and invoke some evolutionary psychology to tell the story. These would be familiar ways to earn our naturalistic bona fides (i.e., naturalistic in the sense of relying on, referring to, or predicating only properties continuous with or of a kind with the properties of final fundamental science) in a story about the origin of the target concept. Even so, it is dubious that any of these accounts would deliver an important animating metaphysical ambition of the present project, namely, showing that our spatial concepts plausibly predicate only naturalistically respectable properties. First, it might be that we don't have a grip on what relevant properties would figure in those accounts. Second, we might worry that those introducing the concept need not track a stable set of properties. Perhaps they project extra-naturalistic properties onto a variegated set of kinds in a misdirected effort to pick out something stable in the world. (Because 'non-naturalistic' is already used in a variety of ways, let *extra-naturalistic properties* be those that are not continuous with or predicated by final fundamental science.) Third, there could come to be a gap in the acquisition or origin of a concept and what the concept picks out. So, more is required.

One way we might try to make some progress by showing how the acquisition of the target concepts support the predication of naturalistic properties. Call this a positive reconstructive genealogy. (It is positive in the sense of not being a negative or debunking genealogy, in the vein of Nietzsche or Foucault.) To do this, we might begin with a counterfactual, but putatively naturalistic story about how something very much like humans—call them quasihumans—acquired concepts of spatial relations.

On this reconstructive approach, we endeavor to adduce conditions that would have made it nearly ineluctable that quasihumans acquire spatial concepts. If we could tell that story in a way that adheres to reasonable standards of compatibility with naturalism, and if we could respect various important constraints on theory generation, we might produce an account that would give us reason to think that the properties predicated by the spatial concepts of quasihumans would be naturalistically reputable. That is, they would be continuous with, or of a kind with, the properties that figure in final fundamental science, and not anything ruled out by such a science. If we thought that our spatial concepts play similar conceptual roles, that they would have arisen from similar practical and theoretical pressures, and that they were expressive equivalents, then we would have some reason to infer that, like theirs, our spatial concepts predicate naturalistic properties.

The proponent of the reconstructive approach will rightly insist that quasihumans must be like us in most ways. The chief difference is that quasihumans lack spatial concepts.

They have speech, and concepts of various kinds. Like many animals without spatial concepts, they might reliably orient themselves in space and even have a variety of behavioral regularities that conform to the vagaries of spatial properties. What they would lack is concepts about space and spatial relations. They wouldn't be guided by such concepts. They wouldn't reason in spatial ways. That is, they would be incapable of recognizing that this or that response is required in view of spatial considerations. Their communications would lack spatial concepts.

Among the constraints on the story is that it must be parsimonious; the proponent of the approach will also add that it must avoid relying on lucky flukes in the acquisition of the target concept. The account must only involve processes that would be likely to produce concepts whose role in our practical and theoretical life are approximately necessary for quasihumans. To that end, we might find it helpful to stipulate that quasihumans lack sight. To be sure, in being an unseeing species, they are less like us. However, this difference helps reduce the risk that the resultant origin of spatial concepts is fluky, a byproduct of the ubiquitous utility of vision for navigating space. This is a departure from many humans as we now find them, but there is virtue in this difference. If we can pull off the genealogy's explanatory lift with this ballast, so much the better for creatures like us who enjoy sight.

Here, then, is an outline of a counterfactual genealogy of spatial concepts. Given linguistic creatures with standard biological needs and given the fact that human-friendly environments have relatively stable properties for which spatial concepts would enable navigation and exploitation of those properties, the acquisition of spatial concepts would almost necessarily be developed by quasihumans. Why? Coming to possess these concepts would enable the attainment of various intra-agential goods (reliable food, shelter, mating, wayfinding, threat tracking). Further, such concepts would help secure valuable goods in the context of multi-agent life (e.g., a more sophisticated ability to hunt and plan, given spatial discourse; in making it easier to jointly identify where it is safe to flee to when the forest is on fire, and so on). The particulars of how quasihumans construct such concepts from other notions or how they would construct such notions from the dispositions, beliefs, and interests they already have is, of course, the tricky bit. Perhaps they construct them as analogs of temporal concepts. Perhaps they acquire them as an elaboration of indexicals and demonstratives. For the moment, we can bracket these otherwise crucial details.

Recall the general strategy: if we tell this story in a sufficiently rich way, appealing only to naturalistically respectable properties, we might reasonably conclude that as quasihumans go, so go humans. The counterfactual genealogy of spatial concepts would give us a warrant for thinking our spatial concepts aren't spooky, and that the properties they predicate are naturalistically acceptable. We could claim to have shown not just that our spatial concepts could be compatible with a broadly naturalistic picture of the world, but also

something more ambitious, namely, that the target concepts plausibly predicate only naturalistically plausible properties. We might say that we would have demonstrated that a kind of realism about spatial concepts is true.

3. Speculative accuracy

Could a (positive, reconstructive, and counterfactual) genealogy along the lines sketched in the preceding section do better than more familiar forms of philosophical and scientific theorizing, at least in securing the governing ambition to demonstrate that spatial concepts predicate only naturalistically respectable concepts? Would it matter if the particulars of the account came apart from standard scientific accounts of the development of human thought and practices about the involved concept? Even if the promise of the approach is appealing, one might worry that a genealogy along these lines generates more difficulties than solutions.

Consider the nature of the reconstructive genealogy. In some contexts of theoretical reconstruction, a theorist's ambitions can be satisfied with a "just-so" story, because the aspiration is something like illumination of how some feature currently functions in a system, as opposed to an account of that function's causal-historical origin. For example, Michael Bratman's deployment of a method of Gricean creature construction is explicit that the account is intended as a just-so story, where the payoff is in terms of what it illuminates about the functional differences in particular kinds of agency (2007, 49-50). Nothing turns on the accuracy of the account as a causal-historical account of the acquisition of those forms of agency.

Philosophical just-so stories permit us to say that the target notion could have had such a history, and that (in Bratman's case) our current forms of agency might have been acquired because of pressures that figure in standard naturalistic accounts. If the story of acquisition were erroneous, though, it would tell against the core features of the theory. The theory is not about the origins of agency but about how agency (here and now) functions and what relationships different kinds of agency might have to each other. Indeed, for some purposes, it might be an important insight if a just-so reconstruction of some forms of agency gave us concepts that predicate extranaturalistic factors (Vargas 2013, 35-44, 52-78). In contrast, for the reconstructive methodology at stake here, one that aspires to make it plausible that the acquired concepts only predicated (and continue to predicate) naturalistic properties, one cannot be so indifferent to the historical details.

The thesis that it is plausible or very likely that the predicated properties are indeed of a kind that figure in foundational scientific theories requires more than mere possibility, a possibility that could be the product of a fluke. The properties need to be (1) suitably central to the causal-historical acquisition of the concept, and (2) they must persist in ongoing concept use.

Here, I focus on the acquisition side of the story; later, I will return to the persistence side of the story. On the acquisition side of the story, the candidate properties adduced in the reconstruction need to be the sort of thing that would very likely—perhaps nearly ineluctably—figure in the concept acquisition and its role in daily life. This is one motivation for holding that a fluky path to naturalistic predication would not show that the concept we here and now have is rightly understood in a naturalistic, realist manner. Stipulating the unsightedness of quasihumans was an effort at corralling contingency of the fluky sort. So, one issue of concern is how much flukiness is too much, and conversely, whether the method is too intolerant of contingency.

Second, and relatedly: *how true does the story need to be?* Given that it is counterfactual, the genealogy cannot quite pretend to an actual history. At the same time, the force of the account is that it provides something descriptively adequate to the actual world. On pain of pointlessness, counterfactual genealogy must be a method for illuminating reality-relevant features that can inform our understanding of actual concepts and what they predicate. How does this work?

The genealogist could hold that the approach is analogous to, or even an instance of, scientific models. If we think of scientific models as fictional heuristics, the model tells us about the world, but is itself a bit of pretense or fiction. The reader is invited to imagine things to be as the model describes—that's the fiction—without the imagined elements necessarily being true. On this approach, the fictional model earns its keep as a heuristic, for example, in illuminating the phenomena it describes in merely instrumentalist ways (Frigg and Hartmann 2020, §2.2).

One might object that a standard worry about models-as-fictional-heuristics is that this construal seems to mischaracterize the epistemic significance of models. Models are typically deployed because they capture something like truth or instrumental utility, and fictions are sometimes thought to be poor servants to such aims. Yet, suboptimal service is sometimes good enough. If the fiction is conducive to familiar epistemic aims—e.g., knowledge, explanation, or understanding—then the counterfactual genealogist might insist that the counterfactual method has all it needs.

So far, so good. Still, we might worry that we need a bit more, especially if we aspire to a realist metaphysics of the target concept. Without some ready empirical check on the philosophical model, we will find our philosophical model collapsing into little more than an especially elaborate means for producing a conjecture to be measured in all the usual ways (predictive power, explanatory insight, integrative power, descriptive scope, etc.). This might be okay, but it would forego one appealing feature of the method as it has been so far represented, namely, that it gives us special reason to think that the target concept or concepts predicate only naturalistic properties. Absent some tethering of the fiction to the

relevant causal bits, or some equivalent down payment on reality, the genealogy could float free of the actual world. If the account is supposed to be something more than an elaborate generator of candidates for traditional philosophical analysis, then the reconstruction needs a causal etiology that gets right the functional arrangement of the phenomenon as it now operates, or something equivalently explanatorily robust. That is the grounds on which the fiction earns its explanatory keep.

Here, though, we might worry that a special problem emerges with some of the constraints on the construction. The requirement of parsimony and the injunction against flukes raises a special problem for getting the details right. Consider the idea of path dependence, according to which the nature of something is in an important sense an artifact of the order and conditions of its acquisition. The “no flukes” constraint can keep us from seeing actual path-dependence in concepts, practices, and abilities. This difficulty matters because path-dependence often can and does matter for the evolution of practical and theoretical phenomena. In the real world, path-dependence is pervasive and highly sensitive to flukes. For example, evolution—whether biological or cultural—is fluky in exactly this way. An accidental mutation or some unintended variation in a practice can be unexpectedly advantageous for a community (Henrich 2016).

Relatedly, an overly narrow conception of parsimony in our explanations of the development of naturalistic phenomenon can preclude us from capturing the actual causal elements our reconstructions need to get the phenomenon right. Bats and whales both echolocate. Strict parsimony would maintain that this must be a case of shared origins for this distinctive ability. Yet this is not so. Echolocation is a case of convergent evolution, with bats and whales employing different physiology to echolocate with no relevantly shared evolutionary history (el-Showk 2013). What goes for biology goes for conceptual innovations in the history of science. Wallace and Darwin both got to the theory of natural selection by independent paths, as did Leibniz and Newton in their respective inventions of calculus. The natural world is oftentimes fluky and sometimes indifferent to the theorist’s impulse to ontological and methodological parsimony.

The injunction against flukiness and in favor of parsimony is intended to buy confidence that the resultant concepts predicate naturalistic properties. We have seen that this is purchased at the cost that the counterfactual reconstruction closes off certain kinds of possibilities that can matter for evolution of things and concepts in the actual world. The genealogist could reply in the way sketched in the presentation of the genealogy: if a theory can succeed under a more demanding set of standards (e.g., no flukes, no quasihuman vision) then so much the better under more generous conditions (i.e., real world occasional flukiness, possession of vision).

Such a reply misses the point. First, for complex phenomena, there is often no single unified parsimony metric by which the theoretical presumptions can be measured as more and less demanding. Some posits are more demanding along one axis, and less demanding along another. Second, even when there is a sensible unified metric, the methodological stakes are whether the account has the right presumptions to tell us something about the actual world, and not whether the explanatory lift is greater or lesser in some general sense.

For example, we could imagine the evolution of spatial concepts in a world filled with aggressive human-eating plants, flying sharks, or a competitor species that hunts humans by tracking self-aware consciousness. Each of these further stipulations would plausibly involve a more demanding set of starting conditions for a genealogy of spatial concepts. Each would be more demanding, in some readily recognizable sense. All the same, each is plausibly a distortion of the conditions relevant for the genealogy. Evolutionary space, whether cultural or biological, is sensitive to starting conditions. What responses emerge and what follows from them and what concepts come to be developed will vary between reconstructions that include and reconstructions that do not include those differences in starting presumptions. In short, for the counterfactual genealogy's explanatory power to transfer to the actual world, fidelity to the (causally, functionally, explanatorily) relevant elements is what matters. Call this the requirement of relevant speculative accuracy, or more concisely, speculative accuracy.

On the proposal at hand, then, a necessary condition on a counterfactual genealogy telling us whether a given concept plausibly predicates only naturalistically reputable properties is that it has speculative accuracy in ways relevant to the causal, functional, or explanatory features that matter for the acquisition and current operation of the concept in the actual world.

4. Pettit's ambition

The picture of counterfactual genealogy that has been the subject of this discussion has been inspired by Philip Pettit's important regimentation of the method in his *The Birth of Ethics* (2018). There, Pettit offers a counterfactual genealogy of the origins of morality in the service of a defense of naturalistic moral realism. On his account, a plausible, parsimonious, and non-fluky naturalistic reconstruction of the origins of moral concepts can demonstrate that moral concepts predicate only naturalistically respectable properties. To do so, he imagines a counterfactual society (Erewhon) that has creatures like us (Erewhonians) but where those creatures lack normative concepts. From there, he makes an inspired and rigorous case that various practical pressures would almost ineluctably give rise to our familiar normative concepts.

Pettit's account is the best and most detailed version of the methodology, so its successes or shortcomings in grappling with questions about the approach will be as

probative as we are likely to find. It is also an important departure from a still too-common quietism about philosophical methodology. He is forthright about his theoretical ambitions, the methodological presumptions of his account, and the basis on which the methodology is to be deployed. It is a model for doing self-consciously reflective theorizing in a way that does not hide from some of the hardest questions in philosophical theory-building.

In what follows, I focus on two questions: (1) whether Pettit's particular account of counterfactual genealogy avoids worries about its speculative accuracy, i.e., whether it identifies the conditions that are causally, functionally, or explanatorily relevant to the acquisition and current operation of our moral concepts; (2) whether the method can secure the naturalistic bona fides it seeks to secure. With answers to those questions in hand, we'll turn to a broader assessment of the prospects for other counterfactual genealogies.

Addressing these questions requires some interpretive choices. The first concerns the putative uniqueness of Pettit's method, and the second, its explanatory ambitions.

Regarding the first, recall that on my account, the method of counterfactual genealogy is distinctive. It is intended to secure particular theoretical goods. This is faithful to one aspect of Pettit's account, especially threads that treat the method of counterfactual analysis as an alternative to standard forms of philosophical analysis (2018, 22-32). In other places, though, the methodology is presented as a (perhaps more self-conscious) version of a common philosophical methodology. Citing Grice, H.L.A. Hart, Sellars, Lewis, Kripke, Bratman, Railton, Williams, Fricker, and others, Pettit suggests that his approach is just another instance of standard philosophical methodology (2018, 52-4).

Although these two thoughts may be reconcilable, they pull in different directions. First, it is not obvious that the other accounts he cites in fact maintain, as Pettit does, that there is something about their approaches that provides a special license for the naturalistic bona fides of what is predicated by their target concept. Second, the putatively companion accounts are not of a methodological piece. Even when putatively methodological companions are plausibly reconstructive or genealogical, they are not all obviously *counterfactual* in any salient way. Consider Hart's account of law. Many social communities really did have to organize themselves in ways that had rules by which new rules are rightly introduced. Or, as in the causal theory of reference, there have and will continue to be things that get named for the first time, potentially anchoring a causal-historical chain of references.

To be sure, some of these accounts can be understood as offering a model that purports to capture the key historical causal mechanisms of the world in a way that is intended to capture and make plausible that the ongoing contemporary function of the concept continues as the speculated history describes. In this, those accounts share ambitions and methods with Pettit's project. However, some of the accounts he lists are better understood as conjectures that invite acceptance or rejection conditional on their subsequent

fruitfulness. Still others offer analyses of our concepts, the meanings of terms, or necessary constraints on our coming to have various practices, concepts, and concerns. Yet others have pursued theoretical constructs sufficient for explaining the phenomenon as it currently functions, explicitly eschewing claims to be identifying necessary conditions (e.g., Bratman 2007, 50). Given the fact of methodological diversity and the distinctive ambitions of Pettit's approach, I am inclined to think that we do better to accept the picture according to which counterfactual genealogy is a distinctive methodological proposal with potentially distinctive advantages (as in Pettit 2018, 24-28).

The second interpretive choice concerns the ambitions of the method. Sometimes, Pettit writes as though it would be enough to offer a proof of possibility, showing that the conceptual developments crucial to his account "could" or "might" emerge in the way he depicts (24, 27, 299). In other places, the ambition is considerably greater. In these latter passages, he claims that the account endeavors to show how it was "more or less inevitable" (5, 29) and "relatively inescapable" (5) that the identified conceptual innovations would occur.

Given the wider project of defending moral realism, the more ambitious construal better fits both the spirit and particulars of Pettit's project. If his ambition is only to show the conceptual possibility of morality's being naturalistically respectable, then it would be enough to show that we *could* acquire normative concepts that *could* predicate naturalistic properties. Naturalistic compatibility is not nothing, perhaps especially with respect to naturalistic moral realism. Critics of naturalistic moral realism tend to think it is in some deep way inadequate to central features of moral thought and talk, so it would still be an important achievement to give a proof of possibility.

Pettit frequently indicates that he wants more. On the more ambitious reading, his goal is to show that what moral concepts currently predicate is indeed naturalistically *plausible* and not just potentially compatible with naturalism. That is to say, the account maintains that the balance of wider considerations favors the truth of its core theoretical commitments (for more on the distinction between naturalistic compatibility and naturalistic plausibility, see Vargas 2013, 58-60). This stronger ambition is an appealing one on its own terms, and it guides much of what follows. At the end, I'll return to consider whether the weaker claim might be enough for many purposes.

With these two interpretive decisions in hand, we can now turn to some challenges for any account of the genealogy of morality that wishes to avoid worries about speculative accuracy.

5. Interdependence and animality

Above, I introduced the idea of speculative accuracy, or the idea that a counterfactual genealogy must capture the causal, functional, or otherwise explanatory features that matter for the actual concept. There are at least two places where Pettit's account may be insufficiently accurate in its speculative commitments. The first is in its picture of the moral psychology of Erewhonians, and the second is in his account's commitment to methodological anthropocentrism. These issues are of interest in two ways, both as a matter of the appeal of Pettit's particular account, but also as an illustration of how difficult it is to secure speculative accuracy when constructing naturalistic counterfactual genealogies.

Pettit's Erewhonians are depicted as mostly opportunistic, self-interested rational calculators sensitive to reputation (60-3, 245, 302). They are typically described in terms that suggest that they are the classical individual negotiators of rational choice theory. Notably, their lives are not obviously structured by pervasive vulnerability and mutual dependence requiring concerted cooperation and protection.

In all of this, Erewhonians are remarkably unlike human beings as anthropologists and psychologists have often noted (Boehm 2012; Tomasello 2018). Humans are everywhere born into dependence and vulnerability, into reliance on the efforts, concern, and yes, the kindness of others (Okin 1989; MacIntyre 1999). This dependence lasts for years at the outset, and often for years towards the end, punctuated in between with episodes of variable but always imperfect independence. What is true of individuals in this case is true of us as a species. Although not uncontested, it is now a widely accepted view that we have always been interdependent, reliant on ties of affection and altruistic motivations to offset our extreme vulnerability as individuals (Kropotkin 1910; Sober and Wilson 1998; Boehm 2012). The self-interested individualistic negotiators that are the paradigmatic inhabitants of economic and philosophical contracts are, on a broadly naturalistic picture of human history, always only possible against an occluded substrate of altruism and interdependence.

Pettit acknowledges that both humans and morality are social in deep ways. In the context of this recognition, though, it is puzzling why it is supposed to be explanatorily advantageous to delete the fact of our interdependence, altruism, and cooperative impulses from the genealogy of morality. To posit beings for whom the psychology involved in the acquisition of moral concepts is exclusively or at least overwhelmingly self-interested is to posit beings for whom the trajectory of needs is fundamentally different. It is to posit beings fundamentally different from us.

This difference cannot be dismissed on the grounds that if self-interested agents can arrive at morality, then so much the better for more altruistic agents. As we've seen, conceptual innovations are frequently path dependent. It might well be that emphasizing a fuller suite of affective, prosocial, and cooperation-enabling psychological endowments might make an anti-realist, sentimentalist, or non-cognitivist story about the emergence of morality

among interdependent beings seem more appealing than an emergence story that emphasizes the rational calculations of comparatively independent self-interested bargainers. More generally, different background assumptions catalyze different solutions and impose different necessities on the human predicament.¹

A second place one might balk at Pettit's particular counterfactual genealogy of morality concerns its methodological anthropocentrism, according to which morality is not something found in nonhuman animals. Pettit is, again, admirably forthright about this: "The altruistic responses of animals toward their young need hardly count as moral responses. And that, presumably, is because considerations articulated in moral terms play no part in prompting or regulating those responses" (13). On Pettit's account, animals aren't the right place to look for morality as they lack moral concepts. They lack the linguistic abilities that, according to Pettit's genealogy, are crucial for the acquisition of moral concepts.

Perhaps this is right. Yet, it is one thing to note a difference in the articulation of moral reasons and terms and another to treat human morality as a great disjunction or saltation, rather than an especially elaborate instance of accumulation of individual and collective psycho-social technologies that overlap to greater and lesser degrees with those of other animals. This is to say that there is a methodological choice point here between (a) treating morality as a distinctively human activity about which we can learn only incidental things by comparison with other animals, and (b) treating morality as a phenomenon of pro-social dispositions, behaviors, attitudes, norms, and practices emphasizing group binding and social organization that need not but often contingently do involve linguistic tools. On the latter approach, much of our moral life is plausibly a function of wider forms of animality. An account of the origin of moral life and practice that aspires to naturalistic merit should be informed by what we know of those continuities and discontinuities with other animals.

The former (anthropocentric) approach closes off the possibility that our cognitive, theorized conception of morality is itself a species of a wider phenomenon. It closes off the

¹ Pettit's position on this matter changes over the course of *The Birth of Ethics*. In a reply to remarks by Tomasello, after the main text, Pettit (2018, 350-1) concedes that humans are likely more cooperative than his Erewhonians. He then reiterates the putative methodological virtue of a "worst case" methodology that posits a more self-interested psychology (that thought is, of course, what the present discussion contests). On the final page of the book, he pivots to a conciliatory conclusion that holds that he might be able to retain a good deal of his machinery without relying on the self-interested rationality of Erewhonians (358). While I welcome this final twist, my concern remains the same. Starting conditions matter, and until a more altruistic genealogy is told, it is not obvious that what is a "nearly inescapable" conceptual trajectory for self-interested agents is the same thing as it would be for agents who are, from the start, interdependent and cooperative. That's a claim that would need to be earned. If I am reading him correctly, Pettit appears to acknowledge this at the very end of *The Birth of Ethics* (358).

idea that our linguistically saturated practices are a species of a more general phenomenon, the contours of which become visible only when human practices are compared to non-human. In contrast, in studying animal practices we could find that whatever conceptual requirements there are on moral concepts can, at least sometimes, be satisfied by some non-human animals. On this latter approach, the answers to these questions are at least partly empirical matters, not rightly settled by fiat.

For example, Monsó and Andrews (2022) report that empirical studies have found in non-human animals a range of behavior that, at least when found in humans, is regarded as evidence of moral capacities. In great apes, monkeys, and canids, this includes consolation behavior, mourning behavior, helping behavior, and inequity aversion. Moreover, varied sets of three of those four types of behavior have been found in elephants, rodents, and cetaceans (whales, dolphins, and porpoises). More generally, they argue that contemporary empirical findings give us good reason for thinking that capacities for care, complex forms of self-governance, and normativity can be found in non-human animals.

Of course, the interpretation of animal behavior is not always straightforward, but contemporary empirical studies are sophisticated in method and conceptualization. For those of us with a naturalistic bent, these considerations suggest that any story about the emergence of normativity and moral concepts might do well to suspend the assumption of a radical disjunction between us and the rest of the animal world. We might do better to begin with the thought that we are not unique in being affective, social creatures capable of temporally extended organization, and with capacities for individually and collectively managing our psychologies. These particular building blocks seem relatively widely shared, and they are hardly unique to humans. Whether they are sufficient to support morality, or parts of morality, is a question that the naturalistically-minded perhaps ought to regard as open.

Recall that on Pettit's account, language plays a central role in enabling the achievement of moral concepts. One might insist that methodological anthropocentrism can be preserved on this basis. However, this too has been subject to dispute by those studying multimodal communication in animals (Eva Meijer 2019). For example, there is evidence of recursion in birdsong and of syntax in bonobo communication. Further, expanding empirical findings about animal communication are fueling better theories about the evolution of language (Andrews and Monsó 2021, §3.5), which suggests that linguistic abilities may themselves be a scalar matter, rather than a discrete and radical difference in communicative kind.

I have no firm convictions about whether non-human animals have moral cognition, nor about the extent to which their putative moral psychologies overlap with ours. Whatever the truth of the matter turns out to be, the issue seems to me the kind of thing that merits a reckoning in a naturalistic account of the origins of morality, as opposed to exclusion by fiat.

I also worry that in simply adopting a picture of morality as involving a saltation by humans, we are stacking the deck in a way that favors a particular story about the nature or content of morality, and perhaps relatedly, that we may be misconstruing the nature of morality by mistaking part for whole.

Both issues—whether morality is rightly understood as restricted to humans, and whether assuming that it is violates the spirit of naturalistic inquiry—are in themselves interesting questions. Reasonable people might disagree about whether Pettit’s positions on these issues amount to stacking the deck in favor of the naturalistic moral realist. But the fact that we can reasonably disagree about these things matters for our confidence in the account’s speculative accuracy.

To be sure, Pettit can rightly reply that it is possible that these things could make a difference, but that it is one thing to raise the specter that these things *might* matter and another to show that they *do*. Nothing I have said here shows that these discrepancies, or potential discrepancies, undermine the integrity of the counterfactual reconstruction on offer. But the point here is methodological: the less confident we are about the account’s speculative accuracy, the less confidence we should have in the genealogy’s capacity to deliver a theory adequate to its aims.

6. Money problems

Getting the foregoing objections off the ground requires that we accept a reasonably demanding version of a principle of speculative accuracy. On the present approach, these departures matter, if they do, because they show how speculative accuracy is hard to come by. However, a proponent of counterfactual genealogies might object that I’ve exaggerated the importance of speculative accuracy. After all, one might think that there are cases in which speculative accuracy plays no real role in a nevertheless powerful counterfactual account. Pettit points to one such genealogy—the genealogy of money—as motivation for his own account.

On the standard story given in introductory economics texts, tracing back to at least Adam Smith’s *The Wealth of Nations*, money emerges as a flexible solution—a technology or medium of exchange—in a society hampered by the constraints of one-to-one bartering. Although this genealogy of money has been doubted in many of its particulars, it seems to earn its keep in illumination (Pettit 2015, 6-7, 27, 50-53). If that’s right, the counterfactual genealogist could rightly insist that speculative accuracy matters less than I’ve argued. So long as we are scrupulous in appealing to naturalistically credible phenomena, the resultant genealogy can still give us confidence that the reconstructed phenomena predicate naturalistic properties even if the genesis, motivation, role, and sequence of particulars is in error.

I agree that the case of money is illuminating, but I contend that its, uh, cash value is somewhat different than Pettit contends. We don't need a detailed history of money to make it plausible that the concept of money arose in the context of social practices. Where a genealogy does work is in animating a particular picture of the relationship between the concept and the practices from which it sprang. That's what the money example is supposed to show. On the textbook version of the story, money was invented as an efficient solution to the problem of exchanging goods in a world of barter.

Here, though, naturalists have a reason to object. Anthropologists have emphasized that the standard story about money emerging from barter is not just incorrect, but that it has led us to badly misunderstand important features of credit, debt, and the nature of money (Strauss 2016; Graeber 2011; Hockett and James 2020).² If anthropology's revisionism about economics' founding myth is right, money didn't start out as a means for facilitating barter exchange that subsequently produced credit as an even more efficient way to enable the exchange of goods. Instead, accounting of credits and debts was already in place. People eventually invented money as a way of tracking the antecedent phenomenon of credit and debt. Indeed, barter seems to emerge only among people who already have money, and it is what they temporarily turn to when cash is tight. As a matter of historical record, though, money was *not* introduced to solve problems in a barter context. As anthropologist Caroline Humphrey put it, "No example of a barter economy, pure and simple, has ever been described, let alone the emergence from it of money; all available ethnography suggests that there never has been such a thing" (1985, 48).

On this revisionist retelling of the history of money, the erroneous genealogy common to economics textbooks has had pernicious consequences. First, within economics, it created the illusion that the economy somehow floated free of institutional ecologies, that it had nothing essentially to do with taxation and institutionally regulated debts (Hockett and James 2020, 17-34). The false genealogy enabled the ready acceptance of a fictional pre-institutional market, with its own autonomous laws ("the invisible hand"), that was grounded in a pre-governmental, pre-regulatory substrate of barter relations. On the eventually canonical mythology, intervention by governments or institutions was always an imposition on an economic state of nature, rather than a condition on its very existence (Graeber 2011, 43-45). Second, the advent of money and the myth of a pre-institutional barter system seems to have fostered a pair of misleading views about the human predicament: (a) a false picture of human agency as lone contractors animated exclusively by self-interest (the

² Pettit's (2018, 50) discussion of the genealogy of money cites Graeber (2011), but curiously he says nothing about Graeber's argument that the standard genealogy is both erroneous and pernicious in its effects.

homo economicus that figures in Pettit’s genealogy of ethics), and (b) a picture according to which markets, credits, and debt enjoy some pre-institutional, moralized authority (Graeber 2011; Strauss 2016). Third, thinking of money as merely a transaction technology, as it is portrayed on the barter story, has sometimes disastrous consequences for markets, even in the contemporary world (Hockett and James 2000, 109-112).

As suggested above, the better story, at least on a family of alternative accounts, is that money is a measure of credit and debt, and these things can sometimes emerge in private contexts. In the ordinary case, though, money is produced by institutional bodies (typically, states) that enforce and dictate the juridical terms of that measure (Graeber 2011, 52, 54; Hockett and James 2020). Credit and debt don’t enjoy some pre-institutional, moral bedrock authority. They are everywhere partial products of fungible social arrangements. That’s something much harder to appreciate if one accepts the myth of barter origins for money, and much easier to appreciate if one accepts the credit origin story instead.

If all of this is right, the genealogy of money is not a counterexample to the importance of speculative accuracy. It is a case study for why it matters, one that goes through the heart of the counterfactual strategy. As Pettit characterizes that strategy, “The role of the simulation in reconstructive analysis means that with ethics as with money, it ought to give us a simulated sense of the patterns articulated in moral concepts” (2018, 27). If we get it wrong, if we misdescribe the concept, or mischaracterize the social context, or fail to note some pressure on how agents experienced the context, then the “simulated sense of the patterns articulated” run a high risk of error. Just as it did in the case of money.

(Indeed, the temptation to populate genealogies with atomistic, non-dependent, self-interested individual bargainers suggests that we may have been taken in by the money mistake twice over: once as a persistent foundational myth about *homo economicus* and its accompanying moral psychology, and a second time as a model for philosophical methodology.)

There is a further cautionary note about methodology worth sounding here. It wasn’t armchair reconstructions within economics that unraveled the false model of money. Instead, it was engagement with careful empirical work from adjacent disciplines (anthropology, history) that unearthed the relevant historical facts.³ So, speculative accuracy

³ Roughly a century after Smith’s genealogy, Kant proposed a different genealogy in his *The Metaphysics of Morals* ([1797] 1996, §6:288, 434-435) that emphasized the role of institutional and tax-based origins of money. In the early 20th century, British economist Alfred Mitchell-Inness (1913) explicitly argued for a credit/debt view of the origin of money. Neither sufficed to overturn the barter story’s grip on the imagination of economists. Tellingly, economist John Maynard Keynes’ adoption of Mitchell-Inness’s credit/debt view was itself a product of his detailed study of the Babylonians. When it comes to human practices, anthropology often laughs last. For discussion of these historical antecedents and a brief history of money see Hockett and James (2000, 21; 87-104).

matters, but it may be that the best way to get it is by attending to empirical details that are not captured by armchair reflection. If that is right, then anthropology, animal moral psychology, and historical and evolutionary accounts might matter a good deal for a naturalistically credible account of the acquisition of moral concepts. If naturalistic ontologies benefit from naturalistic methods, then the genealogist's prospects for success may turn on the degree to which the account drinks deeply from the relevant contemporary sciences.

7. Predicative drift

Suppose a counterfactual genealogist is scrupulous about the naturalistic bona fides of her account, that the picture is suitably indebted to the relevant sciences, and that her account is generally cautious to preserve speculative accuracy in her counterfactual genealogy. Suppose, too, that the application of the method yielded a detailed and cautious account of the almost ineluctable acquisition of a concept in response to naturalistically respectable pressures. Would all this suffice to show that the target concept predicates only naturalistically respectable properties, as Pettit claims for his account of the genealogy of moral concepts?

I am doubtful that the method, even when well-executed, provides adequate grounds for thinking that the target concept predicates only naturalistically respectable properties. Here, I focus on two considerations: the first concerns whether *any* story about acquisition can guarantee predication of naturalistic properties; the second concerns what I will call the possibility of *predicative drift*, that what is predicated can change over time, including in ways that drift away from naturalistic respectability over time.

First, let us suppose that we accepted the putatively naturalistic story of spatial concepts outlined in the outset. Even if a full genealogy of spatial concepts gets right all the details of our concept acquisition, and even if the forces that produced it and regulated it were all naturalistic in nature, this does not guarantee that what we predicate of those concepts will survive in the final accounting. We can see this by recalling that prior to 1904, the idea of absolute simultaneity—that events can happen at the same time in some non-relativistic sense—seemed naturalistically plausible to many physicists and laypersons. It was an entirely pedestrian feature of our spatiotemporal concepts. Yet, experimental evidence for general relativity overturned it, implicating a wide range of properties plausibly predicated of natural spatial concepts—including the nature of vacuums, the motion of falling objects, the impermeability of solidity, and so on. The case of absolute simultaneity is proof that naturalistically acquired concepts can fail to be disciplined in what they predicate. As folk physics goes, so goes a genealogy of moral concepts. In neither case are strict naturalistic genealogies able to guarantee that what is predicated of either is real.

One could protest that this is too quick. For example, you could think that our folk physics never predicated properties not predicated by or continuous with final fundamental science, but that we erroneously thought that they did. On this view, the crucial advance for scientific physics was our learning that those falsely-believe-to-be predicated properties we associated with spatial concepts were not really predicated by our folk concepts.

That's a view, but it is neither obvious nor uncontested. It is at least as natural to think that we were predicating properties that have no place in final fundamental science, and it is in light of scientific discoveries that we acquire new replacement concepts that do the work of the older concepts in newer and better ways. However, the reasonableness of disagreements about these matters is enough to show that either the methodology cannot do what it promises, or else it requires controversial commitments to avoid counterexamples. Either way, this undermines the authority with which a counterfactual genealogy can claim to provide us more than a proof of mere possibility of naturalistic predication for its target concept.

I don't see how communities—actual or counterfactual—can reliably guard against the possibility of useful fictions, errors, heuristics, or projections from sneaking into their predication. This seems especially so if what is predicated enables the successful functioning of a concept in theoretical and practical life, as useful fictions are wont to do. Perhaps the nub is just this: the case of folk physics suggests that concepts can serve a naturalistic function while ascribing what, in the end, turn out to be extra-naturalistic properties (i.e., properties not predicated by final fundamental science). It is not obvious why we should think any story of naturalistic concept acquisition, no matter how detailed and speculatively accurate, can provide a sufficiently strong warrant for thinking that the properties ascribed by the target concept are naturalistic (on this point, see also Cuneo 2020, 578). If this is true of folk spatial or physical concepts, it seems equally true of moral properties, or indeed, any properties for which we might endeavor to provide counterfactual genealogies.

Let's turn to a second problem for the methodology's aspirations to demonstrate naturalistic predication: why can't predication wander? That is, even if we can demonstrate that predication must initially be restricted to naturalistic properties, why can't a concept predicate something at one time and predicate something else at another? (If we define a concept in terms of what it predicates, then we can rewrite the challenge in this way: why can't we acquire a successor concept that predicates differently than its predecessor? A genealogy may give us an account of the predecessor, but it can't preclude the possibility that it was merely a starter concept, and that we now live with the successor to that prior concept.)

Absent some very specific and extensive commitments about how reference or predication operates, there are a range of possibilities here. These possibilities include cases

where a naturalistically acquired concept (a) ascribes only naturalistic properties (the favored case for the naturalistic genealogist); (b) ascribes only extra-naturalistic properties; (c) ascribes some combination of naturalistic and extra-naturalistic properties; (d) comes to shift from ascribing naturalistic properties to extra-naturalistic properties; (e) exhibits different reference conventions, ascribing naturalistic properties in some cases and extra-naturalistic properties in others; and so on. Under the right conditions, any of these variations might occur in a concept that arose in response to naturalistic, practical pressures.

How might referential or predicative drift happen? Pettit's discussion of what he calls the "Cheshire Cat" fallacy suggests one mechanism: people come to see something as basic because they ignore or fail to see the conceptual framework that makes the notion possible (2018, 276-277). A similar path might lead to a predicative or referential shift when people lose track of the naturalistic origins of some concept, shifting to a (potentially local, contingent) conceptual innovation that doesn't respect our concerns for naturalistic decency. The conceptual innovation might be sticky if it serves some persistent function in that milieu.

To sum up: the case of folk physics suggests that first, even in the case of naturalistically disciplined concept acquisition, concepts can predicate extra-naturalistic properties. Second, even if a genealogy gives us confidence that terms under the concept exclusively predicate naturalistic properties during the initial concept formation, this does not preclude the possibility of predicative drift.

8. Options

We are now in a better position to directly address the question asked at the outset: Can armchair reconstructions of explicitly fictional pasts teach us about the nature of things? The considerations advanced here suggest that the answer is complicated.

Recall the distinction between more and less ambitious versions of the method of counterfactual genealogy. The less ambitious form aims to show that the properties predicated by the target concept are compatible with naturalism. The more ambitious attempts to show that the concept in fact predicates only naturalistically respectable properties. The ambitious form of counterfactual genealogy endeavors to be instructive about the nature of things in a particularly bold way. However, consideration of existing counterfactual cases of genealogy—for example, the cases of money and simultaneity—show that the method cannot ensure what it promises. Even when the relevant speculative bits are accurate, overlooking some other relevant factor (e.g., a relevant piece of the causal-explanatory nexus, or some contingent change in conceptual commitments or predication) can undo the reliability of the genealogy as a guide to our current commitments. A true account of moral or spatial concepts must capture what is in fact determinative of reference

for our actual concepts. However, concept acquisition is one thing and contemporary reference is another.

A less ambitious version of the method of counterfactual genealogy might still attempt to show something about the world, but the basis on which we evaluate the success of less ambitious versions suggests that there is room for considerable variation in aims and forms of less ambitious efforts at counterfactual genealogy. Four less ambitious but nevertheless interesting possibilities are worth mentioning.

First, recall that one less ambitious version of counterfactual genealogy might focus on providing a proof that some concept is compatible with naturalism. For the naturalist, again, this is not nothing. More generally, if we could show that a given concept might predicate naturalistic properties—for example, that free will could be compatible with determinism, or that race may pick out a social kind that need not depend on biology—this will be enough to declare various sorts of local victories. In this deployment, the methodology offers more than just another proposal for philosophical analysis: it offers a just-so story that constitutes a proof of possibility, whether under naturalism or some other set of philosophical constraints.

Second, one might hold that the methodology itself does not secure anything particularly notable (like a proof of naturalistic realism in a given domain) but that it is nevertheless a useful tool for generating philosophical theories that are themselves measured in all the usual ways. On this construal, the genealogical method is incidental (and in-principle severable) from the final proposal. Still, genealogy might earn its keep as a particularly fecund source for theories, even if nothing about the genealogy itself funds its value.

A third possibility involves leaning into the counterfactual element as a tool for generating a range of not-necessarily veridical but nevertheless potentially useful insights. Where the second possibility involves theories that claim to illuminate some truth about the phenomena, this possibility holds that fictional genealogies might earn a place in our constructive toolkits in virtue of their ability to help us think more creatively and more broadly—and, ultimately, by helping us see better the possibilities that might have eluded us thus far (Landy 2012). These are valuable things for philosophy, quite apart from whether they directly secure naturalistic plausibility or produce veridical theories from the outset.

A fourth, and to my mind, especially interesting possibility maintains that the work of the genealogy is not so much to capture some antecedent metaphysical fact or the scope of the existing concept's predication of properties. Instead, the genealogy might be a method for identifying a new target concept (whether this is understood to be a revised or replaced concept), or alternately, a new set of properties to which it would be rationally or practically preferable to have as referents. In pursuing new theoretical and practical ends, the genealogist could understand her proposal as in the vein of explicitly ameliorative, or more

generally revisionist theorizing (e.g., Haslanger 2000; Vargas 2013; Plunkett 2016; Vargas 2023).

On this latter approach, the possibility of predicative drift is less a problem than an opportunity. The naturalist can seize on the possibility of conceptual change, using the genealogy as a basis for identifying a suitable candidate to regiment usage going forward. On this approach, the genealogy would highlight or make plausible that there is some concept, or some cluster of properties, that could earn its keep in the manner given by the genealogist. If that's right, then it is open to the genealogist to hold that our current construal of our (potentially extra-naturalistic) concept or any naturalistically dubious properties predicated by it are either (a) connotative content or else (b) a reason to explicitly change what we denote (Vargas 2023). In either case, the naturalistic counterfactual genealogist might secure the relevant bona fides of a positive proposal without needing to deny the possibility of predicative drift, and relatedly, without needing to insist that the concept would have ineluctably been arrived at in thus and so manner. The power of the reforming proposal would turn on its capturing genuinely worthwhile causal, functional, or explanatory elements that we find in the real world.⁴

A worry here, though, is that the genealogy is gratuitous, and that the real work is done in the identification of naturalistically appealing concept that can make sense of existing thought and talk. Still, if the genealogist can motivate some distinctive or promising reason for thinking genealogy is especially good at identifying a naturalistically appealing concept, irrespective of it being the one we currently employ, that would be enough to justify adoption of the methodology.⁵

⁴ In later work, Pettit (2020) has indicated some sympathy with a version of this approach, although an explicit openness to the revisionary element is not obviously present in *The Birth of Ethics*. If we were to read *The Birth of Ethics* in that way, different questions might loom large, including those standardly raised in response to broadly revisionist accounts, concerning for example, the nature and scope of the revision, whether the proposal has enough in the way of theoretical payoffs to justify its conflicts with folk convictions, the basis on which the revisions are undertaken, whether the proposed revision is not what was meant all along, and so on.

⁵ Thanks to Tristram McPherson, David Plunkett, Dan Speak, Clinton Tolley for helpful discussions and detailed feedback on a prior draft of this paper.

References

- Andrews, Kristin and Susana Monsó. 2021. "Animal Cognition." in *The Stanford Encyclopedia of Philosophy*, edited by Edward Zalta.
<https://plato.stanford.edu/archives/spr2021/entries/cognition-animal/>. Accessed June 15, 2021
- Boehm, Christopher. 2012. *Moral Origins: The Evolution of Virtue, Altruism, and Shame*. New York: Basic Books.
- Bratman, Michael E. 2007. *Structures of Agency: Essays*. New York: Oxford University Press.
- Cuneo, Terence. 2020. "Morality, Money, and Method: Pettit's *The Birth of Ethics*." *Analysis Reviews*. 80 (3):575-583.
- el-Showk, Sedeer. 2013. "An Example of Convergent Evolution in Whales and Bats." *Accumulating Glitches: Exploring the Grandeur of Evolution*. Accessed June 15, 2021.
<https://www.nature.com/scitable/blog/accumulating-glitches/an-example-of-convergent-evolution/>
- Frigg, Roman and Stephan Hartmann. 2020. "Models in Science." in *The Stanford Encyclopedia of Philosophy*, edited by Edward Zalta. Accessed May 12, 2021.
<https://plato.stanford.edu/archives/spr2020/entries/models-science/>
- Graeber, David. 2011. *Debt: The First 5,000 Years*. Brooklyn, N.Y.: Melville House.
- Haslanger, Sally. 2000. "Gender and Race: (What) Are They? (What) Do we Want Them to Be?" *Nous*. 34 (1):31-55.
- Henrich, Joseph. 2016. *The Secret of Our Success*. Princeton, NJ: Princeton University Press.
- Hockett, Robert and Aaron James. 2020. *Money From Nothing*. Brooklyn, NY: Melville House.
- Humphry, Caroline. 1985. "Barter and Economic Disintegration." *Man*. 20 (1):48-72.
- Kant, Immanuel. 1996. "The Metaphysics of Morals." In *Practical Philosophy*, edited by Mary J. Gregor, 353-604. Cambridge, U.K.: Cambridge University Press.
- Kropotkin, Petr Alekseevich. 1910. *Mutual Aid: A Factor of Evolution*. London: William Heinemann.
- Landy, Joshua. 2012. *How to Do Things With Fictions*. New York: Oxford University Press.
- MacIntyre, Alasdair C. 1999. *Dependent Rational Animals: Why Human Beings Need the Virtues*. Chicago, Ill.: Open Court.
- Meijer, Eva. 2019. *When Animals Speak: Toward an Interspecies Democracy*. New York: New York University Press.
- Mitchell-Innes, Alfred. 1913. "What is Money?" *Banking Law Journal*. 30:273-296.
- Monsó, Susana and Kristin Andrews. 2022. "Animal Moral Psychologies." in *The Oxford Handbook of Moral Psychology*, edited by Manuel Vargas and John Doris, 388-420
New York: Oxford University Press.
- Okin, Susan Moller. 1989. *Justice, Gender, and the Family*. New York: Basic Books.
- Pettit, Philip. 2018. *The Birth of Ethics: Reconstructing the Role and Nature of Morality*. New York: Oxford University Press.

- Pettit, Philip. 2020. "Analyzing Concepts and Allocating Referents." In *Conceptual Engineering and Conceptual Ethics*, edited by Alexis Burgess, Herman Cappelen, and David Plunkett, 333–57. New York: Oxford University Press.
- Plunkett, David. 2016. "Conceptual History, Conceptual Ethics, and the Aims of Inquiry: A Framework for Thinking about the Relevance of the History/Genealogy of Concepts to Normative Inquiry." *Ergo*, 3 (2):27-64. <https://doi.org/10.3998/ergo.12405314.0003.002>
- Sober, Elliott and David Sloan Wilson. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, Mass.: Harvard University Press.
- Strauss, Ilana E. 2016. "The Myth of the Barter Economy." *The Atlantic Monthly*. <https://www.theatlantic.com/business/archive/2016/02/barter-society-myth/471051/>. Accessed May 12, 2016.
- Tomasello, Michael. 2018. "Commentary on Philip Pettit's *The Birth of Ethics*." Pp. 333-346 in *The Birth of Ethics*, edited by Kinch Hoekstra. New York: Oxford University Press.
- Vargas, Manuel. 2013. *Building Better Beings: A Theory of Moral Responsibility*. Oxford, U.K.: Oxford University Press.
- Vargas, Manuel. 2023. "Revisionism." In *A Companion to Free Will*, edited by Joseph Campbell, Kristin M. Mickelson, and V. Alan White, 204-220. Oxford: Wiley-Blackwell.