

3

Universal Darwinism and Human History*

David Christian

Abstract

This essay discusses Universal Darwinism: the idea that Darwinian mechanisms can explain interesting evolutionary change in many different domains, in both the Humanities and the Natural Sciences. The idea should appeal to Big Historians because it links research into evolutionary change at many different scales. But the detailed workings of Universal Darwinism vary as it drives different vehicles, just as internal combustion engines differ in chain-saws, motor cycles and airplane engines. To extend Darwin's ideas beyond the biological realm, we must disentangle the biological version of the Darwinian mechanism from several other forms. The paper focuses particularly on Universal Darwinism as a form of learning, a way of accumulating information. This will make it easier to make the adjustments needed to explore Darwinian mechanisms in human history.

Keywords: *Universal Darwinism, collective learning, information, Big History.*

Countlessness of livestories have nether-fallen by this plague, flick as flowflakes, litters from aloft, like a waast wizard all of whirl-worlds. Now are all tombed to the mound, isges to isges, erde from erde.

Finnegans Wake, Ch. 1

James Joyce's strange masterpiece, *Finnegans Wake*, is fractal. You can read it at many different scales, but you always have the eerie feeling that you are hearing a story you have already heard somewhere else. A mathematician might say the stories are 'self-similar'. You may think you are reading about the wake for a drunken bricklayer who fell to his death from a ladder; but you are actually reading about the fall of humanity and the expulsion from Paradise; and then again the story is *really* about Dublin and the many rises and falls of that city's history, people and landscapes. Something similar happens in the

* My thanks to David Baker, Billy Grassie, Nick Doumanis, Ji-Hyung Cho, and Seohyung Kim for reading suggestions and comments on earlier versions of this paper.

emerging discipline of Big History (see Christian 2004, 2010). Big History surveys the past at the scales of cosmology, physics, geology, biology and human history. Each discipline tells its own story, but as you get to know the stories, they start to overlap, and we begin to see each discipline refracted in the others. Like *Finnegans Wake*, Big History is ‘self-similar’. And like *Finnegans Wake*, Big History derives much of its power from the synergies that arise when you glimpse unexpected connections across different scales and domains.

This paper explores one of these fractal phenomena: ‘Universal Darwinism’. In biology, the Darwinian paradigm describes a distinctive form of evolutionary change that generates adaptive change through repeated copying of selected variants. Universal Darwinism is the idea that similar mechanisms may also work in many other domains. If so, do they always work as they do in biology? Or can we distinguish between a core machinery and the modifications needed to drive it in different environments?

Universal Darwinism

Richard Dawkins coined the phrase ‘Universal Darwinism’ in an essay published in 1983. If we find life beyond this earth, he argued, it will surely evolve by ‘the principles of Darwinism’ (Dawkins 1983: 403). But there will also be differences. For example, the replicators may not be genes. Dawkins suggested that human culture might offer an example in the ‘meme’, an idea or cultural artifact such as a song or fashion that varies, that replicates through imitation, that travels in sound or images, and colonizes human minds when selected from a population of rival artifacts (on meme theory see Blackmore 1999). More generally, he suggested that, ‘Whenever conditions arise in which a new kind of replicator *can* make copies of itself, the new replicators *will* tend to take over, and start a new kind of evolution of their own’ (Dawkins 2006: 193–194).

Universal Darwinism treats natural selection as one member of a family of evolutionary machines that generate adaptive change through repetitive, algorithmic processes. Always we see variation, selection and replication. Some variations are selected, then copied and preserved with slight modifications, after which the process repeats again and again.

Here is a description of the basic machinery by a physicist, Lee Smolin,

To apply natural selection to a population, there must be:

- a space of parameters for each entity, such as the genes or the phenotypes;
- a mechanism of reproduction;
- a mechanism for those parameters to change, but slightly, from parent to child;

- differentiation, in that reproductive success strongly depends on the parameters (Smolin 2005: 34).

And here, to illustrate slight variations in our understanding of the basic machinery, is a description by a psychologist, Susan Blackmore:

Darwin's argument requires three main features: variation, selection and retention (or heredity). That is, first there must be variation so that not all creatures are identical. Second, there must be an environment in which not all the creatures can survive and some varieties do better than others. Third, there must be some process by which offspring inherit characteristics from their parents. If all these three are in place then any characteristics that are positively useful for survival in that environment must tend to increase (Blackmore 1999: 10–11).

Repeated many times, these simple rules yield interesting evolutionary change. Variation creates diversity, but by selecting some variations over others you steer diversification in a particular direction. You ensure that surviving variations will fit the environment that selected them, so they will be 'adapted'. In this way, the Darwinian machinery steers change away from the random mush ordained by entropy and the second law of thermodynamics. And if by chance some selected variants are slightly more complex than others, then we have, in Universal Darwinism, a way of increasing complexity. Indeed, Lee Smolin argues that natural selection provides the *only* scientific way to explain how complexity can increase against the tide of entropy (Smolin 2005: 34). (As I write this paper, I watch myself selecting some ideas, words and metaphors, and rejecting others; and I know that eventually the paper itself will have to take its chances in a competitive world populated by many other academic papers.)

So powerfully does the Darwinian machinery steer biological change that many find it hard to avoid imagining that there must be a designer. Surely, organs as beautifully designed as wings or brains must have been, well, designed! Yet natural selection needs no cosmic project manager. This is what Daniel Dennett called 'Darwin's Dangerous Idea': operating without purpose, the Darwinian algorithm creates the appearance of purposefulness (Dennett 1995). From camels to chameleons, species fit their environments so precisely that they seem to transcend the laws of entropy. Yet they need no teleology and no driver. Darwin's ideas threatened theism because they explained the appearance of direction without needing a divine director (*Ibid.*).

Even in Darwin's time, some wondered if the same machinery could work outside the domain of biology. In a section on language in Chapter 3 of *The Descent of Man*, Darwin wondered if languages evolved like living organisms. After all, he noted, languages vary, they are reproduced, and their components – words, grammatical forms and even particular languages – are sub-

ject to selection for their 'inherent virtue'. Darwin concluded that, 'The survival or preservation of certain favoured words in the struggle for existence is natural selection' (Darwin 1989: 95). Darwin's friend, Thomas H. Huxley, suggested that there might be evolutionary competition between different bodily organs, while William James extended the idea of evolution to learning in general (Plotkin 1994: 61–64).

But it was in biology that Darwin's ideas really triumphed. In the 1930s and 1940s, several lines of research converged in the 'neo-Darwinian synthesis', which fixed several weaknesses in Darwin's original theory. For example, Darwin assumed that inheritance was blended, an idea that threatened to eliminate successful variations by driving all variation towards a mean; Darwin also feared that natural selection worked too slowly to generate today's biodiversity, particularly on a planet he believed to be less than 100 million years old. The neo-Darwinian synthesis used the work of Gregor Mendel to show that inheritance works not by blending but by copying discrete alleles. August Weismann showed the importance of distinguishing between phenotype and genotype, between characteristics acquired during an organism's lifetime, and those inherited through the germ line, which ruled out intentional or 'Lamarckian' forms of evolution; and this suggested that genetic mutations had to be random rather than purposeful. Finally, population geneticists such as Ronald A. Fisher and John B. S. Haldane proved mathematically that successful genes could spread fast enough to generate all the variety we see today, and geologists showed that the earth was almost 50 times older than Darwin had supposed (Mesoudi 2011: 40–51). Just as James Watt's modified steam engine made it industry's standard prime mover, so the neo-Darwinian synthesis turned Darwinism into biology's standard explanation for biological change. The discovery of DNA and the evolution of genetic research consolidated Darwinism's paradigm role within biology.

Paradoxically, the success of the neo-Darwinian synthesis inhibited its use in other fields by creating the impression that all Darwinian machines had to be neo-Darwinian. Replicators had to be particulate; they had to be distinct from the entities in which they were tested (phenotypes or bodies); and variation had to arise randomly. Outside of biology, the neo-Darwinian model worked much less well than it did within biology. Historians and social scientists resisted Darwinian models for another reason: applied carelessly or too rigidly, they seemed to encourage Social Darwinism. The idea of Social Darwinism attracted scholarly attention after the publication of Richard Hofstadter's, *Social Darwinism in American Thought*, in 1944 (Hofstadter 1944). For Hofstadter, Social Darwinism's primary meaning was 'biologically derived social speculation'; but others associated it more closely with racist theories, though even Hofstadter had warned that '[Darwinism] was a neutral instrument, capable of supporting opposite ideologies' (Leonard 2009: 41–48). These fears helped pre-

serve the gulf between the humanities and the natural sciences that Charles P. Snow bemoaned more than 50 years ago (Snow 1959).

In the late twentieth century, scholars in several fields returned to modified Darwinian models of change. They found them at work in immunology, in economics, in the history of science and technology, and even in cosmology, where Lee Smolin has proposed a theory of 'cosmological natural selection' (Smolin 1998; Nelson 2006; Campbell 2011). In Smolin's model, new universes are born in black holes. Information about how to construct universes resides in basic physical parameters, such as the power of gravity. Reproduction generates variation because daughter universes may inherit slightly different parameters. Variations are 'selected' and preserved because they will survive only if they generate universes complex enough to form black holes and reproduce. So cosmological natural selection does not generate a random mix of universes, but only those universes with just the parameters needed to create complexity. Our own existence proves that some universes will be complex enough to yield planetary systems, and life and creatures like us. Here we have a Darwinian explanation for the existence of a universe such as ours whose parameters seem exquisitely tuned for complexity.

Wojciech Zurek and his colleagues at the Los Alamos National Laboratory have even detected Darwinian mechanisms in quantum physics (Campbell 2011: 89ff.). When a quantum system interacts with another system, perhaps by being measured in a lab, just one of its many possible outcomes is selected and launched into the world, in the process known as 'decoherence'. We have variability of the initial possibilities, a selection from those possibilities, and a copying of the selected possibilities from the quantum to the non-quantum domain. 'This Darwinian process allows a quantum system to probe its environment searching for and selecting the optimal low entropy states from all those available, thus allowing greater complexity to be discovered and survive' (*Ibid.*: 154). (The author of this paper makes no claim to understand these processes except in the most superficial way. The point is that Darwinian mechanisms may be at work even at the quantum level.)

Darwinian ideas have also returned to the humanities and social sciences, attracting the attention of anthropologists, linguists, psychologists, game theorists and some economists, political scientists and historians of technology (see Mesoudi 2011 on cultural evolution; Fitch 2010 on language origins and Nelson 2007: 74 on Darwinian models in other fields). Such explorations may get easier because the neo-Darwinian synthesis is loosening its grip within the core territory of biology. When the human genome was deciphered in 2003, it turned out that humans have far fewer genes for the manufacture of proteins than had been expected, little more than 20,000, fewer than in the rice genome. This discovery reminded biologists and geneticists that DNA is not a lone autocrat; it rules through a huge biochemical bureaucracy, whose agents often manage

their ruler, as civil servants manage politicians. Mechanisms within cells control how and when the information in DNA is expressed, and occasionally they even alter DNA itself, if only to repair it. Even more striking, some of these changes seem to be heritable. Through this modest backdoor, Lamarckian inheritance is creeping back into biological thought. In a recent survey of these changes, Jablonka and Lamb write that ‘there is more to heredity than genes; some hereditary variations are nonrandom in origin; some acquired information is inherited; evolutionary change can result from instruction as well as selection’ (Jablonka and Lamb 2005: 1).

These debates within biology may help us stand back from the biological form of the Darwinian machinery and see how different variants work in other realms, including human history.

Information and Universal Darwinism

Darwinian machines run on information: they replicate patterns, and that means replicating *information* about those patterns. So to understand their general properties, we need the idea of information. But information is a mysterious and ghostly substance that sometimes appears to float above reality, so we must define it carefully (accessible surveys include Floridi 2010; Gleick 2011; Lloyd 2007; Seife 2007).

The idea of information presupposes the existence of differences that matter. To an antelope it matters if the animal behind the tree is a tiger or another antelope. Information reduces uncertainty by selecting one of several possible realities. This is why Donald MacKay described information as ‘a distinction that makes a difference’ (Floridi 2010: 23). A difference matters if other entities can detect and react to it. They may be able to detect it directly; but if not, they can often detect it indirectly, by secondary differences that correlate with the initial difference. This is where information steps in. When two differences are correlated, the second can carry a message from the first to any receiver able to interpret the message. In this way, causal chains carry potential information, whether or not there is a mind at the end of the chain. An antelope may detect a nearby lion by its shadow, and that should remove uncertainty about the danger. Run! But an electron can also be said to detect and react to a proton through its electric charge. Inserting a conscious entity into the chain simply adds one more link. It may add uncertainty, but all links do that. In this way information can travel along causal chains because we infer differences that are hard to detect from others that are easier to detect. Information is embedded in chains of cause and effect. ‘[It] is not a disembodied abstract entity; it is always tied to a physical representation. It is represented by an engraving on a stone tablet, a spin, a charge, a hole in a punched card, a mark on paper, or some other equivalent’ (Rolf Landauer, cited in Seife 2007: 86).

When information travels through long causal chains, it can lose precision. The second, and third and fourth differences are not, after all, the same as the first. So we can judge a message by how well it represents the original difference. Faulty genes trick cells into making cancer cells, and an antelope can take a trick of the light for a tiger's shadow. But some chains transmit information more efficiently than others. As a general rule, digital or particulate information carriers detect differences better than continuous or analogue carriers, because they *have* to discriminate. That is why DNA employs genes, languages use words, and computers prefer on/off switches. Effective transmission systems can partition the smoothest of changes.

We can also judge a transmission system by the amount of information it carries. Claude Shannon, the founder of 'Information theory', showed that information increases precision by reducing uncertainty (Floridi 2010: 37ff.). You can measure the amount of information in a message by the number of alternative realities it excludes. 'There is a tiger behind the bush' is helpful advice; it reduces uncertainty. But if a friend adds that the tiger is hungry and in a bad mood, that should eliminate any doubts you had about running away. If, from all the possible things that might have happened, a message selects a tiny, not-easily-predicted sub-set, then it eliminates a vast number of other possibilities and a huge amount of uncertainty. Each rung on a molecule of DNA can exclude three out of four possible futures; so the entire molecule, with billions of rungs, can exclude a near infinity of possible creatures. It tells you how to build just one, say, an armadillo. Not an amoeba, or an archaeopteryx, but an armadillo. In information theory, 'the amount of information conveyed by [a] message increases as the amount of uncertainty as to what message actually will be produced becomes greater' (Pierce 1980, Kindle edition, location 461).

We have seen that information does not need minds. However, words like 'meaning' make sense only when the causal chain *does* include a mind. Only then can we describe information as *semantic*. And when the information is complex it makes sense to call it knowledge. Luciano Floridi writes,

Knowledge and information are members of the same conceptual family. What the former enjoys and the latter lacks ... is the web of mutual relations that allow one part of it to account for another. Shatter that, and you are left with a pile of truths or a random list of bits of information that cannot help to make sense of the reality they seek to address. Build or reconstruct that network of relations, and information starts providing that overall view of the world which we associate with the best of our epistemic efforts (Floridi 2010: 51).

We needed this digression on information because Universal Darwinism builds complexity by accumulating, storing and disseminating information about how to make things that work. Darwinian machines generate unexpected

outcomes, like armadillos or human brains, because they accumulate information that is *not* entropic mush. So wherever they are at work, unexpected things happen – whether in the immune system or in DNA, or in human history or entire universes (Blackmore 1999: 15). Darwinian machines learn (a classic summary is Campbell 1960: 380). This is why Karl Popper described the growth of knowledge as: ‘the result of a process closely resembling what Darwin called “natural selection”, that is, *the natural selection of hypotheses*: our knowledge consists, at every moment, of those hypotheses which have shown their (comparative) fitness by surviving so far in their struggle for existence’ (Plotkin 1994: 69).

Three Darwinian Learning Machines

Seeing Darwinian machines as learning machines will help us understand how they may shape human history. On this planet, living organisms learn in three distinct ways. All are Darwinian, but they use different variants of the same basic engine.

Genetic Learning and Natural Selection. The first variant is natural selection. Biologists have studied this engine for a long time and they understand it well. It explains how molecules of DNA accumulate adaptively significant information. DNA codes information about how to manufacture proteins using four nitrogenous ‘bases’: Adenine, Thymine, Guanine and Cytosine. Differences in the order of the letters really matter. Exchange one A for a T in the code for a protein with 146 different amino acids and you get sickle cell anemia. DNA stores information that is rich because it is specific, impossible to generate randomly, and therefore it is unexpected. Over time, billions of new genetic recipes for building proteins and whole organisms accumulated in the world's stock of DNA to generate the species we see today.

Generation by generation, packets of DNA are sieved as their products enter the world. Mutations, copying errors and recombination during reproduction create random variations in genes and in the organisms they give rise to, so that slight modifications on the original instructions are continually being tested. Only those packages that produce viable organisms will survive and reproduce. Much of the information they contain tells cells how to choose the tiny number of biochemical pathways that resist entropy. For example, it may include recipes for enzymes that steer biochemical reactions along rare but efficient pathways, or that help export entropy outside the organism (Campbell 2011: 102). In each generation, that information can be updated. This explains why living organisms have an uncanny ability to track changing environments.

DNA preserves information because it acts like a ratchet (on the ‘ratchet effect’ in human history, see Tomasello 1999). Mechanical ratchets allow a gear-wheel to turn in only one direction because the ‘pawl’ catches on the cogs and prevents the wheel from turning backwards. By only copying infor-

mation that works, DNA ensures that the gear wheel of evolution normally turns in the direction that accumulates viable variations. Without an information ratchet, the wheel of evolution could turn in either direction, viable variations would survive no better than any others, and biological change would drift with the flow of entropy. That is why it makes sense to suppose that life itself began with DNA or its predecessor, RNA. Before the evolution of DNA or RNA, parts of the Darwinian machine already existed: there was plenty of variation within pre-biotic chemistry, and variations could be selected for their greater stability. But only after DNA evolved (possibly preceded by RNA) could successful variations be locked in place so that genetic information could accumulate. With DNA preventing any backsliding, life was off and running.

To summarize key features of genetic learning: information accumulates as it is locked into the biochemical structures of DNA molecules. Most variations arise randomly during reproduction. Variations survive only if the DNA molecules they inhabit are copied. Genes are particulate, but when working together, they can create the impression of a 'blending' of characteristics. Because most variation arises during reproduction, genetic learning is non-Lamarckian; it does not preserve 'acquired variations', variations generated during an individual's lifetime. Random variations are tested, one by one, surviving only if they create organisms that fit their environment. These are the rules of the neo-Darwinian synthesis.

Individual Learning. The other two forms of learning have been studied less closely than the genetic machine, and we do not understand them as well.

I will call the second machine 'individual learning'. It works not across species or organisms but within the neurological system of a single individual. It is at work in species as varied as cephalopods, crows and chimpanzees. It works even in simple organisms, which can learn to detect and react to gradients of light or warmth or acidity. But individual learning is most impressive in animals with brains. Imagine our antelope glimpsing a lion near a waterhole. Was that really a lion? Should it make for another waterhole? With no guidance, it might have to choose randomly, as young animals often do. It will soon find out if its gamble succeeded. But intelligent animals also have better ways of choosing. They accumulate memories of past experiences associated with pain, fear, anxiety or with a sense of pleasure and ease. If any of those memories are similar to what is happening right now, they may provide guidance. Trying out possibilities in memory is less dangerous than trying them out in the real world, and the accompanying sensations, installed over time by genetic learning, will provide better than random criteria for repeating or avoiding particular experiences. Alasdair MacIntyre reports that if a young cat catches a shrew, it will eat it as if it were a mouse. It will then become violently ill, which is an unpleasant experience. But it has learnt a difference that matters and from now on it will avoid

shrews (MacIntyre 2001: 37). A memory that should help the cat survive has out-competed a memory that once caused it misery.

Put more generally, an intelligent organism undergoes experiences that carry information about the outside world, if they can be stored and interpreted. Memory provides an information ratchet as it encodes experiences in neurological networks. It accumulates useful information within an individual's lifetime. Faced with an important choice, the organism can refer to its memory bank and look for experiences that had happy or unhappy outcomes. As it replays memories with their associated experiences of pleasure or pain or fear or comfort, it learns to make better choices. Significant memories are selected by being reinforced (through repetition or association with other strong experiences), while memories that are not reinforced will fade away (Campbell 2011: 119–120). The criteria for selection – repeated reinforcement or strong association with experiences of pain or pleasure – will have been built into the organism by genetic learning, which teaches you to cherish parents and shun predators. Here we have the complete Darwinian cast: varied experiences that are encoded in memories, only some of which are selected for preservation.

So individual learning is a Darwinian machine. But it does not work quite like the machinery of the neo-Darwinian synthesis. Its arena is the individual brain, rather than the outer world. Individual learning preserves useful memories acquired during an individual's lifetime, but those memories can also change; unlike genes, memories are not fixed from the moment of their birth. So individual learning can be Lamarckian. It contains no simple analogue to the neo-Darwinian separation of genotype (which does not change during an individual's lifetime) and phenotype (which can change within a lifetime). Variation arises mainly from the diversity of individual life experiences, though some may arise from mistakes in coding or assessing those experiences. In individual learning, the primary information carriers are neurological networks, and memories, their psychological correlate. Both are more diffuse and variable than genes and subject to constant minor changes as they join or separate from other networks and memories. Selection occurs through reinforcement rather than reproduction, as networks are selected for their strength and connectedness, which depend on the number and strength of the synapses from which they are constructed. Networks that are reinforced strongly because they are repeated often ('*that* waterhole is safe') or are particularly shocking ('nearly got caught that time!'), will survive, while the rest will dwindle and fade. The criteria for selection do not reside in the outer environment, but are built into the organism by genetic learning. But selection is not purely mechanical. Sometimes it demands a judgment call '*that* waterhole is safe but the water does not taste as good, Hmmm'. At this point we may conclude that animals ponder alternatives before selecting consciously and with intent. Selection is beginning to look purposeful.

So here we have a Darwinian machine that lacks the bells and whistles of the neo-Darwinian synthesis but can still generate new, non-random and significant information. It also sports some glossy new features. It is very fast; it can accumulate new information in seconds, while genetic learning gets to test new variations just once in a lifetime. Individual learning is also specific; instead of producing generic adaptive rules for millions of individuals, it tells a particular individual how to live in a particular time and niche. But individual learning is also ephemeral; it cannot survive outside the arena of the individual brain. A lifetime of learning evaporates on the death of each individual, so every generation starts from scratch. Individual learning is Sisyphean; it cannot accumulate information at time scales larger than a lifetime, so it does not lead to a long-term change. That is why it cannot generate what we humans call 'history'; change at scales larger than a single lifetime.

Darwinian Machines in Human History: Collective Learning

Our third Darwinian machine *does* generate long-term change. I call it 'collective learning', and it seems to be unique to our species, *Homo sapiens* (for brief discussions see Christian 2004, 2012).

Collective learning happens when you join individual learning to a sufficiently powerful system of communication. It depends on the ability of individual learners to share what they have learned with others, and to do so in such volume and with such precision that new information accumulates at the level of the community and even the species. As Merlin Donald writes, 'The key to understanding the human intellect is not so much the design of the individual brain as the synergy of many brains' (Donald 2001: xiii).

Collective learning uses a new and more powerful information ratchet. Unlike individual learning, it stores information in many minds over many generations, so that information can outlive the individuals who created it. If a fraction of that information improves how individuals exploit their environments, collective learning will tend to increase the ecological power of whole communities. Like all animals, humans exploit their environments to extract the energy and resources they need to survive; but only humans keep discovering and sharing new ways of exploiting their environment, so that over time they can extract more and more energy and resources. Our ecological creativity explains why humans are the only species that has a history of long-term changes in behaviours, social structures and ecological adaptations. Like individual learning, collective learning also works much faster than genetic learning. That is why, within just a few hundred thousand years we have become more powerful than any single species in the 3.8 billion year history of life on earth, so powerful that some geologists argue we have entered a new geological epoch, the 'Anthropocene' (see Steffen *et al.* 2007).

By sharing ideas, information, gossip and beliefs, collective learning creates human ‘culture’, which Mesoudi defines broadly as ‘information that is acquired from other individuals via social transmission mechanisms such as imitation, teaching, or language’ (Mesoudi 2011: 2–3; for a similar definition see Distin 2011: 11). Of course, humans are not alone in having ‘culture’ in this sense. Songbirds, chimps and whales all share information. The difference is in the degree of sharing, but that small difference really matters. Animal languages lack an efficient information ratchet, so in the animal versions of ‘telephone’, information leaks away within a few exchanges and has to be constantly relearned. This is why knowledge accumulation has little impact on any species except ours, and that is why no other species has a history of long-term change over many generations. Alex Mesoudi sums up a broad consensus among those who study animal culture:

Although numerous species exhibit one-to-one social learning and regional cultural traditions, no species other than humans appears to exhibit cumulative culture, where increasingly effective modifications are gradually accumulated over successive generations. This might therefore be described as the defining characteristic of human culture (Mesoudi 2011: 203).

There is a narrow but critical threshold between individual and collective learning. To appreciate its significance, imagine pouring water into a bathtub with no plug. A trickle of water will deposit a thin film at the bottom of the bathtub. But the level will not rise because water leaks away as fast as it pours in. Increase the flow and the water level will rise and settle at a new level. (We see something like this in species such as *Homo erectus*, or in some species of primates.) Increase the flow just a bit more and suddenly the level starts rising and keeps rising as water enters faster than it leaves. You have crossed a critical threshold beyond which there appears a new type of change because now the water level will keep rising without limit (until it overflows the bathtub).

How did our ancestors cross the threshold to collective learning? We do not really know, though we have plenty of suggestions. Many changes led our ancestors towards the threshold of collective learning (for recent discussions, see Tattersall 2012; Fitch 2010). They included larger brains; insight into the thinking of others (a ‘theory of mind’); some ability to cooperate; the ability to control vocalizations and interpret the vocalizations of others; the use of fire to cook and pre-digest food, which, as Richard Wrangham points out, gave access to the high quality foodstuffs needed to grow brains. Many other species share some of these qualities and abilities (Tomasello 1999, 2009; Wrangham 2009; MacIntyre 2001: chs 3, 4). So, as Richerson and Boyd put it, we can imagine several species gathering at the barrier before collective learning, until eventually one broke through (Richerson and Boyd 2005: 139). Our own histo-

ry suggests that the lucky species would then deny passage to its rivals: ‘humans were the first species to chance on some devious path around this constraint [the difficulty that culture works only within a community of skilled social learners], and then we have preempted most of the niches requiring culture, inhibiting the evolution of any competitors’ (Boyd and Richerson 2005: 16). Since humans broke through, our closest hominine relatives, from Neanderthals to Denisovans, have perished and our closest surviving relatives, the chimps and gorillas are approaching extinction. Even if several related species arrived almost simultaneously at the barrier to collective learning, there was apparently room for only one species to sneak past it.

But the speed of the change – we, humans, began our climb to world domination less than 500,000 years ago, a mere second in paleontological time – suggests that a single push shoved us through. Perhaps, it was a glitzy new neurological gadget, some form of Chomsky’s ‘grammar’ module, or a new form of the FOXP2 gene that pushed us through. Or perhaps, as Terrence Deacon has argued, it was symbolic language (Deacon 1998). Some have argued for a slower transition. But, as a recent article argues, even if human language evolved 500,000 years ago, in evolutionary terms, that is a ‘flash in the pan’, implying that ‘language abilities were relatively rapidly cobbled together from pre-adapted cognitive and neurophysiological structures’ (Dediu and Levinson 2013: 10). Whatever the explanation, we should expect to find a single, critical change, because it defies reason to suppose that all the necessary pre-adaptations could have converged simultaneously on a single point in paleontological time. As Michael Tomasello writes, ‘This scenario [of a single switch] solves our time problem because it posits one and only one biological adaptation – which could have happened at any time in human evolution, including quite recently’ (Tomasello 1999: 7).

Suddenly, humans began to communicate not just in semantic fragments (‘Tiger!’), but in organized and contextualized strings of information (‘Yup, it’s got the same markings as the one that got Fred, and it’s behind the same bush!’). They began to use large, coherent packets of symbolic information, words like ‘family’ or ‘gods’ that compressed a world of experience into a few sounds, and linked those sounds into precise relationships using grammar (Deacon 1998). Human language locked up cultural information as tightly as DNA molecules locked up genetic information. As Tomasello puts it, ‘The process of cumulative cultural evolution requires ... faithful social transmission that can work as a ratchet to prevent slippage backward – so that the newly invented artefact or practice preserves its new and improved form at least somewhat faithfully until a further modification or improvement comes along’ (Tomasello 1999: 5). That is why some anthropologists describe cultural accumulation as ‘cultural ratcheting’ (Pringle 2013).

Once the switch for collective learning was thrown, our ancestors could start building new knowledge, community by community, accumulating local knowledge stores that steered each group in different directions to generate the astonishing cultural variety unique to humans. At the same time, our inner world was transformed as ideas washed from mind to mind. We do not just learn collectively; we *experience* collectively. The anthropologist, Clifford Geertz, described this realm as, ‘that intersubjective world of common understandings into which all human individuals are born, in which they pursue their separate careers, and which they leave persisting behind them after they die’ (Geertz 2000: 92). A simple thought experiment illustrates the power of this mental sharing. Look inside your head and do a quick census of everything that is there. (It takes just a few seconds.) Then ask the question: how much of that stuff would be there if you had never had a conversation with another human? Most will agree that the correct answer is: ‘Very little’. And that ‘very little’, mostly produced by individual learning, hints at the inner world of chimps. While chimps learn alone or in ones and twos, humans learn within teams of millions that include the living and the dead.

When did our ancestors cross the threshold to collective learning? In paleontological time, the crossing took an instant, but in human time it was probably smeared out over tens of thousands of years (a paradox captured in the title of McBrearty and Brooks 2000, ‘The Revolution that Wasn’t’). And even when the engine of collective learning spluttered into action, it took time to pick up speed. So we cannot easily judge when human history began. But we do know what to look for. We should look for sustained evidence of humans adding ideas to ideas to form new ideas. We should look for sustained innovation and ever-increasing cultural diversity. We should look for new and more diverse tools, and signs that humans were exploiting many new niches. And if, as Terrence Deacon and others have suggested, the breakthrough was the acquisition of symbolic language, then we should also look for evidence of symbolic thinking in art, body painting or signing (Deacon 1998).

The first speakers of a fully human language may not have belonged to groups normally classified within our own species, though they were surely very similar to us (Dediu and Levinson 2013). If they did belong to our species, we can date human history to at least 200,000 years ago, because that is the date of the oldest skull generally assigned to *Homo sapiens*. It was found in Omo, in Ethiopia in the 1960s (Tattersall 2012: 186).

But what we really need is evidence of new behaviours. In a comprehensive survey of African evidence from the Middle Stone Age, published in 2000, Sally McBrearty and Alison Brooks found hints of collective learning from as early as 250,000 years ago (McBrearty and Brooks 2000; and for a brief update see Pringle 2013). The Acheulian stone technologies associated with *Homo ergaster* were replaced by new, more delicate and more varied stone tools,

some of which may have been hafted. The new tools are associated with species that few anthropologists would classify as *Homo sapiens*, so the technological speed up may have preceded our own species. By 150,000 years ago, when members of our species were surely around, McBrearty and Brooks find hints that some groups were using shellfish and exchanging resources over long distances. We also see evidence of regional cultural variations. Ecological migrations are important because they show a species with enough technological creativity to move further and further from its evolutionary niche. Early in our history, new knowledge counted most at the edge of a population's range, where people faced the dangers and opportunities of testing new plants or animals. Before 100,000 BCE, we have tantalizing hints that some humans had entered deserts and forests (McBrearty and Brooks 2000: 493–494). After 60,000 such evidence multiplies; humans appear in Europe, in Australia and then in Ice-Age Siberia and, by at least 15,000 years ago, in the Americas.

Language leaves no direct traces, but archaeologists have found many hints of symbolic thinking. More than 260,000 years ago, early humans near Twin Rivers in modern Zambia used hematite (red iron oxide), possibly to paint their bodies (Stringer 2012: 129). Later evidence is less equivocal (for a good survey see Pettit 2005; on Blombos cave see Henshilwood *et al.* 2011). At Pinnacle Point in South Africa, in sites dated to about 160,000 years ago, we find the earliest evidence for the use of shellfish, along with signs of composite tools and lots of hematite, of a particularly brilliant red, which points to symbolic uses (Stringer 2012: 129). By 115,000 years ago, similar evidence turns up in modern Israel, where, in Skhul cave, archaeologists have found evidence of symbolic burials. But the best evidence of all for rich symbolic activity comes from the marvellous South African site of Blombos cave, whose remains date from almost 100,000 years ago. Here, Chris Henshilwood and his team have found delicate stone tools, seashell beads, and lumps of ochre carved with wavy lines that could almost be an early form of writing (*Ibid.*: 129–130).

Evidence for early signs of collective learning will surely come into sharper focus, but in the meantime, these hints suggest that if human history began with collective learning then something had cranked up the motor certainly by 100,000 years ago, perhaps, as early as 250,000 years ago and possibly 500,000 years ago (Dediu and Levinson 2013).

Collective Learning as a Form of Universal Darwinism

Collective learning launched and sustained our species on its astonishing journey towards planetary domination. If this argument is right, it seems that some form of Universal Darwinism has driven human history. We see *variation* in the ideas and information of different human societies, from their technologies to their religious rituals, from their art and clothing to their cuisine and enter-

tainment. Individuals and whole societies *select* some variants and reject others. And selected variations are *preserved* as they flow between minds.

But in detail, collective learning works differently from genetic learning and individual learning, and any Darwinian accounts of human history must take these differences into account. As Alex Mesoudi writes,

...many of the details of biological evolution that have been worked out by biologists since [*The Origin of the Species*], such as particulate inheritance (the existence of discrete particles of inheritance, genes), blind variation (new genetic variation is not generated to solve a specific adaptive problem), or Weismann's barrier (the separation of genotypes and phenotypes such that changes acquired in an organism's lifetime are not directly transmitted to offspring), may not apply to cultural evolution (Mesoudi 2011: x).

Why does collective learning work so much faster than genetic learning? In part because it builds on the machinery of individual learning, which works with neurological impulses rather than entire organisms. A genetic mutation must wait a generation before it effects change; a suddenly triggered memory can have you swerving in a second. Collective learning also copies fast. It can transmit new ideas on the fly, as they evolve, and can broadcast them to many brains at once because it works with sound waves (in speech) or light waves (in signalling and imitation). Like genetic learning, collective learning is auto-catalytic, so it has generated better ways of storing and transmitting information, from writing to printing to the telegraph and internet. Auto-catalysis explains why collective learning generates not just change, but *accelerating* change. Finally, collective learning, like individual learning, builds on acquired as well as inherited variations. While genetic learning gropes randomly in the dark, collective learning can probe more purposefully.

How do variation, selection and reproduction work in collective learning?

In collective learning, as in genetic learning, some variation is blind, arising from mutation and drift; but these variations arise from misunderstandings or simple blurring of meaning rather than from biochemical glitches. Much more important is another source of variation: deliberate innovation. Richerson and Boyd call this 'guided variation' (see the taxonomy of cultural evolutionary forces in Richerson and Boyd 2005: 69). Individuals deliberately add what they have learnt to the common pool of knowledge, or tweak and modify existing ideas. A little more salt in the soup, or tautness in the bowstring, or even a separate boiler for the steam engine. Moment by moment, and often with a sense of purpose, individual learning adds new information to a shared pool of knowledge, whereas genetic learning receives its variations at random.

Selection, too, can be conscious and purposeful in collective learning. Richerson and Boyd describe purposeful selection as 'biased transmission'. We

select using ‘content-based’ biases when we choose an idea or cultural variant on its merits, for its beauty or precision, perhaps. Other forms of selection are deliberate but less thoughtful. In a conformist or lazy mood, we often choose the most accessible idea or behaviour, or we choose ideas or behaviours associated with admired role-models. In the taxonomy of Richerson and Boyd these are called ‘frequency-based biases’ or ‘model-based biases’. Either way, selection is trickier in collective learning because cultural variations are fuzzier than genes, though often, when we choose one word or another or vote for one political party rather than another, we chop up the cultural flow.

Reproduction is fuzzier and more complex than in genetic learning. Ideas have many parents. They can also replicate in their thousands at religious festivals or political rallies or through mass media. Most important of all, in collective learning reproduction is less tightly bound to the reproductive success of particular individuals than in genetic learning. This is why humans often select variations that are *not* adaptive under the rules of genetic learning. For example, they may choose to have fewer children than possible, thereby reducing their reproductive success (Richerson and Boyd 2005: ch. 5). This makes no sense under the rules of genetic evolution, which measure success by the number of genes passed on to the next generation. Even worse, humans sometimes risk their lives for others who are not even close kin. Genetic reproduction can just make sense of sacrifices on behalf of close kin (who do, after all, share genes with you). But it cannot explain sacrifices on behalf of strangers or people you may never have met. Collective learning can explain such behaviour, because collective learners live within shared flows of ideas, information and motivation that create a sense of shared meaning and purpose, and magnify the importance of reciprocity. We inherit ideas and values from dead strangers and living teachers as well as from parents and grandparents, and we cannot always distinguish clearly between the two types of inheritance. So collective learning allows behaviours that, from the perspective of genetic learning, seem like errors, such as the choice of a group of ducklings to treat Konrad Lorenz as their mother. Symbolic thinking blurs the line between genetic and imagined kinship. And where meanings are shared so, too, are their emotional charges. Flags and national anthems can motivate us as powerfully as family, particularly if cultural differences sharpen our sense of shared community. Richerson and Boyd have shown that in such environments models predict the rapid spread of altruistic behaviours. This is particularly true where cultural selection is ‘conformist’, where people choose values because they are normal within their community (*Ibid.*: ch. 6).

In short, a sense of shared meaning blurs the distinction between individual and group success. In collective learning, the viability of ideas (and sometimes of the humans who carry them) depends as much on the reproductive success of entire groups as on that of individuals. So where collective learning is at work, group selection may be as important as individual selection, because with the

flourishing of human culture, genes are no longer the primary shapers of behavioural change. Group mechanisms including shared cultural norms and social structures clearly play a profound role in explaining human behaviour. So we should not be surprised to find that humans collaborate so effectively in bands, tribes and nations as well as in families. Though the idea of group selection is fiercely contested at present (for two different positions see Pinker 2013 and Wilson 2007), something like group selection is surely at work in the evolution of human culture.

Finally, and most mysteriously, collective learning generates an entirely new form of change, cultural change. Like information, cultural change often seems to inhabit a limbo between the physical and mental worlds. John Searle, who has spent much of his career trying to explain cultural phenomena, argues that the cultural realm arises from ‘shared intentionality’, or the shared sense of meaning created by collective learning (not his term) (Searle 2010: 3–8 and *passim*). ‘Shared intentionality’ explains why only humans can assign conventional meanings or functions to people and objects. It matters if they agree to call a piece of paper a twenty-dollar bill. The agreement creates rights, obligations and possibilities; it motivates behaviours that go well beyond our sense of individual wants or needs. Searle argues that such agreements are the foundation of all social relations and institutions. They are what make human societies different.

Conclusion: Different Versions of the Darwinian Machine

Wherever we see change swimming against the flow of entropy, we should suspect that a Darwinian machine is at work. Human history represents a spectacular example of this kind of change, so we should expect to find a Darwinian machine lurking somewhere within the discipline. Most historians have rejected this possibility, partly from fear of Social Darwinism, partly because the neo-Darwinian synthesis fit human history so poorly. But as we have seen, Darwinian machines come in different versions. A clearer appreciation of these differences may encourage historians, too, to explore the possibility that Darwinian mechanisms of some kind can help us explain the remarkable trajectory of human history. But they may also help us see human history itself as part of a much larger story of increasing complexity, most of which (perhaps all of which) was driven by Darwinian mechanisms of some kind.

‘Mutt.—Ore you astoneaged, jute you?’

Jute. — Oye am thonthorstrok, thing mud’ (Finnegans Wake, Ch. 1).

References

- Blackmore S. 1999.** *The Meme Machine*. Oxford: Oxford University Press.
- Boyd R., and Richerson P. J. 2005.** *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.

- Campbell D. T. 1960.** Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes. *Psychological Review* 67(6): 380–400.
- Campbell J. 2011.** *Universal Darwinism: The Path of Knowledge*. Create Space Independent Publishing Platform.
- Christian D. 2004.** *Maps of Time: An Introduction to Big History*. Berkeley, CA: University of California Press.
- Christian D. 2010.** The Return of Universal History. *History and Theory, Theme Issue* 49 (December): 5–26.
- Christian D. 2012.** Collective Learning. *The Berkshire Encyclopedia of Sustainability: The Future of Sustainability* / Ed. by R. Anderson *et al.*, pp. 49–56. Gt. Barrington, MA: Berkshire Publishing.
- Dawkins R. 1983.** Universal Darwinism. *Evolution from Molecules to Men* / Ed. by D. S. Bendall, pp. 403–425. Cambridge: Cambridge University Press.
- Dawkins R. 2006.** *The Selfish Gene: 30th Anniversary Edition*. Oxford: Oxford University Press.
- Deacon T. 1998.** *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: Norton.
- Dediu D., and Levinson S. C. 2013.** On the Antiquity of Language: The Reinterpretation of Neanderthal Linguistic Capacities and its Consequences. *Frontiers in Psychology* 4 (July): 1–17.
- Dennett D. C. 1995.** *Darwin's Dangerous Idea: Evolution and the Meaning of Life*. London: Allen Lane.
- Distin K. 2011.** *Cultural Evolution*. Cambridge: Cambridge University Press.
- Donald M. 2001.** *A Mind so Rare: The Evolution of Human Consciousness*. New York – London: W.W. Norton and Co.
- Fitch W. T. 2010.** *The Evolution of Language*. Cambridge: Cambridge University Press.
- Floridi L. 2010.** *Information: A Very Short Introduction*. Oxford: Oxford University Press.
- Geertz C. 2000 [1973].** *The Interpretation of Cultures*. New York: Basic Books.
- Gleick J. 2011.** *The Information: A History, a Theory, a Flood*. New York: Pantheon.
- Henshilwood, C. S., et al. 2011.** A 100,000-year-old Ochre Processing Workshop at Blombos Cave, South Africa. *Science* 334(6053): 219–222.
- Hofstadter R. 1944.** *Social Darwinism in American Thought*. Philadelphia, PA: University of Philadelphia Press.
- Jablonka E., and Lamb M. J. 2005.** *Evolution in Four Dimensions*. Cambridge, MA: MIT Press.
- Leonard T. C. 2009.** Origins of the Myth of Social Darwinism: The Ambiguous Legacy of Richard Hofstadter's Social Darwinism in American Thought. *Journal of Economic Behavior & Organization* 71: 37–51.
- Lloyd S. 2007.** *Programming the Universe: A Quantum Computer Scientist Takes on the Universe*. New York: Vintage.
- MacIntyre A. 2001.** *Dependent, Rational Animals: Why Human Beings Need the Virtues*. Chicago, IL: Open Court.

- McBrearty S., and Brooks A. 2000.** The Revolution that Wasn't: A New Interpretation of the Origin of Human Behavior. *Journal of Human Evolution* 39: 453–563.
- Mesoudi A. 2011.** *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*. Chicago, IL: University of Chicago Press.
- Nelson R. 2006.** Evolutionary Social Science and Universal Darwinism. *Journal of Evolutionary Economics* 16: 491–510.
- Nelson R. 2007.** Universal Darwinism and Evolutionary Social Science. *Biology and Philosophy* 22: 73–94.
- Pettit P. 2005.** The Rise of Modern Humans. *The Human Past: World Prehistory and the Development of Human Societies* / Ed. by Ch. Scarre, pp. 124–173. London: Thames & Hudson.
- Pierce J. R. 1980.** *An Introduction to Information Theory: Symbols, Signals & Noise*. 2nd rev. ed. New York: Dover.
- Pinker S. 2013.** The False Allure of Group Selection. *The Edge*. URL: <http://edge.org/conversation/the-false-allure-of-group-selection>
- Plotkin H. 1994.** *Darwin Machines and the Nature of Knowledge*. Cambridge, MA: Harvard University Press.
- Pringle H. 2013.** The Origins of Creativity. *Scientific American* 308(3): 36–43.
- Richerson P. J., and Boyd R. 2005.** *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago, IL: University of Chicago Press.
- Searle J. R. 2010.** *Making the Social World: The Structure of Human Civilization*. Oxford: Oxford University Press.
- Seife C. 2007.** *Decoding the Universe: How the New Science of Information is Explaining Everything in the Cosmos from our Brains to Black Holes*. New York: Penguin.
- Smolin L. 1998.** *The Life of the Cosmos*. London: Phoenix.
- Smolin L. 2005.** The Case for Background Independence. [Talk for the Perimeter Institute for Theoretical Physics, Waterloo, Canada.] [arXiv:hep-th/0507235v1]. URL: <http://arxiv.org/abs/hep-th/0507235>
- Snow C. P. 1971 [1959].** The Two Cultures and the Scientific Revolution. *Public Affairs* / Ed. by C. P. Snow, pp. 13–46. London – Basingstoke: Macmillan.
- Steffen W., Crutzen P. J., and McNeill J. R. 2007.** The Anthropocene: Are Humans Now Overwhelming the Great Forces of Nature? *AMBIO* 36(8): 614–621.
- Stringer C. 2012.** *Lone Survivors: How We Came to be the Only Humans on Earth*. New York: Times Books.
- Tattersall I. 2012.** *Masters of the Planet: The Search for Our Human Origins*. New York: Palgrave/Macmillan.
- Tomasello M. 1999.** *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Tomasello M. 2009.** *Why We Cooperate*. Cambridge, MA: MIT Press.
- Wilson D. S. 2007.** *Evolution for Everyone: How Darwin's Theory Can Change the Way We Think about Our Lives*. New York: Delacorte Press.
- Wrangham R. 2009.** *Catching Fire: How Cooking Made Us Human*. New York: Basic Books.