Analyse & Kritik 01/2012 (© Lucius & Lucius, Stuttgart) S. 63–80

Matthew Braddock/Alexander Rosenberg

Reconstruction in Moral Philosophy?

Abstract: We raise three issues for Kitcher's *Ethical Project*: First, we argue that the genealogy of morals starts well before the advent of altruism-failures and the need to remedy them, which Kitcher dates at about 50K years ago. Second, we challenge the likelihood of long term moral progress of the sort Kitcher requires to establish objectivity while circumventing Hume's challenge to avoid trying to derive normative conclusions from positive ones—'ought' from 'is'. Third, we sketch ways in which Kitcher's metaethical opponents could respond to his arguments against them.

1. Introduction

Philip Kitcher has sought to do, at least for moral philosophy, what John Dewey aimed to do for philosophy as a whole in *Reconstruction in Philosophy*: identify the errors in its established tradition, and offer a naturalistic (and a pragmatist) alternative to the tradition, one which preserves as much of it as is worth saving. What Kitcher hopes to save, as so many other latter-day naturalists have tried to save, is the objectivity of ethics. Unlike most of the other philosophers and many non-philosophers with the same ambition, Kitcher does full honor to Hume's challenge. This acknowledgement, together with Kitcher's other achievements in reconstructing the philosophy of science, the philosophy of mathematics, and the relation of both to our political institutions, make his *The Ethical Project* the best current test case of naturalism's ability to ground the objectivity of ethics without violating Hume's strictures on inferences from the descriptive to the normative.

In this paper we advance three considerations that Kitcher's reconstruction may need to address: First, we offer an alternative account of the origin of ethical norms. Second, we raise questions about whether Kitcher has provided an adequate account of moral objectivity as global progressiveness. Third, we sketch ways in which proponents of more traditional ethical justifications could respond to Kitcher's arguments against them.

2. An Alternative Analytical History

Kitcher offers an "analytical history" that includes "hypotheses about how the ethical project actually began" (*EP*, 6). Kitcher's genealogy of morals begins with psychological altruism, a disposition not just to align one's behavior with the interests of others, but to align one's desires with the desires of others. Ethics emerged presumably through a process of Darwinian cultural selection for solutions to a 'design problem': psychological altruism is by itself too weak or not extensive enough to secure outcomes good enough to ensure human survival in the Pleistocene. Before there was psychological altruism, there could have been no ethics. Ethics in turn requires a capacity for normative guidance. Indeed, *The Ethical Project* reads in a way that suggests normative guidance is what ethics consists in.[1]

Here we offer an alternative 'analytical history' of ethics. One in which ethics emerges not as a remedy for altruism failures, but as a body of solutions to other equally fundamental 'design problems', ones that had to be solved long before the date Kitcher identifies for the appearance of ethics. Our hypothesis casts doubt on Kitcher's history of how ethics actually began. It suggests that ethics has multiple sources, giving rise to it before and independent of its role in mitigating the failures of psychological altruism 50,000 years ago.

Kitcher's analytical history makes ethics a very late arrival in the emergence of human culture. How late is his dating? "My estimates here are speculative. I suppose that the ethical project began with the acquisition of full language, at least fifty thousand years ago, and that human societies were small until, at least, fifteen thousand years ago." (*EP*, 97, fn. 37) Moreover, when it emerged the process was remarkably civilized: "The first ventures in ethical practice probably involved group discussions, on terms of rough equality, directed towards issues of sharing and intragroup aggression."

So, the emergence of ethics turns out to be part of the evolutionary anthropologist's "problem of behavioral modernity" (Mellars 2006). The problem is easy to state: anatomically modern humans date from no later than 200,000 years ago, perhaps even substantially earlier. The *Homo sapiens* line of descent separated from the Neanderthal (and the Desenovian, not to mention *Homo florences*) at least 400,000 years ago. The problem of behavioral modernity is 'why did it take so long'? What kept anatomically modern humans from hitting upon culture for at least 150,000 years? Of course, as evolutionary biologists know well, absence of evidence is not the same as evidence of absence. This problem is particularly acute in evolutionary anthropology since culture does not fossilize well. So one solution to the problem is to assert that culture was present and developing right from the start but didn't begin to leave traces until 50,000 years ago. Perhaps we simply have not yet located the artifacts which establish the simultaneous onset of anatomical and behavioral modernity. Another alternative is that the *capacity* for culture emerged early but was not called upon until a combination of changed ecologies and population size began to select for its

[1] See chapter two, especially section 11, and p. 80.

expression. Kitcher's theory about when ethics emerges allies him with those who hold that the problem of behavioral modernity is indeed real.

The trouble with Kitcher's view about the recency of ethics, if not about the rest of the armamentaria of culture, is that *Homo sapiens* could not have gotten to the point of inventing culture 50,000 years ago unless, long before, they had already acquired and put into operation the capacity for normative guidance that on Kitcher's view is sufficient for the ethical project to commence. In a capacity for normative guidance *Homo sapiens* managed to find the solution to several prior design problems long before psychological altruism failures begin to be a design problem. The evolutionary predicaments solved by humans long before culture kicked in already required what Kitcher holds to be sufficient for the ethical project to begin. That makes problematical Kitcher's unicausal account of the late origin of ethics in failures of psychological altruism.

There are at least two achievements that must have been accomplished well before the attainment of behavioral modernity 50,000 years ago, which both required well-developed capacities for normative guidance of a plainly ethical kind. The first is language and the second involves complex tool making and using.

Start with language. As quoted above, Kitcher locates ethics as emerging with the acquisition of 'full language'. But the sequence of developments from gestures and grunts to pidgins already requires a significant nesting of desires and beliefs about them, and even more evidently a set of conventions that provide normative guidance. Grice (1957) famously limned the number of iterated and nested desires and beliefs that contribute to speaker's meaning. What he discovered was the structure of the shared intentionality required for language. The demands of psychological altruism are modest in comparison with the amount of shared intentionality required for 'full language'. They are still modest in comparison to the demands of the much earlier rather basic pidgins, ones that had to emerge either at the onset of *Homo sapiens* life on the African savanna, or relatively soon after it and long before the last 50,000 years. But the degree of coordination and cooperation that makes language possible at all, still less solve the major design problems facing humans for the last 2 million years or so, required a capacity for normative guidance. It had to have appeared long before altruism failures emerged to threaten further human evolution.

There was a really pressing problem that faced our ancestors for a long time after their arrival on the African savanna was not psychological altruism failure. At the bottom of the food chain, *Homo* was burdened by three distinct traits that made their survival highly problematical: First, compared to other primates, individually they were living long lives and procreating throughout their lives; second, the birth interval between off-spring was so short that there were always far too many mouths to feed; third, owing to post-natal neural development, off-spring were almost completely dependent for many years. Converting these traits from a survival threat to a means of eventual dominance required using the long dependence to train increasing numbers of highly teachable children to cooperate and collaborate in the manufacture and use of new technologies. Almost every

step—simultaneous or successive—in the construction of this solution to the haupt-design problem facing *Homo* requires normative guidance.

Hunting mega fauna, and even before that, protecting scavenged carcasses from mega fauna, puts demands on normative guidance. It puts selective pressure on the emergence of signaling conventions in strategic interactions (Skyrms 2010). These signaling conventions are where pidgins and eventually creoles and full language come from. But from their very start they require individuals to have fairly rich theories of other minds, and the ability to convert adventitious local regularities into norms that guide their illocutionary and soon enough their perlocutionary acts. Amongst the most important of these is the norm of truth-telling, which it is frequently tempting to speakers to violate. The temptation combined with a very small number of speakers with a very restricted vocabulary of incipient symbols, must have repeatedly aborted the start of language before it finally took hold among our ancestors.

Language's speaker-meaning prerequired a further development of the theory of mind we shared with the other primates. But it also required a commitment to informational collaboration and coordination that primates conspicuously lack (Warnecken/Tomasello 2009). This capacity for normatively guided collaboration and coordination that language requires cannot initially emerge between mother and child, which Kitcher suggests is the source of psychological altruism. It must also be much older than 50,000 years before the present.

It took more than 2,000,000 years for *Homo* to move from somewhere near the bottom of the savanna's carnivorous food chain to the very top. The bottom was scavenging corpses felled by predators initially higher up in the food chain than our ancestors. As noted above, the first problem scavenging raises is how to chase away the predators that brought down the prey. The second is how to extract nutrients from the felled prey. Inability to solve the first problem exacerbates the second, since little will be left to scavenge if the first problem can't be solved. It can't be solved when the predator is a carnivorous feline, for example, and groups of scavengers are small in number, tool-less and disorganized. At this point in the early Stone Age, the only thing *Homo* had going for itself was that those individuals who scavenged together were probably closely related kin whose foolhardy willingness to scare megafauna away from carcasses was kin-selected. That was not enough.

Small groups of Lower Paleolithic *Homo* could only have survived if they solved the second problem: securing nutrients left by the predator. This is where serendipity or genius introduced a good idea: the hand axe—basically a rock of a certain size and shape—to break marrow bones and skulls. Initially populations were low, bands were small, and extinction threats great and prevalent. But it is safe to assume that lines of *Homo sapiens* arrived in the savannah with the ability to make and use rudimentary hand axe, since it is something other primates can do. They had only to hit upon the idea of using them to break into brain cases and marrow pockets the mega fauna could not reach. If the idea of a hand axe to break such bones had been hit on only once, it would probably have died out with its creator's lineage, owing to the high extinction rates of individual isolated lineages on the savanna. Low populations under

continual extinction threats cannot solve the knowledge preservation problem for long periods. Accordingly, the lower the population, the more times the practice of scavenging with the hand axe had to independently emerge. Once it became sufficiently common among a larger population, hand axe use's chances of becoming extinct, along with any one particular lineage that possessed it, were reduced. If the hand axe solved the problem of scavenging well enough, then it led to a population expansion in the lineages that possessed it which increased its likelihood of survival, and it led to an increase in that populations source of brain-supporting protein.

Paleoarcheology suggests that it took over a million years for *Homo* to move from the hand axe to the compound (two part) handle hand axe and longer to get to any other complex tools. Could the problem have been a matter of 'rocket science'? Probably not, but what was lacking for a million years or so were the institutions—sets of norms required to prevent the repeated extinction of good ideas.

Consider the succession of innovations in hand axe technology: material-choice, core shape choice, flaking, bi-facial flaking, heat treatment, fixing axe-stones to handles. No Individual can risk specializing heavily in discovery, preservation, improvement, transmission (teaching) of good new ideas in hand axe technology when the opportunity costs in food, shelter, warmth and predator avoidance are high. Innovators and inventors need a 'market' in which to trade these improvements for life's immediate necessities. At the outset, when good ideas are easy to come by without much time or trouble, and when every one is 'in the same small firm'—the nuclear family, this is less of a problem: if every one is a close kinsman, all are knowledge-providers/bearers/teachers, and exclusive knowledge is hard to protect, there is no inclusive fitness reduction in the spread of good new ideas. But the low population size of the innovator's family makes preservation, accumulation, augmentation of the suite of hand axe ideas—both its manufacture and its use components—precarious.

Long before psychological altruism kicks in, what was needed to make possible language, the division of labor and the emergence of technology as well, was normatively guided cooperation. Where did it come from? One scenario is biologically attractive.

As Hrdy (2009) has pointed out there are a small number of relatively unrelated mammalian species that engage in collaborative childcare: tamarinds, dogs, dolphins, and humans, but not other primates. Collaborative childcare is evidently an available variation. Once hit upon by one or more small groups, it is likely to persist, for it has great evolutionary advantages.

Collaborative child-care will solve the triple problem of *too* many *too* dependent offspring produced *too* quickly. In fact it will gradually convert it into an adaptation that moves *Homo sapiens* from the bottom of the savanna food chain to the top in a matter of a million years or so, long before psychological altruism becomes possible or its failures become serious. Cooperative child care frees mothers to forage, it leads to divisions of domestic labor with consequent productivity enhancements, it provides adults with the time, the opportunity and eventually the skills to teach all the group's long-dependent off-spring the

complex skills there was previously no time to acquire or transmit. And co-operative childcare kick-starts a co-evolutionary cycle with the incipient theory of mind we shared with the other primates that develops both into much more powerful adaptations. But collaborative childcare already requires a capacity for normative guidance. Once such a capacity can be relied upon, language and other complex human institutions become possible.

Early in human prehistory, technological innovation along with improvements in collaborative childcare began demanding the division of labor. But as Adam Smith pointed out in *The Wealth of Nation*, in small populations, opportunities for even low levels of any kind of labor-division are rare. Specializing in any one consumption good requires that others produce and exchange all the other consumption goods needed for survival. Specializing in an intermediate production good raises the further problems of, first, assuring that others provide the other components it 'dove tails' with to make a consumption good, and second, that someone combine the specialized inputs, trade the finished consumption good and pay all intermediate good producers for their work. The coordination problems that need to be solved by the division of labor in chattel property or services are evidently considerable.

Simple versions of these problems, and similar ones, will have to be addressed almost from the start of human evolution to ensure the domestic division of labor between males and females. Even the most imperfect approximate suboptimal solutions will require that individuals engage in normatively guided behavior. In small populations and within nuclear families there may not be much of an enforcement problem. But norms and their enforcement are still required. On the other hand, while populations remain small there are fewer opportunities for productivity enhancing divisions of labor and more chances that any such innovations will become extinct along the lineages that hit on them.

Technological innovation introduces a further design problem for *Homo sapiens*, one much more severe when the division of labor is between those who produce chattel goods and those who produce, improve, or teach good ideas. The problem is well understood in economics.

The nature of good ideas as abstract objects is a major barrier to their optimal provision under any circumstances, including a market economy. To begin tinkering to come up with a good idea, doing R and D, can have heavy, indeed sometimes fatal, opportunity costs. Secondly, good ideas are close to public goods: like full-fledged public goods, good ideas can be consumed 'non-rivalrously': if I adopt crop rotation in my fields, that does not reduce your use of this idea in your fields. Full-fledged public goods must also be 'non-excludable'. I can't consume nighttime street lighting without my neighbors getting the advantage too, even if I am the only one paying for it. Unlike public goods, the consumption of good ideas is excludable: I can take care to employ a good idea only when no one who could copy it is watching, or I could build a wall around the place I use it. But these steps are often costly or so impractical. (Crop rotation is a huge competitive advantage, but how can I hide it from my neighbors?) Owing to these two features—non-rivalrousness and quasi-non-excludability—a free market will not provide an optimal supply of good ideas.

No one will invest in 'R and D'. Every one will be on the look out for good ideas they can copy. Anyone who hits upon a good idea by accident will use it privately and at greater cost or less advantage than otherwise. The solution to this problem in an advanced market economy is a new institution, a set of norms—the patent system or something like it.

It may now be more evident why so simple an advance as the compound axe did not become a permanent feature of human life for a million years or so after we moved on to the African savanna. The combination of small vulnerable populations and the difficulty of establishing norms of exchange of any kind, let alone trade in abstract objects, made the emergence, improvement and persistence of technological innovations a very rare and difficult accomplishment.[2]

Improvements in the hand axe like bi-facial flaking may have been hit upon by accident. The thrusting spear requires a minimum amount of point working, and when fitted with a stone-point requires technology as complex as the compound axe. What norms must have been honored by members of a kin-group in which the idea of the thrusting spear and its use was preserved long enough to come down to us? Whatever norms they are, they may not have been sufficient to preserve the thrusting spear among small populations of close kin honoring unenforced norms of exchange. Most probably 'luck'—drift—was involved in the survival of complex tools as well, since the idea occurs multiply and apparently independently in the archeological record.

One norm we might expect to be honored by such groups involves horizontal and oblique transmission of the good idea—those who transmit it must do so directly or indirectly to all the group's children, not just their own, if the idea is not to be lost. They must have opportunities to innovate, experiment, and take some design-risks in doing so. Teaching and experimenting can't reduce their individual fitness, even when it impinges on their available scavenging, and/or hunting time. Accordingly others will have to observe norms of reciprocity—sharing resources or otherwise compensating teachers and researchers with chattel goods for non-rivalrously consumed and difficult to exclude abstract objects and for services in delivery of abstract objects.

The thrusting spear is a good idea that opens a wide range of protein resources to a lineage hitherto limited to scavenging corpses in which predators are

---

[2] Notice that trade in chattel property is itself a good idea, as is the very notion of the division of labor, and the division of labor between production of good ideas and chattels. Insofar as the practice of exchange of chattel property, the division of labor or the division of labor between chattels and good ideas was itself an informational innovation, and not just the emergence of a practice that had serendipitous benefits, they themselves presupposed a prior solution to an economics of information problem—inventing, introducing, preserving, and figuring out how to enforce the norms that constitute these institutions in the absence of prior institutions of trade of any kind. The solution to the problem of providing good ideas is some sort of enforced rules of exchange, but this is itself a good idea whose provision in the absence of norms ensuring compensation for it will be suboptimal. The same goes for the ideas of the division of labor and the division of cognitive and physical labor, etc. When one considers how formidable were the obstacles to the establishment of patents rights, low long and how many nations did without them, and how poorly they are enforced even today, the dimensions of the problem of insuring anything like the optimal provision of even relatively good ideas in early human evolution must be very great. And every available solution to it looks like it is going to have to make use of normative guidance, plus coercion to impose and enforce the guidance.

no longer interested. Once on the scene, the thrusting spear will enable a group of scavengers to defend themselves against predators, and to force predators away from their kills before all the available flesh has been consumed. This makes for another substantial increase in protein and so accelerates the co-evolutionary cycle in which it is a factor. But the thrusting spear's use requires that the lineage solve the problem of coordinated use, along with the problem of preserving and transmitting the technology of spear production. There is independent evidence that the shared intentionality required for such coordination is beyond the cognitive capacities of chimps, but is available in human infants as early as the age of 2 (Herrmann/Call/Lloreda/Hare/Tomasello 2007). There must have been strong selection against those *Homo* species that lacked this capacity, and strong selection for the enhancement of shared intentionality even independent of language (as the doubtfully linguistic Neanderthals use of thrusting spears reflects). Minimal levels of this capacity are required in the coordinated use of thrusting spears and in teaching their use along with making them. And these achievements both presuppose prior or simultaneous development of norms that enforce collaboration in the solution of the complex design problem posed by the need to divide labor, or posed by the advantage of doing so, and of dividing labor between production of abstract objects (good ideas) and concrete objects (axes and spears). The level of coordination involved in these levels of division of labor between tool making, tool use-teaching, and actual use is probably higher than that required for the coordinated use of the thrusting spears that the division of labor provided!

It appears that modal size of hunter-gather groups is 150 individuals or so. Larger groups appear to fission. At such sizes, kin-relatedness is too low to reward those specialists who provide good ideas to all members or teach all the group's children. Increases in their inclusive fitnesses are not great enough for such specialization to persist. What is needed of course are norms that encourage or reward the specialist's knowledge by enhancing his or her individual fitness. Many different schemes will achieve this end, including according status to savants or shamans, and providing them with offerings of concrete goods that propitiate their conveyance of good ideas or their products to others. But it is evident that once the division of labor in the innovation of good ideas has gone even a little way towards complex tools, teaching their use will exploit and in its turn select for more shared intentionality and its normative governed use in teaching, learning, practice and ultimate practical application in the field.

The solution to these division of labor problems is obviously trade between non-kin related individuals. This institution and the norms that guide it appear to have sprung up among *Homo sapiens* no later than the middle Paleolithic period, 120,000 years ago.

Trade of course requires norms of reciprocity and requires enforcement. When population survival turns on trade in technologically strategic commodities, say obsidian arrows or bone needles, which requires special knowledge and training in the manufacture of, then the need for normative guidance is further strengthened. Besides the norms governing exchange, the norms that make investment in the transformation of a good idea into its concrete instances become crucial

to the survival of whole groups. And all this must have been assured to make human survival possible long enough for altruism failures to actually even get started.

The upshot of all these considerations is that a capacity for normative guidance and with it ethical norms must have emerged earlier than and independent of those which function to remedy failures of psychological altruism.

3. All Change?

Unlike too many naturalists Kitcher is sensitive to David Hume's challenge about all inferences from facts to values, from 'is' to 'ought'. "[S]omething is uncontroversially correct about the Humean worry." (*EP*, 255) He neither interprets it as an observation that such inferences can't be deductive and so must be inductive (Dennett 1995), nor does he offer arguments that deny the distinction between the descriptive and the normative. Nevertheless, ethics is an objective matter that can and sometimes does move in the direction of global progress. The development of ethical norms is not as Kitcher says, a matter of "just one thing after another". This is where a pragmatism that has been articulated elsewhere in Kitcher's philosophy provides the fulcrum on which to lever his analytical history into a theory of what is morally right, correct, true. To deal "with Hume's challenge, the key is to recognize the effect of substituting progress for truth as the fundamental notion" (*EP*, 267).

Recall the pragmatist's way with truth and especially its application in the philosophy of science: scientific realists insist that at least some of the most important claims theories make are knowably true and made true by the world in spite of the fact that these very claims about the world cannot be established by observation. Antirealist skeptics draw attention to this apparent inconsistency. Pragmatists provide an alternative: at the end of inquiry science and scientist will converge on a single set of conclusions. That set is the final truth about the world in virtue of the convergence on it. Claims made before the arrival at that final set are true just in case they or recognizable successors figure in the final package of claims at the end of inquiry. In the words Kitcher recalls more than once from William James, "truth happens to an idea" (*EP*, 7, 210)—the idea acquires it by fitting into a system of ideas with mainly predictive virtues in application. Progress in formulation of this set is conceptually prior to truth.

Realists respond that if and when science converges on a single set of successful ideas, there will remain the question of why the package succeeds. They will go on to insist that the only explanation of cumulation is that the claims are individually made true by some fixed fact about the world. Realists accept that if the fundamental laws of working governing reality keep changing over the history of the universe, science will not converge on any single set of theories but will continually change, tracking changes in reality. Keep these points in mind.

Like the pragmatist philosopher of science Kitcher grounds the objectivity of ethics on an account of progress as convergence and treats truth, rightness, correctness, as a matter of membership in the set of norms at which ethical

progress culminates. A change is locally progressive, or as Kitcher also calls it, weakly progressive, just in case it leads to greater fulfillment of those functions of ethical practice that have emerged at the pertinent stage—the historical period in which the innovation is introduced. The functions of ethical practice are understood in terms of the notions of biological function that go back to Wright and Cummins and which Kitcher has reconciled in Kitcher (1993). The base function, recall, is remedying altruism failures. Over time norms that reduce these failures in various contexts raise new design problems that select for new norms, building on one another into institutional structures of human culture. Global or "strong" progress occurs "just in case it [a change in the norms] introduces elements that will be preserved in any indefinitely extended sequence of [locally or weakly] progressive changes" (*EP*, 261). And moral truths are the ones that are so preserved, world without end.

Notice how this strategy circumvents Hume's problem.[3] The original step from altruism failure to a norm that avoids it is purely factual, or if normative at all, a matter of uncontroversial 'instrumental norms' or 'hypothetical imperatives', ones with no more normatively controversial content than 'You ought to use a Philips head screwdriver if you are driving a Phillips screw'. Similarly, 'In the absence of storable resources, you ought to adopt the norm of equal division in order to avoid an altruism failure'. The grounding of such norms on purely factual considerations cannot fall afoul of Hume's strictures on not inferring ought from is.

Many historians may conclude that since the onset of the Holocene, 10,000 years ago, changes in the local social and economic ecology made much of human history a matter of oriental despotism, slavery, patriarchy, feudal serfdom, industrial exploitation. Under their prevailing conditions of great scarcity, many norms imposed by *force majeure* did not qualify as weakly progressive. They exacerbated altruism failures instead of avoiding or reducing them. Kitcher may be more sanguine in his view of actual human history than this: "The standards

---

[3] The relevant passage reads, almost in full:

> "[...] it is worth stating explicitly how Hume's challenge has been met. Hume worries that any naturalist account of ethics is committed to forms of interference incapable of justifying the ethical conclusions obtained. Pragmatic naturlism understands notions of ethical truth and justification in terms of the fundamental notion of progress, conceived as functional fulfillment and refinement. Introducing ethical novelties, whether at the beginning of ethical practice or in subsequent modifications, is justified when those who make the change do so by following a process likely to lead to better functional fulfillment. Any inferences they make are thus evaluated by appeal to fundamental criteria, which either may make allowances for their historical circumstances [weakly progressive justification] or may consider the indefinite evolution of the ethical project [strongly progressive justification]. Many of the processes through which historical actors made the changes they did are likely to satisfy the less demanding criterion, and later analysts can meet the more exacting [strong progressive justification] standard for some (perhaps all) core themes in ethics. [...] Once ethics is viewed as a social technology, directed at particular functions, recognizable facts about how those functions can better be served can be adduced in inferences justifying ethical novelties. The mystery that worried Hume disappears." (*EP*, 262)

for weak justification [of a change as locally or weakly progressive] were probably met quite frequently in the progressive transitions our ancestors made." But he immediately asks, "Can a stronger conclusion be defended?" His answer: "Perhaps."

> "Honesty exemplifies the 'vague core themes' that serve as the best candidates for ethical truth [...]. Analysis can offer lines of reasoning for thinking that some elements of our current practice—like rules enjoining honesty—would be preserved in any indefinitely proceeding [weakly] progressive sequence of transitions. The injunction to truthfulness remedies a class of altruism failures that will continue to be important for us as long as human beings communicate with one another. By identifying permanent human needs, analysis can conclude that some elements in ethical practice will probably figure in any progressive development of what we now have. When they do, their judgments are [...] justified as [strongly, globally] progressive." (*EP*, 261)

Honesty as a strongly or globally progressive 'vague core theme' is an interesting case just because, as the passage above recognizes, enjoining this norm is required for proto-linguistic communication to get off the ground, stay airborne and develop into full fledged language. It must therefore long antedate the emergence of the first remedies for altruism failure.

Be that as it may, why suppose that the indefinitely extended sequence of weakly progressive changes that begin from today will include any norm that we count as weakly progressive? Ethical norms have selected functions. Let's assume what we disputed in the previous section, that they are all selected for remedying altruism failures in the environment in which humans operate. However, as Darwin recognized, environments change, sometimes so radically that one environment's adaptation is another environment's maladaptation: white coats work well for the polar bear in the arctic, but make them increasing conspicuous as the planet warms up, moving the temperate zone north. Altruism was of course itself presumably selected for in the immediate pre-Holocene environment of evolutionary adaptation, and remains so in the modern environment of human beings.

The question of whether there will be global progress of the sort Kitcher requires for the objectivity of ethics thus turns on the continued constancy of the environment in which altruism failures may occur. An illustration will make the problem clear. Conceive of a future in which, owing say to cold fusion or something like it, no further scarcities of any sort burden humanity and the consequences of altruism failure become negligible. Do the norms that continue in force retain their status as matters of morally justified local and global progress? Without altruism failures to remedy, they have no function. Do they still have ethical force? Most people will be inclined to answer this question in the affirmative. The answer reflects people's beliefs that moral norms are right, correct, justified, true, and would remain so even when there are no longer any altruism-failures to remedy. Kitcher repeatedly diagnoses these convictions as hankering

after a sort of moral truth that doesn't exist (and is not required either). The conclusion, to which Kitcher should subscribe, is that under these circumstances, moral norms become adaptationally neutral, vestigial traits. They might even be forgotten, disappear, become extinct, for want of occasions on which they are needed to remedy altruism failures. If the set of norms that remedy actual altruism failures is the null set, do the ones that solve merely possible ones retain their status as globally progressive, are they still the morally right norms?[4]

Consider a less far-fetched and more disturbing scenario. Hume observed, and Rawls concurred, that ethical norms have force for creatures like us under conditions of moderate scarcity. But what of conditions of extreme scarcity? Will the norms that remedy altruism failures during periods of moderate scarcity, continue to be preserved in the indefinitely extended sequence of locally progressive changes? Some norms that are effective in making individuals behaviorally altruistic will almost certainly be fitness-reducing even as they continue to function to 'remedy' psychological altruism failures. Other norms may well cease under conditions of extreme scarcity to remedy altruism failures altogether. For example, as between two individuals who cannot both survive given available resources, norms of equal sharing will make both parties worse off and prevent both from taking steps that insure the survival of either of them.

Suppose the environment changes over the future in ways as substantial as it has changed over the past. This is a modest assumption given the ability of human kind to modify its own environment over evolutionary time-scales and to do so to its own disadvantage by overpopulation, pollution, improvident resource consumption, etc. Under some of the changes possible or likely, the norms that remedy altruism failures (even if psychological and behavioral altruism continue to be adaptations) may never converge on a fixed set, they may cycle through various sets, or they may never show any discernable pattern in the succession of norms. Given the possibility of radical (natural and social) environmental change, can we be confident there will be global or strong progress? More concretely, can we be confident that some norms we now endorse, will continue under all circumstances to remedy altruism-failures? For Kitcher this becomes the question, can we be confident that our most cherished altruism-failure remedying norms are morally right, correct, true?[5]

Return to the parallel with scientific theories. Could the trajectory of our 'best' science show no accumulation, no convergence on a single package of theories with ever increasing predictive power and range? Well, it's possible, though

---

[4] Under circumstances in which altruism failure need not be remedied, we might be tempted by another challenge to naturalism, G. E. Moore's challenge. We might ask the question, 'Is psychological altruism morally good?' If the question makes sense, if it is, as Moore said, 'an open question', we are going to be dissatisfied with the theory that morality consists in following altruism-failure remedying norms. For it will make sense to ask the question, 'Is remedying altruism-failures morally good?' when the answer to that question should on Kitcher's account already be given. Kitcher's reply to this challenge is to ask the question, 'Do things go better on alternative pictures?'—the Kantian or constructivist or contractarian justifications of ethics (*EP*, 272). His answer is no, pragmatic naturalism is no worse than the leading alternatives when it comes to a challenge like the one usually attributed to Moore.

[5] Notice this worry turns out to make Moore's 'open question argument' even more pressing for conditions of extreme scaricty than for ones of complete superfluity.

no one would bet on it. Why? Because the 'environment' against which the predictions are tested remains the same in the sense that its fundamental laws of working are presumably fixed over the whole history of the universe, and it is convergence of the package of theories we create on these fundamental laws that global scientific progress consists in. This is why even nonpragmatists are prepared to tolerate a pragmatist's account of 'truth' as something that happens to an idea. Of course pragmatists about the objectivity of science themselves will reject any demand for an explanation of the successful convergence in terms of semantic truth. All they need is global progress.

Pragmatism about ethics substitutes global progress for truth as the ground of its objectivity. But the environment selecting for ethical norms changes over time. Its vicissitudes may preclude convergence on a set of norms that after a certain point continue to remedy altruism-failures time without end. Not only is there no assurance of global progress, but a lot of local regress—which Kitcher himself documents—may sustain pessimism about the prospects for strong global progress, even if over a very long period norms remedying altruism failure remain unchanged or are subject only to refinement instead of replacement. Without the assurance of progress, the ethical project lacks the objectivity that Kitcher hopes to confer on it while avoiding the naturalistic fallacy.

The alternative groundings for the objectivity of ethics that Kitcher canvasses —intuitionism, Kantian constructivism (but probably not contractarianism), have the feature that they will endorse roughly the same norms as objectively right come what may in the ecology in which human history is played out, including the one our own institutions shape for us. This must lead us to doubt Kitcher's defense of pragmatic naturalism's grounding for the authority of ethics:

> "*The accounts proposed by pragmatic naturalism and by the critics are on a par.* If any of them can deliver an explanation of the 'real objective authority' of ethics, pragmatic naturalism can do so. If something more is wanted, all are deficient—and it will be reasonable to judge that the sort of authority demanded is a myth." (*EP*, 269)

Something more is wanted, and needed, even by pragmatic naturalism to show the ethical project is globally, strongly progressive and therefore has a right to arrive at the truth.

4. Kitcher's Challenge to External Moral Constraint Theory

In Part II of *The Ethical Project*, Kitcher develops an argument against competing metaethical models. We conclude by assessing the force of his objections against these more traditional views.

Kitcher collects his targets—prominently, standard moral realist and constructivist views—under the label 'external constraint' models. On this view, there are objective moral constraints or moral truths to which we have epistemic access. Moreover, on this view moral truth is given priority in the analysis of

moral progress and knowledge: moral progress is analyzed as the acquisition of innovative moral knowledge and moral knowledge is analyzed as the reliable acquisition of moral truth (*EP*, 178–9, 209–10, 265).

Kitcher imposes an explanatory burden on external constraints theorists: to show that the best explanations of the processes that led historical innovators like the Quaker abolitionist John Woolman and women's rights advocate Mary Wollstonecraft to their apparently progressive moral judgments supports the reliability of those processes: "What psychological processes go on in innovators, and how are the constraints on ethical progress registered in their thinking or feeling? How exactly do reason, intuition, and perception work to generate new ethical insights?" (*EP*, 6, 180, 182–183, 186) Kitcher argues that if external constraints theorists cannot meet this explanatory burden, then they cannot account for innovative moral knowledge and thus genuine moral progress—they are forced to view apparent moral progress as 'mere change', one damn thing after another.

Kitcher's 'debunking challenge' proceeds as follows.[6] On the external constraints model, there are objective moral truths to which we have epistemic access. If so, genuine moral progress (as opposed to 'mere change') consists in the arrival at new moral knowledge of these truths by historical ethical innovators such as Woolman and Wollstonecraft or by contemporary thinkers (*EP*, 187, 193–5). However, the best explanations of the moral judgments of historical ethical innovators do not support their reliable formation (*EP*, 178–86). If the best explanation of a judgment does not support its reliable formation, then we are justified in thinking that the judgment is not reliably formed. Moreover, contemporary thinkers have no reliable ways of moral judgment-formation that were unavailable to historical ethical innovators (*EP*, 205–7). So, neither historical ethical innovators nor contemporary thinkers have arrived at new moral knowledge. As a result, on the model of external constraints, there has been no genuine moral progress, only 'mere change'.

Kitcher's challenge turns on a fairly neutral epistemic principle, namely the premise that the accessible unreliability of a judgment *defeats* the judgment's knowledge status. Moreover, Kitcher forestalls the anticipated 'expertise' response to historical debunking—that is, the response that we experts today are morally reliable even if our forebears were not. He does so by reasonably denying an epistemic schism between the judgment-forming processes of historical ethical innovators and those of contemporary thinkers.

But Kitcher does not show *why* external constraints theorists bear the explanatory burden he assigns them. Without a convincing argument in favor of

---

[6] Kitcher sketches its core briefly in the Précis:

> "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 [...] 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." (Précis, 9)

imposing this demand, Kitcher's challenge to such metaethical theories lacks force against them.

The explanatory demand against external constraint theories appears to be grounded on the epistemic principle that if we are justified in thinking that the best explanation of a judgment does not support its reliable formation, then we are justified in thinking that the judgment is not reliably formed. But consider: if the best explanation of our philosophical (e.g. epistemic, metaphysical) judgments does not support their reliable formation, then this epistemic principle would provide a defeater for them as well. Other philosophical theories would be 'companions in guilt' with the ethical constraint theories. This companion in guilt worry can develop into serious *tu quoque* objections. If the best explanation of Kitcher's philosophical judgments (in his challenge and throughout his book) does not support their reliable formation, then these judgments face defeat too. Kitcher could reply of course that the best explanation of his judgments vindicates their reliability, but this would demand argument. For the same sort of non-vindicating psychosocial explanations he gives of, say, the philosophical judgments of G. E. Moore and Kant (*EP*, 206–7) could be sketched for many of our equally situated philosophical judgments, including Kitcher's own. Moreover, it is no light burden to deliver vindicating explanations of our philosophical judgments. It is hard enough for philosophers of science and historians to show that the best explanation of standard scientific judgments (e.g. concerning subatomic particles) vindicates their reliability (*EP*, 206–7). The availability of analogous stories for each of the confident philosophical, mathematical, and scientific judgments that we want to sustain appears rather dim at present, in part because we know little about the sources of many of these judgments. All this suggests: even if the current best explanations of our judgments do not support their reliability, this *by itself* does not defeat them.

Kitcher perhaps could revise his epistemic principle to avoid the potential *tu quoque*: justified belief that the best explanation of our moral judgments does not support their reliability gives us *prima facie* (rather than decisive) reason to doubt their reliability. An additional premise: this *prima facie* reason amounts to decisive reason to reject moral reliability *given* the absence of independent reasons favoring moral reliability. The revised epistemic principle is plausible—the fact that moral reliability does not cohere well with our current best explanations of our moral judgments gives us some reason to doubt their moral reliability. With the revised principle Kitcher's challenge could also avoid some 'companions in guilt' objections this way: our moral judgments are defeated because we lack independent reasons in favor of their reliability, but our confident mathematical, scientific, and philosophical beliefs are not defeated because we possess independent reasons favoring their reliability. Considerations of, say, indispensability and predictive success may vindicate the reliability of our confident mathematical and scientific judgments. But it is to say the least unclear whether good arguments exist for the reliability of our purely philosophical judgments. If there are such arguments for philosophical reliability (say, based on appeal to our autonomous rational capacities), then similarly patterned arguments might very well work for the reliability of moral judgments.

Moreover, advocates of the ethical constraint view could argue that Kitcher's challenge does not undermine moral reliability *to any extent* (*prima facie* or otherwise). For the lack of coherence between the best explanation of our moral judgments and moral reliability to undermine moral reliability to any extent, the best explanation must not be merely best, but rather *probably true*. If Kitcher's explanations of, say, Woolman's innovative moral judgments or his explanations of the innovative moral judgments of our ancient forebears are only best among a set of improbable explanations ('a poor lot'), then it is difficult to see why the failure of moral reliability to cohere well with such epistemically suspect explanations is any mark against moral reliability. Given legitimate concerns about the relative merit of Kitcher's explanations of the moral judgments of our forebears,[7] it is arguable whether his explanations can bear the load required for the success of his debunking challenge.

Finally, exponents of the ethical constraint approach can develop explanations of the reliability of important moral judgments out of Kitcher's own explanations of how they arose. Consider how Kitcher describes his central cases of Woolman and Wollstonecraft. Was Woolman justified in his abolitionist proposals? Was Wollstonecraft justified in her call for women's emancipation? According to Kitcher's pragmatic naturalism, this depends on whether the "underlying processes are likely to lead to ethical progress" (*EP*, 250). The question of Woolman's or Wollstonecraft's status as a justified agent of moral progress thus becomes for Kitcher whether the processes underlying their proposals were likely to lead to the mitigation of altruism-failures.

Here is Kitcher's answer to this question:

> "Empathetic understanding, especially when sensitive to [...] desires that are typically suppressed can foster societal discussion to remedy existing altruism failures. People who call attention to the frustrated desires, beginning a broader deliberation about them, are likely to move the ethical project toward refinements of its original function. Wollstonecraft could not frame the issue in this (analyst's) way; but her sensitivity to women's suppressed aspirations would be expected to refine the original function of ethics. Woolman is a more ambiguous figure. His account of his route to rejecting slavery reveals the role his own anxieties about personal salvation played. Insofar as these were dominant, we cannot view the psychological processes that generated his proposed reform as reliable. Nevertheless, some considerations underlying his hesitation about writing the bill of sale—concerns for the pains and abuses visited upon slaves, worries that official interest in their well-being (their salvation) is a sham—indicate the operation of an empathetic capacity. To the extent that he resembles Wollstonecraft, we can view him, too, as a justified agent of ethical change." (*EP*, 250)

[7] Kitcher concedes that "there is much even about John Woolman's psychological development that remains unknown" (*EP*, 184). Kitcher's psychosocial explanation of Woolman's innovative rejection of chattel slavery appears to depend centrally on a somewhat strained reading of a passage from Woolman's Journal. For the reading, see *EP*, 159–62, 184–5.

In this passage Kitcher suggests that to the extent Wollstonecraft and Woolman are justified agents of moral progress, this was due to the operation of their *empathetic capacities*, even if they were not aware that their empathetic sensitivity to the frustrated desires of others would likely lead them to moral progress. But if the operation of impartial empathy confers reliability (in Kitcher's pragmatic sense) on the proposals of ethical innovators, why can't the external constraints theorist legitimately claim that impartial empathy confers reliability (in the traditional epistemic sense) upon the processes leading to innovative moral judgments? Impartial empathy *tends* to get us to take into consideration the interests and concerns of others. So, if moral truths are grounded fairly impartially in human interests, or in principles that nobody could reasonably reject, or in the responses of an empathetic ideal observer, it is open to argue that the operation of our empathetic capacities (all else equal) would be likely to lead us directly to *moral truth*.

Pace Kitcher, the external constraints theorist need not posit crude episodes of 'moral insight', direct moral perception, analogs to scientific discovery, or a heady awareness of the truth conditions of moral judgments (*EP*, 196–7). Rather they could invoke impartial empathy to explain the reliable acquisition of new moral knowledge. All this is to say that if impartial empathy vindicates moral progress on Kitcher's pragmatic naturalism, it could apparently also vindicate moral progress on an external constraints model.

Another vindicatory route suggests itself: empathy could be supplemented by morally reliable interest-based reasoning. Kitcher observes that "Woolman recognizes, in perfectly straightforward ways, that slaves are subjected to all kinds of pain, and Wollstonecraft understands that her own desires, as well as the wishes of women she knows, are thwarted". Woolman and Wollstonecraft also "do not find anything strong enough" in the views and arguments of their opponents "to counterbalance the perceived pain relief and desire satisfaction" that would accompany the abolition of chattel slavery and the education of women (*EP*, 196–202). It appears then that these ethical innovators reasoned to moral reform in part from the balance of interest-based considerations. Even if they did not reason in this way, other ethical innovators have. If moral truths are grounded fairly impartially in human interests, on principles nobody can reasonably reject, or the responses of an empathetic ideal observer, then such reasoning could very well be morally reliable. If so, external constraints theorists possess another way of explaining moral reliability and hence another way of vindicating moral progress.

Kitcher might respond that even if an innovator's judgments are reliably formed through empathy and interest-based reasoning, nonetheless the fact that nearly everybody around *disagrees* with her defeats the knowledge status of her innovative judgments. So even if moral reliability can plausibly be explained, moral *knowledge* (conceived as *de facto* reliability plus the satisfaction of a 'no defeat' condition) cannot.[8] But, arguably, disagreement only defeats knowledge

[8] See *EP*, 200–203. Also: "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." (Précis, 7)

under certain conditions, perhaps some epistemic peerhood condition. Given that Woolman and Wollstonecraft seriously engaged with the objections of their opponents and (rightly) found them wanting and frequently based on factual mistakes, empathy failure, partiality bias, and so on, it is doubtful that disagreement defeats their innovative judgments.

The last few considerations strengthen the suspicion that pragmatic naturalism may possess no advantage over the external constraints model with respect to making sense of moral progress. External constraints theorists will aim to meet Kitcher's explanatory burden by offering vindicating explanations of the reliability of our moral judgments. With Kitcher, we doubt the merit of such explanations, but much more needs to be said in defense of this doubt and about why the present unavailability of good vindicating explanations poses an important skeptical problem for external constraints theorists.

5. Conclusion

We have raised doubts about Kitcher's genealogy of morals and about his metaethics. We have been silent about its substantive normative claims. From these we do not dissent. Indeed, as critics we have been entirely silent on the great virtues of *The Ethical Project*. But these are so obvious as to not require rehearsal by anyone.

Bibliography

Dennett, D. (1995), *Darwin's Dangerous Idea*, New York

Grice, H. P. (1957), Meaning, in: *Philosophical Review* 66(3), 377–388

Herrmann, E. / J. Call / M. Lloreda / B. Hare / M. Tomasello (2007), Humans Have Evolved Specialized Skills of Social Cognition: The Cultural Intelligence Hypothesis, in: *Science* 317, 1360–1366

Hrdy, S. (2009), *Mothers and Others*, Cambridge/MA

Kitcher, P. (1993), Function and Design, in: *Midwest Studies in Philosophy* 18(1), 379–397

— (2011), *The Ethical Project*, Cambridge/MA

—, Précis of *The Ethical Project*, in: *Analyse & Kritik* 34, *this issue*, 1–19

Mellars, P. (2006), Why Did Modern Human Populations Disperse from Africa ca. 60,000 Years Ago? A New Model, in: *Proceedings of the National Academy of Science* 103(25), 9381–9386

Skyrms, B. (2010), *Signals: Evolution, Learning, and Information*, Oxford

Warneken, F./M. Tomasello (2009), Varieties of Altruism in Children and Chimpanzees, in: *Trends in Cognitive Sciences* 13(9), 397–402