

## 18 Is human nature natural?

All the *sentiments* of the human mind, gratitude, resentment, love, friendship, approbation, blame, pity, emulation, envy, have a plain reference to the state and situation of man, and are calculated for preserving the existence and promoting the activity of such a being in such circumstances.

Hume, *Dialogues Concerning Natural Religion* (1779), part 3, §13

It is quite common for philosophers, if not biologists, assessing the impact of Darwin on our views about human nature to associate him with Hume. Owen Flanagan even talks of a ‘Humean-Darwinian’ picture of human nature.<sup>1</sup> In this chapter I want to chart the relation in some detail. The result is not to disrupt the marriage but to suggest, in more detail than usual, how close and how fertile it actually is.

## I HUMAN NATURE

By ‘human nature’ we generally denote not our bare animal construction—two eyes, four limbs, one head, upright gait and so on—but our rather more exciting psychological traits. Theorists of human nature are particularly concerned with our cognitive and motivational dispositions and capacities, the subjects of Hume’s *Treatise of Human Nature* (1739). Nowadays, on each of these topics, people turn to Darwinian theory, or speculation, to discover whether it is in our nature to be rational, emotional, selfish, altruistic, short-sighted, prudent, aggressive, pacific, promiscuous, monogamous, murderous, or moral. I shall here mainly discuss motivational states, although our cognitive capacities will also make an appearance.

Once a Darwinian take on our motivational states is in view, it can immediately be doubted whether the idea of human nature is worth saving, or whether it is merely a stranded remnant of the Aristotelian idea that everything has a Natural State: an essentialism which was itself one of the casualties of Darwin’s revolution, destroyed by the twin ideas of mutability of species, and variation within populations.<sup>2</sup> The second of these is in front of us all the time: it is after all commonly observed that people differ in respect of selfishness, emotionality, aggression and the rest. We can if we like suggest that such differences as we observe are only superficial, but that is a ploy that needs careful handling.

For, put slightly more precisely, the problem with one human psychological nature is, first, that there is variation in the human genome – indeed, most writers believe that fomenting this variation is the function of sexual reproduction and its accompanying

genetic recombinations—and, second, that in fine detail (and the making of brains requires fine detail) the journey from genome to phenotype shows no one fixed relationship. It shows only a variety of ‘norms of reaction’, as genes express themselves differently in different epigenetic environments. The importance of those environments firmly brings culture into the picture. We even know something about neurological substrate showing its influence. The neuroscientist Michael Meaney found in rats that maternal licking and grooming resulted in offspring with better-developed hippocampi, which released less of the stress hormone cortisol when the rats were startled.<sup>3</sup> The mother rats had shaped the brains of their offspring by activating serotonin receptors in the hippocampus, which in turn sent transcription factors to turn on a gene that inhibits stress responses. Here is a nice case of an epigenetic intervention altering the neurological and hence the psychological nature of offspring. And as with rats so, we can be sure, with us: if maternal behaviour (and in other experiments, even certain dietary changes, such as those involving the methyl radical) can make these changes, then so much for any brash confidence that our genes are our destiny, or that culture is causally inert. It is noteworthy, as well, that the difference is able to generate lineages of rats with differing stress reactions: rats that have been licked and groomed make less stressed-out mothers, and so the trait is passed on.

Other epigenetic variations are known to persist through generations. The molecular geneticist Emma Whitelaw has described the consequence of epigenetic inheritance thus:

It changes the way we think about information transfer across generations. The mind-set at the moment is that the information we inherit from our parents is in the form of DNA.

Our experiment demonstrates that it's more than just DNA you inherit. In a sense that's obvious, because what we inherit from our parents are chromosomes, and chromosomes are only 50 percent DNA. The other 50 percent is made up of protein molecules, and these proteins carry the epigenetic marks and information.<sup>4</sup>

In the face of all this potential for variation, someone might still suggest that there is a unique human nature 'in the raw', perhaps gesturing at the kind of environment in which our ancestors most probably lived as they gradually became something different from our hominid predecessors. This can be said, but the language is unfortunate. It easily suggests that anything that differentiates us now from them then is 'superficial' or even in some sense a veneer or disguise of our true natures. This is no more sustainable than saying that since, in order to flourish in Africa, these ancestors were probably beneficiaries of darkly pigmented skin and tightly curled hair, those of us who now have neither, bear merely 'superficial' phenotypical appearances, as if underneath a veneer of fair skin and straight hair our 'real nature' is to be dark skinned and curly haired. If the idea of the 'raw' is attached to that of one real nature, then there is no raw to be found. Anything suggested as raw is just one environment that our ancestors occupied at one time during the change from something in the Cambrian to ourselves, and it will be an open question whether any behavioural dispositions that existed in that environment is more than a poor and partial guide to what exists around us in ours.

A much better suggestion, therefore, is that, as with Michael Meaney's rats, there are interesting constancies across a variety of environments, just as there are other phylogenetic constancies in animal development. There would be a finitely storable, lawlike relationship between recurrent features of environments and particular traits, and

our shared human nature would be described by that function. This may be so. It may be that one day we could in principle write the Book of Human Nature describing a complete and lawlike function from environmental factors to psychological natures. But it may be that there is no such book to write, for time and chance, the engines of mutation and change, might affect the function itself just as readily as they affect the overt psychological outcomes. This would not be scepticism about human nature from the discredited standpoint of a 'blank slate', but from the standpoint of a combinatorial explosion of differently primed slates.

We may remain more optimistic than all this suggests about a unique human natural endowment. After all, Hume's list of sentiments, in the quotation at the beginning, is instantly recognisable. We are indeed creatures prone to those sentiments, amongst others, and they seem to bubble up in most normal environments. We are also creatures who do things; and one of the inescapable facts of human existence is our talent for evolving concepts and strategies for dealing with the natural world. We did not make that world; but our conceptual repertoires, which in turn influence our motivational repertoires, are things we have made, or, better, which have made themselves. By that I mean that they themselves may evolve without conscious design, just as our phenotypical features have done. This brings us directly to Hume, but before putting him at the centre of the stage, a detour is necessary.

## II WHAT IS DARWINISM?

Richard Dawkins defines core Darwinism as the ‘minimal theory that evolution is guided in adaptively nonrandom directions by the nonrandom survival of small random hereditary changes’.<sup>5</sup> The notion of evolution being guided at all might disconcert some, if they fear that the ghost of Paley’s clockmaker can be discerned in it, perhaps in the guise of some anthropomorphic version of Mother Nature. Such, of course, would be the reverse of Dawkins’s intention. It would perhaps have been better, then, just to have talked of the minimal theory that evolution *moves* in adaptively nonrandom directions through the nonrandom survival of small random hereditary changes. This is a nice definition for several reasons. It does not imply, for instance, that nothing else could affect the direction that evolution takes. It is thus compatible with random drift, ‘trembling hands’, or catastrophic external impacts upon varieties of life. It does not imply that the result of the process will ever be ‘optimal’. It does not imply that an organism will be the inflexible owner of fixed routines, but leaves open that it may be flexible and elastic in the ways it responds to environmental stimuli and change. It does not ask us to think of evolution ‘selecting for’ one trait or another, unless, as it should be, that phrase is understood as a harmless label for the specific reproductive advantage enjoyed by an organism that bears some trait as the result of heritable variation. In fact the tidied-up minimal theory implies so little that by itself it might scarcely count as a piece of falsifiable scientific theory. The punch comes when the processes of small random hereditary change are identified, and the advantage they give specified. Then science gets underway.<sup>6</sup>

The adaptively nonrandom direction that Dawkins describes is due to the bigger number or better survival rate of descendants of organisms – an increase due to whatever properties the small random hereditary change gives those organisms. In obvious scenarios, it may make them bigger, or faster, or give them better armament, or make them more resistant to some pathogen or other, or give them a better sensory adjustment to whatever it is that they need from their environment. It may make them more intelligent, or more cooperative, or more selfish, or equally, less selfish. For contrary to the common misapprehension of the Darwinian jungle as favouring only the big beasts, the predators and psychopaths, I take it that it is now well understood that in many circumstances restraint and moderation bring reproductive advantage. Among parasites, for example, a lethal variety that quickly kills its host is set to spread less efficiently than an otherwise identical variant that takes enough from the host to make it ill, but no more, enabling it to continue to move about and thence spread the parasite's offspring to other hosts. Contrary to the prevailing spirit in biology a generation ago, increased understanding of iterated games and of evolutionary dynamics in general has made it abundantly clear that sometimes nice guys finish first.

Thus imagine a room with a set of tables, at each of which a number of persons are engaged in some version of a prisoners' dilemma game.<sup>7</sup> In this structure, the total gathered by the table is greatest if each cooperates with the others. But everyone can do better for themselves, on each play, by 'defecting' or cheating, breaking the implicit agreement. If we imagine plays repeated, then the social goods accrue to the cooperating tables, while the tables with defectors stay impoverished. If we then imagine a dynamic whereby cooperators gravitate to cooperating tables, and defectors are banished to

defecting tables, the room will tend to split into two kinds: all cooperating and all defecting. If we make this an evolutionary dynamic—suppose the traits are inherited and people leave offspring in numbers proportional to their chips—gradually the room fills with cooperating tables.<sup>8</sup> If the environment is tough and livelihood depends upon cooperation, it is especially clear that we must hang together or we all hang separately, so hanging together it has to be.<sup>9</sup>

Advantage requires a moment's thought. It is of course relative to context: a trait that gives a higher reproductive success rate in some environments may not do so in others, just as the *dernier cri* in last year's fashion may fail to attract a mate in this year's dating market. But the notion of an advantage, that is, a trait being causally responsible for the numerical success in question, hides no intrinsic difficulty, although it has been suggested that it does. Taking Lewontin and Gould's famous example of the co-occurrence of spandrels and arches, Jerry Fodor argues that when traits are coextensive it makes no sense to say that it is one that evolution has 'selected for' rather than the other: this kind of remark could only be made relative to the intentions or purposes of a designer, so that unless we again postulate Paley's watchmaker or Mother Nature, the notion loses its application.<sup>10</sup> But there is no real difficulty here. Causation is discriminating. Two traits may be found together in nature, but one can be causally responsible for an increased number of descendants when the other is not. It may be that all and only vertebrates with eyes carrying around various proteins that go to making up eyeballs. But the sensitivity to light is what give the advantage, not the carriage of those proteins. If we imagine a close-up movie of the evolution of an eye, we would expect to find some initially minuscule sensitivity to light just enabling a creature to find food or

avoid predators or find a mate, when its blind competitors cannot. And that initial setup is all that is needed that creature a reproductive advantage due to light sensitivity, and not due to carrying little lens-shaped pods of proteins. (If a mutation generated creatures burdened with little lens-shaped pods of the same proteins that were not in any way connected to an ability to respond to light, then they would be merely weighed down—the reverse of advantage.) If we can detect any such benefit due to light sensitivity, then in principle we can anticipate and explain the differential in reproductive success, and if the trait can be inherited, then that is all evolution through natural selection needs.<sup>11</sup>

Fodor may not have intended his scepticism to embrace all cases of coextension, which is easily compatible with different causal powers attaching to different properties, but only to a different phenomenon which is slightly more complex. His illustration of the problem he has discovered is the whiteness of polar bears, and the question he asks is whether nature selects for the whiteness or for their property of being the same colour as their environment? Being white, and being the same colour as the environment are indeed two different features because lots of animals have the latter property but not the former. The idea is that neither answer could be other than arbitrary, and hence that, in the absence of the ghost of a designer, evolutionary theory's privilegings of one trait as an adaptation and another as a mere passenger are arbitrary across the board. That, however, would be an unwarranted generalization, since the example is quite untypical. In the case of the polar bear, it is tempting to say that being white *is what it is* to be the same colour as the environment—in the arctic. Whereas coextension does not in general deliver this temptation: nobody would say that having a heart *is what it is* to have a liver, just because all creatures that have one have the other.

And that in turn leaves the evolutionary theorist two compatible options for dismissing the difficulty. Consider this parallel case. Suppose my beloved child telephones with the bad news that his account is \$10,000 in the red; he is being pursued by the bank's tame thugs, and is naturally perturbed. I kindly put \$10,000 in his account, relieving his mind. A Fodorian question would be: which feature of my action actually relieved his mind? Was it my *paying off all his debt*, or was it my *putting \$10,000 into the account*? Like the whiteness and the camouflage, these are two different features: they can come apart. If he had owed more, I could have done one without doing the other. But the robust answer is that we simply don't want to choose: in this context putting \$10,000 into the account simply *is* paying off all his debt, and that's all we need to say. We do not need to refine further. If one psychologist says that I relieved his mind by paying off all his debt, and another says I did so by putting \$10,000 in his account, they are not at odds with each other. They might become so, if they added different microscopic detail to the exact route from my child finding what I had done, to his being relieved, but as it stands they are not.

A different answer, motivated by familiar philosophy of science, is that if Fodor holds a gun to our heads and insists that we do have to choose, it may be better in general to say that it is *paying off all his debt* that relieves my son's mind, because there is a simplicity and generality to that answer that accords with good scientific practice. If his debt had been greater, as it easily might have been, paying it all off would still have done the trick, whereas \$10,000 might not have. The application to camouflage is obvious. Being white only works in some contexts; being the same colour as the environment much more generally confers an advantage. But we do not have to choose to answer, and

an equally good response is gently to push the gun away altogether. A rather different charge, which also hovers in Fodor's article, is that there can be no general science of selection, any more than there can be a science of getting rich: there are simply too many underlying ways of doing it, too much 'variable realization' of the phenomena of advantage for there to be anything interesting and general to say about it. There may be some truth in this, but it is far from any kind of rejection of core Darwinism. It just means that each application of Darwinian reasoning stands on its own feet, as it were. But this is something that biologists certainly recognise; for one kind of advantage to one kind of organism will of course not look much like another kind of advantage to a different kind of organism.

A reasonable caveat about advantage is that it is not always easy to identify just one beneficiary of any advantage. An advantage may be multi-level: a new treatment of a football injury may benefit someone's ankle, the player himself, his team, the spectators and the doctor who invented it. Here there is a clear arrow of causation; for example, the treatment benefits the team by benefiting the player, and not the other way round. But in other cases it may be the other way round: better seating may benefit the players by benefiting the spectators, and generating a bigger gate. Advantage does not have to be 'bottom up', and neither does it have to be the advantage of one unique recipient only.

As Dawkins himself notes, core Darwinism gives us the template of an historical explanation of why something is the case. The template can be filled in genetically, in terms of mutations in DNA, for example, but in principle it could be filled in quite differently. This flexibility gives us a choice. We might want to say that only strict genetic mutation and variation with ensuing differential reproductive rate counts as a

Darwinian explanation of anything human. Or, we might want to say that if other processes throw up differences, and then those differences generate different rates of persistence, then those count too.

Dawkins's own famous example of extragenetic inheritance is that of the 'meme', thought of as a replicating cultural unit, capable of spawning more or less accurate copies of itself which then in turn have differential rates of reproductive success. If we take this more relaxed view of what can count as a Darwinian replicator, we might talk of the evolution, say, of post-Shakespearean English vocabulary as a genuinely Darwinian process. Shakespeare's time was one of remarkably fertile generation of lexical mutations and adaptations of old vocabulary to new uses. Some, such as 'eyeball' or 'retirement' did very well in the environment of listeners and readers and what they wanted, and thrived. Others such as 'skyish' or 'maugre' did not, and have gone extinct.

If we like to see this selective preservation and extinction as a Darwinian process, we must be aware that one crucial element of the more traditional Darwinian explanations of change is so far missing. In the case of polar bears' colouring or the evolution of the eye, we have a definite conception of the precise advantage responsible for the frequency change: the ability to blend into the environment and the sensitivity to light. In the case of English we do not. Nobody can say why some coinages succeed and others fail. We can, if we like, say that there must be 'something'—some need, some reason for catching on—but we do not know what it was, and perhaps nobody could have predicted that there was or was not any such feature. After all, attempts to identify an unfilled niche in a language and then fill it are notoriously quixotic: everyone knows that English really needs a gender-neutral singular pronoun, and at various times people have

campaigned for ‘thon’, ‘hesh, ‘hu’ and others, but with absolutely no success whatever. More generally, all we have in many cases of change is the post-hoc and quantificational judgment that ‘there must have been some advantage/disadvantage to it’. It is this that makes us suspicious that the idea of a ‘meme’ only yields the near-tautology that whatever catches on, must have caught on for some reason. It may be an invitation for us to look for the reason, but is not as it stands a piece of falsifiable science.

Nevertheless, the abstract or structural nature of Dawkins’s definition remains an advantage. For it opens up the possibility of seeing variation by natural selection in other cases where the advantage *can* be specified, and where we can have a good idea of the reason why it might have had an impact on numerical success. As we shall shortly see, specificity edges Darwinian explanation closer to Hume.

### III HUME’S NATURALISM IN DARWINIAN RETROSPECT

There are four main headings under which Hume’s title to being a Darwinian before his time can be highlighted. They are (1) his naturalism, or general emphasis on nature, not reason; (2) his particular emphasis in ethics on passion as against rational cognition; (3) his ‘ecological’ account of particular traits; and (4) most directly, his explicit genealogy of particular abilities, and in particular those associated with cooperation. There is space here only to sketch the kind of consilience with Darwinism that these themes in Hume illustrate.<sup>12</sup>

Hume’s emphasis on natural belief pervades all his writings. By ‘natural belief’ is meant the way our natural propensities bend or force the way we interpret the world. His

human being is an animal like others, bound to see and to believe. But reason cannot underwrite the process nor ever do more than tinker with it at the margins. In all the essential business of life, nature is too strong for reason. Thus any attempt to show that when events have fallen into some pattern then they *must* go on doing so, or even *will* go on doing so, fails. Any attempt to understand the notion of a causal power in things, thought of as a kind of straightjacket ensuring that the order of nature is stable and will forever remain so, fails. Any attempt to validate the senses, showing that they lead us to see things as they are, fails—indeed fails particularly spectacularly, since what happens instead is that when we reason things through, we find that our ordinary interpretation of the world around us cannot possibly be correct. Hence, if we relied upon our reasoning capacities to form or sustain our beliefs, we would be left with no confidence in anything, the situation of Pyrrhonian scepticism. But we are not extreme sceptics, because nature – our nature – is too strong for scepticism: ‘the great subverter of Pyrrhonism or the excessive principles of scepticism, is action, and employment, and the occupations of common life’.<sup>13</sup> Throughout all this he stresses the continuity with the animal world: ‘Though the instinct be different, yet still it is an instinct, which teaches a man to avoid the fire; as much as that which teaches a bird, with such exactness, the art of incubation, and the whole economy and order of its nursery’.<sup>14</sup>

His naturalism does not mean that Hume has no standpoint from which to criticize some common habits or dispositions of belief formation. He thinks that we are often unduly swayed by ‘trivial’ properties of the imagination, that we take inexact views of things, that we are influenced by passions that have nothing to do with the case, and so on. These are qualities of mind that are not adaptive, not useful, or at least not so when

they surface indiscriminately. This imperfection in turn gives him an excuse for his own philosophical activity. He is well aware that having denigrated reason so thoroughly, he is in a somewhat unstable position, writing books on philosophy, using the very tools of reason and reflection that he has apparently denounced. But he answers that there is still a role for ‘mitigated scepticism’, a position that accepts with a shrug the general caveats about the scope of reason, but goes on to use it with caution, and industry, in gathering data and reflecting on human nature and human existence.

Hume is here an ancestor of modern pragmatism, or the view that our cognitive machinery is an adaptation to enable us to cope, not primarily to enable us to represent the world, for as soon as we investigate whether it does that job of representation, scepticism is bound to assail us. Pragmatism is itself often thought of as a post-Darwinian philosophy, one determined to stick closely to the adaptive function of different parts of our cognitive architectures, but Hume is unmistakably on the same side. Indeed, he goes so far as to see belief itself in quasi-dynamic terms. A belief is distinguished by its ‘force and vivacity’, and its force and vivacity refer to its capacity to determine intentions and action.

#### IV HUME, DARWINISM AND THE PASSIONS

The second Darwinian element in Hume is his account of motivation in general, and moral motivation in particular. He entirely reverses the philosophical picture from Plato and Aristotle onwards, whereby unruly passions are to be governed and directed by the light of reason. His rejection of this is flamboyant: ‘reason is and ought to be the slave of

the passions, and can never pretend to any other office than to serve and obey them'.<sup>15</sup> He is not, of course, denying familiar processes of self-control, as when we summon up fears of getting fat in order to fend off the temptation of the second helping, or in more serious contexts where we think things through, bringing all our reserves of knowledge and experience to bear on practical decisions. But the point is that reason is here indeed serving the passions: by identifying alternatives and consequences and aspects of the situation we might have missed, reason clears the field, as it were, for our passions to operate. It nevertheless remains the case that we will only avoid the second helping if indeed we are afraid of getting fat, and the knowledge and experience we bring to bear is only relevant if it too reveals aspects of a situation that matter to us, that engage our passions.

None of the above meshes directly with Darwinism. But it begins to do so when applied to moral motivation. Hume's bold stroke is to assert that moral motivation is nothing special. It is no exception to the general picture of motivation by the passions, and it is certainly nothing with a tinge of the supernatural about it. The contrast here is not only with a Platonic view of a strange realm of laws, eternal fitnesses of things or principles of justice and right, available to a properly trained elite, but also with the religious or supernatural embodiment of the Platonic idea, seeing moral law as some kind of edict of the divine lawgiver. These are profoundly un-Darwinian ideas, since once the moral truth is divided from the natural world there can be no account of why 'seeing' that truth should give any particular advantage, nor is there any explanation of why we should be drawn to whatever we 'see', nor any apparent connection between what is mysteriously 'seen' and what matters to us here and now.

Hume's response is to short-circuit the whole thing. A moral motivation is essentially no different from any other. There are certain qualities of things, particularly of other people, which excite positive reactions in us: love and admiration, benevolence and pleasure. There are others which do the reverse. These tendencies are quite open: indeed they are so clear they are often simply given in our language. We all know if we are being criticized or applauded, whether we are called mean, boring, selfish, pedantic, rash, and foolish, or alternatively generous, cheerful, hard-working or thoughtful, or if we are called any of a thousand other things.<sup>16</sup>

Attributed to ourselves, these qualities produce shame or pride, and motivate us accordingly. Applied to others, they denote aversion or admiration. But how do qualities get onto these lists of merits or demerits, virtues or vices? After sifting the evidence, Hume determines that it is when they are 'qualities of mind useful or agreeable to ourselves or others'. It is a contingent fact about human beings—a brute fact about our natures—that we notice these qualities and are affected by them, even when they are not being exercised to our own personal benefit or cost (we admire or condemn historical agents and even fictional characters). But of course this fact about us is what makes social life and social coordination possible. Our caring as we do about moral qualities makes possible a 'common point of view' with those who are or were in the orbit of the agent. Hearing of a benevolent person, we feel a kind of love, just as those who were in his family or his circle of friends must have done. Hearing of a malevolent or even just a foolish person, we feel a corresponding annoyance or dislike. And it is this annoyance or admiration, aversion or respect, that is shared and absorbed as we become socialised, and that is voiced in our endless activities of gossip and evaluation.

By bringing ethics within the sphere of the passions Hume demythologises it. He represents it as a natural facet of our social lives, the expression of passions that are themselves advantages to us, useful and often agreeable. If at this point we worry that there is nothing much in this picture about obligation, duty, justice or principle, then we should wait for the further story about justice, which I shall come to shortly.

## V TRAIT ECOLOGIES IN HUME

Before turning to the genealogy of justice however, we should notice the third of Hume's proto-Darwinian interests, which is the particular functional or adaptive story he gives about individual traits and passions. Hume worked with confidence in a shared human nature: the complexities of variation and epigenetics were unknown, and little enough anthropological evidence of human variations was available. What was known and well-evidenced, however, was history; and Hume knew through his extensive work as a historian that against any background uniformity there existed considerable variation. In several of his essays (especially 'Of National Characters' and 'Of Some Remarkable Customs') he dwells on just this theme.<sup>17</sup> Other writers, such as Montesquieu, had speculated that national differences somehow arose from differences of climate.<sup>18</sup> Hume looks instead for what he calls 'moral causes'; and in so doing he offers what are, in effect, *functional* explanations of why a particular trait might be adaptive or useful in one set of circumstances but not in others. In fact, his targets for explanation are often extremely droll, such as the spendthrift nature of army officers, or the hypocrisy of priests, or the asymmetric attitude many cultures have towards male and female chastity.

He locates this double standard, as does modern speculation, in the asymmetric roles of men and women in reproduction. However, unlike modern evolutionary accounts, he also believes it is a fragile and induced asymmetry, culturally reinforced and sustained. He has no great faith in the ‘natural’ monogamous nature of women, and this doubt surely gives him an advantage over evolutionary psychologists who believe that such a nature is cemented into the genes. For cultures would not put in the enormous efforts they often do to encourage or enforce female chastity if it came as naturally as growing hair or feet.<sup>19</sup>

A final example of this work that I shall mention, although we have just scratched the surface, is the second part of the famous chapter on miracles, in Hume’s *Enquiry Concerning Human Understanding* (1748). In the first part of the chapter, Hume shows that when we are given a report of a miraculous event, an event absolutely contrary to the hitherto uniform course of nature, it should only be credible insofar as the falsity of the report would be as great a miracle as the event alleged to have happened. In the second part he goes on to urge that this condition is far from being met in any of the actual testimonies that different religions wave in front of us. But this leaves a puzzle: what is it in our cognitive economies that makes us susceptible to these kinds of story? Hume – here following in the tradition of Francis Bacon – sets out to theorise about cognitive dysfunction, and particularly the way in which such things as vanity and love of the marvellous distort our natural mechanisms of belief formation. He is here an ancestor of the pursuit of similar themes in contemporary psychology, such as those concentrating on the mechanism – well-known to politicians, priests and gurus – underlying the surprisingly contagious nature of other peoples’ avowals or confidence. Our surprising degree of gullibility, or ‘cognitive conformity’, held no surprises for Hume.

The above is but a small sample of the various phenomena of human nature that Hume considers. But the pattern is always the same: naturalistic, economical, and based in a strong sense of what proves adaptive and useful, and what the reverse. Here too we might mention that where the ecological setting is lacking, Hume is utterly scornful of any attempt to impute a psychology. Thus the quotation at the head of the chapter is part of a passage mocking the religious believer's attribution of human emotions to the deity. For our sentiments 'are calculated for preserving the existence and promoting the activity' of ourselves in our situations. Since the monotheist's deity doesn't inhabit any particular natural or social niche, Hume drily points out that 'it seems unreasonable to transfer such sentiments to a supreme existence or to suppose him actuated by them'.

## VI HUME'S DARWINIAN GENEALOGIES

The fourth and in some ways the most telling example of incipient Darwinism is Hume's genealogical account of justice and obligation. First, a caveat. Hume puts under 'justice' some quite specific things, notably the obligations associated with the ideas of property, promising, and government or law. He is not concerned with immediate perceptions of fairness, of the kind that excite small children anxious about their share of the cake, or even capuchin monkeys, outraged at being rewarded less for the same task for which a fellow monkey was rewarded more.<sup>20</sup> Hume would have classified such cases with natural desire or aversion, triggered by the perception of a contemporary getting more in return for doing the same or less. This divergence does not detract from the interest of

Hume's account, and for that matter its economic and political importance. But the structures Hume is interested in are more complex. He is interested in cases where a person is motivated to do something initially disadvantageous, but which turns into an advantage only on the expectation that another plays their desired part in the enterprise. Or, as he puts it, he is interested in structures that are not like heaps, in which each stone adds its bulk regardless, but like arches, in which each stone plays its supportive role only on the condition that the other stones do.<sup>21</sup>

The outline is well-known and I shall not rehearse it here: it has frequently been acknowledged in modern writings on the evolution of convention and cooperation.<sup>22</sup> One of Hume's own summaries is admirably brief:

All *moral* duties may be divided into two kinds. The *first* are those, to which men are impelled by a natural instinct or immediate propensity, which operates on them, independent of all ideas of obligation, and of all views, either to public or private utility. Of this nature are, love of children, gratitude to benefactors, pity to the unfortunate. When we reflect on the advantage, which results to society from such humane instincts, we pay them the just tribute of moral approbation and esteem: But the person, actuated by them, feels their power and influence, antecedent to any such reflection.

The *second* kind of moral duties are such as are not supported by any original instinct of nature, but are performed entirely from a sense of obligation, when we consider the necessities of human society, and the impossibility of supporting it, if these duties were neglected. It is thus *justice* or a regard to the property of others, *fidelity* or the observance of promises, become obligatory, and acquire an authority over mankind. For as it is evident, that every man loves himself better than any other person, he is naturally impelled to extend his acquisitions as much as possible; and nothing can restrain him in

this propensity, but reflection and experience, by which he learns the pernicious effects of that licence, and the total dissolution of society which must ensue from it. His original inclination, therefore, or instinct, is here checked and restrained by a subsequent judgment or observation.<sup>23</sup>

From the Darwinian point of view there is a crucial point concealed in this passage, and in Hume's longer treatments of the issue. Hume presents the obligations of justice in respect of property and fidelity in keeping promises as 'artificial', or the results of the emergence of convention. But he also presents that emergence as the result of 'judgment or observation'. It is with 'reflection and experience' that we come to want to do better than any pre-social, non-cooperative version of a war of all against all.

We might want to query the introduction of ratiocination at this point. After all, Hume explicitly says that it is by the same mechanism that other pillars of social life emerge, and he cites the structures of language and money.<sup>24</sup> Yet he might have reflected that prototypes of language, in the signalling systems of other animals, emerge without reflection and ratiocination. There are even signals that are the functional equivalent of promises, such as the 'canid bow' that dogs make to signal non-aggression or the intention to play, and whose misuse, in some species, can cause severe repercussions, such as exclusion from the pack.<sup>25</sup> So how should we think of the contrasting versions? On either version of the story, it is the inconvenience of the war of all against all compared to the overwhelming advantage of cooperative stability that is the engine driving the emergence of convention. On Hume's version the engine works because we are aware of these things. We have a sense of common interest in doing better. On the other version it would work by less cognitive mechanisms. If we suppose a small random

change, a ‘trembling hand’ inducing some cooperation somewhere, then whether it is recognized or not, the advantage this would give would be in principle ready to drive the usual Darwinian dynamic.

Hume’s version has certain advantages. He is clear that in conforming ourselves to these conventions we need a motive to restrain our immediate self-interest. The only motive can be a conditional expectation: the confidence that *if* we do our part, others will do theirs, but that *if* we do not, then they will not either. And, at least for human beings, it is not easy to imagine the kind of behaviour generated by *this* kind of confidence happening without that confidence being cognitively engendered and sustained. In actual practice, of course, it is; and when things are working properly—when the agent ‘intends to live on good terms with mankind’—the disposition is then cemented into place by the shame and fear of contempt that attend even the thought of transgression.

Is the Humean story Darwinian in essence? There are good reasons to think so. Firstly, given an initially non-cooperative situation, it might still take a small random event, a trembling hand, to initiate the first faltering steps towards any cooperative equilibria. And then, secondly, our awarenesses and cognitive functioning make up just one part of our animal natures; so, from an abstract point of view, natural selection due to the operation of these abilities is just another case of natural selection. And Hume is clear that it takes time: the rule of stability of property, for instance, ‘arises gradually, and acquires force by a slow progression, and by our repeated experience of the inconvenience of transgressing it’<sup>26</sup>. Finally we should notice that Hume’s genealogy escapes the charge made against the promiscuous use of the notion of a ‘meme’. The advantages he is talking about are the stability of property and the fidelity given to

promises; and it is very clear indeed why any unfortunate hominids, or societies, who could not manage either would be poor rabbles, set to do badly in any competition with us, who can manage both.

Unlike Nietzsche's more notorious genealogy of morals, Hume's is a story that can leave us quite satisfied with ourselves.<sup>27</sup> It is the reverse of a debunking story. It shows humans having a problem and solving it. Or, if we prefer, it shows an adaptation—a set of traits that are advantageous, and that are heritable in the sense in which epigenetic features are so. The children of a culture in which norms of cooperation are entrenched will be much more likely to grow into cooperative adults than children who start in a world of the war of all against all. On the other hand, as with female chastity, we must be clear that it is a political and social achievement to sustain such a culture. We cannot rely on our natures to do it for us, as the unhappy descents into bellicose equilibria of many parts of the world show.

We might therefore justly call Hume's account Darwinian. And we might want to prefer it to recent 'modular' evolutionary psychologies, which claim that there is an inbuilt moral faculty that determines our responses to each other's actions, in much the same way that a native universal grammar is postulated to underlie our linguistic skills.<sup>28</sup> If the idea of modularity implies that the outputs of this 'module' are in themselves impermeable, like the perception of what we know to be visual illusions, unaffected by what else we think, perceive, or say, then I see little evidence for it in the human world.<sup>29</sup> Unlike our sensitivity to the syntax of our mother tongues, our ethics are relentlessly subjects of reflection, conversation and persuasion, and discursive pressure changes peoples' minds. Furthermore, if the outputs of the module are supposed to be

motivations, rather than inert judgments, then the catastrophic equilibria already mentioned stand as counterexamples. A person finding that others do not reciprocate or play their role is quick enough to think of himself as, in Hume's words, a 'cully of his integrity', and change his behaviour accordingly.<sup>30</sup> Sadly, enough people are not motivated to behave remotely decently in any case for us to doubt whether decency sits in our genes, even in whichever way language does. Finally, of course, stories about modules are only stopgaps from Darwin's perspective, since they throw the issue of understanding our moral capacities back onto the evolutionary story behind such a thing, which in effect will reintroduce the general mechanisms which more economically do the same job.

#### VII CONCLUSION: NATURE HUMANIZED (AND HUMEANIZED)

There is, I hope, a final moral to all this drawing out of affinities between the Humean and Darwinian programmes. I began by worrying about the very concept of human nature. With such worries around, we may lose much appetite for the familiar nature-versus-nurture dialectic. But I take it the genealogies I have been sketching also help to kick any simple opposition into touch. It is natural to us to build conventions, just as it is natural to build shelters, wear clothes or support roofs with arches. Hume saw this as well, and by the time of he published his *Enquiry Concerning the Principles of Morals* (1751), he had himself substantially abandoned the earlier vocabulary of 'natural' versus 'artificial' virtues, while still giving the same theory of the same structures of convention. Whatever may have been the case among our ancestors back in the Pleistocene and beyond, today, among adult humans in societies that work, artifice lies in our natures.

---

<sup>1</sup> See Flanagan, this volume.

<sup>2</sup> For some of the complexities here, see Sober 1980.

<sup>3</sup> Weaver et al. 2004.

<sup>4</sup> Quoted in ‘DNA is Not Destiny’, *Discover*, Nov. 2006.

<sup>5</sup> Dawkins 2003, p. 81.

<sup>6</sup> It may be a little churlish but I think appropriate to raise the same complaint against Daniel Dennett’s substitution of ‘cranes’ for ‘skyhooks’; see Dennett, this volume. Cranes may stand on the ground, but they also tug in a particular direction, whereas the great feature of evolutionary theory is that living things grow, as indeed do mountains, without being tugged in any direction at all. They chug along, with more of some kinds and fewer of others over time.

<sup>7</sup> For an introduction to game theory in evolutionary discussions of morality, see Rosenberg, this volume.

<sup>8</sup> Notice that if a defector can disguise his true nature and skip from table to table, he can still do best of all, as crime in cyberspace illustrates.

<sup>9</sup> Sober and Wilson 1998.

<sup>10</sup> See Fodor 2007. Pained correspondence was continued in the *London Review of Books*, by Philip Kitcher, Stephen Rose, Jerry Coyne, Tim Lewens and myself.

<sup>11</sup> Parker 2003.

<sup>12</sup> Darwin in his notebook period was an enthusiastic reader of Hume, which may go some way to account for the congruities described here between Darwinian and Humean approaches. On Darwin’s reading of Hume, see Richards 1987, 106, 109.

<sup>13</sup> Hume 1748, Section XII, para 22, p. 206.

<sup>14</sup> Hume 1748, Section IX, para 6, p. 168.

<sup>15</sup> Hume 1739, II.iii.3, p. 415.

<sup>16</sup> Hume 1751, Section VI, part 1, p. 125.

<sup>17</sup> Hume 1985.

<sup>18</sup> Montesquieu 1748.

<sup>19</sup> Hume 1739, III.ii.12. For a fuller account see Blackburn 2004, ch. 13.

---

<sup>20</sup> Brosnan and de Waal, 2003.

<sup>21</sup> Hume 1751, Appendix 3, p. 171.

<sup>22</sup> For example, David Lewis 1969, p.3, allows that the theory of games, which provides the technical heart of his book, is but ‘scaffolding’, and the theory of convention that emerges is in effect that of Hume.

<sup>23</sup> Hume 1985 (‘Of the Original Contract’).

<sup>24</sup> Hume 1739, III.ii.2, p. 490; Hume 1751, Appendix 3, p. 172.

<sup>25</sup> Bekoff 1974, 1977.

<sup>26</sup> Hume, 1739, III.ii.2, p. 490

<sup>27</sup> Nietzsche 1887.

<sup>28</sup> Hauser 2006.

<sup>29</sup> See also Sterelny (forthcoming).

<sup>30</sup> Hume 1739, III.ii.7, p. 535.

[For the list of references:]

Bekoff, Marc. 1977. Social communication in canids: evidence for the evolution of a stereotyped mammalian display. *Science* 197: 1097 – 1099.

Bekoff, Marc. 1974. Social play in coyotes, wolves, and dogs. *BioScience* 24(4):225-230.

Blackburn, Simon. 2004. *Lust*. New York: Oxford University Press

Brosnan, Sarah F. and de Waal, Frans B. M. 2003. Monkeys reject unequal pay. *Nature* 425: 297-299.

Dawkins, Richard. 2003. *A Devil’s Chaplain*, Boston & New York: Houghton Mifflin.

---

Fodor, Jerry. 2007. 'Why Pigs Don't Have Wings', *London Review of Books*, 18 Oct.

Hauser, Marc. 2006. *Moral Minds: How Nature Designed our Universal Sense of Right and Wrong*. New York: Ecco.

Hume, David. 1739. *A Treatise of Human Nature*, ed. L. A. Selby-Bigge. Oxford: Oxford University Press, 1964.

Hume, David. 1748. *An Enquiry Concerning Human Understanding*, ed. Tom L. Beauchamp, Oxford: Oxford University Press, 1999.

Hume, David. 1751. *An Enquiry Concerning the Principles of Morals*, ed. Tom L. Beauchamp. Oxford: Oxford University Press, 1998.

Hume, David. 1779. *Dialogues Concerning Natural Religion*, ed. Dorothy Coleman, Cambridge: Cambridge University Press, 2006.

Hume, David. 1985. *Essays, Moral, Political, and Literary*, ed. Eugene F. Miller, Indianapolis: Liberty Classics.

Lewis, David. 1969. *Convention*. Cambridge, MA: Harvard University Press,

Montesquieu, Charles-Louis de Secondat, baron de. 1748. *De l'esprit des lois*. translated and edited by Anne M. Cohler, Basia Carolyn Miller, Harold Samuel Stone, Cambridge: Cambridge University Press, 1989

Nietzsche, Friedrich. 1887. *The Genealogy of Morals*. trans. Douglas Smith, Oxford: Worlds Classics, 1998.

---

Parker, Andrew. 2003. *In the Blink of an Eye*, Cambridge, Mass: Perseus Publishing.

Sober, Elliott. 1980. 'Evolution, Population Thinking, and Essentialism', *Philosophy of Science*, Vol. 47, No. 3, pp. 350-383

Sober, Elliott and Wilson, David Sloan. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behaviour*, Cambridge, MA: Harvard University Press.

Sterelny, Kim (forthcoming) 'Moral Nativism: A Sceptical Response'.

Weaver, Ian C. G. et al. 2004. 'Epigenetic programming by maternal behavior', *Nature Neuroscience* 7, 847 – 854.