PRÉCIS OF *THE ETHICAL PROJECT*

Philip Kitcher

Columbia University

Zwei Seelen wohnen, ach! in meiner Brust,
Die eine will sich von den andern trennen:
Die eine hält in derber Liebeslust
Sich an die Welt mit klammernden Organen;
Die andre hebt gewaltsam sich vom Dust
Zu den Gefilder hoher Ahnen.

Goethe, *Faust I*

Introduction

Most people who have ever lived have subscribed to principles about how to act and doctrines about what is valuable, ethical views they regard as handed down from an authoritative source, usually from a being or beings far more powerful and insightful than themselves.   Most philosophers, by contrast, have supposed that any approach to ethics along these lines is erroneous, and have substituted the thought that insights into right action and the good life must be discovered by some special sort of investigation, perhaps by using reason alone, perhaps by mixing reason with some exercises in ordinary or extraordinary types of perception.   (These investigations are typically viewed as importantly different from standard – broadly scientific – inquiries; they are the special province of philosophers.)   In my view, both of these perspectives on ethics are fundamentally mistaken.

The Ethical Project offers a different vision.   Ethics is something human beings have been working out together for most of our history as a species.   The needs that prompt the cooperative activity of the ethical project lie deep in our human characteristics, and were focused sharply in our human past.   Over tens of thousands of years, different human societies have conducted "experiments of living", in Mill's apt phrase, trying to find ways of attending to the difficulties inherent in a form of social life to which evolution inclined our pre-human ancestors.   As Dewey says, "Moral conceptions and processes grow naturally out of the very conditions of human life" (Dewey and Tufts 1932, 343).

I am in sympathy with those recent scholars who believe that important preconditions for the ethical life evolved in our primate or mammalian forebears (De Waal 2007, Churchland 2011), but I am principally interested in understanding how the complex practice of ethical life, not only as it exists today but as it has existed from the time of the earliest historical records, emerged (or could have emerged) from the social life that those evolved preconditions made possible. I want to explore how we got from there to here. Since the evidence is often insufficient to favor a single scenario, the explanation of full ethical life cannot always be "how actually" but must sometimes settle for "how possibly". When the clues are too scanty to say how a particular type of advance was made, it is important to demonstrate how it might have happened – for otherwise suspicion will linger, whispering that one of the perspectives I want to reject is the only option, that we cannot manage without divine revelation or philosophical discovery.

The Ethical Project comes in three parts. First, I offer an analytical history, one specifying the starting point for the ethical life, as I conceive it, showing how the capacities we reasonably attribute to our hominid and human ancestors give rise to recurrent problems in living together, and exploring the ways in which their first attempts to deal with those difficulties initiated a process of cultural evolution that was able to engender the forms of ethical life visible at the time of the invention of writing (5000 years B.P.) and present today. Second, that analytical history is deployed to provide a meta-ethical perspective, one that seeks to address questions about whether we can talk of any kind of objectivity in ethics, or whether we must settle for viewing the evolutionary process out of which our ethical ideas have emerged as a sequence of "mere changes". Because this question cannot be completely answered without adopting a substantive view about how the ethical project should continue, the second part leads to a third, and to the *proposal* of a normative stance. It is important to understand, from the beginning, that this third part can only propose. No philosopher can do more – for there are no experts here, and there never have been any.[i] The ethical life belongs to human beings, living together in ever larger groups, and working out their shared lives with one another. Philosophy's task is to facilitate that working out.

Analytical History

The social lives of the apes are interestingly various, ranging from the relatively solitary patterns found in orang-utans and gibbons to the chimpanzee and bonobo groups mixed by age and sex. Paleoanthropological evidence suggests that hominids conformed to the chimp-bonobo lifestyle. I argue that living in that way requires a capacity for psychological altruism.

A psychological altruist is an animal that modifies its wishes, emotions and intentions in response to its perception of the wishes, emotions and intentions of others. Imagine two circumstances that are almost alike, differing only in the presence or absence of a conspecific. When the conspecific is present (or taken to be present) the altruist modifies its psychological attitudes to align them more closely with those it attributes to the other, doing so on the basis of its perceptions of the other and not in expectation of some future satisfaction of its erstwhile (solitary) wishes. Primatologists (Goodall 1988, De Waal 1996) have provided striking examples to support the thesis that our evolutionary relatives are capable of psychological altruism. I suggest that there are reasons to think that a tendency for altruism towards offspring could evolve to spread more broadly, first to kin, and eventually to unrelated conspecifics. I trace the latter movement to the demands of a game more fundamental to cooperation than the iterated Prisoner's Dilemma – the *Coalition Game*. For primates to engage in the kinds of interactions they do, there must be a propensity to form alliances and coalitions that ultimately structure the payoffs. Animals who were able to spread their altruistic dispositions towards non-relatives, particularly under conditions where they and those with whom they bonded were weak, were better able to survive in a harsh world. Chimp-bonobo-hominid social life was made possible by an extended capacity for psychological altruism.

Yet psychological altruism is a multi-dimensional notion, varying in intensity, in the individuals towards it is directed, and in the contexts in which it is aroused. Chimpanzee and bonobo societies would not exist in the form they do if their members lacked a capacity for psychological altruism – but they would not be as they are if that capacity were always aroused and intensely directed towards every fellow band member in every context. The *limitations* of psychological altruism are apparent to everyone who has observed a troop for a few hours. Animals fail to respond to the wishes, emotions and intentions of others, perhaps because those others belong to a different subcoalition, perhaps because some tempting reward is in the offing, so that conflict arises, and peace is only restored after prolonged activities bring mutual reassurance (grooming huddles). Psychological altruism allows chimpanzees and bonobos to live together in groups of a particular size range, but it does not allow them to do so smoothly. Peacemaking is often necessary (De Waal 1987).

There is every reason to think that was once the hominid prediament. Yet, zooming forward through the generations, it is clear that at least by 7000 years B.P. human beings had been able to assemble in far larger groups, and, by 5000 years B.P., they had written down some rules for joint living (rules that almost certainly are extensions of all sorts of principles, stories, and maxims that had guided the ancestors for generations). So there is a starting point: a tense social life, made possible by limited tendencies to psychological altruism, but requiring time-consuming activities of

peacemaking, and restricted group size.   Combining the endpoint (explicitly rule-governed conduct) with our knowledge of the groups who live most closely to the Paleolithic human condition (hunter-gatherer bands), there is an obvious explanation of what happened.   Our forebears invented the ethical project, working out together, initially (and probably for 80% of the project's history) in small groups the constraints that should govern their behavior.

Those forebears, the proto-ethicists, must have had a capacity for *normative guidance*.   In its earliest form, that capacity was surely primitive: a disposition to check one's behavior in advance when the action considered was recognizable as a type that would cause trouble.   Normative guidance started with the ability to give oneself commands – typically "Don't do that!" – in situations where the anticipated consequences were disliked, even feared.   It was refined as successful groups found ways of bringing other emotions into play: not only fear of the actions of other group members but of beings typically unobserved (ancestors, spirits, deities), awe and respect for these beings, feelings of group solidarity, pride in aligning oneself with fellows, reassurance in mutual warmth, and many more.   Without conscious insight into the possibilities of motivation, the most culturally successful groups were able to increase conformity to the agreed-on guidelines for conduct, weaving many strands together to build a conscience.

Because the problems afflicting these small ancestral societies stemmed from the limitations of psychological altruism, the earliest rules sought remedies for *altruism failures*, situations in which one group member failed to adjust wishes, emotions and intentions to the clearly perceivable attitudes of others.   Our knowledge of the difficulties that arise for bands of chimpanzees and bonobos, as well as our understanding of the lore articulated by hunter-gatherers, strongly suggests that the earliest parts of proto-ethics were very simple: injunctions to share and prohbitions of violence were the most obvious starting points.

How could the complex ethical practices we can identify in the present and in the written historical record ever have emerged from such simple beginnings?   There are several specific versions of the question: How were ethical ideas extended outside the local group? How did people come to have ideas of individual virtues, displayed not only in interactions with others but also in self-regarding conduct?   How did human beings attain the "ethical point of view"?   How did ethical maxims come to be differentiated from matters of etiquette, law and religion?   My naturalistic account needs to address all these issues and more besides.   That is, it needs to show how some evolutionary process – a *cultural* evolutionary process involving variation among ethical experiments and differential success in cultural transmission – *could have* led from the admittedly primitive starting point to the intricate proceedings of the past five

thousand or so years.   Because much of this cultural evolution took place during the Paleolithic, so that clues are scanty, it would be presumptuous to make claims about how the process *actually* went.   Nevertheless, it is possible to turn back the challenge that it required something naturalists cannot allow, something "spooky" like a Divine Revelation or an Apprehension of the Realm of Value, by providing a "how possibly" explanation.

Sometimes the explanation will dispute the alleged phenomenon.   Consider the idea, popular among philosophers, of a special ethical point of view, one that consists in the apprehension of overriding obligations.   My historical account concedes that normative guidance begins in fear, as proto-ethicists inhibit actions in which they would otherwise engage because they recognize consequences for themselves that they strongly dislike.   It also recognizes that, as the ethical project proceeds, the psychological causes of normatively guided conduct become more complicated – all sorts of emotions may be recruited to the task of creating propensities for following the constraints on behavior on which the group agrees.   As different societies became more skilled at socializing their members, they reaped the benefit of increased compliance.   Conscience-building has assembled all sorts of dispositions, and those who celebrate human ability to abide by ethical rules should be glad of that fact.   This is not because there is some supreme propensity, embodied in an ethical point of view, that was somehow discovered along the way from there to here, but simply that a host of mutually reinforcing mechanisms is better than any of them, taken in isolation.   We can, and should, abandon the myth of the ethical point of view, recognizing that, like other parts of our evolved biology, our capacity for acquiring tendencies towards normative guidance benefits from redundancy, from the presence of all sorts of back-up systems that can do the job if others fail.

Yet there are other steps in cultural evolution that were clearly taken.   The record of human deposits at resting sites shows how separated groups sometimes came together; the existence of tools at places far away from the nearest sources of the raw materials testifies to Paleolithic experiments in trade.   How, then, could ethical principles be extended to cover at least some sort of activities involving "outsiders"? One possibility starts from the primitive rules for sharing, responses to recurrent problems of altruism-failure.   The command to share already recognizes the desirability of a situation in which all band members have the resources they need. Because that situation is desirable, there is pressure to institute social arrangements that help achieve it.   Division of labor is born, and employed to store up resources for hard times.   The possibility of extending the division of labor in productive ways, by trading with the band across the river, requires the setting-up of restraints on behavior in an enlarged context: rules for avoiding some sorts of actions are extended in scope, affording temporary protections to human beings who are not members of the local

group.   The small societies that learn to exchange peacefully thrive, and the idea of expanded ethical precepts prospers under cultural selection.   *Plainly*, this cannot claim any status as *actual* history, but that is not the point: the story just sketched shows how a particular type of ethical evolution *can* occur, without supposing ethical revelation or ethical discovery.

Once the division of labor is developed, and as the range of projects pursued by the group expands, it is possible to recognize the need for serious attention to the demands of some roles.   Training the young for these roles may require the elders to spot those who are most apt for the part, and may require the apprentices to work hard to master established techniques.   Particular qualities emerge as important for the band's continuation of its successful practices – industry, obedience, and courage are to be fostered, initially for the few who undertake complex tasks, but eventually for all. As the vision of a desirable life expands, from simple satisfaction of bodily needs to a sense of fitting into patterns of joint activity and of contributing to the group's success, those virtues can come to seem central to living well.

We lack extensive knowledge of the "experiments of living" that were carried out during the late Paleolithic and early Neolithic, but, as the last paragraphs suggest, it is possible to find accounts, compatible with the fragmentary evidence available about patterns of human life during that period, that will lead from the proto-ethics of the first pioneers to the practices, conceptions, and questions discernible in the ancient world. Plato is a footnote to the history of ethical practice.

Once written records are available, however, it might seem that the dynamics of the cultural evolution of ethics would be more easily ascertained.   Perhaps we can understand how ethical changes are made – and whether they involve anything that any extant theory in meta-ethics could count as a *discovery* – by studying historical episodes?   Unfortunately, our ability to probe the psychological and social details of the most striking ethical changes that have occurred is quite limited.   Nevertheless, attention to the history of ethical practice yields two important conclusions.   To modify a famous sentence: "History, if viewed as more than a repository of anecdote and chronology, might reform the image of ethics by which we are possessed."   (Kuhn 1962 1, slightly amended).

The first point is that it is hard to resist the thought that some changes in ethical practice have been progressive.   That is not to say that the history we can reconstruct shows a constant progressive tendency – indeed, ethical progress may be very rare – but rather that the notion of progress seems to make sense in this domain.   Looking at the early version of the *lex talionis*, in which murderers are punished through the death of a relative of theirs corresponding to the victim (if a "senior" kills the child of another "senior", a child of the former is to put to death), it seems a progressive step when the

sentence falls directly on the perpetrator.   (Not that this is the last word on the subject, but it looks very much like an improvement.)   Studying the transition from a state in which women are treated as the property of fathers and husbands to one in which they can make important decisions for themselves, or contemplating the shift from blanket condemnation of sexual activity between members of the same sex to conceiving the value of relationships as determined not by the organs that are brought into contact but by the psychological attitudes expressed (the same in both homosexual and heterosexual cases), it's hard to think that this is "mere change".   Most striking, perhaps, because of the appalling suffering inflicted on slaves, is the apparent progress achieved in condemning the practice of allowing some people to own others.

Because the last example has been so carefully studied by historians, and because the debates between champions and critics of the "peculiar institution" are recorded in surviving texts, we might think that this case would permit a clear delineation of what is involved in ethical progress.   *How* exactly were the early pioneers who resisted slavery brought to their advanced insights?   Unfortunately, even the most extensive records we have leave many details indeterminate.   Nevertheless, they are full enough to cast doubt on the views about ethics that dominate religious and philosophical traditions.

Early opposition to chattel slavery in the New World is particularly prominent among the Quakers, and one of the pioneers was John Woolman, a devout "Friend" who devoted most of his adult life to the abolitionist cause.   Woolman left us a journal in which he described the occasion on which he first came to feel "uneasy" about the practice of slavery.   That journal is remarkable both for what it does and does not say. It is full of conscientious probing of the notion of a Christian duty, and of opposition of Woolman's own conception of Christian duty to the rival views of his contemporaries. *In this respect, it is, effectively a conversation with representatives of other points of view on the disputed issue of slavery.*   On the other hand, there are no graphic portrayals of the abominable treatment of slaves, nor even any individual slaves who emerge as fully-characterized individuals.

According to his journal, Woolman felt squeamish when his employer requested him to draw up a contract for the sale of a female slave.   Although he did as he was asked, he was subsequently unhappy with his own conduct.   In the weeks and months that followed, he reflected further and discussed his doubts with Friends.   Out of these reflections and discussions, as well as arguments with convinced apologists for slavery, came a determination to oppose slavery, an attitude that intensified through years of traveling, preaching and writing on the issue.

Compare Woolman's change of mind with cases of discovery in other areas. Röntgen famously discovered X-rays by observing fluorescence on a screen; Mendel

discovered genes ("factors", as he called them) by crossing and counting pea plants. In instances like these we can understand the discovery by connecting the psychological changes in the discoverer to the phenomena discovered. *Nothing analogous is available in Woolman's case.* There is simply no plausible story to show how a previously unrecognized ethical "phenomenon" registered on John Woolman, but not on the thousands of his contemporaries who daily performed actions similar to his. Instead, Woolman renewed a conversation, of the kind once central to the ethical project. He raised for his contemporaries a question about the way in which they should live together, whether a licensed feature of social life should be preserved. That conversation was refracted through a conceptual medium that had become central to the ethical project: for, with the idea of a transcendent policeman firmly in place, Woolman's tradition had absorbed the thought that particular writings testified to the will of this deity, and hence prescribed the features of proper human conduct. Conversation was distorted by that conceptual medium. Nevertheless, however distorted, it was the character of the conversation that kept Woolman going, that led him to his tireless advocacy of abolitionism, and, eventually, gave birth to a movement that would eventually prevail.

Progress was apparently made. It came, however, not from any moment of revelation (Woolman would surely have reported that, had it occurred), nor from anything analogous to scientific discovery. How then is *progress* possible? How can the view that the evolution of ethical practice is a succession of "mere changes" be resisted?

Meta-Ethical Perspective

The central meta-ethical questions that emerge from the analytical history I favor (or anything importantly similar to it) seem to turn on the possibility of truth and knowledge in the ethical domain. It is tempting to believe that opposing the "mere change" view requires making sense of these notions: that we can only understand the progressiveness of the crucial episodes if we recognize them as movements from falsehood to truth. Most of Woolman's contemporaries believed, falsely, that owning slaves is permissible – some, like Cotton Mather, even thought that acquiring slaves was obligatory – but Woolman worked his way to apprehension of the truth.

Tempting, but incorrect. If one tries to develop a notion of ethical truth *prior* to that of ethical progress, one that support the view of progress as the accumulation of truth, serious difficulties emerge. In particularly, there is no way of connecting the treatment of ethical truth with the thoughts and actions of those we take as making the

great ethical advances, the abolitionists and the early feminists, those who opposed intolerance of same-sex relations, and all the unknown people who contributed to the expansion of ethical maxims beyond the local group or to the reform of the *lex talionis*. Frame standard philosophical accounts of the nature of the good and the right as you will, opt for moral realism or constructivism or non-cognitivism of any sort that seems most promising, you will be baffled by the behavior of the historical actors who proceed in ways entirely disconnected with your preferred stories of ethical truth, ethical justification or the aptness of ethical emotions (Gibbard 1991).

To see this, consider a generic account of ethical truth that initially appears one of the most plausible candidates. Ascription of some ethical property (goodness, wrongness) to some action or state of affairs is true just in case the action or state would elicit a particular reactive attitude or would be endorsed by some particular process of reflective thought. The wrongness of the deeds of some brutal dictator consists in their arousing our revulsion or in the failure of the dictator's maxim to accord with the test of the categorical imperative (for example). On the face of it, proposals of this sort, assuming that they can be developed beyond my blunt formulations, seem capable of accounting for the ways in which ethical advances are made.

For consider John Woolman, a young man with a penchant for taking his own emotional temperature. In his description of the episode in which he acquiesced in his master's request to draw up a contract transferring a slave, Woolman reports that he felt "afflicted in his mind" (Woolman 1961 14-15). Perhaps his feelings, generated in response to the situation, can be seen as the means through which detection of wrongness proceeds. If wrongness simply is the disposition to generate "uneasiness", there will no longer be any mystery about the judgment.

Yet the difficulties of accounting for Woolman's evolving attitudes towards slavery have only been postponed. People feel all sorts of emotional responses to the events they observe. Woolman's untroubled fellows, for example, reacted to the transfer (and similar occurrences) with different feelings, and Woolman was aware of this. What convinced him that the specific reaction he experienced warranted a judgment that the transfer is wrong? The natural thought is that *some* emotional responses are the *right* ones – these emotions are the "moral sentiments", representing "the party of humanity in the breast" – rather than idiosyncratic deviations. *But the emotion is not accompanied by some reassuring voice whispering "I am a genuinely MORAL sentiment."* Quite reasonably, Woolman described his situation with modest perplexity – he felt "afflicted in his mind". Only later, after thorough discussions with others, did he become sufficiently confident to insist on the wrongness of slavery. As his journal and his other writings reveal, it was important to him to work through the religious questions raised by the defenders of slave-holding, to relate his own emotions to the

religious texts he saw as divinely inspired, to consider various ways of reading them – in short, to enter into a debate with alternative points of view represented in his society. *Conversation was crucial.*

Or consider the different ("constructivist) proposal that wrongness applies to states and events that individuals who followed a particular procedure would thereby judge as wrong. So, for example, 'wrong' picks out actions that fail the test of the categorical imperative (their maxim cannot be thought as a universal law without contradiction). Concede, for the sake of argument, that the envisaged procedure (the categorical imperative test) can be delineated clearly enough to allow us to understand its demands. Application to Woolman's decision-making remains problematic. The more obvious, but less important difficulty, is that there is no evidence Woolman (or any other ethical innovator) ever arrived at a new insight by following any such procedure. If he pursued an approximation to the favored procedure, it is one so rough as to be useless. To substitute some vague requirement of universalizability for the rigorous test of the categorical imperative would be entirely inadequate to settle the live eighteenth century debates about slavery, for these turned crucially on the legitimacy of certain exceptions to strict universality, and vague thoughts about universalizability are powerless to endorse those exceptions or to repudiate them.

More critical is a point that runs exactly parallel to the objection leveled against the view that wrongness is grounded in dispositions to produce particular emotions. Any ethical innovator, however clear-headed about the use of some constructive procedure to discriminate the wrong from the right, will know that the ambient society differs in its classification of the focal case(s). Under these circumstances, there are ample grounds for wondering if the procedure pursued is an apt one, and if it is, whether it has been carried out correctly. Even if Woolman thought he had some touchstone for making ethical distinctions, he would have been entirely unreasonable to make a judgment on that basis. Modesty, the kind of modesty he actually exhibited, would have been more appropriate.

There are, I contend, no moments of sudden insight in the history of ethics. Since I also maintain that there is no useful notion of truth without some explanation of how truth (understood in the proposed way) is apprehended, and that we who come later in the unfolding of the ethical project have no special ways of apprehending ethical truth that were unavailable to our predecessors, out of whose efforts what we take as ethical truth emerged, I conclude that appeals to ethical truths, sometimes discovered in human history, should be abandoned. If you take the historical record seriously, the thought that individual people have successively discovered ethical truths their forebears did not know has to go.

My conclusion looks disastrous.   Without ethical truth, we cannot make sense of ethical progress, and without a concept of ethical progress, we are forced to the "mere change" view.   The reintroduction of slavery in the New World and its later abolition have to be treated symmetrically.   The historical approach debunks ethics.

Not so!   The picture of progress as the accumulation of truth holds us captive. It can profitably be abandoned.   Think instead of progress as the prior notion, and take ethical truth – to the extent that we might want to talk about it – to be what stably emerges as we make progress.   "Truth", in James' pregnant phrase, "happens to an idea" (James 1987 823).   Indeed, progress is the more important notion for our ethical deliberations, since, when we look into the past or confront our own ethical ideas with those of different groups, or with rival possibilities we imagine, the important question is whether we would make ethical progress by adopting features of the practices we recognize or envisage.   Since not all the pertinent features are propositional, thinking in terms of truth is not always apt – instead, we have to ask what potential changes would be progressive.

Escaping in the way just proposed evidently requires an account of progress that is not based upon a prior notion of ethical truth.   Fortunately, one is readily available from other areas of inquiry: in technology, for instance, we understand progress, not in terms of truth, but as the introduction of devices that fulfill particular functions and as the refinement of their functioning.   Moreover an account along these lines fits readily with my analytical history (or any analytical history that preserves the gross structure – the parts that can confidently be defended – of the one I have offered).   Our propensities for psychological altruism committed us to a form of social life, and the limitations of those propensities made that form of life problematic.   *The ethical project is a social technology, one that originated against that problem background.*   The initial function of ethical practice was to solve the problems generated from recurrent altruism-failures.

The notion of function I rely on here is one I have articulated and defended elsewhere (Kitcher 2004 Chapter 7).   Functions arise against problem backgrounds, sometimes specifically recognized by cognitively sophisticated organisms, sometimes posed by Darwin's "hostile forces" for organisms that have only limited cognition (or none at all), sometimes only vaguely and partially appreciated.   The problem background that spurred our ancestors to begin the ethical project was the prevalence of altruism-failures: they would have felt the trouble in recurrent social tension and the need for peacemaking, without any ability to identify the source of their difficulties.

Remedying altruism-failures is the *original function* of the ethical practice.   But, like other forms of technology, this social technology acquires new functions from efforts to discharge the original function.   A comparison with the technology of

transportation proves useful here.   The original function is to enable people to move to previously inaccessible places, and to destinations they could already reach, with greater ease and speed.   The introduction of various forms of motorized transport – land vehicles, water craft and airplanes – fulfills that function, but the devices introduced cause further problems and thus generate new functions.   Roads, waterways, and landing strips need to be built; traffic needs to be coordinated; standards for avoiding interference with others need to be set and to be enforced.   So there grows up a host of derivative technologies involving the improvement of surface materials, the invention of devices that give clear signals, programs of training, regimes of inspection and policing, insurance, and so forth.   The evolution of transportation technology shows progress as old problems are solved or as old solutions are improved – there is progress *from* various problematic states, but there is no progress *to*, no ideal system that beckons the inventors on.

Progress is easy to understand, though not necessarily to effect, when the functions so far generated are in harmony with one another, that is when there are ways of fulfilling one or more of them (or fulfilling some functions better) without any loss with respect to other functions.   Life becomes more complex when the only apparent ways to fulfill some functions would diminish the fulfillment of others.   When that happens, we face *functional conflict*.   Transportation technology must cope with the tension between shortening journey time and promoting the safety of travel.   Similar difficulties arise in the ethical case.

The original function of ethics was the remedying of altruism failures.   To that end, the ethical pioneers introduced primitive rules against initiating violence and requiring the sharing of scarce resources.   Effectively, they endorsed the desires of band members to satisfy certain basic needs – for food, water, shelter, and safety.   As my "how possibly" story suggests, that endorsement created pressure to increase the supply of basic resources, leading to division of labor, introduction of special roles, differentiation of contributions, and a far more extensive set of aspirations.   Solutions to the problems generated along this route produce a situation in which people are no longer contented with the satisfaction of their basic needs.   As is evident from the reflective writings of the ancients, the conception of the good life embraces the prospect of contributing to the surrounding society through the most extensive development of one's talents.   Retrospectively, we can applaud this unfolding of the ethical project, recognizing a derivative function for ethics in the proliferation of human possibilities.

Yet that function is in tension with the original function of remedying altruism-failures.   The first ethical ventures were directed towards maintaining the equality of all band members, securing adequate shares of basic resources for all.   As human possibilities proliferate, however, inequalities are inevitably introduced: the emphasis

shifts from responding to the wishes of all to enabling the truly talented to flourish. *Functional conflict* arises here and complicates the account of ethical progress.

Ethical progress consists in fulfilling the functions of ethics, original and derived, more thoroughly and more efficiently. In the absence of functional conflict, enhanced fulfillment of one function does not interfere with fulfillment of others. When functions conflict, however, transitions in ethical practice might enhance some functions at cost to others. One approach to ethical progress incorporates a Pareto-style condition: progress occurs only if none of the functions is compromised. A more inclusive attitude allows for progress so long as *some* function is enhanced, no matter what losses are incurred with respect to others. I adopt neither of these polar positions. Instead, I suggest, the successive problem backgrounds provide a basis for deliberation as to how to assess the advantages of different packages of gains and losses. *This means that the meta-ethical account of pragmatic naturalism is entangled with a normative stance.*

My account of ethical progress elaborates the Deweyan idea that "moral conceptions and processes grow out of the very conditions of human life". Those conditions are constituted by our social predicament and by the limits of our capacities for psychological altruism. Altruism-failure is a deep fact about human existence, and, in its origins, the ethical project is a response to that.

In criticizing attempts to resist the "mere change" view by taking ethical progress to consist in the discovery of ethical truth, I laid great emphasis on the difficulties of fitting the would-be accounts to the history of the ethical project. How does my attempt to develop a conception of progress fare in this regard? Return to the case of John Woolman. To the extent that he had a moving insight, it lay in the possibility of regarding the selling of human beings as a form of altruism-failure. Woolman, however, modestly recognized his contemporaries' denial that failing to respond to the desires of slaves constituted altruism-failure. *In the manner of the ethical pioneers, Woolman began a conversation.* Given the unfolding of the ethical project over tens of thousands of years, and the hierarchies and doctrines introduced in that unfolding, the original conditions of conversation were distorted. Some people (slaves, for example) were excluded, and there were appeals to ethical authority that could never have been registered around the Paleolithic camp-fire. Yet, for all that, a flawed conversation, one in which different perspectives held unequal power, culminated in the endorsement of the neglect of slave desires as an altruism-failure. *Progress is achieved through the conversation.* Woolman's great achievement was to initiate it and to keep it alive.

My opposition to the conception of ethical progress as the accumulation of ethical truth, rejects the thesis that truth is the prior notion, but it allows for a derivative (Jamesian) account of truth. Prescriptive statements have descriptive counterparts:

corresponding to "Refrain from *A*!" is "*A* is wrong". Ethical truths are the descriptive counterparts of prescriptions that would be stable under progressive transitions. That is: *S* is true just in case *S* is the descriptive counterpart of a prescription *P*, introduced in a progressive transition, and satisfying the requirement that it would be maintained in an indefinite sequence of further progressive transitions. It is reasonable to suppose that there are some ethical truths. The most likely candidates are vague generalizations of the sort we all learned at our parents' knees: "It is right to tell the truth", "It is wrong to initiate violence". Like many valuable generalizations in the natural sciences, these are surely not exceptionless. Even if our successors make progress forever, they may still fail to classify all the pertinent difficult cases. Nevertheless, we have reason to suppose that the vague generalizations (or something like them, framed perhaps in different language) will endure.

To recognize some instances in which ethical statements have truth-values does not entail that the notion of truth is applicable across the board. We should recognize the possibility of ethical pluralism. For there may be statements such that neither they nor their negations meet the conditions required for ethical truth. When that occurs, rival traditions whose ethical practices have embodied the alternatives may progress indefinitely without ever converging. Behind their different trajectories will lie an irresoluble functional conflict.

I close my meta-ethical discussion with a response to the familiar worry that naturalism – of which my approach is plainly a species – must inevitably commit a fallacy. It is worth remarking at the outset that there is some difficulty in identifying precisely the fallacy that naturalism is supposed to commit (Joyce 2006). If the concern consists in disputing the possible bearing of factual premises on ethical conclusions, the core of an answer is that, when functional conflict does not arise, factual information can provide insights into how the functions of the ethical project can be better fulfilled. If the worry is that understanding ethics as an evolving project deprives it of any authority it might have, the first part of an adequate response emphasizes the importance to the ethical project of programs of socialization that induce members of a society to feel the authority of the local code. When the critic goes on to explain that the serious demand is to account for the *real*, not the *felt*, authority of ethics, to understand not why people should feel the force of the rules of their group, but how those rules are constrained by *ethical truth*, the second part of the answer will reformulate the demand in terms of the fundamental notion of progress: my naturalism honors a distinction between acquiescing in the precepts you happen to acquire and adopting those that would make progress from what you have been taught.

The charge of a fallacy can be dramatized by asking naturalists to reply to the famous troublemakers who crop up here and there in the history of ethical theory,

proposing that justice is the advantage of the stronger, or recommending sensible knavery, or urging that "free spirits" be liberated from the confinement of morality. Objectors of these sorts might be questioning a particular ethical practice, in force at a time and place, but not the ethical project itself.   If so, they are properly invited to join the conversation.   Or they might be rejecting the project *in toto*.   In that case, it would be appropriate to remind them of the history, specifically of the fact that the ethical project offered an escape from the uneasy and limited predicament of chimp-bonobo-hominid social life.   What alternative to the ethical project do the troublemakers have in mind?

The replies I have outlined to three versions of the charge of a naturalistic fallacy make a promising start, and turn back some forms of the objection.   They do not, however, settle all the important issues.   For the explanation of how facts bear on ethical conclusions depended critically on supposing that functional conflict did not arise, and similar presuppositions are involved in the attempts to bring troublemakers into line.   Inviting them to join the conversation may not put an end to the trouble – for they may harbor different ideas about which functions are to be given prominence. Lacking an account of how functional conflicts might be resolved, the meta-ethical perspective is incomplete.   Or, to put the point differently, that perspective may not only allow for pluralism, but let pluralism run rampant.   Are there as many varieties of "ethical progress" as there are ways of satisfying *some* function(s) generated in the ethical project?   Does even the original function have any enduring hold on us?

Answering these questions – and thus completing the meta-ethical perspective – requires the articulation of a normative stance.

Normative Stance

When we think about ethics as a human project, something we work out together, that evolves and is never finished, the goals of a normative inquiry must be framed differently.   Instead of thinking of a set of principles we are out to discover, which, once found, can be graven in stone, or even of a body of ethical truth to which we seek an approximation, always aware that our judgments are revisable and that precision and accuracy may forever elude us, the task at hand is to *improve* the version of the project we have inherited.   Progress is made by solving problems, and ethical inquiry should attempt to identify the places in our ethical practice where the most urgent difficulties arise and then consider what can be done to overcome them.   A normative stance consists in an appraisal of what we have, and some suggestions about how it could be progressively modified.

Fundamental to the perspective I have adopted is the denial of ethical expertise, so, to repeat, any presentation of a normative stance can only be a *proposal*. Philosophy's task is to facilitate conversation, and one way to achieve that is to offer some suggestions about how to go on from where we are, coupled with considerations about their various merits and drawbacks. The proposal I offer comes in three parts: first, a conception of the good towards which, given our current state, we might strive; second, a suggested method for ethical discussion – rules for deliberative conversation; third, a preliminary survey of the loci in contemporary practice where conflicts are most apparent, and some opening thoughts about how they might be resolved.

It may well seem that proceeding in this way is inadequate to meet the challenges that emerged from the meta-ethical perspective. Assuming that people can reach agreement on pertinent facts, there seems no great difficulty in deliberating about modifications to ethical practice when functional conflict does not arise. Yet when functions conflict, different interlocutors come to the discussion with different priorities. They are apparently free to reject a conception of the good that highlights some functions and gives less emphasis to others, and to refuse to abide by methods of ethical discussion that would favor priorities different from their own. Hence the rejection of ethical expertise proves disastrous, for all that is left is conversation under no fixed rules, and that will inevitably allow pluralism to run rampant.

Although this challenge harbors an important insight, matters are nowhere near so dire as it suggests. Once the normative stance is conceived as consisting in a package of three proposals – one about the good, one about method, and one about the current situation – it is subject to a requirement of internal coherence: *the parts have to fit together*. Part of philosophical midwifery, the proper activity in which conversation-facilitating proposals are advanced, lies in making it clear which packages of potential suggestions cohere, and which do not.

There is a connection between the view just outlined and the Rawlsian idea of reflective equilibrium (originally derived from Nelson Goodman; Goodman 1956, Rawls 1971). Rawls can be read as offering a method for going on from where we are, one that looks for the fit between particular judgments and general principles and tries to bring them into agreement. From my meta-ethical perspective, this counts as an approximation to a more reliable method, one that analyzes the functions of the ethical project and seeks principles, at all levels of generality, that fulfill them. When functions conflict, that analysis proves difficult, but we can reasonably demand that an ideal of the good that proposes to resolve a functional conflict must be ratified by a conversation according with the rules of method proposed, that the method itself conform to the ideal of the good, and that the particular diagnoses about the contemporary situation accord with both the ideal and the method.

16

In focusing on an ideal of the good as the core of the first proposal in the package, I recognize two important points.   First, the ethical project can be conceived as a series of ventures in which people seek to make the world better than it currently is, where the conception of what counts as "better" is guided by an ideal of the good. *Dynamic consequentialism*, which takes ethical principles (and more generally all the resources of ethical practice) to be grounded in the search to make things better, which allows *many* potential conceptions of the good, including some that question all the reductive assumptions required to derive standard utilitarianism from consequentialism, which permits features of the good world to depend on matters of distribution, on a broad class of psychological states, and on much else besides, is a flexible ethical stance that can represent and propose to resolve functional conflicts. The history of the ethical project can be viewed as a sequence of experiments in specifying the good (a sequence that will almost certainly never be finished).

The first part of my proposed normative stance is an egalitarian conception of the good.   Faced with the functional conflict between the original function of remedying altruism-failures and the derivative function of enhancing human possibilities, I suggest renewed emphasis on the original function, and advocate integrating the types of human flourishing made possible through the ethical project with a commitment to responding to the wishes of all.   *In effect, we should scale up the original approach to the ethical project.*   The pioneering proto-ethicists faced the problem of pervasive altruism-failures within their small bands, and, now that human beings belong to far larger communities, the prevalence of altruism-failure is even more extensive (and potentially dangerous).   Proto-ethics thought in terms of basic desires that should be satisfied for all members of the group.   We should think not only of a species-wide population, but of a far broader collection of desires and aspirations, many of them products of the evolution of the ethical project, that include wishes for lives that express enduring relationships and that make contributions to joint enterprises.   The emergent ideal of the good centers on serious equal opportunity for all of lives that attain genuine worth, where worth is partly constituted by free choice of one's own project (Mill's pursuit of one's own good in one's own way), partly by success in that project, and partly by the importance of rewarding relationships and active cooperation to any well-chosen project.

Because this ideal could not even be temporarily attained, let alone sustained for our descendants, were the human population to expand indefinitely, questions of the appropriate size of that population are fundamental. Skeptics may claim that equality of opportunity for a worthwhile life is a pipe-dream, or realizable only if coercive measures are introduced, restrictions that undercut the thesis that the worthwhile life is a freely-chosen one.   When seriously probed, this skeptical complaint proves not to be cogent: limiting the human population need impose no constraints on valuable lives.

The ideal of the good is grounded in an examination of the ethical project in its original form, and attempts to preserve the most basic features while stripping away distortions that have since ocurred.   So it rejects distinctions drawn among types of people and recommendations attending to the aspirations of only some of these types, and abandons any thought that ethical prescriptions are grounded in the will of some transcendent being, an entity to which some, but not others, have special access. Similarly, in considering ethical method, my normative stance seeks to emulate the early stages of the ethical project.   It envisages an ideal conversation as one in which all group members – that is, all human beings, including those who will come after us – are included.

How can that be?   Any actual conversation would be a huge cacaphony of discordant voices, most of them asserting that ethical practice must conform to the deliverances of some preferred traditional lore or scripture.   No such conversation can be realized: we can no longer sit down together in the "cool hour" as our ancestors did (probably for the first 80% of the history of the ethical project) and as a few small bands of hunter-gatherers continue to do.   Ethical deliberations can, however, be carried out by groups of people with dissimilar ways of living, backgrounds and traditions.   Those deliberations will go well if they accord with the *standards of ideal conversation*.   The hard task of proposing a method in ethics consists in suggesting what those standards might be.

As in the case of psychological altruism, there are cognitive and affective conditions.   Conversation is defective when it embodies mistaken ideas about the contents of the natural world – when it supposes, for example, that there is a transcendent being who prescribes and proscribes what we should do.   Similarly, it goes awry when people operate in error about the consequences of actions for others, when they misrepresent the preferences of others, and when they do not see how those preferences would be altered through discussions that made the wants of each apparent to all.   Besides these cognitive conditions, it's also important that ideal deliberators adjust their own attitudes to the aspirations of their fellows.   The first step in the process consists in filtering desires, by eliminating those that conflict with agreed-upon elements in our ethical practice. Beyond that comes *primitive mirroring*: you see me as having specific wants, and modify your own preferences to give some priority to the satisfaction of mine.   The final stage, *extended mirroring*, approaches residual conflicts among preferences through taking into account an indefinitely complex sequence of reactions on the part of others: finding that my wants are in tension with those of some third party, you consider the reactions of many other people to our wants, assessments of those reactions, assessments of the assessments, and so forth.   To engage in extended mirroring, you need a scheme of balancing, something that generates an overall evaluation from the array of iterated assessments.   If ideal deliberators cannot

18

find consensus, their last task will be to make their schemes of balancing explicit and to discuss them, specifically considering how they might be viewed as responding to the predicaments of all.

Many people who reflect on ethics believe that discussions of outstanding ethical questions are unlikely ever to reach consensus, or even to find the weaker agreement that some potential answer, while not optimal, is acceptable. Since approximations to the conditions of ideal deliberation are extremely rare, we should not be too hasty in supposing that achieving either of these outcomes is impossible or improbable. Eliminating errors that are currently identifiable – for example, the thesis that certain maxims embody the will of a transcendent being who is to be obeyed – would release many of the most fraught contemporary discussions from the predicaments in which they are so frequently stuck. Moreover, we should be encouraged by the thought that people who most closely realize the conditions of mutual engagement, members of loving families or of tightly-knit communities, for example, are often capable of working through difficult problems and finding strategies for common action. Yet any account of ethical method ought to allow for the possibility that consensus may not be reached, and even that the discussants may divide into groups, each of which finds unacceptable the favored answers of the others.

When that occurs, what can be done? Sometimes there will be much to be said, by way of comparison of schemes of balancing and expanding the capacities for sympathetic imagination. Sometimes there will be important differences with respect to matters of fact, currently unsettled. Part of my proposal for method in ethics allows for practices of trial, further "experiments of living", willingly undertaken by groups who are inclined to a particular answer of an unresolved issue and who wish to organize their social relations on its basis. Their envisaged experiment should be assessed, for not all such explorations are permissible, but even those who disagree with them may recognize its legitimacy. For groups that disagree may, nonetheless, view one another as being in a state of *conditional mutual engagement*: they see that an issue is unresolved, and they concur in thinking that, if matters were as the others think they are, those others would have met the requirements of ideal deliberation. Some differences under acknowledged ignorance should be tolerated, and the notion of conditional mutual engagement delineates the sphere of proper tolerance.

I follow Dewey in thinking that an attempt to work out a method for ethical inquiry, comparable to the methods of investigation developed for probing the natural world, is one of the central challenges for contemporary philosophy (Dewey 1984, Chapters 9-10). Yet Dewey had an extremely rosy vision of the methodological successes of early modern science. In fact, the great seventeenth century pioneers made fruitful, but imprecise, suggestions about how inquiry into nature should go, and

the far more refined views of method we now have are the product of the successful investigations their inchoate proposals inspired, together with subsequent reflection on how the successes were attained.   It would be foolish to expect, in the case of ethics, that the beginnings of a specification of method will look much better.   My proposal should be taken as a first crude effort, one that may be combined with, or compete with, others, so that, generations hence, our descendants, inspired to renew the ethical project according to different methodological approaches, may distill from whatever successes have been achieved far more precise and illuminating methodological ideas.

Let me now return to the challenge of rampant pluralism.   Earlier, I suggested that a normative stance consists in a package of proposals that can be assessed for their internal coherence.   Now that two elements of my favored stance have been sketched, I can give substance to the thought that my package is coherent and that rivals that downplay the original function of ethics are not.   Arguably the conception of the good as equal opportunity for a worthwhile life would be accepted by a discussion under conditions of mutual engagement – although it would be worth trying to approximate those conditions, and verifying that this was so.   By contrast, any proposal for ignoring large classes of altruism failures, and concentrating, say, on the enhanced possibilities for the especially privileged, must be supplemented with some further methodological suggestion that will allow something recognizable as the ethical project to be continued – for, as the discussion of the famous philosophical troublemakers emphasized, there is no evident human alternative.   I claim that it is very hard to find any such methodological suggestion, or even to introduce rules for delineating the distinction between the privileged and those whose wishes are to be ignored.

The root of the difficulty in evading – or "transcending" – the original function of ethics lies in the features of the human condition that prompted the ethical project in the first place.   Our capacity for psychological altruism brings us together, but its limitations require attention to the pervasive altruism-failures among people who causally interact with one another, unless we are to remain stuck in, or revert to, the state of the chimpanzees, the bonobos, and our hominid forebears.   Because the evolution of the ethical project has opened up possibilities the ethical pioneers could never have imagined, the richer domain of desires that frame our conceptions of a worthwhile life needs to be integrated with the focus on altruism-failure.   We cannot, however, be so beguiled by the idea of developing those possibilities and desires that we neglect the original problem, and fail to discharge the original function on a much-enlarged – global – scale.   Human lives are framed by the fact that we are incomplete psychological altruists: two souls live in each human breast …

My account of the ethical project concludes with some preliminary discussions of how the normative ideas just outlined bear on our contemporary situation.   Most

evidently, they support maxims directing the fortunate to do what they can to bring about a situation in which the fundamental goods – the material conditions for a worthwhile life – are enjoyed by all people.   Beyond that, they can be deployed on a small scale, to reflect on the habits that guide us daily, and on a larger scale to scrutinize the roles and institutions we accept, often without thinking.   Here, my normative stance echoes themes found not only in Dewey, but in social revisionaries like Marx and Foucault.

For anyone who thinks, as I do, that ethical inquiry is a matter for joint deliberation, that there are no experts, any substantive proposals, even about the agenda for discussion, must be tentative.   Given the factual constraints on method, any individual question requires the assembly of the best pertinent information.   So, in my final chapter, I only point in some directions ethical conversation might go.   I hope, however, to join those thinkers who have returned philosophy to the burning issues of our time, who have not been content to flourish their "intuitions" about artificial cases (often cases so idealized that our everyday skills in ethical judgment are inapplicable, so that the verdicts rendered are queasy and unconfident).   Philosophy should, I believe, return to its proper role as midwife to the broad discussions central to ethical inquiry and practice, and thereby contribute to a project that has made us human.

REFERENCES

Churchland, Patricia   2011   *braintrust* (Princeton NJ: Princeton University Press)

De Waal, Frans   1987   *Peacemaking among Primates* (Cambridge MA: Harvard University Press)

De Waal, Frans   1996   *Good Natured* (Cambridge MA: Harvard University Press)

De Waal, Frans   2007   *Primates and Philosophers* (Princeton NJ: Princeton University Press)

Dewey, John   1984 *The Quest for Certainty* (Later Works of John Dewey, vol 4; Carbondale IL: University of Southern Illinois Press)

Dewey, John and Tufts, James   1932   *Ethics* (second edition; New York: Holt)

Gibbard, Allan   1991   *Wise Choices, Apt Feelings* (Cambridge MA: Harvard University Press)

Goodall, Jane   1988   *The Chimpanzees of Gombe* (Cambridge MA: Harvard University Press)

Goodman, Nelson   1956   *Fact, Fiction, and Forecast* (Indianapolis: Bobbs-Merrill)

James, William   1987   *Writings 1902-1910* (New York: Library of America)

Joyce, Richard   2006   *The Evolution of Morality* (Cambridge MA: MIT Press)

Kitcher, Philip   2004   *In Mendel's Mirror* (New York: Oxford University Press)

Kuhn, Thomas S.   1962   *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press)

Rawls, John   1971   *A Theory of Justice* (Cambridge MA: Harvard University Press)

Woolman, John   1961   *Journal* (New York: Citadel)

---

[i]   That does not mean that the great figures in the history of philosophy have contributed nothing to ethics, but simply that their ideas cannot be understood as authoritative *discoveries*.   Many of them are valuable as *proposals* that might be kept in mind as people continue to work out the ethical project together.