

Philip Kitcher

Afterthoughts. Reply to Comments

Abstract: I attempt to respond to the many questions and objections raised by the commentaries. The responses are grouped by themes, rather than focusing on the essays in sequence. So there are sections on worries about my analytical history, concerns about my meta-ethical perspective, doubts about my normative stance—and complaints about my perpetration of a ‘naturalistic fallacy’.

1. Preliminaries

I am extremely grateful to the authors of these essays, especially because the misunderstandings are few, while the questions and objections are diverse and interesting. Although I suspect that my replies will not always settle doubts, I hope that they will further develop and strengthen the approach begun in *The Ethical Project (EP)*.

Like Gaul, my book is divided into three parts, and, rather than treat the individual essays seriatim, it seems better to consider related themes together. The next three sections move through the three main divisions of *EP*: first considering my ‘analytical history’, then turning to my ‘meta-ethical perspective’, and focusing after that on my ‘normative stance’. Because questions about the ability of any version of naturalism to avoid fallacies are so frequent, and because my attempts at defense are spread over different chapters, I add a final section in which objections of this common type are taken up (I would like to say: ‘decisively rebutted’).

2. Doubts about the History

The commentaries voice three different sorts of concerns about my analytical history. Some of them worry about whether it is possible to reconstruct a history of human ventures in ethics that deserves serious consideration. Others are more confident about my aspirations to do that, but think I have some important details wrong. A last criticism charges that my account is not simply wrong in detail, but that it is the wrong sort of history, that the framework I have adopted is somehow askew. I shall take up these different kinds of objection in turn.

There are some things we know about the history of human ethical life. We have evidence from early written documents of a substantial body of rules governing many different spheres of human conduct. Historical sources reveal subsequent periods during which the ethical views current in a group undergo change. We can also be certain that, tens or hundreds of thousands of years ago, our ancestors had no body of ethical lore that they could instill in their offspring, and that would govern, generation after generation, the actions of members of a social group.

My analytical history advances hypotheses about the changes that link our pre-ethical ancestors with the members of Mesopotamian societies, with their complex collections of norms. Sometimes it is possible to marshal evidence from archeology, anthropology, primatology, psychology and history to defend some hypothesis as an account of the changes that *actually* occurred. When the evidence is insufficient for that, it is still important to provide an account of how we could *possibly* have proceeded from one stage to a hypothetical subsequent stage. For the aim of the analytical history is to avoid moments of revelation and to be *compatible with* the best general accounts we have of the development of human social life (including, of course, evolutionary theories, biological and cultural). Notice that, for large parts of *EP*, the aim is not to *specify the causes of changes* but to *trace the path* our ancestors followed. I am thus often interested not in pinning down the selective advantage of a particular innovation, but only in describing a sequence of changes compatible with well-grounded ideas about the action of selection.

Chrys Mantzavinos (in the guise of his ‘student’) worries that this history is too ‘conjectural’ to teach us much about ethics, and Cailin O’Connor et al. raise a similar charge about ‘speculative genealogy’. In both instances, there’s a failure to distinguish those occasions on which I am explaining how something *could have* occurred from those instances in which I offer evidence for a hypothesis about how they *actually* occurred. If all the first part of *EP* offered were a ‘how possibly’ story, then it would be quite illegitimate to use its features to ground meta-ethical claims—and thus difficult to understand what relevance it might have. Michael Baurmann, too, offers a number of observations about the limited significance of my analytical history that depend crucially on not distinguishing those parts of the history that I take to be well-supported by empirical evidence (‘how actually explanations’) and those parts for which the evidence is insufficient to discriminate among alternative hypotheses (‘how possibly explanations’). The better version of kindred complaints focuses on features of my narrative for which I attempt to provide evidence, and to argue that the evidence is insufficient to exclude rival hypotheses.

O’Connor et al. move to this superior version when they emphasize the possibilities of cooperative behavior arising without any explicit agreement on rules (although, along the way, they make a telling mistake, in praising me for attending to “the selective advantages conferred by the ethical project” (*this issue*, 54): to repeat, I am tracing a path, not identifying selective advantages). Quite possibly our hominid ancestors achieved quite a lot in coordinating their activities without agreed-on rules: ‘nice behavior’ may have become more prevalent. But,

contrary to what many authors on the evolution of ethics tend to assume, ‘nice behavior’—that is conduct we would retrospectively appraise as conforming to our explicit ethical norms—is not the same as behavior about which people reflect, and to which they apply rules or paradigms of conduct. To put the point bluntly: the innovative and illuminating work in evolutionary game theory to which O’Connor et al. refer (work I also greatly admire) doesn’t address the phenomenon to be explained, to wit, how human beings came to have these abilities to reflect and to appraise. Five thousand years ago there were lots of explicit rules about, and it’s reasonable to think that explicit rule-governed behavior had been going on for a very large number of generations. My task is to understand the origins of *that*.

The insight behind the O’Connor et al. point is that the initial state for the emergence of normative guidance (in my specific sense) may differ from the chimp-like condition I foist on our last pre-ethical ancestors in virtue of *some* increased cooperation achieved through reaching evolutionary equilibria. Just as I allow for further evolution of altruistic behavior (for example in alloparenting within human groups), this insight can be absorbed within the history I propose. The only serious challenge would come if one thought that the further refinement of altruism or the attainment of cooperation simply removed the problem of altruism-failures. Since altruism-failures persist in human societies, even given all our non-rule-mediated coordination and all our further development of altruistic tendencies, even in the presence of rules enjoining responding positively to others, it seems very hard to see how this challenge might be sustained.

Braddock and Rosenberg offer a more extensive set of reasons for amending my analytic history. They begin by misreading my conservative suggestion of fifty thousand years as a minimal age for the ethical project. I am glad to concede to them, as I concede to O’Connor et al., that coordinated behavior of various sorts emerged before our ancestors acquired normative guidance (again in my sense of giving and following explicit rules and paradigms), and also that such behavior was important to the evolution of full language. I am also happy to suppose that the early stages of the evolution of language are intertwined with the formulation and following of rules, and that the ethical project began far earlier than fifty thousand years before the present (that was a *minimal* age, and, for my meta-ethical purposes, the longer we have been ethical creatures the better). Where Braddock and Rosenberg go astray, I think, is in underestimating the seriousness of the problem of altruism failures. They point out, quite correctly, that our hominid ancestors faced a lot of challenges and difficulties. What they overlook is that the problem of altruism failures is a constituent of all these problems. To the extent that early hominids could work together, they could meet the challenges—or at least, muddle through. Even supposing that they worked together more successfully than contemporary chimpanzees and bonobos, their ability to cooperate, and thus to become as successful as *Homo* eventually became, was limited by the tendencies of individuals to fail to support other members of the band, and the consequent propensity for groups to fall apart.

Part of the disagreement here consists in what may either be a misunderstanding or a different interpretation of primatological findings. Braddock and Rosenberg suppose that all sorts of exciting things happen “[l]ong before psychological altruism kicks in” (*this issue*, 67). Chapter 1 of *EP* argues that psychological altruism is a disposition human beings share with other extant species, and is almost certainly inherited from some common ancestor of humans and chimpanzees. But the range and scope of that ancestral psychological altruism were limited (possibly refined by the development of tendencies to cooperative child care, but nonetheless limited). Because of the limits, social stability required prolonged peace-keeping. Because of the frequent failures to respond altruistically, and because of the time invested in peace-keeping, cooperative ventures were less extensive than they would eventually become (in post-ethical *Homo*). And that reduced ability to cooperate lies behind the large budget of problems to which Braddock and Rosenberg gesture. Hence, the essential parts of the analytical history stand, although I should recognize that some of the ideas favored by Braddock and Rosenberg point to possibilities of providing different ‘how possibly’ explanations (since those are quite candidly labeled as not forced by the evidence, plurality here is welcome).

In his thoughtful and wide-ranging essay, Allen Buchanan makes a smaller point about the analytical history I propose, claiming that there is good reason to think that the societies in which the ethical project first emerged were gender-inegalitarian, and some grounds for doubting their commitments to equality among males. To say the least, the sources he cites on gender inequality have been controversial, particularly when the societies under consideration are taken as representative of the ancestral condition. As *EP* makes clear, my ideas about the egalitarianism of early societies are dependent on three different anthropological studies: a pioneering essay by Knauff, Christoph Boehm’s detailed investigations of hierarchy, and the wide-ranging research conducted on the !Kung (arguably one of the very best living models for the lives of our ancestors) (Knauff 1991; Boehm 1999; Lee 1979). I am impressed not only by Richard Lee’s thorough documentation of the elaborate strategies used by the !Kung to maintain equality, but also by ethnographic accounts of the lives of !Kung women (for example, Marjorie Shostak’s *Nisa*). These studies are reinforced by my account of the emergence of cooperative societies (see particularly §9 of *EP*), in which I argue that the habitat would be divided among different large coalitions (social groups of the order of 80 members) in which the members would need to unite to resist predations from rival bands. So I am not moved by Buchanan’s doubts.

Michael Baurmann presents a wide-ranging critique of the analytical history I offer, some of which I take to be vitiated by a mischaracterization of important features of my account, but part of which offers a valuable insight. I start with the points of disagreement, before focusing more closely on the serious challenge he poses.

Baurmann’s initial description of collective decision-making is unfortunate, at least as an account of the ancestors who began the ethical project. As *EP* is as some pains to note, the ethical pioneers could not have framed the issues as later analysts do—they could simply not have known what Baurmann supposes

me to think they knew (see *EP*, 221–2, for example). The Deweyan conception behind *EP* is of ethics as a deep response to problems posed by the ‘conditions of human life’, in which progress is occasionally made, but typically made blindly. The analytic hope is that by understanding what has been going on, we can learn how to proceed more systematically and more reliably (see my further response to Baurmann in the final section below).

Baurmann is dubious about the idea of Paleolithic human beings sitting around the campfire in the cool hour, figuring out the rules that are to govern their lives together. A significant part of my contrary belief that this is a plausible scenario for the earliest ethical ventures stems from the records of ethnographers who describe similar things in the lives of hunter-gatherer societies. Once you strip away the kinds of advanced understanding and self-awareness that Baurmann wrongly assumes to be part of my picture, the conversations I envisage are well-supported by the evidence.

Nor should we replace my characterization of the problem with the one Baurmann seems to prefer, in which rules come in to address difficulties in coordination or weaknesses of commitment. Here, I think, he is reading my concept of altruism-failure in the same narrow way favored by Braddock and Rosenberg. To repeat: the root difficulty, when coordination fails or when commitments are not followed through, is typically a lack of responsiveness to others—just what I count as an altruism-failure. Perhaps I should have explained that altruism-failure is *generic* in ways that the surrogates proposed by some of my critics are not.

I thoroughly agree with Baurmann’s thought that hierarchies emerge when social groups are ‘on the path to growth’—indeed this is part of the story of *EP*, §§19–20. I follow prominent anthropologists in thinking that inequality increases in the very late Paleolithic, as animals are domesticated and societies become larger. So, on my account, a more lengthy period in which the ethical project was conducted on terms of equality gave way to a shorter time interval (of the order of ten thousand years) in which the initial form of the project was changed. Changed, moreover, in ways that affected the ability to respond to the problem of remedying altruism failures, the problem that began it all. (As I suggest in the final section below, this can easily be viewed as a distortion of the project.)

Now to Baurmann’s important insight. At the heart of his essay is a complaint that my thoughts about the early stages of normative guidance are too simple. To caricature my views: the first ethicists perceive that certain types of behavior regularly cause trouble; that perception inhibits some individual impulses to action, and also issues in a collective decision to announce that those forms of behavior will be punished. If I read Baurmann aright, he thinks that a lot of social work has to be done before people can get to that collective decision. Was it possible for the groups I envisage to carry out that work?

That seems to me to be a very good question. I am inclined to think that the possibility of discussion rests on the two sides of the human condition to which I point. Our ancestors were able to sit down together and discuss the problem (once they had language, of course!) because they were limited *altruists*; the problem was there to be discussed because they were *limited* altruists. Even

before they developed ‘rules of recognition’ and other advanced bits of machinery, there were some contexts in which they could talk together about past difficulties in their social lives. My extensive discussion of altruism and its dimensions is intended to pave the way for drawing this conclusion.

It is worth juxtaposing Baurmann’s complaint with the suggestion of Braddock and Rosenberg to the effect that my starting point for the ethical project takes pre-ethical social life to be much closer to the lives of chimpanzee troops than it actually was. In essence both criticisms chide me for telling too simple a story about the arrival of normative guidance. In both cases, that general complaint is probably correct. I contend, however, that, while various kinds of non-verbally-mediated coordination and possibly further refinements of psychological altruism come into play before the ethical project gets started, these only make small advances on the generic problem of altruism failure. Our current inability to respond to the needs of others in cases where there are not well-established norms is surely a reminder of how narrowly limited was the altruism that pervaded pre-ethical societies. I agree with Braddock, Rosenberg, and Baurmann that a more complex account of the emergence of socially embedded normative guidance needs to be given, an account that allows for further evolution of cooperative tendencies and one that appreciates the conditions that need to be in place for the conversations around the campfire I envisage. What I have attempted to do here is sketch some reasons for thinking that the improved account would not subvert the parts of the history I need for drawing my philosophical conclusions.

I turn now to the masterpiece of *chiaroscuro* painted by my old friend Kim Sterelny (confession: although I might have guessed that Kim had an inner Marx, ready to be channeled, who would have suspected an inner Nietzsche?). To put Sterelny’s view in its simplest terms, the explicit rules whose origins I hope to trace have been, on the whole, a thoroughly bad thing; the good stuff stems from the flowering of the sympathetic emotions. I go astray in telling a *Kantian* story of the evolution of ethics; Sterelny prefers to throw in his lot with Hume and Smith.

Interestingly, the overlap in sources between Sterelny and me is very large, even though the ‘sunny’ story about morality he attributes to me, and the ‘dark’ one he relates are radically different. How can this be?

Part of the answer lies in the fact that this really is *chiaroscuro*, while *EP* uses an extensive color palette. Sterelny wants a sharp opposition, and, to this end, he ‘constructs my views’ by ignoring all sorts of qualifying discussions. Many ethicists would find his comment that my account is too Kantian quite perplexing (indeed, Buchanan insightfully asks whether I have given sufficient place to ethical reasoning—see *section 4* below). Moreover, Sterelny seems not to have taken seriously my attempt in *EP* §11 not to *take sides* in the debate between Rationalists and Sentimentalists (between Kant and Hume/Smith) but to *transcend* it. That section denied the existence of any special ‘moral point of view’, and suggested that it is a mark of well-crafted ethical traditions that they deploy a wide variety of forms of ethical motivation (our entangled psychologies are well-buffered systems). It would have been better if I had not talked so

much about rules in giving my account of the ethical project—I do mention the importance of stories that provide exemplars of conduct, and that often shape our emotions, but, for all the significance of stories and ideals, it is easier to focus on the simpler machinery of explicit precepts and rules. Perhaps my simplifications misled Sterelny; if so, they will surely mislead others.

A more important part of the answer is that I do not claim that the ethical project, in the forms it has taken, is something worthy of celebration through and through. From early in the discussion of normative guidance (*EP*, 94–5), it is made completely clear that the project can ‘warp’ socialized individuals. Many of the developments of the project are seen as mixed in character. So, for example, invoking a deity solved some problems while manufacturing many others (compare §§17; 26; 27), and attempts to satisfy endorsable desires led to the abandonment of the commitment to equality (§§19–20; 36–7). I had thought that echoes of Rousseau and Marx would have been impossible to miss in these discussions—but I should probably have been more emphatic about those connections.

With these preliminary explanations, Sterelny’s *chiaroscuro* can be compared with the picture I actually painted. On his account the sympathetic emotions, forged through cooperative foraging, are key to the flowering of human societies; normative guidance comes late, and typically gets in the way. On his version of my story, the sympathetic emotions are weak, and cooperation flourishes solely because groups are able to reason their way to norms and procedures for enforcing them. My actual narrative does not suppose that norms are introduced by clear-headed reasoners (Sterelny offers the same overinterpretation here as Michael Baurmann—a fact that suggests to me that the first paragraphs of *EP*, §14 should already have introduced the points made at 221–2). Further, it envisages a sequence of interactions between explicitly adopted norms and stories and the development of our sympathetic emotions (*EP*, 81–5; 134–7).

Sterelny’s *chiaroscuro* misses this important part of my narrative. On his account, I find no place for collective intentions. To the contrary, I am happy to include them in the evolution of ethical practice, and even devoted a section to the expansion of altruism (*EP*, §21). So I am happy to celebrate Kim and Melanie’s walk up the hill—and to view it as an instance of higher-order altruism. Perhaps that walk was prefigured in some occasions of chimpanzee joint foraging, but I am inclined to think that Kim and Melanie bring to their co-peregrinations a reflective perspective that depends on a capacity for normative guidance.

My analytical history supposes that the emotional capacities available to human beings evolved through *both* practice in cooperative behavior and normatively-guided reflections on that practice. Our ancestors started with limited capacities for cooperation. By solving a few problems of altruism failure, they began to cooperate a bit more. Increased cooperation expanded their emotional repertoire, and normative reflections on what they had achieved (and the problems still facing them) led them to introduce more rules and stories. The new rules and stories induced further cooperation and developed the sympathetic emotions further. So far, a happy account of increasing regularity in social behavior and more intense social sympathies. Yet, as *EP* is at pains to recognize,

as human societies grew (and the growth was made possible by the virtuous cycles of expansion just noted), some of the rules and stories interfered with the prosocial tendencies, creating divisions and focusing sympathy more parochially. The normative stance of chapters 8 and 9 is a direct response to the darker parts of this history, to the ways in which the virtuous cycles have been compromised and distorted.

Sterelny concludes by conceding that his picture is probably too dark. I suspect that, as with so many debates about human behavior (recall ‘genetic determinism’ versus ‘social determinism’), the sensible views concern the relative contribution of factors that extreme views identify as all-powerful. *EP* was less extreme, and more sensible, than Sterelny portrays it as being. In the end, our disagreement is probably one about the weight to be assigned to different kinds of causes in episodes in which human social life was modified. I hope my remarks here provide a basis for continuing our conversation.

Let me turn now to the more systematic concerns about the framework of my analytical history. My goal was to specify the psychological capacities and the social conditions that were required for the emergence of human ethical life, and then to trace paths that connect primitive ethical practices with the ethics of the present (identifying actual processes where I can, and, in other instances, showing how connections might have occurred). The concept of psychological altruism is central to my treatment of psychological capacities, and some of my critics have challenged my treatment of it.

The principal disagreements between Buchanan and me all stem, I believe, from a difference about the characterization of psychological altruism. He writes:

“The notion of being responsive to the desires of others here is susceptible of a stronger and weaker interpretation: on the stronger interpretation it means that each gives equal weight to the desires of all affected in a particular interaction, including his own; on the weaker interpretation a psychological altruist is someone who gives weight to the desires of all, but not necessarily equal weight.” (*this issue*, 82)

This introduces a dichotomy, where I see a multi-dimensional continuum of forms of altruism. The root notion is that of *A* being altruistic towards *B* in context *C*, and my account explicitly distinguishes the intensity of the response (how much weight *A* gives to the desires attributed to *B*), the range of altruism (which individuals elicit any altruistic response), and the scope of altruism (the contexts in which *A* responds altruistically). Even before considering the complications of paternalistic altruism and higher-order altruism, there’s a three-dimensional space in which individuals possess altruism profiles. Buchanan’s two types of altruism, the ones between which I’m supposed to be undecided, are two very particular profiles drawn from this vast space. The root notion of an altruism failure is that of some failure of responsiveness, and the task of the ethical project is to promote greater responsiveness by singling out which altruism profiles should be encouraged. That might lead to privileging one of the special profiles Buchanan considers, but there is no advance reason for supposing that.

Because the potential resources for responding to generic altruism failure are so much larger, developments of the ethical project can be enormously richer than the options to which Buchanan takes me to be limited. As we'll see in later sections, this basic disagreement about altruism lies behind our diverging views about whether my account can meet the important challenges he poses.

Christine Clavien also questions my treatment of altruism (not only in its psychological forms), but ours is more of a family quarrel. I completely agree with her that two different notions of biological altruism can be found in the evolutionary literature: a strong one that demands that the losses for the altruistic agent must endure, and a weaker conception (perhaps now less common) that allows for Trivers' reciprocal altruism. *EP* should have been clear on this point. She is also correct to emphasize the various notions of behavioral altruism that seem to be at play in experimental economics, and to recognize the difficulties of deciding whether to attribute psychological altruism in many human behavioral contexts. Precisely this difficulty led me to defend the possibility of psychological altruism by citing non-human examples where some of the complicating factors are absent.

The most radical challenge to the framework in which I develop my history comes in Ken Binmore's essay. Binmore is not only one of the most distinguished game theorists of our times, but rightly esteemed for his extensive attempts to apply game-theoretic ideas to the study of human social interactions. If I read his essay correctly, he regrets the fact that *EP* doesn't pursue a project more akin to his own (similar regrets can also be found in the complaints of Braddock and Rosenberg, and O'Connor et al. that I have slighted non-linguistic coordination of behavior). At an early stage of developing my ideas, I assumed that the right approach would be along similar lines to those pursued by Binmore and by Brian Skyrms.¹ I allude to my departure from this general program in the acknowledgments section of *EP*, but perhaps I should have explained my reasons more extensively.

Binmore distrusts my emphasis on psychological altruism (which he labels 'emotional altruism'), taking me to regard this as central to moral motivation. This last attribution is a mistake, since §§11–13 are quite explicit about the diversity of the ways in which people are morally moved. I emphasize psychological altruism because I take it to be a precondition of the type of sociality that sets the problem for ethics to solve. Because they were limited *altruists* our ancestors could live together in small groups mixed by age and sex; because they were *limited* altruists, they were unable to do so easily. Hence, the starting point for my inquiries has to be that of saying what psychological altruism means, providing grounds for thinking that human beings and some of our evolutionary relatives are capable of it, and quelling doubts that psychological altruism could not originate and be maintained under natural selection.

That is where the game theory comes in—and the only place it needs to come in. Once human beings have the capacities to start the ethical project, the

¹ I first met Ken Binmore when I was in transition. He served as commentator for a lecture I gave in London, in which I was struggling to develop some of the ideas of the early parts of *EP*. I am grateful to him for his extremely constructive suggestions on that occasion.

evolutionary dynamics are not only more complicated, but also more permissive of a wide variety of options. The Folk Theorem I need is the one I cite from Boyd and Richerson: Punishment allows the evolution of cooperation (or anything else) in sizable groups (*EP*, 87 n. 22). Since (recall!) my task is to trace a path, not to identify the selection processes that took us down that path, what's required is some assurance that the hypothetical path accords with the evolutionary constraints. The Boyd-Richerson theorem delivers just that.

But the game theory I construct is quite different from Binmore's. His agents are rational egoists who engage in repeated interactions. Part of my story about the evolution of psychological altruism depends on showing how certain types of interactions—optional games (in which players can choose partners, and even whether to play at all)—might come about in the first place. Behind the Prisoner's Dilemma (and all its sisters and its cousins and its aunts) lies a more basic *explanandum*—how did groups of primates become able to live in the same places at the same times and thus set the stage for repeated interactions? In §9 of *EP*, I suggest that our remote primate ancestors were forced into playing a complex game, the *Coalition Game*, in which large coalitions, with subcoalitions, emerged to divide the habitat. Because the Coalition Game is intractable, there's no reason to think that even the cleverest rational agents will do better at it than psychological altruists—perhaps there are even reasons for thinking they will do worse.

So I can now sum up the principal differences between Binmore and me. He wants to explain the emergence of coordinated behavior, that might eventually be captured by explicit rules, through giving analyses in standard game theory. I want to explain the emergence of psychological altruism, seen both as the precondition of a certain type of sociality and (because of its limits) as setting the problem for the ethical project; I do so, by extending the altruistic tendencies towards young and towards kin, through the demands of a mathematically intractable ur-game, the Coalition Game.² Once coalitions are in place, the payoffs in interactions are modified both by (limited) altruistic adjustments and by the coalition structure within the society (you have to keep your friends). The crucial explanatory task is to explain how one might have come from a pre-ethical state of this general form to the complex ethical life we now have. That requires tracing a path compatible with the evolutionary constraints, but *it does not require identifying the mechanisms of selection*. Hence, at this stage, evolutionary game theory drops out of the picture, in favor of an understanding of how normative guidance might have originated and might have been extended.

If what you want to explain is the history of *ethics*—as opposed to 'nice behavior' or coordinated action—then it seems to me that my framework is preferable. Part of the reason for that is that it focuses on the linguistic devices that are the expressions of reflective ethical life. Another part is that it accords better with the data available from primatology, archeology, and anthropology. My original defection from standard game-theoretic approaches was prompted

² I suppose that recognizing this game is important for understanding the different forms of societies found among the Great Apes. Sterelny's charge that it is irrelevant to the evolution of 'hominin social worlds' seems not to understand the role I assign to the Coalition Game.

by recognizing how difficult it is to square the agents of game theory with the animals of the present (and the past).

3. Objections to the Meta-Ethics

I divide the critique of my ‘Meta-Ethical Perspective’ into four main parts. First come challenges to my approach to ethical progress, focused on claims about functions for the ethical project. Next are claims that I err in thinking that my analytical history undermines more traditional approaches to truth and progress in ethics. The third type of objection complains that I have failed to provide an explanation of ethical progress. Finally, there are doubts about whether I am entitled to talk of ethical truth.

Mantzavinos (or his ‘student’) charges that the notion of function is ‘crypto-normative’. Although appeal to a concept that conceals its normative character is supposed to be ‘the main trick’ of my approach, the complaint is not elaborated at any length (indeed, ‘student’ seems then to head off in a rather different direction). Nevertheless, this type of criticism is an instance of an important family: naturalists can easily be cast as conjurers who pull normative rabbits out of descriptive hats, inviting their opponents to identify the exact way in which the rabbit was inserted in the first place. So it is worth my being explicit about why I take the objection to be mistaken.

On my account, functions are grounded in problems (more exactly, problem backgrounds: *EP*, 218–21). It is tempting to think that a problem arises for a person when that person has desires that cannot be satisfied as things currently stand (or perhaps, goals that cannot be achieved under the present state of affairs). Viewed in this way, the approach would appear to veer towards a subjective notion of progress, of the sort that chapter 5 of *EP* opposes. That worry is explicitly addressed in §39 of *EP*, where I suggest that problem backgrounds (of the type with which I am concerned) have both a subjective and an objective side. The tensions of chimp-hominid societies do interfere with the aspirations of the members of those societies (subjective side), but they would interfere with the aspirations of almost any hominid or human being placed in those conditions—in other words, there are features of the situation itself, independent of the various idiosyncratic desires of the affected people, that threaten satisfying whatever desires those people may have (objective side). The reference here to ‘almost all people’ is important, for there might be a few exceptional individuals whose desires would not conform to this regularity. Imperfect though the generalization probably is, it holds widely enough to support Dewey’s proposal (the epigraph of *EP*) that “[m]oral conceptions and processes grow out of the very conditions of human life”.

It is now possible to see why Mantzavinos thinks the notion of function is ‘crypto-normative’. The conditions of chimp-hominid society are problematic for almost all human beings—or, as we might better say, for *normal* human beings. People who were unaffected by living in such a fragile and unreliable society would be abnormal, lacking a particular type of sensitivity, in the ways that

some severely autistic children are abnormal, in the ways that the color blind or the deaf are abnormal, in the ways that sufferers from type 1 diabetes are abnormal, and so forth. There are many species of normative concepts, and, if we choose, we can think of *problem background* and *function* as belonging to the genus. Yet it would be quite wrong to think that this unmasks a ‘trick’—for, even though a rabbit may be pulled out of the hat, what has been shown to be inserted at the beginning isn’t a *rabbit* but something interestingly different. The kind of normativity that emerges is not the ‘normativity’ of the notion of function. More significantly, whether or not the concepts *problem background* and *function* count as normative, they can be applied through processes of observation and reasoning we readily understand. Nobody thinks there is any great mystery in gaining evidence for the conclusion that normal human beings have 23 pairs of chromosomes (even though some unfortunate people do not), or the conclusion that normal human beings can do things that severely autistic children cannot, or any of a host of similar conclusions.³ Because whatever normativity my key concepts have proves epistemologically unproblematic, we achieve a naturalistic account of ethical progress, one that solves one of the principal difficulties of meta-ethics. Once that point is appreciated, whether or not you apply the label ‘crypto-normative’ doesn’t matter.

The clarifications I have tried to make help address a second worry about my use of the concept of function, one present in the essays of Clavien and of Braddock and Rosenberg, but most fully articulated by Buchanan. Buchanan sees me as making a serious equivocation between two attributions of the original function of ethics.

“According to the first, avoiding social costs is the ultimate function, and remedying altruism failure is a function only so far as it contributes to that; according to the second, the ultimate function is to remedy altruism failures, regardless of whether doing so is the best way to avoid their social costs.” (*this issue*, 82)

He is fully aware of my use of a medical analogy, recognizing that I think of the social tensions as symptoms of an underlying disease (a widespread tendency to altruism failure), but, supposedly, my commitment to a particular type of altruism (one of the altruism profiles Buchanan selects—see *section 2* above) commits me to a “moralized understanding of what it is to remedy altruism failures” (*this issue*, 84).

This is a very important objection, for Buchanan is claiming to find a real *rabbit* placed in the hat at the outset. I think his argument goes astray because he modifies my approach to altruism and thus misses my concept of altruism failure. It’s worth taking this slowly.

To say that a widespread tendency to altruism failure is part of the fabric of chimp-hominid society, or, equally, of Dewey’s ‘very conditions of human life’, is

³ Here there is a point of kinship between the approach originated by Philippa Foot in her *Natural Goodness* (and developed further by a number of philosophers influenced by her). Of course, there are many substantial differences between my account and Foot’s, most of them stemming from my preference for post-Darwinian biology over Aristotle’s.

to recognize two things: first, that dispositions to psychological altruism made living together in small groups mixed by age and sex possible for our ancestors; second that the limits to those dispositions gave rise to frequent failures to respond to the aspirations and emotions of others, and consequent frustration of human desires. Experiencing those frustrations is akin to feeling the pain caused by some underlying structural condition. Because the effects of that underlying condition are so pervasive, extending across a vast range of possible projects and aspirations and arising with such frequency, the problem has to be attended to at its source. Like many parasitic infections, it has to be dealt with by dealing with the underlying cause.

This underlying cause is not identified as our lack of some specific altruism profile—say the one that Buchanan attributes to me as the remedy for altruism failure. Rather, the source of the trouble is the fact that our evolved dispositions to psychological altruism are so limited. The ethical project finds its original function in a response to that fact, *and the form of that response is not set in advance as some commitment to equal responsiveness to the desires of all*. I argue that, as a matter of empirical fact, the conditions of Paleolithic human life, made it impossible to solve the problem posed by our limited altruism without attending to the wishes and aspirations of all members of the band. So, on my account, the problem of altruism failure consists in widespread tendencies not to respond to the wishes of others, and the solution to that problem had to be worked out through negotiations among group members (through identifying some desires as endorsable and others as contaminated: *EP*, 223–5). But, because the group members need one another, because altruism failures can arise within any subset with two or more members, and because of the frequency with which limited altruism is overridden, all members have to be considered. Because of his misinterpretation of my concept of altruism, Buchanan has been led to inflate my notion of altruism failure, and thus to make an illicit conflation of things that are conceptually separate. Properly understood, my notion of altruism failure is not already moralized. Once again: no rabbit.

Several commentators believe that the analytical history is not as inhospitable to traditional accounts of ethical truth as I take it to be. Mantzavinos, for example, thinks that some version of non-cognitivism can survive the challenges posed in *EP* chapter 5. For non-cognitivism to succeed, however, would require an account of progress for the emotions taken to be expressed—an account of what is meant by saying that the emotions elicited at later stages are more ‘apt’ than those appearing at earlier times. Despite my admiration for Allan Gibbard’s work, I cannot find any such account in his writings, nor in those of any other non-cognitivist I know. So, if the challenge is to be met, some serious work will have to be done.

The defense of a more traditional picture of ethical truth is most cogently articulated by William Rottschaefer, who has done a lot of serious work in trying to find a version of moral realism compatible with the evolution of ethics. I applaud Rottschaefer’s detailed account of how psychological capacities we actually have might put us in touch with the moral facts realists want to invoke, and I am grateful for his sympathetic attempt to forge connections between his moral

realism and my pragmatic naturalism. In the end, however, I want to stick with the unvarnished pragmatic naturalism of *EP*, without realist additives. Perhaps I am missing something in the details of Rottschaefer's moral realism, but I think there is a crucial lacuna in the story.

Start with the debate between Harman and Sturgeon (Sturgeon 1985; Harman 1986). Sturgeon claimed, ambitiously, that people can sometimes see the wrongness of an action—as when onlookers respond to the torturing of a cat by judging that the action is wrong. For Harman, and for me, the judgment is socially mediated: precisely because the onlookers have been educated in the mores of their societies, they apply the predicate 'wrong' without hesitation. For Sturgeon to defend his supposed epistemological connection to moral reality, he would have to explore the *phylogeny* of the social practice—to show that the categories and applications of them that now come so fluently grew out of some prior detection of the moral facts. So I disagree with Rottschaefer about the outcome of the debate. Moral realism has a very specific task to discharge. It must show how there is a history behind our current social practices, one that is like the history behind current usages of the scientific instruments we take for granted. The argument of chapter 5 of *EP* is not only that nobody has demonstrated the pertinent sort of history, but that features of my analytical history undermine the possibility.

Rottschaefer's story of Sonia focuses on *ontogeny*, rather than phylogeny. And that is important, because the sorts of processes to which he points in the moral growth of children already presuppose the social backdrop whose status with respect to the apprehension of moral reality is at issue. No doubt, given the prevalent ideas of her community, Sonia can learn how to apply the moral concepts as others do, in just the kinds of ways Rottschaefer suggests she does. All that can happen whether or not the community's lore tracks moral reality.

Rottschaefer and I both emphasize empathic capacities, although we give them rather different roles. I agree with him that perception of particular situations might excite our feelings in ways that move us to particular types of action. I can even concur in the thought that those situations provide affordances for certain types of actions. The worry is that, *unless some ethical framework is already presupposed*, there is no basis for applying any ethical category to those actions. Moral realists underrate the difficulties of those who make innovations. Even supposing Woolman (or some other revisionary) felt empathy for the sufferings of a black slave, it is not clear how the judgment that the response is a *moral advance* is to be justified. The emotion does not come with its progressive moral character clearly labeled. On my account, ethics is the product of collective human deliberation about the problems we face in living together, and it is not hard to appreciate how an empathetic response can be the starting point

for renewed conversation.⁴ What I do not see is how moral realism can supply more justification than that.

The version of moral realism offered by Rottschaefer adds further details to the simplified set of resources for justification I have considered, but, so far as I can see, the elaboration does not close the crucial lacuna. What is required is an account of moral reality that shows how the processes taken to disclose it can put us in touch with its features. As we learn more about the physical world, we can explain how we are able to detect some of its aspects, both through unaided observation and by means of the instruments we design and build. The root trouble is that there's nothing parallel to that for the moral realist to rely on. Moral reality is simply so under-described that it is quite mysterious how (say) empathetic reactions to other people could put us in touch with it.

My response to Braddock and Rosenberg's suggestion that the situation is no worse in ethics than in other areas of knowledge is simply to deny that our predicament in other instances is so dire. Various types of knowledge may be difficult for us, but, on any moral realist account I know, it appears that moral deliberation is simply cut off from the supposed moral reality. I would draw similar conclusions about any area in which the alleged subject matter could not be apprehended by beings with the kinds of capacities people have. One response to this predicament is to inflate our cognitive resources—hence the popularity of appeals to 'intuition' in ethics and other instances. I prefer Rottschaefer's admirably honest attempt to show how the abilities we actually have deliver the ethical goods, even though I doubt that they can actually detect any external ethical reality.⁵

Turn now to the third line of opposition to my meta-ethical perspective, Buchanan's suggestion that, at the end of the day, I have failed to explain ethical progress. He is completely correct when he claims that I have not explained how progressive 'transitions come about'. With respect to some episodes I offer a partial account of what the innovators did (as, for example, in my discussions of Woolman and Wollstonecraft), but it would be sheer folly to attempt any serious explanatory narrative for the transitions that occurred in the remote past. With respect to those, I can only sketch a how-possibly story, indicating one way in which things might have gone (as I do, for example, in sections of *EP*, chapter 3).

Does this vitiate my account of progress? No. There are two different explanatory tasks. The first is to explain what ethical progress consists in. The second is to explain how human beings come (when they do) to make ethical progress. All that needs to be done with respect to the second, for my purposes at least,

⁴ As *EP* notes, Woolman, like other revolutionaries, would have known that large numbers of apparently knowledgeable and pious members of his society did not think that black suffering mattered morally. Unlike the figures envisaged by moral realists, he recognized the need to engage the complex web of arguments invoked to defend the idea that the suffering of slaves was not morally significant.

⁵ I think mathematics poses similar troubles for realism. In this instance, I try to supply a non-realist account of mathematical statements, so as to show how the processes undergone by mathematicians during the historical evolution of mathematics can actually deliver justifications. See 'Mathematical Truth?', in my forthcoming collection *Preludes to Pragmatism*.

is to rebut doubts that the concept of progress I deploy makes ethical progress impossible or incomprehensible—and that task is discharged by the discussions of the actual contributions made by Wollstonecraft and Woolman (*EP*, 249–51) and the presentations of possibilities of *EP*, §§19–21. The crucial question for my approach to ethics is the first. That is the topic of *EP*, chapter 6, and the question can be completely addressed without venturing into the territory Buchanan views me as neglecting.

Finally, Braddock and Rosenberg raise an interesting objection to my attempt to defend the thesis that some ethical statements are true (according to my pragmatist conception of ethical truth). I take an ethical statement to be true just in case it would be adopted in a progressive transition and retained under an indefinite sequence of future progressive transitions (for a slightly more exact version, see *EP*, 246). As Braddock and Rosenberg see, the grounding of the notion of progress in problems to be overcome seems to introduce an environment-relativity. If the environment changes sufficiently dramatically, will ethical statements that were ‘once true’ cease to be so? Or should we say that, because there are environments in which the problems go away, that such statements were never true? And if we do say that, will there turn out to be any ethical truths?

The best way to come to terms with this worry is to consider my central example of ethical truth, the case of honesty. I claim (*EP* §46) that a vague generalization to the effect that truth-telling is good and lying bad was introduced progressively and is likely to endure through a sequence of progressive transitions. How can I be so confident about the future, and about the environments in which human beings will find themselves? In fact, I am making some assumptions about those environments: I am taking it for granted that people will continue to live together, to exchange information with one another, and that they will continue to need information from others if they are to realize the goals that matter to them. I would regard these as part of Dewey’s ‘conditions of human life’. If our species were ever to be freed from them, the ethical project would have come to an end—for fully self-sufficient angels there is no such project. To support the attribution of truth to the vague generalization about honesty, what is required is to show that, so long as the ethical project continues, progressive transitions would retain something like the judgment that truth-telling is good and lying bad. So long as the project continues, truthful testimony from others will tend to promote the satisfaction of the recipient’s desires. Refusal to tell the truth and deliberate lying will thus be a general type of altruism failure, so that maintaining the generalization (and its associated rule) will fulfill the original function of ethics. If I am right in thinking that the original function will continue to be important in human life (*EP*, §56), the generalization (or something akin to it) will endure in progressive transitions. Hence, on my account of truth, it can be hailed as true.

Of course, one instance, even an important one, does not settle the general worry presented by Braddock and Rosenberg. Perhaps it will turn out that possible environmental variations doom the ascription of truth to some of our common ethical maxims (‘mother’s best advice’). That remains to be seen. In

the meantime, the objection is valuable in showing how deeply my approach to ethics conceives of it as a *human* project—one resulting from the fundamental conditions of human life.

4. Normative Issues

The material of the final third of the book elicits less attention from my commentators, most probably because they think enough mistakes have been made earlier to undercut the treatment of the last three chapters. Zed Adams is an exception, however. His principal concerns about my picture focus on the possibility of a naturalistic fallacy (see *section 5*), and on my claim that there are no ethical experts. My present concern will be with this latter issue.

Adams asks a good question: why, if ethics is a social *technology*, aren't there ethical experts? Part of the answer to this is to acknowledge that, when there is some function (or functions) that can be identified, when there is a recognizable defect in discharging the function(s), and when there are ways to remedy the defect without compromising other functions, the problem is one that can be addressed completely by anyone with the ability to find an appropriate means (or to discover the best means). So, for some kinds of ethical advances, there are experts, and they play just the role that experts do in other areas of technology.

What makes ethical life, and ethical discussion, hard is the existence of functional conflict. Adams recognizes this, I think, when he suggests that our most difficult ethical question is “figuring what, if anything, has *become* of the function of ethics” (*this issue*, 164). I agree with his conclusion that my genealogy is ‘incomplete’—and that is prefigured in the problem that emerges at the end of Part II of *EP*, and which occupies the first two chapters of Part III. I would put it slightly differently: a continuation of the ethical project requires sorting out how to go on when functions conflict, but I can see how that might involve coming to decide ‘what the ethical project has made of us’.

The Normative Stance of Part III proposes that individuals, however well-informed they may be, can only make proposals about the weights to be assigned to different functions. The actual assignment of weights depends on a collective decision. Even with respect to the proper making of the decision, individuals can only contribute procedural proposals. Just as there are no religious experts in ethics—for there is no source to which such people might have access—there are no philosophical experts either. A consequence of my genealogy (even if incomplete!) and of the meta-ethics I draw from it, is that there is no source of ethical truth outside the decisions of the community that pursues the ethical project. Chapters 8 and 9 of *EP* present two proposals, one about the conception of the good that might guide us at the present stage of our evolution, and one about the community of discussants and the form of their conversation. By endeavoring to show that these proposals mutually cohere, I try to rebut a certain type of skeptical challenge (for more on this, see the next section).

Part III reduces the role and the claims of philosophy, but it continues to recognize the value of philosophy in continuing the ethical project. Although

I deny a certain type of philosophical expertise, I affirm another. Because the decision for how to go on must be collective, no philosopher, however learned, wise, and virtuous, has the authority to pre-empt the collective discussion. On the other hand, people who have grappled with the systematic approaches that have been offered in the history of ethical theorizing, people who have tried to think through hard problems and to introduce and refine distinctions and concepts, people who have a broad sense of the possibilities of human living and of the extent to which those are often compromised, might have bodies of knowledge and intellectual and empathetic skills that are valuable in facilitating ethical conversation. When experts are conceived as people who have abilities that make them useful in advancing a collective project, it's even possible to agree with Adams' suggested answer to his question: in a sense, there are ethical experts (philosophers among them). It is important, however, not to confuse this sort of expertise with the stronger idea of special people who can discover—all by themselves—the Ethical Truth.

Buchanan closes his penetrating critique with a discussion of the 'cosmopolitan' approach of my normative stance. In my judgment, his label is entirely apt. What puzzles him is the way that a project originally founded in the remedying of altruism failures can be extended in the ways I suggest. Part of the explanation of his perplexity lies in his misidentification of the original function: to repeat, that function consists in the remedying of altruism failures, where these are conceived, not as the absence of some specific altruism profile, but as a general tendency not to respond to the aspirations of others. Altruism failures were originally problematic within small groups, and the failures addressed through the early phases of the ethical project were those affecting group members. We live in a world of dramatically enlarged causal dependencies, and a central *proposal* of Part III is that this demands a wider notion of responsiveness. The ideal conversation has to involve all of us (including our descendants).

Yet I think there is a deep insight in Buchanan's approach throughout his essay, one that discloses a significant omission in *EP*. The focus on collective decision and conversation gives rise to a notion of *ideal* conversation, in Part III, that fails to incorporate dimensions of *actual* discussions, aspects that surely emerged at some point in the evolution of the ethical project. Perhaps early on, our ancestors voiced their discontents and suggestions, but, as a body of maxims and paradigms of conduct accumulated, they would surely have thought about how new ideas might fit within the framework they already had, how potential solutions of new problems might be defended or rebutted. The conversations envisaged in *EP* do not explicitly portray the collective decision-making in this way: the discussants are not shown as *reasoning* with one another. In my zeal to portray ethics as a human social practice, ethical *theorizing* is held at a distance. As I read him, Buchanan identifies this as a defect of my approach in *EP* (see, in particular, his comment that I neglect the role of moral reasoning in moral progress), and here I take him to be correct. Besides the conditions chapter 9 imposes on ideal conversation, there should also be requirements on reasoning. The set of cognitive constraints (*EP*, 344–5) should thus be extended, to include conditions that specify when attempts at justification need to be undertaken,

and what features justifications should have or should lack. The method outlined in chapter 9 can be further developed along these lines, and, so far as I can see, that further development would strengthen the Normative Stance I propose.

5. The Inevitable Question

We come at last to the issue I expect to emerge, not only in discussions of my own version of naturalism, but in reactions to any attempt to work out a naturalistic program for ethics. One of my fellow naturalists, Richard Joyce, has offered an excellent treatment of the diversity of the charges that naturalists fall into fallacy, and I dedicated an entire chapter to responding to some important variants (Joyce 2006). Nor was my work over at the end of that chapter, because a particularly challenging mutant remained to be tackled only in §56. But the worry will not go away. It recurs (gently!) in Clavien's essay, and underlies some of Buchanan's worries. It is developed by Adams in a different form, and it is the central theme of the critique of O'Connor et al. As I'll try to explain, my response grows out of the thought that we can learn from history—but precisely because I am an avid fan of that thought, I do not have any high hopes that my explanation will scotch forever the Inevitable Question.

Adams adapts an interesting thought-experiment of Mill's to suggest that my conception of the good must be incomplete. Focusing on my claim that the original function of ethics is the remedying of altruism failures, he asks if it would be 'a great joy and happiness' if all altruism failures were prevented. I am inclined to think that a world in which all altruism failures were eliminated⁶ would be a splendid thing, something truly worth rejoicing over, but quite apart from that, my account is well-designed to answer the Millian question about the unending quality of human aspirations. For integral to my approach is the idea that new functions are constantly emerging in the articulation of the ethical project, and I echo Mill's profound point that the evolution of our ethical practices makes possible new forms of life for us, and shows our earlier conceptions of the good to be 'puerile'. Chapter 8 of *EP* emphasizes the open-ended character of our search for the good. Because we develop new, and refined, ideas of what it is to be human, the ethical project is never finished.

O'Connor et al. present the most sustained attempt to demonstrate that, like so many previous efforts at ethical naturalism, mine too fails to avoid fallacy. I have 'tried and failed to somehow wring an *ought* from an *is*'. The authors begin their essay by quoting a famous passage from Hume (one I also cite), as if to suggest that that particular version of the problem dooms my enterprise.

This is, I think, an important error. Hume's problem can be formulated in the following terms: specify a cogent mode of inference that will lead from purely factual premises to an action-guiding conclusion. (This slightly amended

⁶ In a world of this type, there would be no occasion on which anyone would perceive others as failing to make an appropriate response to them. For those whose desires were not taken into account in a given course of action would understand the considerations that led to that neglect, and would identify with them.

characterization of the statement to be inferred is needed if one is to avoid trivializing the problem by allowing ‘normative conclusions’ that are conditional in form or that are wide-scope negations of claims about what ought to be done; see *EP*, 254–5.) The significance of Hume’s problem lies in its challenge to the idea of *justifying* ethical claims, given a particular epistemological picture. According to that picture, we have ways of discovering factual statements about the world (and, some would add, truths of a different kind—logical, analytic, conceptual, or what you will). If naturalists are to explain the justification of ethical claims, they must do so on this basis. Hence the need for specifying the cogent mode (or modes?) of inference.

Yet, as I hope *EP* makes clear, that epistemological picture is rejected. Ethics starts in response to the fragility of hominid (or early human) social life, and, throughout the ethical project, people acquire a mix of ethical and factual statements, that they then modify throughout their lifetimes. The right analog of Hume’s question, concerns *change of belief and commitments*. The account of *EP* is thoroughly pragmatist, following the dynamic approach to epistemology common to Peirce, James, and Dewey. So a better version of the issue is how, given an existing corpus of beliefs and ethical commitments, and given a new set of factual discoveries, should people *modify* that corpus: what modes of transition are cogent? Given my proposal that ethics is social, something to be worked out together, an even better formulation would ask for the specification of how *groups of people* cogently revise their shared beliefs and commitments in light of new factual information. Those who have read chapter 9 should recognize that this is the task my discussion of method in ethics attempts to address. Hume’s classic challenge and the idea of ‘somehow wringing an *ought* from an *is*’ have been already considered and superseded.

Despite the fact that O’Connor come *closer* to recognizing the real question as their essay proceeds, they never fully succeed in shaking the foundationalist Humean epistemology from their arguments. I am supposed to be deriving ‘substantive conclusions about what our own approach to ethics should be’ from my genealogy; my normative stance is taken to be justified solely by its coherence; and so forth. The best way to address their critique is, I think, to substitute the characterizations of the last paragraph from their faulty specifications of what needs to be done, and then to explain why I think I have met the real challenge.

Start with a point related to Adams’ question about the possibility of ethical experts. When ethical functions do not conflict, it is entirely possible that factual information could teach us how to discharge a function (or set of functions) more completely than hitherto. This is utterly unmysterious, and not particularly interesting. Further, as §42 of *EP* argues, there are good reasons for taking us to be committed to continuing the ethical project. The deep worry, then, is whether there are ways of reasonably limiting disagreement about ways of going forward, so that rampant pluralism is avoided. This is the worry raised in §43, that gives rise to the exploration of a normative stance and of ethical method in chapters 8 and 9, and that is finally addressed head-on in §56.

The source of the worry is the presence of functional conflict, and I approach it by looking at a particular example of it: can one find any basis for responding to

someone who believes that the original function of ethics has been transcended, that we should now give priority to different functions? At this point coherence enters the picture. Its role is not to justify a particular elaboration of ethical practice *ab initio*. Rather, the aim is to take an existing body of factual knowledge and ethical commitments, to absorb the previous conclusion that we are embedded within the ethical project, to consider possible further developments of it, one that subordinates or rejects the original function of ethics and one that gives it priority, and to argue that the latter approach possesses a coherence the former lacks. Coherence is here playing a far more limited role, one I believe it can properly fill. Of course, if you are haunted by a foundationalist picture, if you overlook the pragmatist character of my naturalism, and if you glide over my references to Neurath's boat, you will continue to read me as making the absurd claim that considerations of coherence can wring an *ought* from an *is*.

I shall close by elaborating the pragmatism, and trying to explain more concretely what genealogy can show. As *EP* notes, the track record of progress in the ethical project is spotty. Moreover, like Dewey, I take what progress we have actually achieved to have been to a significant extent a matter of fortunate power relations—the voices of people who have hitherto been excluded became too loud to ignore. Dewey's hope was for a more fully-developed ethical *method* (which he thought of as analogous to the 'experimental method' he supposed to have emerged in the seventeenth century), that would enable future progress to occur more regularly and more reliably. Part of my normative stance consists in using my genealogy and my meta-ethics both to point out a possibly progressive conception of the good and a possible method for continuing the project. Given my emphasis on the social embedding of the project, and the importance of widely-representative discussions, these proposals cannot be fully justified, independently of their consideration and endorsement by others. Yet, I want to say, it is reasonable, given the current state of the ethical project, and given my understanding of the genealogy and its meta-ethical significance, to bring them forward *as proposals for wider discussion*.

What underwrites that—carefully circumscribed!—claim of reasonableness? I already confessed my answer: we can learn from history. Good histories often lead people, *reasonably* lead them, to see their institutions, roles, social relations, and lives differently. Genealogy can show you that you are not doing what you thought you were doing. It can provide ways of viewing some ideas and institutions as distortions of a coherent project. It can offer you analogies between historical situations and processes and current conditions. In particular, it can enable you to see the ethical project differently, and, in light of some commitments you cherish, to undertake revisions of some parts of your ethical practice. Although I concede that the modes of reasoning in play here cannot be characterized with the precision philosophers rightly prize in deductive logic, they are not entirely opaque (see, for example, Gentner 2001 on analogical reasoning).⁷

Consider two examples of transitions in beliefs and commitments. One comes from the history of the natural sciences, the decision by the community of

⁷ This book contains not only essays by Dedre Gentner and her collaborators, but also contributions from others influenced by her ideas.

chemists, over a period of two decades, to abandon phlogiston chemistry in favor of Lavoisier's 'new chemistry'. The other is my proposal that the problem of remedying altruism failures now arises for us on a global scale. In both instances, the arguments originally motivating discussions of the novel idea rest on judgments about the urgency of particular problems and on suggested analogies between one set of phenomena and others. In the chemical case, it is possible to examine closely the reasoning offered by individual protagonists and by the community as a whole, and, although I can give no formal account of it, I think it can be presented in ways that make the reasonableness of the transition apparent.⁸ I claim that the considerations advanced in chapter 8 of *EP* are analogous to those offered at early stages of the debate about the 'new chemistry'. We can learn from genealogy, as scientists sometimes learn at moments of great upheaval. Specifically, we can come to see the tensions across contemporary human life as analogous to the tensions within the small human groups in which the ethical project first began. Perhaps the principal difference between ethics and the natural sciences lies in the fact that the difficulties underlying the occasional revolutions in the sciences arise more frequently in deciding how to go on with the ethical project.

At this point, it is possible to address Baurmann's skepticism about whether we should take our ancestors seriously. Baurmann's closing suggestion that "we can make our life a bit easier" (*this issue*, 50) misses an important part of the strategy of *EP*, namely that the historical reconstruction is needed to expand our meta-ethical possibilities, liberating us from the range of options philosophers have taken to be on offer. The analytical history is supposed to be the basis not only of Part III but of Part II, as well. Beyond that, I think Baurmann falls into the same trap as O'Connor et al., in that he underestimates the power of genealogy to reveal to us what we have been up to. If my Deweyan hopes for a more systematic exploration of ethical issues are to be realized, then it will be important to identify the places in which the ethical project has been distorted in the growth of larger societies. In effect, my normative proposals rest on attempting to depict the history of our ethical practices as one in which the human situation (Dewey's 'conditions of human life') posed an important problem for our living together, in which our ancestors devised strategies for coping with a few very simple instances of that problem, in which their successes enabled their social lives and their aspirations to develop in ways they could never have conceived, and in which that process of development removed some important parts of the strategies for tackling the instances of the basic problem that emerged in the wider societies. If you see things that way, you will perhaps be led to conceive our ethical life differently, and to entertain the proposals I make in my 'Normative Stance'. Genealogy gives a gestalt on ourselves we didn't previously have, and the change of gestalt can—reasonably—be consequential.

Despite my conviction that their critique arises from a misformulation of the challenge for ethical naturalism, I am especially grateful to O'Connor et al. for their determined efforts to raise the Inevitable Question, in that I hope my remarks will enable future discussions of it to begin in the right place (as noted,

⁸ See the discussion in chapter 7 of *The Advancement of Science*.

I dare not hope that the question will disappear). And let me end as I began, with an expression of gratitude to all the essayists for taking my proposals as seriously as they have.

Bibliography

- Boehm, C. (1999), *Hierarchy in the Forest*, Cambridge/MA
- Gentner, D./K. Holyoak/B. Kokinov (2001) (eds.), *The Analogical Mind*, Cambridge/MA
- Harman, G. (1986), Moral Explanations of Natural Facts, in: *Southern Journal of Philosophy* 24, 57–68
- Joyce, R. (2006), *The Evolution of Morality*, Cambridge/MA
- Kitcher, P. (1993), *The Advancement of Science*, New York
- (2012), *Preludes to Pragmatism*, forthcoming
- Knauff, B. M. (1991), Violence and Sociality in Human Evolution, in: *Current Anthropology* 32, 391–428
- Lee, R. (1979), *The !Kung San*, Cambridge
- Sturgeon, N. (1985), Moral Explanations, in: Copp, D./D. Zimmerman (eds.), *Morality, Reason, and Truth*, Totowa