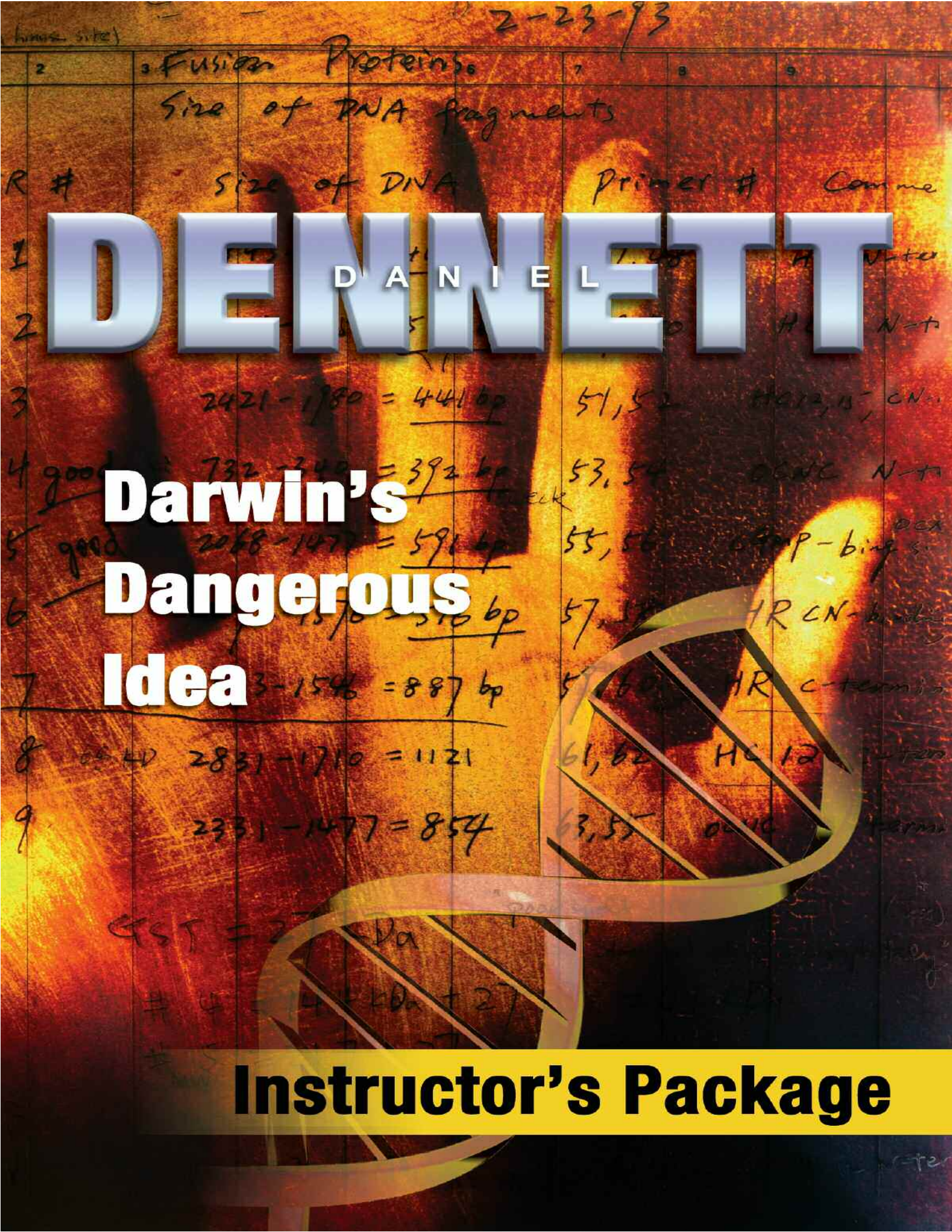

DANIEL
DENNETT
Darwin's
Dangerous
Idea
Instructor's Package

About This Guide

This guide is intended to assist in the use of the DVD *Daniel Dennett, Darwin's Dangerous Idea*.

The following pages provide an organizational schema for the DVD along with general notes for each section, key quotes from the DVD, and suggested discussion questions relevant to the section.

The program is divided into seven parts, each clearly distinguished by a section title during the program.

Contents

Seven-Part DVD

Articles by Daniel Dennett

1. A Dangerous Idea

Dennett considers Darwin's theory of evolution by natural selection the best single idea that anyone ever had. But it has also turned out to be a dangerous one. Science has accepted the theory as the most accurate explanation of the intricate design of living beings, but when it was first proposed, and again in recent times, the theory has met with a backlash from many people. What makes evolution so threatening, when theories in physics and chemistry seem so harmless? One problem with the introduction of Darwin's great idea is that almost no one was prepared for such a revolutionary view of creation. Dennett gives an analogy between this inversion and Sweden's change in driving direction:

I'm going to imagine, would it be dangerous if tomorrow the people in Great Britain started driving on the right? It would be a really dangerous place to be because they've been driving on the left all these years….

But you know, not so long ago, the people of Sweden did exactly that. And it worked beautifully, of course, because they planned for it, and they did it all in unison. They held hands, and they took the step together.

Now, if the Swedish civil service could have organized and planned for the reception of Darwin's great book on the origin of species, maybe it wouldn't have been so dangerous. That, in fact, the book landed on people's minds and people's lives like a ton of bricks. And from the very outset, there was a tremendous amount of anxiety and hostility and fear. And of course, fear often leads to dangerous reactions, and that's why Darwin's idea is dangerous…. You're used to one way of seeing the whole world, and now you're told you have to sort of invert.

The theory of how life on earth developed seems much more relevant to our own lives than theories of physics and chemistry. If no intelligent creator was necessary to design and produce these intricate living creatures, then we have lost one of the most compelling reasons to believe that an intelligent creator, that God, exists. If science can explain all these amazing natural phenomena, and God no longer plays a role in this explanation, in the words of Friedrich Nietzsche, "God is dead." And this is a frightening, panic-inducing idea, because if there is no God it seems that our lives are without purpose.

Discussion Topics:

- Why might Darwin's theory of evolution be threatening to those who have previously believed that an intelligent creator designed life on earth?

- Ask students if they can think of any way that someone could accept Darwin's theory without also seeing it as a challenge to their belief in a God figure.

2. Darwin's Inversion

Darwin's great idea is a radical inversion of common sense. Living beings, humans, giraffes, protozoa, are incredibly complicated creations, and any ordinary complicated creation requires an even more complicated, more intelligent, creator to design that creation.

You never see a pot making a potter. You never see a horseshoe making a blacksmith. It's always the other way around, big fancy smart things making things that are rather less wonderful than themselves.

Dennett labels the common-sense view of creation the "trickle-down theory." This is a completely obvious fact about the order of the creative process, so obvious that Intelligent Design proponents have created a "test" to "prove" to anyone with common sense the absurdity of complicated creatures designed by an ignorant process like evolution.

Here's a page from a propaganda pamphlet that a student sent me some years ago, it was a creationist pamphlet. "Do you know of any building that didn't have a builder? Yes, no. Do you know of any painting that didn't have a painter? Yes, no. Do you know of any car that didn't have a maker? Yes, no. If you answered, "yes" for any of the above, give details." Take that, you Darwinians.

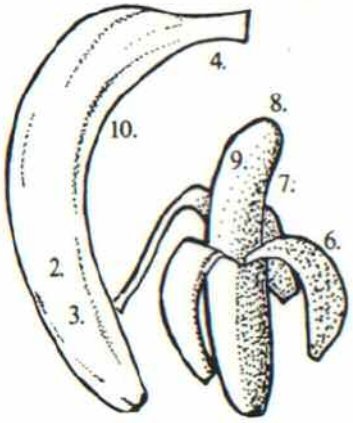
TEST TWO

1. **Do you know of any building that didn't have a builder?** ☐ **YES** ☐ **NO**
2. **Do you know of any painting that didn't have a painter?** ☐ **YES** ☐ **NO**
3. **Do you know of any car that didn't have a maker?** ☐ **YES** ☐ **NO**

If you answered "YES," for any of the above give details:

_____

The banana – the atheist's nightmare.



Note that the banana:

1. *Is shaped for human hand*
2. *Has non-slip surface*
3. *Has outward indicators of inward contents: Green – too early, Yellow – just right, Black – too late.*
4. *Has a tab for removal of wrapper*
5. *Is perforated on wrapper*
6. *Bio-degradable wrapper*
7. *Is shaped for mouth*
8. *Has a point at top for ease of entry*
9. *Is pleasing to taste buds*
10. *Is curved towards the face to make eating process easy.*

Darwin's theory takes this picture and puts it directly on its head—the most complicated creations in the world, all living creatures, are designed by an absolutely blind, entirely unintelligent process! Dennett describes Darwin's inversion of the ordinary creation process as the "bubble-up theory" of creation.

Now we want to compare that trickle-down theory of creation with what we might call the bubble-up theory of creation, which says that wonderful things can actually be created and improved and made better by a process which itself isn't intelligent at all and has no purpose, which is just a mechanistic mindless algorithmic sorting process.

The amazing feature of the natural selection process is that, given certain conditions, it cannot fail to work! Dennett's quote from *The Origin of Species* claims that if organisms compete for resources that aid their survival, differ even slightly from one another, reproduce, and produce offspring that are likely to be similar to their parents, natural selection *cannot fail* to improve the general design of organisms, very slowly, over successive generations.

Darwin's greatest idea

If during the long course of ages and under varying conditions of life, organic *beings vary at all* in the several parts of their organization, *and* I think this cannot be disputed; *if there be*, owing to the high geometric powers of increase of each species, at some age, season, or year, *a severe struggle for life*, and this certainly cannot be disputed; *then*, considering the infinite complexity of the relations of all organic beings to each other and to their conditions of existence, causing an infinite diversity in structure, constitution, and habits, to be advantageous to them, I think it would be a most extraordinary fact if no variation ever had occurred useful to each being's own welfare, in the same way as so many variations have occurred useful to man. But *if variations useful to any organic being do occur, assuredly individuals thus characterized will have the best chance of being preserved in the struggle for life; and from the strong principle of inheritance they will tend to produce offspring similarly characterized.* This principle of preservation, I have called, for the sake of brevity, Natural Selection.

Origin of Species, end of chapter 4

"If beings vary at all"—that is, if they're not all alike—"and if there be a severe struggle for life"—if there's finite resources so that not everybody can have everything that they need—"if variations useful to any organic being do occur"—that is, if the differences between these individuals are not all neutral, some of them actually provide a benefit of some sort—"assuredly individuals thus characterized will have the best chance of being preserved in the struggle for life."

That's practically by definition. "And from the strong principle of inheritance they will tend to produce offspring similarly characterized." The strong principle of inheritance is simply the idea that the offspring will be more like their parents than like their parents' rivals for those resources. If those conditions are met, Darwin says, we're going to have a process of gradual automatic improvement of the characteristics of those organic beings. And this, he says, is a principle of preservation. This is accumulating design improvement.

If offspring retain those attributes that made their parents successful in reproducing, these offspring will be likely to reproduce in greater numbers than *their* competitors. And as tiny mutations occur in the genes of organisms between one generation and another, some of the new generation will be better "designed" than others who compete for the same resources. Over great periods of time, the general design of organisms will improve, because, as they say, only the strong will survive. Or at least the stronger will survive in greater numbers than their competitors. This is what Dennett refers to as "accumulating design improvement."

This process is amazingly, incredibly simple, and seems to make perfect sense. And, as Dennett reminds his audience, the process requires no intelligence whatsoever; it has no purpose at all. All the wonderful, complex organisms that populate the planet are designed by the random process of genetic mutation and inheritance. And since the consensus has been, for the entire history of human thought, that an intelligent creator is needed to design and produce such intricate living beings, the theory of evolution by natural selection does seem unbelievable.

As soon as Darwin's theory was presented, it was met with a backlash from people employing common sense, the common sense of the creationists who argue from cars that require makers to humans whose existence requires the skill of an intelligent creator. Dennett quotes Robert Beverly McKenzie as a prime example of this reactionary thinking:

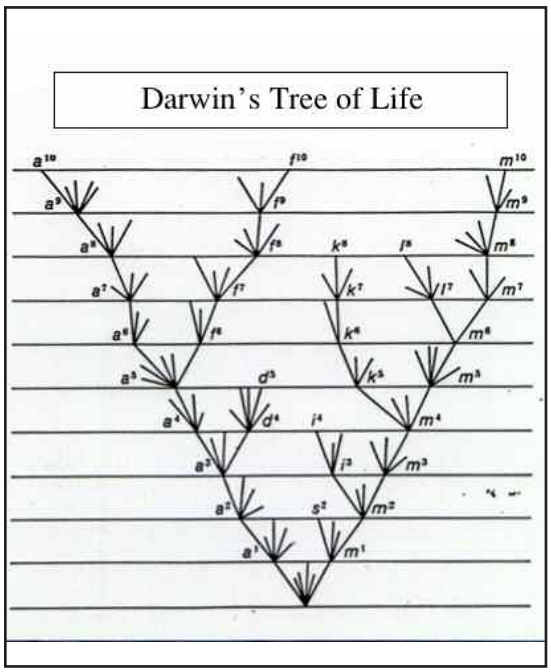
Darwin's greatest idea

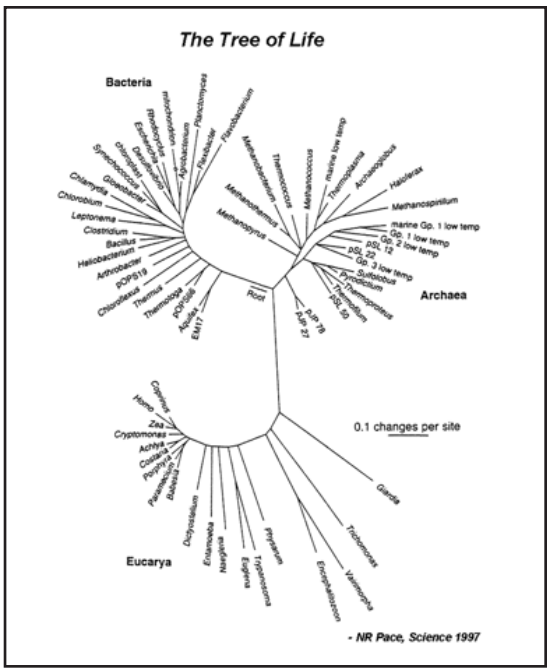
In the theory with which we have to deal, Absolute Ignorance is the artificer; so that we may enunciate as the fundamental principle of the whole system, that, IN ORDER TO MAKE A PERFECT AND BEAUTIFUL MACHINE, IT IS NOT REQUISITE TO KNOW HOW TO MAKE IT. This proposition will be found, on careful examination, to express, in condensed form, the essential purport of the Theory, and to express in a few words all Mr. Darwin's meaning; who, by a strange inversion of reasoning, seems to think Absolute Ignorance fully qualified to take the place of Absolute Wisdom in all the achievements of creative skill. --Robert Beverley MacKenzie, 1868

"In the theory with which we have to deal, absolute ignorance is the artificer; so that we may enunciate as the fundamental principle of the whole system, that, in order to make a perfect and beautiful machine, it is not requisite to know how to make it. This proposition will be found on careful examination to express in condensed form the essential purport of the theory and to express in a few words all Mr. Darwin's meaning, who by a strange inversion of reasoning seems to think absolute ignorance fully qualified to take the place of absolute wisdom in all the achievements of creative skill."

He's exactly right. That's the purport of the theory, that absolute ignorance, a process which has no purpose, no intention, no foresight, is just a mechanical sorting process, that this will automatically compete against absolute wisdom. It can do the work that tradition says requires absolute wisdom.

The process of mutation and accumulating design improvement, over vast stretches of time, is what produces the "tree of life," widely varying but ancestrally related species of organisms that all diverged from the same very first living entity. Dennett's diagrams show how present species, and some extinct species, are all related, and how all diverged from the first living organism.





And what you see is, thanks to this postulated process of natural selection, they grow farther and farther apart. And that's not in space. That's in design space as it were. They diverge in their characteristics enough so that eventually we get a fissioning of lineages into distinct species. Now this is the tree of life as Darwin envisaged it.

As the random, tiny genetic mutations that lead to the design improvement of a particular species build up over time, some organisms may have different slight mutations than other similar organisms, but they may be equally, or nearly equally, capable of acquiring resources of their environment, but in different ways. The genes of each of these different organisms will pass down generation to generation, and every so often genetic mutations will occur, which will create even greater differences between organisms of the same generation. Eventually, these organisms will become so different that we identify them as different species.

Discussion Topics:

- How is Darwin's theory of evolution by natural selection an "inversion" of ordinary common sense?
- What does Dennett mean by "accumulating design improvement"?
- How can natural selection result in the improvement of organisms over time?

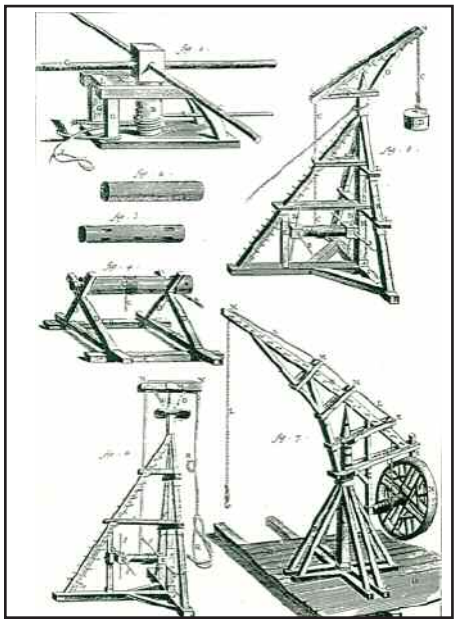
3. Cranes: Getting Here from There

Dennett identifies two competing ideas of how simpler organisms are transformed into much more complex creatures. The first idea is that we just can't get from "there" to "here" through any natural route.

But still, it's quite possible to be skeptical and say, "I still don't see how all that design work could have been accomplished by such a mindless process. I don't think it's possible. I think there may well be some wonders of the biosphere that are just too wonderful to have been generated by natural selection. You can't get here from there, wonder that could not be the result of a long and gradual sequence of small, mindless improvements."

The idea, that tiny mutations couldn't possibly move an organism so far above its parents' design, is that these innovations would require "skyhooks," the intervention of something supernatural, a boost from above that just couldn't come from the natural processes of evolution.

This idea is motivated by a misunderstanding of how gradual the process of evolution is. It takes billions upon billions of years for the process of evolution to stumble across all the mutations that have made organisms more successful at acquiring the resources of their environments. Most mutations are either neutral, neither beneficial nor detrimental, or actually harmful to the organism's chances of survival. But rarely a mutation comes along that is slightly beneficial to the organisms that are lucky enough to have it, and over long periods of time these can be accumulated in descendents that are phenomenally more adept at the survival game.



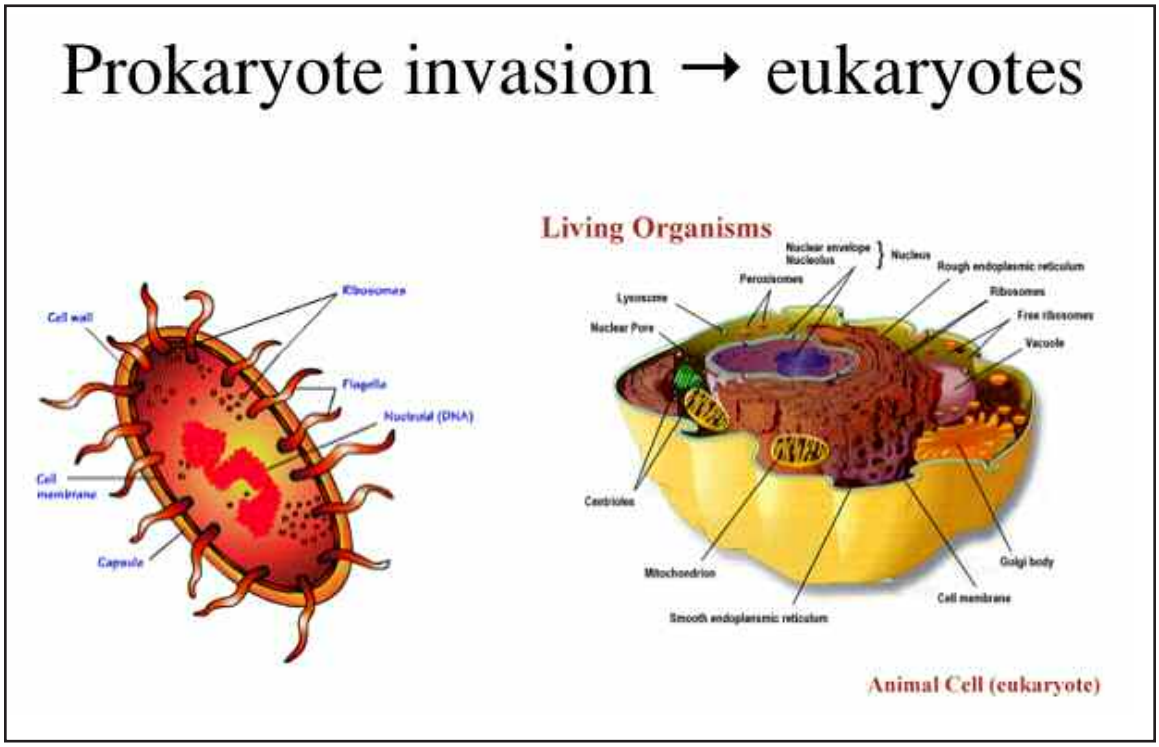
Even more rarely, such mutations lead to the evolution of biological structures or systems in organisms that enhance the process of natural selection itself, speeding it up, or making it more efficient. So the process of natural selection has itself evolved, marked by a series of innovations that have made it a more powerful design-maker. These are what Dennett calls "cranes."

Cranes are not miraculous. You can actually design and build them. They have their own R & D history. And once you've got them, they make possible designs that weren't heretofore possible like Brunelleschi's Dome.

These design innovations, these cranes of evolution, take both time and energy to evolve. But once they exist on the scene, they open up new design possibilities to be explored by the mindless process of natural selection. And new designs, once they are "invented" or "discovered" by natural selection, will be endlessly copied, and improved upon. Darwin called this principle the "Principle of Preservation," the preservation of beneficial mutations. Dennett suggests it could be called the "Principle of Plagiarism"—since a design worth copying doesn't have to be reinvented over and over.

This gradual process of "building" upon these good mutations, these innovations, is what Dennett calls the "design work" of evolution. Creations that have different designs occupy different positions in "design space," abstract positions in the space of all possible designs. Cranes enable the process of evolution to "explore" more widely varying areas of design space. (John Maynard Smith and Eors Szathmary called these cranes the "major transitions of evolution.")

Ever since Darwin, skeptics about evolution have hunted for skyhooks—"irreducibly complex" marvels that could not have evolved gradually by natural selection—but again and again their searches have uncovered cranes—previously unappreciated ways in which evolution can do more design work than had been thought possible. Eyes have evolved from simple light-sensitive cells that could not see but could respond to looming shadows; wings have evolved from structures that provided insulation, not flight; and bacterial flagella evolved from simple jets that served for either propulsion or ejection of toxins.



Dennett cites the eukaryotic revolution and sexual reproduction as two of the cranes of evolution. Once simple, prokaryotic cells became more complicated (in the eukaryotic revolution), they had enough parts to specialize, and these specializations permitted many cells to join into versatile teams creating a single multicellular

individual, which eventually led to organ systems, muscular systems, nervous systems, and specialized reproductive systems—the evolution of sex, which permitted a much more efficient exploration of possible genomes, thanks to the mixing of genomes from two members of the lineage, and the doubling up of genes, creating useful redundancy that could be further exploited by natural selection.

Language is a crane unique to our species. Many species have methods of communicating simple messages from one member to another, but we are the only species to acquire a truly productive language. Language is what allows for the development of human culture, since it enables us to spread complicated ideas from one member of the species to another. Because of the development of language, we are able to explore design space that is inaccessible to any other species. We can create artifacts, beyond the simple tools that we share with other species, to automobiles, skyscrapers, dishwashers, as well as the human artifacts of corporations, governments, art and religion.

Discussion Topics:

- Students may benefit from putting the difference between cranes and skyhooks into their own words.

- Ask students if they can see how something as complex as a wing or an eye might actually develop incrementally from other, perhaps unexpected, adaptations (such as the protowing's insulating properties or light-sensitive cells).

4. Fruits of the Tree of Life

The products of the great tree of life are not only living organisms, but all the creations of the living organisms. Evolution made possible complicated creatures that are themselves creative, and this sort of creativity is not limited to human beings, but is found in other mammals, even insects.

From this perspective, we can begin to look at the fruits of the tree of life and see them a bit differently. On the left, we see a beaver dam. On the right, we see Hoover Dam. They are both fruits on the tree of life. Neither one is alive. Both are artifacts made by living things, and they're both quite complicated, and they serve rather the same function. But the ways in which they're built and the R & D that went into them are, of course, remarkably different in many regards. But they both find their place as products of the tree of life. In fact, products of the tree of life delivered by two species, Homo sapiens and beavers, which are actually quite close together. We're both mammals after all.

The spiderweb and a city power grid are structures that are both woven in ways that fit the strengths of their materials—the strength of a spider's silk is employed to effectively trap its prey, and the productive capacity of electricity is given routes by which to spread itself out in a way that utilizes the very nature of electricity itself.

Natural selection, Dennett claims, can explain not only the living beings on the tree of life, but the inanimate artifacts created by those living beings as well. Some people concede that evolution can design the wing of a bird, but not a poem. But can this really be so? Imagine how difficult it would be for us to figure out how to create the wing of a bird, even a robot bird.

Every living being, and every product of a living being, has a place on the tree of life. The twigs, branches, limbs, and trunks lead to all the living things, related by *descent*, and the products are linked to their living artificers. Some products, such as alpine goat paths and termite mounds, have many makers, of course, with no single author, and the contribution of any one maker has to be based on the work of many predecessors. Similarly, a poem depends, not on the labors of just one poet, but on the products of the many poets who inspired or provoked the poet, on the creation, by many hands, of genres and styles and metaphors, on the creation of printing, and before that, writing, and before that of spoken language itself. This is what makes a poem—quite simple in itself—an even more wonderful creation than a bird's wing. The wing is a product of evolution, but a poem is a product of a *product* of evolution, still part of the grand process of evolution, but requiring even more complex design than a bird's wing.

Genetic engineering, direct and intentional human reorganization of the genetic makeup of living creatures, is itself as much a product of evolution as language or poetry, and this new crane has led to the development of glowing tobacco plants and all manner of mutant mice used in medical research. Could these strange organisms have evolved without the help of the crane of genetic engineering? It is possible but surely hugely unlikely. In any case, now that they are here, their offspring and grandoffspring continue the basic process of evolution by natural selection.

Can the human mind, human culture, be explained by the same process that created the bird's wing and the beaver's tail? If natural selection can explain the dam, the spiderweb, and arguably the written word as well, why couldn't it also explain the human mind? This is where many skeptics, many creationists, dig in their heels and say that this just couldn't be, that *this* innovation *requires* a skyhook, a miraculous intervention by an intelligent designer who designed this element of nature with a purpose.

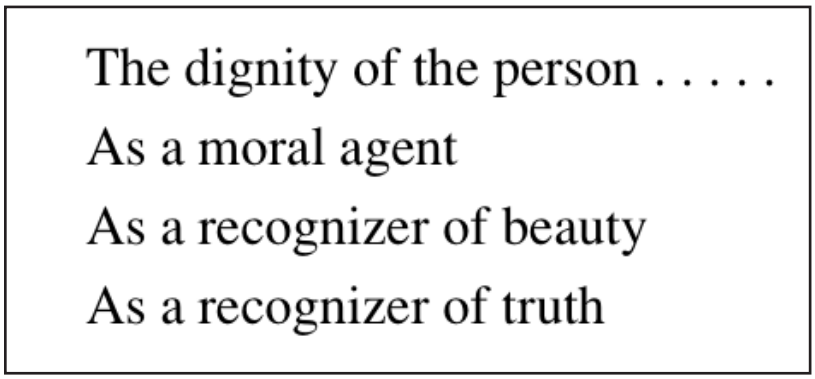
Discussion Topics:

- Ask students if they can explain how a poem might be more "wonderful" than a bird's wing.

- Discuss how the creations of humans and other animals could actually be limbs of the tree of life.

5. Humans without Skyhooks

Does humanity need a skyhook to come into being? Of course Dennett's answer is straightforwardly "no." The idea that we could not be created by the natural process of evolution has deep roots, and the motivations to reject this idea are powerful. Dennett quotes a statement by Pope John Paul II, where the pope claims that evolution is a fact, not a theory, but it cannot explain, or ground, human dignity and the human spirit.



"Theories of evolution which, in accordance with the philosophies inspiring them, consider the spirit as emerging from the forces of living matter or as a mere epiphenomenon of this matter, are incompatible with the truth about man." In other words, yes for the bird, no for the human spirit.

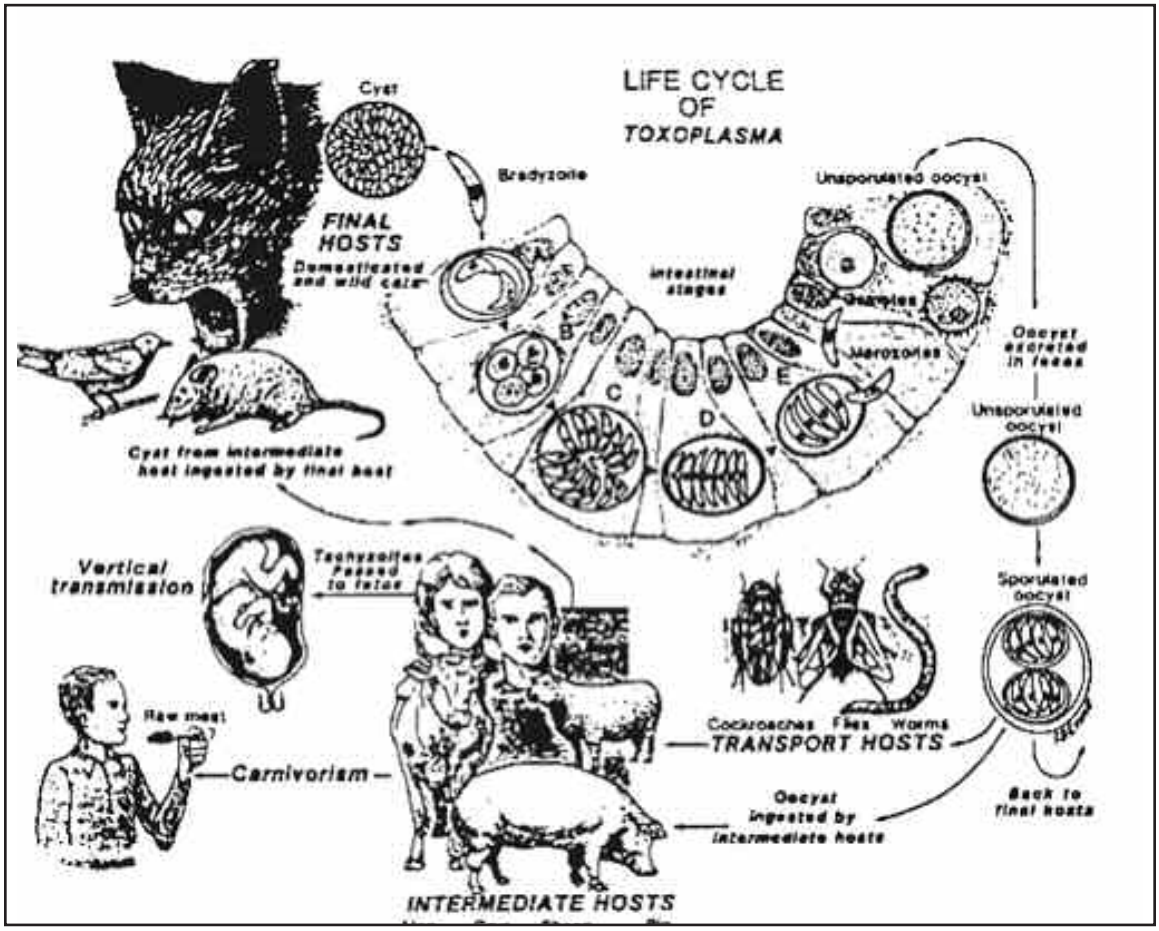
Why is evolution the proper explanation of the bird, but not the human? Because—some think—we have properties that go far beyond any other animal species, properties that seem to require an extraordinary human feature, something like a *human spirit*. Certainly, we do have properties that are a radical departure from the properties of other creatures. We are moral agents; we make distinctions between what's right and what's wrong, and no other creature on earth seems even *capable* of making such a distinction. We appreciate beauty, we recognize truth and deceit, we have rich lives, unlike the kinds of lives of other creatures. The immense difference between our lives and the lives of other animals must be evidence of a skyhook—some think—a miraculous leap from animal to human.

But there are no skyhooks, and so there must be an explanation that involves only cranes, one that employs the process of evolution, just like any other natural design. But the explanation of how we could develop, with all our culture and creations, through a natural process, is mind-bending, because it depends on Darwin's great inversion of the creation process. To understand it, we must keep in mind that evolution by natural selection applies not just to living organisms but to *anything* that has a design and gets replicated over and over with variations. Some designs will be more fit for their environment, and will continue on while others go extinct, forgotten or rejected by those who created them. Such a designed thing may depend on a living thing for its further reproduction, but it may or may not benefit those helpers in some way. Dennett gives the example of the ant infected with a peculiar parasite, the lancet fluke.

You're out in a meadow, and you see a blade of grass, and you notice that there's an ant climbing up the blade of grass. It climbs up the blade of grass and then it falls. And it climbs and it falls. And it climbs again, and you think, "My gosh, this is like the myth of Sisyphus. It just keeps climbing up that blade of grass. Why is the ant doing this? What's in it for the ant?"

And you might wonder, well, maybe it's looking around to see where the food is. Or maybe it's showing off for its mate. If you think that way, you're simply missing another possibility because no benefit accrues to the ant climbing up, expending all of this energy climbing up the blade of grass, none at all. What, you mean it's just a fluke? Yes, in fact, it's a fluke. It's a brain fluke, a lancet fluke.

This is a little parasitic worm that has crawled into the ant's brain because it has to get into the belly of a cow or a sheep to complete its reproductive cycle, and it improves its odds by getting up high on a blade of grass which has to be eaten. [The lancet fluke] simply commandeers a passing ant and drives it like an all-terrain vehicle up a blade of grass and hopes for the best except, of course, that [it] can't hope. It's a mindless little shred of life. But that's what it's doing, and the benefit is to its reproductive success not to the ant.



One great revelation of the theory of evolution by natural selection is that organisms don't need to be aware of what they're doing, they simply have to inherit a genetic mutation that makes them do something that makes them more likely to reproduce than their competition. And given the Principle of Preservation, or Dennett's Principle of Plagiarism, the descendents of the most successful organisms will have that mutation themselves, and pass it on to their descendents in turn.

And so there was an ancestor of the lancet fluke that ended up with the mutation that made it restructure the ant's brain to drive it up the blade of grass. And that organism won the survival competition, and passed that mutated genetic structure along to its progeny, who passed it on to theirs, until there was a species of lancet flukes, all driving ants up blades of grass when they got the opportunity.

Could there be a parasite that infects the human brain the way these parasites control the behavior of their hosts? Of course there could, since we are biological creatures with nervous systems not so unlike the ant's. But because

of a special feature of the human organism, our capacity for language, we are susceptible not only to organic parasites but to parasitic *ideas* as well. The ability to use language gives us the ability to acquire very complicated ideas, very complicated beliefs and desires, and it is these beliefs and desires that drive human behavior. If you make a person believe the right ideas, you can get that person to do just about anything.

As parasites compel ants and rats, and other living creatures, to sacrifice their lives, so some ideas can lead human beings to sacrifice theirs. There are *ideas worth dying for*. Parents would willingly change places with a dying child, or sacrifice their lives for the life of the person they love—and these sacrificial tendencies may be largely genetic, but we are also willing to put our lives on the line for an item of culture: for instance, soldiers are willing to sacrifice their lives for their country, for the *ideals* of democracy and freedom and human rights, or to advance the cause of their religion.

If an idea taps into the right human desires it can compel people to do almost anything, to sacrifice their own lives and the lives of others. Dennett reminds us that the Arabic word *islam* means "submission," in particular, "surrender to the will of Allah." This means that the will of Allah is more important than anything else. Religion is one of many ideas that tap into very powerful, very motivating human desires. Christianity works in this way as well.

Semen est verbum Dei. Sator autem Christus.

The Word of God is a seed. And the sower of the seed is Christ.

Omnis qui audit eum manebit in eternum.

In return for spreading the word, you get eternal life.

"The Word of God is a seed, and the sower of the seed is Christ." What's in it for the vector? "For spreading the word, you get eternal life." It's a nice benefit thrown in there.

Genes help an organism live to reproduce—this is simply what genes do, or at least the genes that are passed from generation to generation, the genes that survive. But because we have language, and therefore we can have *ideas*, we can have goals other than the goals of our genes. For the love of family a parent can sacrifice himself or herself for a child—and this is as true of bears and elephants as it is of us—but we can develop the feeling of love for more than our offspring.

But think... we're the only species that does anything like this. For every other species, the reproductive imperative, that's it. That's the highest goal that can be conceived. Not us. We have become invaded by ideas that we submit ourselves to and treat the furthering of the fate of those ideas as more important even than our own lives. So no other species does it, but it does have a biological explanation. It has to have a biological explanation. It too is a fruit on the tree of life.

Discussion Topics:

- Ask students why it seems to some people that, although some human traits (such as the instinct to protect their young and opposable thumbs) can be explained by Darwin's theory of evolutions, the "human spirit" seems to resist a natural explanation.

- Do students know of any precursors to full-blown human morality that are found in other animal species? (Some species of higher mammals, such as wolves, have apparent social hierarchies and seem to follow certain rules depending on where they fall in the hierarchy.)

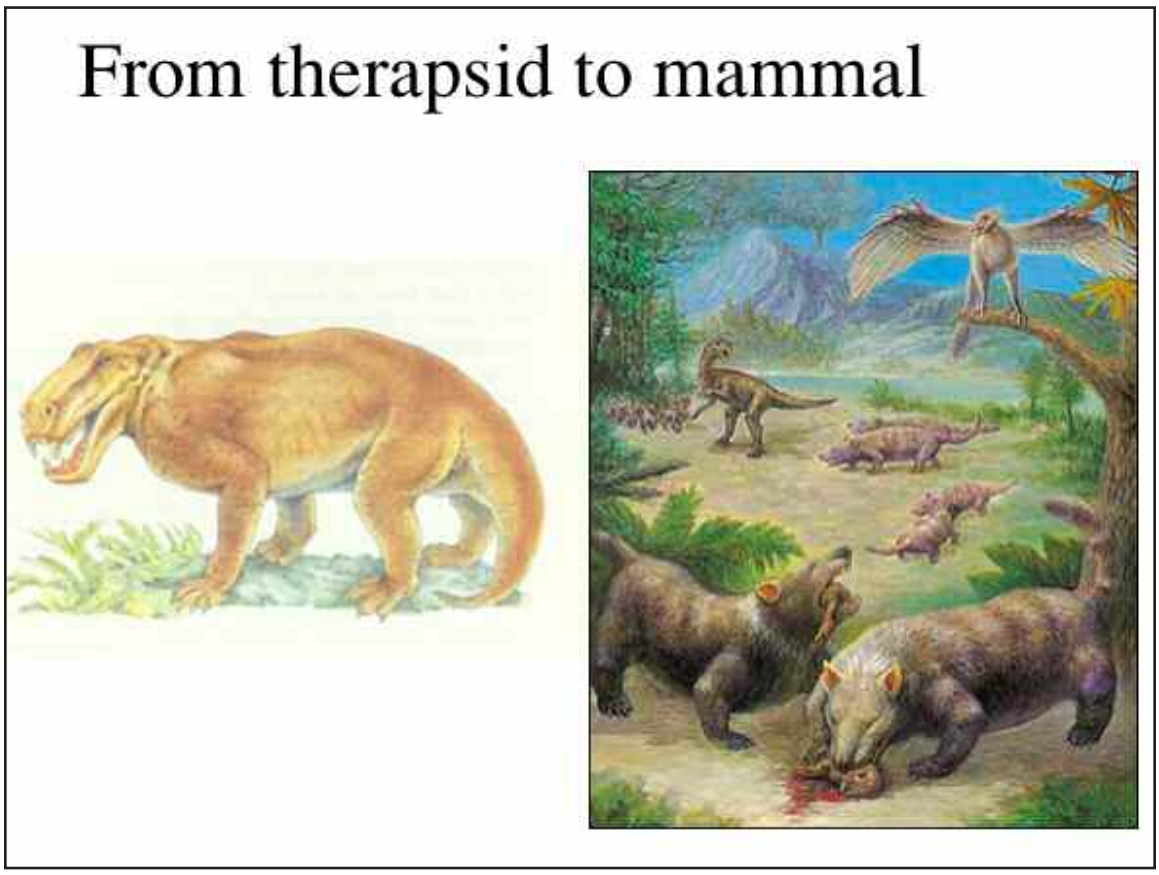
- How might the development of morality be explained as an evolutionary adaptation? (For example, it may have been a solution to the problems of living in close social groups.)

- What other "ideas worth dying for" can students see in human culture? Which of these ideas might be beneficial and which detrimental to society as a whole?

6. Gradualism

"Every mammal has a mammal for a mother." Do you agree? It seems almost definitive of what it is to be a mammal, to be born of a mammal. Well, there have only been a finite number of mammals, right? Now we have a contradiction. It cannot be that both of these propositions are true. It's just a mathematical fact. So something has to give.

So what is going to give? Therapsids are the intervening stage of creatures between reptiles and mammals, and we might say that the first mammal had a therapsid for a mother.



But where should we draw the line between therapsid and mammal? The change between creatures from one generation to the next is so slight that it seems impossible to say, of one generation, that it is a therapsid, and of the next generation, that it is a mammal. How do we make this distinction, find the point at which the mammals began? Dennett tells us not to bother.

We don't have to draw the line. What we can see is that there is a transition from therapsid to mammal, and we don't have to worry about exactly which one was the first and why.

The question of which mammal was the first, distinct enough from the therapsids that it could be considered a different species (leading eventually to a different biological *class*), is a question with *no answer*. The process of evolution is gradual, with numerous features changing at once, so that there is no telling which of the various differences between individuals is most important or "essential." We may be able to identify the individuals that are the

most distinct—and most distantly related—as being of different species, but there will be gray areas in which the species blend together. This is the inevitable result of the gradual process of the evolution of the different species.

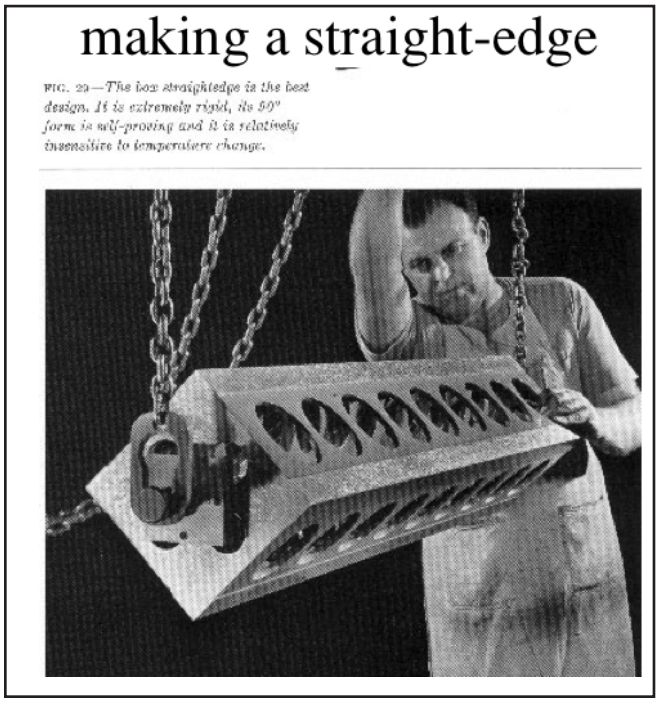
There is always variation among members of a population, and the gradual shift in typical characteristics can be seen, not just over time, but at a single time, spread over space. The example Dennett gives of this gradual change is in *ring species*. A ring species has a population that is spread out quite continuously in a more or less linear way—but one end of the line doubles back and confronts the other end of the line with a quite sharp discontinuity. Dennett discusses the herring gull and the lesser black-back gulls, which circle the Arctic. Are there two species or one? There is a continuous chain of "conspecifics" that connects them all.

If you start here in the British Isles, you have both species. That's where their names were given to them, and they are distinguishable and quite readily distinguishable. As you cross over to Iceland and to southern Greenland and into northern Canada, you find herring gulls too but none of the lesser black-back gulls. But they're a little bit different.

And as you keep going around the North Pole and get into Siberia and coming back, by the time you get back down to the British Isles, you realize that there's a continuity, that the lesser black-back gulls of England are kin to lesser black-back gulls in Finland and Siberia. And as you go around, you realize, well, do you have one species here or two? You have a ring species, and it doesn't really make much difference which way you describe it.

So we don't need a first, original mammal if we accurately understand the process of gradual change. What we need instead of a first mammal is a "bootstrap." As Dennett employs the term, bootstrapping isn't the impossible activity of "pulling yourself up by your own bootstraps," but rather the act of beginning in the generally correct area

and making improvements from there. Dennett's example is the creation of the straightedge. When we began making straightedges, we ground stones against each other until their sides were as straight and smooth as the sides of two stones could be. As our technology improved, we were able to make straighter and straighter edges, until we reached the accuracy of the "box straightedge," the standard for creating machine tools that depend for their effectiveness on the straightness of their edges.



But to reach this very accurate straightedge, we didn't need an original straightedge; instead, we started with a protostraightedge that was fairly straight, somewhere between crooked and straight, and *gradually* improved the straightness of our edges until we ended up with the straightedges we have today. But even these straightedges are not perfect, as Dennett points out; under magnification we can see that even our straightest edges are quite jagged. We can conceive of a perfectly straight edge without having any example of a perfectly straight edge, without any miraculous "skyhook" straightedge to give us the idea. We have built the idea of the straightedge up gradually, from less to more straight edges.

Discussion Topics:

- How does Dennett's example of the development of the straightedge mirror the evolution of species?
  (The straightedge started out as the edge of a stone, somewhere between crooked and straight, and gradually evolved into a straighter and straighter edge. A species such as Homo sapiens that we see today has roots in ancestral primates, just as the straightedge has its roots in a smooth edge of a stone. As the straightedge developed through the gradual advancement of technology, Homo sapiens developed through gradual, beneficial, genetic mutations.)

- Can students explain why this process would be called "bootstrapping"?
  (The development of species requires no aid from the "outside," so to speak.)

7. Memetic Revolution

Dennett identifies language and culture as two of the cranes of evolution that themselves spawned other cranes, and led to the evolution of other design improvements. To understand how language and culture might be the product of evolution, Dennett brings us back to the cellular level.

Symbionts are life forms like that prokaryote that landed inside another prokaryote that lived with other karyotes. And the invaders in the prokaryote invasions were not parasites but what are called "mutualists." That is, they were beneficial to each other. In fact, the standard terminology says that symbionts can be divided into parasites proper—those are the ones that are bad for your fitness, that are actually deleterious to your fitness, such as Dicrocoelium dendriticum *[lancet fluke], which doesn't do the ant any good at all, in fact leads to its early demise.*

And then we have mutualists, which enhance host fitness, such as many of the flora in your gut that you couldn't live without. Be very grateful you have those symbionts. And then there's the commensals, which are just neutral.

Dennett claims that symbionts and ideas share quite a lot in common. Parasites settle into a biological organism, sometimes beneficial to their host, sometimes a detrimental invasion, and other times simply neutral.

Now, what I want to suggest is that something very similar to the prokaryotic revolution happened when we got culture, that a certain group of hominoids' brains got invaded by some symbiont visitors that Richard Dawkins calls "memes." These are items of culture that then reproduce, ideas. The resulting hominoid, his head packed with replicating ideas, was fitter than the uninfected cousin. So now, just as with the eukaryotic cell, we've made a teamwork, a new kind of organism. We might call it a "person."

The Memetic Revolution

A hominoid brain was invaded by some *memes*,
And the resulting hominoid, its head packed with replicating ideas, was fitter!
It could specialize,
making it possible for multi-person groups to organize. . . . .
Creating civilization, more culture, science, technology. . . . .

Ideas—memes—are made of bits of information and are themselves designed. They can become cranes that lead to other design innovations, if they themselves are fit enough to survive over other competing ideas. Just as the eukaryotic revolution—a symbiosis between two prokaryotes—enabled cells to specialize and achieve multicellular feats that unicellular organisms never could, having these invading ideas, or memes, in our minds made us able to specialize, to cooperate and organize into different elements of human society, of human culture.

Not only do memes and symbionts share quite a lot in common; because memes are themselves designed entities made out of the building blocks of language, they *are* symbionts, independent entities subject to the same selection pressure as any other designed entity.

As an extra example, consider the "meme" of junk food. We created junk food, and junk food has remained even

though we know that it's unhealthy for us to eat. But we crave junk food, because it satisfies desires that we originally acquired because they were benefical; they allowed us to identify food that was nutritionally good for us. Junk food is *designed,* it is an independent designed entity, and it will survive because it is fit for its environment, an environment in which people crave food with those qualities. So long as we crave it, so long as the environment maintains these properties, junk food will succeed in the competition for existence.

Memes survive or die out on their own, without any human direction. Words are memes, languages are composed of memes, religions and political ideologies and cultures are all complexes of memes, and will survive or not based on how fit they are for their environment.

These pieces of human culture, these memes, will evolve just as anything that has a design will evolve, and they make possible other designs as well. Memes are the basic elements of human culture. As Dennett says,

And thank goodness we have them. They make a lot of things possible that wouldn't otherwise be possible. They create civilization, more culture, more science, more technology. This is the memetic revolution.

What sort of symbionts are memes? Are they parasites? Mutualists? As we have seen, the junk food meme is certainly parasitic, and it is arguable that the meme of civilization itself may lead to our demise. Dennett cites Paul MacCready's statistics regarding the original human population and our present population. At the dawn of our agricultural societies, humans and our domesticated animals constituted .1% of the terrestrial vertebrate biomass. Today, we constitute 98% of that biomass. Our incredible survival and reproduction rate is all due to the spawn of human culture, technology, of our memes.

> Over billions of years, on a unique sphere, chance has painted a thin covering of life—complex, improbable, wonderful and fragile. Suddenly we humans (a recently arrived species no longer subject to the checks and balances inherent in nature), have grown in population, technology, and intelligence to a position of terrible power: we now wield the paintbrush.
> —Paul MacCready

Which were the first memes? How did the memetic revolution begin? These are questions Dennett does not answer in this lecture, but he discusses these ideas at length in his book *Breaking the Spell*. These are questions that do not yet have answers, but we can point in the direction we need to go to find these answers.

Regarding the meme of religion, Dennett proposes that it was originally a "wild meme," analogous to pigeons and rats and squirrels. We do not own or control squirrels, but they are very good at flourishing in the human environment. "Folk religion," as Dennett calls it, is like these wild but familiar animals. We did not intentionally create the idea of religion, just as we did not domesticate squirrels and pigeons, but religion memes are very fit to live in the human environment. But the early memes of religion did not stay wild. When agriculture evolved, people didn't domesticate just plants and animals; they also domesticated their religions, turning them into *organized* religions.

What happened, I claim, is that some of the wild memes of religion, of folk religion, got domesticated. They acquired stewards. And that created organized religion and changed everything.

This grand picture, the development of not only living species but species of ideas as well, is testimony to the "brilliance" of evolution.

Francis Crick, once as a joke, invented what he called "Orgel's Second Rule": Evolution is cleverer than you are. That doesn't mean intelligent design. In fact, on the contrary, evolution is a perfectly unintelligent process, but it is breathtakingly good at finding brilliant solutions to problems that have their own rationales, what I call the "free-floating rationales" of evolution.

Taking the example of sheep, we can see how domestication was a great "idea" for them, how clever the solution to their problem—the struggle for survival. In exchange for free mating they have gained the protection of human shepherds. How clever! Except that sheep themselves are not the least bit clever. They are extraordinarily stupid creatures.

It's not the sheep's cleverness. It's the cleverness of evolution.

Evolution is phenomenally clever, not because it is an intelligent process but because it is a process that will eliminate any of its products that are less fit than their competitors. It is a process through which only the fittest survive. And this is all the brilliance, all the cleverness that is necessary to produce all the great creatures and cultures of the world. This was Darwin's great idea, and Dennett has turned that great idea into a motto:

Delere	Destroy
Auctorem	the Author
Rerum	of Things
Ut **U**niversum	in order to
Infinitum	Infinite Universe
Noscas	Understand

"Destroy the author of things in order to understand the infinite universe."

Discussion Topics:

- How can memes be successful even though they are harmful to their "hosts," that is, to us?

- Why is it appropriate to describe evolution as "clever" even though it is also a simple, blind, algorithmic process?

Could There Be a Darwinian Account of Human Creativity?

INTRODUCTION

Weaver birds create intricate nests; sculptors and other artists and artisans also create intricate, ingenious constructions out of similar materials. The products may look similar, and outwardly the creative processes that create those products may look similar, but there are surely large and important differences between them. What are they, and how important are they? The weaverbird nestmaking is 'instinctual,' and 'controlled by the genes' some would say, but we know that this is a crude approximation of a more interesting truth, involving an intricate interplay between genetic variation, long-term developmental and environmental interaction and short-term environmental variation—in opportunities and materials accessible at the time of nest building. And on the side of the human creator, a similarly complex story must be told. Genes play some role surely (think of the likelihood of heritable differences in musical aptitude, for instance), but so do both long-term and short-term environmental interactions. The myth of the artist "blessed" by a spark of 'divine genius' is even cruder and more distorted than the myth of the birdnest as a simple product of a gene—as if it were a protein.

Our thinking about human creativity is pulled out of shape somewhat by a famous contrast introduced to the world by Darwin. One of his earliest—and most outraged—critics summed it up vividly:

> In the theory with which we have to deal, Absolute Ignorance is the artificer; so that we may enunciate as the fundamental principle of the whole system, that, IN ORDER TO MAKE A PERFECT AND BEAUTIFUL MACHINE, IT IS NOT REQUISITE TO KNOW HOW TO MAKE IT. This proposition will be found, on careful examination, to express, in condensed form, the essential purport of the Theory, and to express in a few words all Mr. Darwin's meaning; who, by a strange inversion of reasoning, seems to think Absolute Ignorance fully qualified to take the place of Absolute Wisdom in all the achievements of creative skill. (MacKenzie1868)

Darwin's 'strange inversion of reasoning' promises—or threatens—to dissolve the Cartesian *res cogitans* as the wellspring of creativity, and then where will we be? Nowhere, it seems. It seems that if creativity gets 'reduced' to 'mere mechanism' *we* will be shown not to exist at all. Or, we will exist, but we won't be thinkers, we won't manifest genuine 'Wisdom in all the achievements of creative skill.' Whenever we zoom in on the act of creation, it seems we lose sight of it, the intelligence or genius replaced at the last instant by stupid machinery, an echo of Darwin's shocking substitution of Absolute Ignorance for Absolute Wisdom in the creation of the biosphere. Many people dislike Darwinism in their guts, and of all the ill-lit, murky reasons for antipathy to Darwinism, this one has always struck me as the deepest, but only in the sense of being the most entrenched, least accessible to rational criticism. There are thoughtful people who scoff at Creationism, dismiss dualism out of hand, pledge allegiance to academic humanism—and then get quite squirrely when it is suggested that a Darwinian theory of creative intelligence might be in the cards, and might demonstrate that all the works of human genius can be understood in the end to be products of a cascade of generate-and-test procedures that are, at bottom, algorithmic, mindless. Absolute Ignorance? Artificial Intelligence? Fie on anybody who would thus put 'A' and 'I' together!

Besides, wouldn't a Darwinian theory of human creativity be covertly self-contradictory? The Darwinian mechanism of natural selection is famously mindless, purposeless, lacking all foresight and intention—the blind watchmaker. (Dawkins 1986) If natural selection is 'the opposite' of God, a strange inversion of the traditional vision of creativity, then it must be 'the opposite' of us, too, since God is made in our image! Human creative endeavors are obviously both foresighted and purposeful, so, then, they are Darwinian processes. What could be more obvious?

But there is a tension, isn't there? A key part of Darwin's great revolution is that we are part of it.

Human beings are just one species among many, fully biological, and hence capable of no miracles, restricted to the same sorts of processes and methods as the other species. Our creative processes are surely natural (not supernatural!), so in that bland sense they are as biological as the creative processes of the weaverbird and the beaver.

William Poundstone (1985) puts the inescapable challenge succinctly in terms of 'the old fantasy of a monkey typing *Hamlet* by accident.' He calculates that the chances of this happening are '1 in 50 multiplied by itself 150,000 times.'

In view of this, it may seem remarkable that anything as complex as a text of *Hamlet* exists. The observation that *Hamlet* was written by Shakespeare and not some random agency only transfers the problem. Shakespeare, like everything else in the world, must have arisen (ultimately) from a homogeneous early universe. Any way you look at it *Hamlet* is a product of that primeval chaos.'

CREDIT ASSIGNMENT FOR CREATIVITY

Where does all that Design come from? What processes could conceivably yield such improbable 'achievements of creative skill'? What Darwin saw is that Design is always both valuable and costly. It does not fall like manna from heaven, but must be accumulated the hard way, by time-consuming, energy-consuming processes of mindless search through 'primeval chaos,' automatically preserving happy accidents when they occur. This broadband process of Research and Development is breathtakingly inefficient, but—this is Darwin's great insight—if the costly fruits of R and D can be thriftily conserved, copied, stolen, and re-used, they can be accumulated over time to yield 'the achievements of creative skill.' 'This principle of preservation, I have called, for the sake of brevity, Natural Selection.' (Darwin 1865)

There is no requirement in Darwin's vision that these R and D processes run everywhere and always at the same tempo, with the same (in-)efficiency. If we think of design work as lifting in Design Space (an extremely natural and oft-used metaphor, exploited in models of hill-climbing and peaks in adaptive landscapes, to name the most obvious and popular applications), then we can see that the gradualistic, frequently back-sliding, maximally inefficient basic search process can on important occasions yield new conditions that speed up the process, permitting faster, more effective local lifting. (Maynard Smith and Szathmary 1995) Call any such product of earlier R and D a *crane*, and distinguish it from what Darwinism says does not happen: *skyhooks*. (Dennett 1995) Skyhooks, like manna from heaven, would be miracles, and if we posit a skyhook anywhere in our 'explanation' of creativity, we have in fact conceded defeat.

What, then, is a mind? The Darwinian answer is straightforward. A mind is a crane, made of cranes, made of cranes, a mechanism of not quite unimaginable complexity that can clamber through Design Space at a giddy—but not miraculously giddy–pace, thanks to all the earlier R and D, from all sources, that it exploits. What is the anti-Darwinian answer? It is perfectly expressed by one of the 20th century's great creative geniuses (though, like MacKenzie, he probably didn't mean by his words what I intend to mean by them).

Je ne cherche pas; je trouve.—Pablo Picasso

Picasso purports to be a genius indeed, someone who does not need to engage in the menial work of trial and error, generate-and-test, R and D; he claims to be able to *leap* to the summits of the peaks—the excellent designs—in the vast reaches of Design Space without having to guide his trajectory (he searches not) by sidelong testing at any way stations. As an inspired bit of bragging, this is *non pareil*, but I don't believe it for a minute. And anyone who has strolled through an exhibit of Picasso drawings (as I recently did in Valencia, while attending the conference that led to this volume) looking at literally dozens of variations on a single theme, all signed—and sold—by the artist, will appreciate that whatever Picasso may have meant by his *bon mot*, he could not truly claim that he didn't engage in a time-consuming, energy-consuming exploration of neighborhoods in Design Space. At best he could claim that his own searches were so advanced, so efficient, that it didn't seem—to himself—to be design *work* at all. But then what did he have within him that made him such a great designer? A skyhook, or a superb collection of cranes? (I have been unable to discover the source of Picasso's claim, which is nicely balanced by a better known remark by a more down-to-earth creative genius, Thomas Edison: 'Genius is one per cent inspiration and ninety-nine per cent perspiration.' (1932))

We can now characterize a mutual suspicion between Darwinians and anti-Darwinians that distorts

the empirical investigation of creativity. Darwinians suspect their opponents of hankering after a skyhook, a miraculous gift of genius whose powers have no decomposition into mechanical operations, however complex and informed by earlier processes of R and D. Anti-Darwinians suspect their opponents of hankering after an account of creative processes that so diminishes the Finder, the Author, the Creator, that it disappears, at best a mere temporary locus of mindless differential replication. We can make a little progress, I think, by building on Poundstone's example of the creation of the creator of *Hamlet*. Consider, then, a little thought experiment.

Suppose Dr. Frankenstein designs and constructs a monster, Spakesheare, that thereupon sits up and writes out a play, *Spamlet*. Who is the author of *Spamlet*? First, let's take note of what I claim to be irrelevant in this thought experiment. I haven't said whether Spakesheare is a robot, constructed out of metal and silicon chips, or, like the original Frankenstein's monster, constructed out of human tissues—or cells, or proteins, or amino acids, or carbon atoms. As long as the design work and the construction were carried out by Dr. Frankenstein, it makes no difference to the example what the materials are. It might well turn out that the only way to build a robot small enough and fast enough and energy-efficient enough to sit on a stool and type out a play is to construct it from artificial cells filled with beautifully crafted motor proteins and other carbon-based nanorobots. That is an interesting technical and scientific question, but not of concern here. For exactly the same reason, if Spakesheare is a metal-and-silicon robot, it may be allowed to be larger than a galaxy, if that's what it takes to get the requisite complication into its program—and we'll just have to repeal the speed limit for light for the sake of our thought experiment. These technical constraints are commonly declared to be off-limits in these thought experiments, so so be it. If Dr. Frankenstein chooses to make his AI robot out of proteins and the like, that's his business. If his robot is cross-fertile with normal human beings and hence capable of creating what is arguably a new species by giving birth to a child, that is fascinating, but what we will be concerned with is Spakesheare's purported brainchild, *Spamlet*. Back to our question: Who is the author of *Spamlet*?

In order to get a grip on this question, we have to look inside and see what happens in Spakesheare. At one extreme, we find inside a file (if Spakesheare is a robot with a computer memory) or a basically memorized version of *Spamlet*, all loaded and ready to run. In such an extreme case, Dr. Frankenstein is surely the author of *Spamlet* (unless we find there is a Ms. Shelley who is the author of Dr. Frankenstein!), using his intermediate creation, Spakesheare, as a mere storage-and-delivery device, a particularly fancy word processor. All the R and D work was done earlier, and copied to Spakesheare by one means or another. Now look at the other extreme, in which Dr. Frankenstein leaves most of the work to Spakesheare. The most realistic scenario would surely be that Spakesheare has been equipped by Dr. Frankenstein with a virtual past, a lifetime stock of pseudo-memories of experiences on which to draw while responding to its Frankenstein-installed obsessive desire to write a play. Among those pseudo-memories, we may suppose, are many evenings at the theater, or reading books, but also some unrequited loves, some shocking close calls, some shameful betrayals and the like. Now what happens? Perhaps some scrap of a 'human interest' story on the network news will be the catalyst that spurs Spakesheare into a frenzy of generate-and-test, ransacking its memory for useful tidbits and themes, transforming—transposing, morphing—what it finds, jiggling the pieces into temporary, hopeful structures that compete for completion, most of them dismantled by the corrosive processes of criticism that nevertheless expose useful bits now and then, and so forth, and all of this multi-leveled search would be somewhat guided by multi-level, internally generated evaluations, including evaluation of the evaluation...of the evaluation functions as a response to evaluation of ...the products of the ongoing searches.

Now if the amazing Dr. Frankenstein had actually anticipated all this activity down to its finest grain at the most turbulent and chaotic level, and had hand-designed Spakesheare's virtual past, and all its search machinery, to yield just this product, *Spamlet*, then Dr. Frankenstein would be, once again, the author of *Spamlet*, but also, in a word, God. Such Vast (not literally infinite, but Very much more than Astronomical—Dennett 1995, p109) foreknowledge would be simply miraculous. Restoring a smidgen of realism to our fantasy, we can consider a rather less extreme position and assume that Dr. Frankenstein was unable to foresee all this in detail, but rather delegated to Spakesheare most of the hard work of completing the trajectory in

Design Space to one literary work or another, something to be determined by later R and D occurring within Spakesheare itself.

REAL ARTIFICIAL CREATORS

We have now arrived in the neighborhood of reality itself, for we already have actual examples of impressive artificial authors that vastly outstrip the foresight of their own creators. Nobody has yet created an artificial playwright worth serious attention, but an artificial chess player—IBM's Deep Blue—and an artificial composer—David Cope's EMI—have both achieved results that are, in some respects, equal to the best that human creative genius can muster.

Who beat Garry Kasparov, the reigning World Chess Champion? Not Murray Campbell or any of his IBM team. Deep Blue beat Kasparov. Deep Blue designs better chess games than any of them can design. None of them can author a winning game against Kasparov. Deep Blue can. Yes, but. Yes, but. I am sure many of you are tempted to insist at this point that when Deep Blue beats Kasparov at chess, its brute force search methods are entirely unlike the exploratory processes that Kasparov uses when he conjures up his chess moves. But that is simply not so—or at least it is not so in the only way that could make a difference to the context of this debate about the universality of the Darwinian perspective on creativity. Kasparov's brain is made of organic materials, and has an architecture importantly unlike that of Deep Blue, but it is still, so far as we know, a massively parallel search engine which has built up, over time, an outstanding array of heuristic pruning techniques that keep it from wasting time on unlikely branches. There is no doubt that the investment in R and D has a different profile in the two cases; Kasparov has methods of extracting good design principles from past games, so that he can recognize, and know enough to ignore, huge portions of the game space that Deep Blue must still patiently canvass *seriatim*. Kasparov's 'insight' dramatically changes the shape of the search he engages in, but it does not constitute 'an entirely different' means of creation. Whenever Deep Blue's exhaustive searches close off a type of avenue that it has some means of recognizing (a difficult, but not impossible task), it can re-use that R and D whenever it is appropriate, just as Kasparov does. Much of this analytical work has been done for Deep Blue by its designers, and given as an innate endowment, but Kasparov has likewise benefitted from hundreds of thousands of person-years of chess exploration transmitted to him by players, coaches and books. It is interesting in this regard to contemplate the suggestion recently made by Bobby Fischer, who proposes to restore the game of chess to its intended rational purity by requiring that the major pieces be randomly placed in the back row at the start of each game (random, but mirror image for black and white). This would instantly render the mountain of memorized openings almost entirely obsolete, for humans and machines alike, since only rarely would any of this lore come into play. One would be thrown back onto a reliance on fundamental principles; one would have to do more of the hard design work in real time—with the clock running. It is far from clear whether this change in rules would benefit human beings more than computers. It all depends on which type of chess player is relying most heavily on what is, in effect, rote memory—reliance with minimal comprehension on the R and D of earlier explorers.

The fact is that the search space for chess is too big for even Deep Blue to explore exhaustively in real time, so like Kasparov, it prunes its search trees by taking calculated risks, and like Kasparov, it often gets these risks pre-calculated. Both presumably do massive amounts of 'brute force' computation on their very different architectures. After all, what do neurons know about chess? Any work they do must be brute force work of one sort or another.

It may seem that I am begging the question in favor of a computational, AI approach by describing the work done by Kasparov's brain in this way, but the work has to be done somehow, and no other way of getting the work done has ever been articulated. It won't do to say that Kasparov uses 'insight' or 'intuition' since that just means that Kasparov himself has no privileged access, no insight, into how the good results come to him. So, since nobody knows how Kasparov's brain does it—least of all Kasparov—there is not yet any evidence at all to support the claim that Kasparov's means are 'entirely unlike' the means exploited by Deep Blue. One should remember this when tempted to insist that 'of course' Kasparov's methods are hugely different. What on earth could provoke one to go out on a limb like that? Wishful thinking? Fear?

But that's just chess, you say, not art. Chess is trivial compared to art (now that the world champion chess player is a computer). This is where David Cope's EMI

comes into play. (Cope 2001; Dennett 2001c) Cope set out to create a mere efficiency-enhancer, a composer's aid to help him over the blockades of composition any creator confronts, a high-tech extension of the traditional search vehicles (the piano, staff paper, the tape recorder, etc.). As EMI grew in competence, it promoted itself into a whole composer, incorporating more and more of the generate-and-test process. When EMI is fed music by Bach, it responds by generating musical compositions in the style of Bach. When given Mozart, or Schubert, or Puccini, or Scott Joplin, it readily analyzes their styles and composes new music in their styles, better pastiches than Cope himself—or almost any human composer—can compose. When fed music by two composers, it can promptly compose pieces that eerily unite their styles, and when fed, all at once (with no clearing of the palate, you might say) all these styles at once, it proceeds to write music based on the totality of its musical experience. The compositions that result can then also be fed back into it, over and over, along with whatever other music comes along in MIDI format, and the result is EMI's own 'personal' musical style, a style that candidly reveals its debts to the masters, while being an unquestionably idiosyncratic integration of all this 'experience.' EMI can now compose not just two-part inventions and art songs but whole symphonies—and has composed over a thousand, when last I heard. They are good enough to fool experts (composers and professors of music) and I can personally attest to the fact that an EMI-Puccini aria brought a lump to my throat—but then, I'm on a hair trigger when it comes to Puccini, and this was a good enough imitation to fool me. David Cope can no more claim to be the composer of EMI's symphonies and motets and art songs than Murray Campbell can claim to have beaten Kasparov in chess.

To a Darwinian, this new element in the cascade of cranes is simply the latest in a long history, and we should recognize that the boundary between authors and their artifacts should be just as penetrable as all the other boundaries in the cascade. When Richard Dawkins (1982) notes that the beaver's dam is as much a part of the beaver phenotype—its extended phenotype—as its teeth and its fur, he sets the stage for the further observation that the boundaries of a human author are exactly as amenable to extension. In fact, of course, we've known this for centuries, and have carpentered various semi-stable conventions for dealing with the products of Rubens, of Rubens' studio, of Rubens' various students. Wherever there can be a helping hand, we can raise the question of just who is helping whom, what is creator and what is creation. How should we deal with such questions? To the extent that anti-Darwinians simply want us to preserve some tradition of authorship, to have some rules of thumb for determining who or what shall receive the honor (or blame) that attends authorship, their desires can be acknowledged and met, one way or another (which doesn't necessarily mean we should meet them). To the extent that this is not enough for the anti-Darwinians, to the extent that they want to hold out for authors as an objective, metaphysically grounded, 'natural kind,' they are looking for a skyhook.

DOES THE AUTHOR DISAPPEAR?

There is a persistent problem of imagination management in the debates surrounding this issue: people on both sides have a tendency to underestimate the resources of Darwinism, imagining simplistic alternatives that do not exhaust the space of possibilities. Darwinians are notoriously quick to find (or invent) differences in genetic fitness to go with every difference they observe, for instance. Meanwhile, anti-Darwinians, noting the huge distance between a beehive and the St. Matthew Passion as created objects, are apt to suppose that anybody who proposes to explain both creative processes with a single set of principles must be guilty of one reductionist fantasy or another: 'Bach had a gene for writing baroque counterpoint just like the bees' gene for forming wax hexagons' or 'Bach was just a mindless trial-and-error mutator and selector of the musical memes that already flourished in his cultural environment.' Both of these alternatives are nonsense, of course, but pointing out their flaws does nothing to support the idea that ('therefore') there must be irreducibly non-Darwinian principles at work in any account of Bach's creativity. In place of this dimly imagined chasm with 'Darwinian phenomena' on one side and 'non-Darwinian phenomena' on the other side, we need to learn to see the space between bee and Bach as populated with all manner of mixed cases, differing from their nearest neighbors in barely perceptible ways, replacing the chasm with a traversable gradient of non-minds, protominds, hemi-demi-semi minds, magpie minds, copycat minds, aping minds, clever-pastiche minds, 'path-finding' minds, 'ground-breaking' minds,

and eventually, genius minds. And the individual minds, of each caliber, will themselves be composed of different sorts of parts, including, surely, some special-purpose 'modules' adapted to various new tricks and tasks, as well as a cascade of higher-order reflection devices, capable of generating ever more rarefied and delimited searches through pre-selected regions of the Vast space of possible designs.

It is important to recognize that genius is itself a product of natural selection and involves generate-and-test procedures all the way down. Once you have such a product, it is often no longer particularly perspicuous to view it solely as a cascade of generate-and-test processes. It often makes good sense to leap ahead on a narrative course, thinking of the agent as a self, with a variety of projects, goals, presuppositions, hopes....In short, it often makes good sense to adopt the intentional stance (Dennett, 1971, 1987) towards the whole complex product of evolutionary processes. This effectively brackets the largely unknown and unknowable mechanical microprocesses as well as the history that set them up, and puts them out of focus while highlighting the patterns of rational activity that those mechanical microprocesses track so closely. This tactic makes especially good sense to the creator himself or herself, who must learn not to be oppressed by the revelation that on close inspection, even on close introspection, a genius dissolves into a pack rat, which dissolves in turn into a collection of trial-and-error processes over which nobody has ultimate control.

Does this realization amount to a loss—an elimination—of selfhood, of genius, of creativity? Those who are closest to the issue—the artistic and scientific geniuses who have reflected on it—often confront this discovery with equanimity. Mozart (in an oft-quoted but possibly spurious passage—see Dennett 1995, p346-7) is reputed to have said of his best musical ideas: 'Whence and how do they come? I don't know and I have nothing to do with it.' The painter Philip Guston is equally unperturbed by this evaporation of visible self when the creative juices start flowing:

> When I first come into the studio to work, there is this noisy crowd which follows me there; it includes all of the important painters in history, all of my contemporaries, all the art critics, etc. As I become involved in the work, one by one, they all leave. If I'm lucky, every one of them will disappear. If I'm really lucky, I will too.

Resistance to extending Darwinian thinking into human creativity and human culture is not restricted to closet Creationists and anti-scientific humanists. Two highly visible Darwinian spokespersons—Stephen Jay Gould and Steven Pinker—who agree on precious little else, find common ground in their doubts about this:

> I am convinced that comparisons between biological evolution and human cultural or technological change have done vastly more harm than good—and examples abound of this most common of intellectual traps....Biological evolution is powered by natural selection, cultural evolution by a different set of principles that I understand but dimly. (Gould 1991, p63)
>
> To say that cultural evolution is Lamarckian is to confess that one has no idea how it works. The striking features of cultural products, namely their ingenuity, beauty, and truth (analogous to organisms' complex adaptive design), come from the mental computations that 'direct'—that is, invent—the 'mutations,' and that 'acquire'—that is, understand—the 'characteristics.' (Pinker 1997, p209)

Pinker has imputed the wrong parallel; it is not Lamarck's model but Darwin's models of unconscious and methodical (artificial) selection (as special cases of natural selection) that accommodate the phenomena he draws to our attention in this passage (Dennett 2001b). And it is ironic that Pinker overlooks this, since the cultural phenomena he himself has highlighted as examples of evolution-designed systems, linguistic phenomena, are almost certainly not the products of foresightful, ingenious, deliberate human invention. Some designed features of human languages are no doubt genetically transmitted, but many others—such as changes in pronunciation, for instance—are surely culturally transmitted, and hence products of cultural, not genetic, evolution.

CONCLUSION

The cranes of human culture didn't just open up Design Space; they opened up perspectives on Design Space that permitted 'directed' mutation, foresighted mutation, reflective mutation, both in cultural and, most recently, genetic innovation. This nesting of different processes of natural selection now has a new member: genetic engineering. How does it differ from the methodical selection of Darwin's day? It is less dependent on the pre-existing variation in the gene pool, and

proceeds more directly to new candidate genomes, with less overt trial and error. Darwin (1865, p38) had noted that in his day, 'Man can hardly select, or only with much difficulty, any deviation of structure excepting such as is externally visible; and indeed he rarely cares for what is internal.'

But today's genetic engineers have carried their insight into the molecular innards of the organisms they are trying to create. There is ever more accurate foresight, but even here, if we look closely at the practices in the laboratory, we will find a large measure of exploratory trial and error in their search of the best combinations of genes. (In fact, biochemists and molecular biologists are finding that artificial evolutionary processes are more efficient R and D procedures than their foresightful hand-work efforts by orders of magnitude. In other words, they are finding that the *breeding* of domesticated micro-organisms and polymers is the best way to conduct their creative searches.)

Are the products of genetic engineering 'Darwinian' products? They are produced not by blind or random trial-and-error variation, but by highly intelligent, guided, foresightful processes. Nevertheless these processes are themselves the products of earlier design work accomplished by Darwinian R and D, and if we look closely at the microprocesses that compose their current, local search, we will still find plenty of instances of random (undesigned, chaotic) generation of candidates for further scrutiny.

It may seem, however, that we have now passed the Pickwickian limits of Darwinian orthodoxy. Does a Darwinian gloss actually supplement or adjust the traditional intellectualist ways of thinking? I think it does, because without the steady pressure of the Darwinian 'strange inversion of reasoning,' it is all too tempting to revert to the old essentialist, Cartesian perspectives. For instance, there is always the temptation, often succumbed to, to establish 'principled' boundaries, or to erect a polar contrast between insightful and blind processes of search, as we saw in the unsupportable assertion that Kasparov's methods are fundamentally unlike Deep Blue's. If Deep Blue's methods are ultimately 'blind and mechanical,' then so, ultimately, are Kasparov's—his neurons are as blind and mechanical as any circuit board. The foresighted, purposeful breeding of domesticated plants and animals is obviously not a damning counterexample to Darwin's theory of natural selection as a foresightless, purposeless process, because his theory shows (as we are beginning to learn) how such foresight and purpose could itself evolve by blind natural selection. Kasparov's creative genius (or Bach's or Shakespeare's) is for the same reason no counterexample to the Darwinian theory of creativity, but rather one of the most recent blooms on the tree of life that we still need to account for in Darwinian terms.

FURTHER READING

Campbell, D.T. 1960. Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review* 67: 380-400.

Mithen, S. J. 1998, ed. *Creativity in human evolution and prehistory*. London; New York: Routledge, 1998.

Mithen, S. J. 1996. *The Prehistory of the Mind: The Cognitive Origins of Art, Religion and Science*, Cambridge: Cambridge University Press.

Toulmin, Stephen. 1972. *Human Understanding. Volume 1: The Collective Use and Development of Concepts*. Oxford: Clarendon Press

ACKNOWLEDGMENTS

Portions of this paper are drawn from Dennett, 2001a. I have been unable to locate the source of Philip Guston's quote, but I have found much the same remark attributed to the composer, John Cage, a close friend and contemporary of Guston's, who (is said to have) said this about painting: 'When you are working, everybody is in your studio?the past, your friends, the art world, and above all, your own ideas?all are there. But as you continue painting, they start leaving, one by one, and you are left completely alone. Then, if you are lucky, even you leave.' Like all other creators, Guston and I like to re-use what we find, adding a few touches from time to time.

REFERENCES

Cope, D., 2001, *Virtual Music: Computer Synthesis of Musical Style*. Cambridge, MA: MIT Press.

Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection*. London: Murray. (page numbers as in Harvard Univ. Press facsimile edn.)

Dawkins, R. 1982. *The Extended Phenotype*. Oxford and San Francisco: Freeman.

Dawkins, R. 1986. *The Blind Watchmaker*, London: Longmans.

Dennett, D. C. 1971, Intentional Systems, *Journal of Philosophy*, 68, pp87-106.

Dennett, D. C. 1987, *The Intentional Stance*, MIT Press, Cambridge, MA.

Dennett, D. C. 1995. *Darwin's Dangerous Idea*. New York: Simon & Schuster.

Dennett, D. C. 2001a. In Darwin's Wake, Where am I? (Presidential Address, Eastern Division, American Philosophical Association, December 29, 2000) in *Proceedings of the APA*, and also in *Cambridge Companion to Darwin*, ed. Radick et al., Cambridge: Cambridge Univ. Press.

Dennett, D. C. 2001b, The Evolution of Culture, in *The Monist* 84, no. 3, Peru, Illinois (special issue on cultural evolution, ed. D. Sperber), pp 305-324.

Dennett, D. 2001c, Collision Detection, Muselot, and Scribble: Some Reflections on Creativity, in Cope 2001, pp283-291.

Edison, T. 1932, interview in *Life*, ch. 24 (according to the *Oxford Dictionary of Quotations*).

Gould, S. J. 1991. *Bully for Brontosaurus,* New York: Norton.

Mackenzie, R. B. 1868. *The Darwinian Theory of the Transmutation of Species Examined*. (published anonymously 'By a Graduate of the University of Cambridge') London: Nisbet & Co. (Quoted in a review, *Athenaeum*, no 2102. Feb 8, p217.).

Maynard Smith, J. and Szathmary, E. 1995. *The Major Transitions in Evolution.* Oxford: Freeman.

Pinker, S., 1997. *How the Mind Works*. New York: Norton.

Poundstone, W. 1985. *The Recursive Universe: Cosmic Complexity and the Limits of Scientific Knowledge*. New York: Morrow.

From Typo to Thinko: When Evolution Graduated to Semantic Norms

"Natural selection edits with an eye only toward what the message says, not to what it means." Richard Powers, *The Gold Bug Variations,* p546

1. Darwinian Perspectives on Culture

Do we need a Darwinian theory of cultural evolution? In one sense, certainly. It is obvious that there are patterns of cultural *change*—evolution in the neutral sense—and any theory of cultural change worth more than a moment's consideration will have to be Darwinian in the minimal sense of being *consistent with* the theory of evolution by natural selection of *Homo sapiens.* Our species name is well chosen, and it is culture that makes us the knowing hominid, so a minimally Darwinian theory of culture must hold that the phenotypic traits that make cumulative culture possible—mainly, language and the habits of sociality—evolved by natural selection, unaided by what I call *skyhooks*: saltations in Design Space that could not be the outcome of standard evolutionary mechanisms (Dennett, 1995). This minimal Darwinism is simply the denial of the hypothesis that culture is, as it were, a miracle, a gift from God.[1] It maintains in one way or another that natural selection eventually provided the foundations for culture, which then took off, and elaborated itself under some regime that explains the patterns of cultural change, but that regime need not itself be Darwinian in any interesting sense.

For instance, the standard model is an economic one: the theorist says, in effect, if Darwin will give me *Homo economicus*, a social group of rational, self-interested individuals getting and spending, saving and making and trading, I can then use the intentional stance (Dennett, 1971, 1987) as the explanatory framework for describing and accounting for the patterns of cultural evolution. This economic model is used not just by economists, of course; it is tacitly presupposed by historians and anthropologists and all the other theorists who treat culture as composed of goods, possessions of the people, who husband them in various ways, wisely or foolishly. People carefully preserve their traditions of fire-lighting, house-building, speaking, counting, justice, etc. They trade cultural items as they trade other goods. And of course some cultural items (wagons, pasta, recipes for chocolate cake, etc.) are definitely goods, and so we can plot their trajectories using the tools of economics. It is clear from this perspective that highly prized cultural entities will be protected at the expense of less favored cultural entities, and there will be a competitive market where agents both "buy" and "sell" cultural wares. If a new method of house-building or farming or a new style of music sweeps through the culture, it will be because people perceive advantages to these novelties and choose them. If Coca Cola bottles proliferate around the world, it is because more and more people prefer to buy a Coke. Advertising may fool them. But then we look to the advertisers, or those who have hired them, to find the relevant agents whose desires fix the values for our cost-benefit calculations. *Cui bono*? Who benefits? The purveyors of the goods, and those they hire to help them, etc. On this way of thinking, then, the relative "replicative" power of various cultural goods—whether Coke bottles, building styles or religious creeds—is measured in the marketplace of cost-benefit calculations performed, consciously or unconsciously, by the people.

Biologists, too, make good use of the economic model, explaining the evolution (in the neutral sense) of features of the natural world by treating them as goods belonging to various members of various species: one's food, one's nest, one's burrow, one's territory, one's mates, one's time and energy. Cost-benefit analyses shed light on the husbandry engaged in by the members of the different species inhabiting some shared environment.[2] Not every "possession" is considered a good, however. The dirt and grime that accumulates on one's body, to say nothing of the accompanying flies and fleas, are of no value, or of negative value, for instance. These hitchhikers are not normally considered as goods by biologists, except when the benefits derived from them are manifest.

These economic models of culture are consistent with

Darwinism but not more specifically Darwinian. Darwinian evolution provides organisms whose ultimate goal is self-replication, and who then track the rational if myopic trajectory to that end; the interaction of such rational agents determines which features of the shared environment will proliferate, which be contested, which despoiled, and so forth. In these models, cultural traits, however they arise, spread as fitness-enhancers, at least locally considered. Agriculture, cooking, clothing, the wheel, writing, bow and arrow—all these cultural innovations are plausibly viewed as improvements that need not arise from gene mutation and recombination and need not be transmitted genetically. They are, one might say, infectious phenotypic features. These features, it is presumed, pass some sort of quality-control test administered by the agents themselves. They are chosen by evolved organisms and put to use, and if they didn't "pay for themselves" in a fitness boost (or at least an apparent fitness boost, myopically considered) they would soon die out, just like genetically transmitted phenotypic variations. The idea is that if innovations are randomly distributed around neutrality, the pernicious innovations will hasten the demise (or mating failure) of those who adopt them, and the enhancements will do the opposite, and over the not very long haul the enhancers will proliferate. This vision allows the possibility, as relatively rare outliers, of mistakes: either good tricks abandoned by mistake or pernicious tricks persisted in thanks to some local illusion.

More ambitious models (Feldman and Cavalli-Sforza, 1981; Boyd and Richerson, 1985) then address the opportunities for co-evolution, for interaction between the items that come to be present or dominant in the cultural marketplace and the genetically transmitted phenotypes of those transmitting and preserving those items. Clothes do make the man, at least to the extent of diverting selection pressures for weather-hardiness, so the cultural transmission of clothing sends ripples through the evolution of human physiology. Similarly, new practices of food gathering and preparation can reflexively change the fitness landscape for digestive prowess. Lactose tolerance in adults descended from people who herd dairy animals is a well-studied parade case. These models are Darwinian in a more than minimal sense because they extend the perspective of population genetics, the replicator dynamics, to these non-genetically transmitted phenotypic features, exploring the effects of horizontal and oblique transmission, for instance. But they also maintain the basic economic presupposition of agent rationality: cultural traits are *adopted* because they are *deemed worth having*, because they are supposed, rightly or wrongly, to contribute somehow to the achievement on one's more ulterior ends, whatever they be.

When a rational agent or intentional system makes a decision about which is the best course of action, all things considered, we need to know from whose perspective this optimality is being judged. Here things begin to get messy. In nature, genes are the ultimate units of "self"-interest. That is to say, adaptations in plants and animals (and simpler life forms) are, by definition, features that further the *summum bonum* of gene replication, directly or indirectly. Are cultural innovations adaptations in this narrow, genetic sense? It is obvious that many cultural features are deemed by the populace to be beneficial, functional, adaptive, useful, life-enhancing, enabling. It is less clear that these esteemed features play a discernible role in enhancing *genetic fitness,* as contrasted with, say, human happiness, the pursuit of which is curiously orthogonal to genetic fitness. One of the striking trends in human evolution, going back thousands of years, is the gradual diminution in the proportion of human effort devoted in any clearly discernible way to the achievement of the fundamental goals we share with animals: avoiding pain, hunger, and predation, and seeking comfort, security, and mating opportunities. Even if the peculiar human *desiderata* of prestige, power, wealth, beautiful surroundings, recreation, music, toys, … have discernible instrumental rationales (improving one's profile in the contest for mates, enlarging one's harem, one's territory, one's margin of error), they have more or less detached themselves from these inaugural foundations and become ends in themselves. The young man bought the guitar in order to attract young women, but now he has become a guitarist who would rather make music than love.

As Cavalli-Sforza and Feldman (1981) note,

There are people determined to risk their life to reach the top of Mt. Everest, and others that spend their life accumulating money, or attempting artistic or scientific creations, or simply trying to do as little as possible. It is difficult to subsume all of these choices under a common schedule admitting no individual variation. (p342)

As they put it,

Control is delegated to a system of poorly understood internal drives and rewards that direct the activity of the individual, . . . Our very inadequate knowledge of this steering system prevents us from making finer statements, but it is probably true that the system's overall activity is directed towards maximizing self satisfaction of the individual. Important complications arise because we can satisfy ourselves in many different, competing ways, many of which demand careful advance planning. (p364)

Feldman and Cavalli-Sforza thus adopt the default assumption, at least in the Western world, and especially among economists, of treating the agent as a sort of punctate, Cartesian locus of well-being. What's in it for *me?* Rational *self*-interest. But while there has to be something in the role of the self—something that defines the answer to the *Cui bono* question for the decision-maker under examination, there is no necessity in this default treatment, common as it is. A self-as-ultimate-beneficiary can in principle be indefinitely distributed. I can care for others, or for a larger social structure, for instance. There is nothing that restricts me to a *me* as contrasted to an *us.* I can still take my task to be looking out for #1 while including, under #1, not just myself, but my family, the Chicago Bulls, Oxfam, the flourishing of mid-twentieth-century acoustic guitar finger-picking techniques . . . . anything human ingenuity can define and become attached to, making its welfare definitive of the decision-maker's "happiness."

It is not obvious that any other organism strives for its own happiness or anything like it. If human happiness is our *summum bonum* (or at least a *bonum* against which we do in fact often attempt to measure costs and benefits), how did it arise? It is here that the prospect of a still more radically Darwinian theory of cultural evolution becomes attractive. Could the unique varieties of human evaluation that are so distantly and indirectly anchored to any plausible litmus test of genetic fitness be accounted for by supposing that human beings have evolved into a condition where they have become the vectors, the hosts, of a new order of symbionts, competing cultural replicators whose *own* fitness, defined in standard Darwinian terms of relative replicative success, has constituted a new sort of entity? An enculturated human being, according to this proposal, is as important a novelty on the evolutionary front as the eukaryotic cell was at its debut: a unification of distinct replicators into a synthetic organization with a displaced goal or *summum bonum*, no longer just an *organism* bent on self-replication, but a *person*, bent on furthering the particular goals and ideals with which that person identifies. Has our guitarist unwittingly become part of a guitar's way of making another guitar? This is a tantalizingly attractive idea, but for such a perspective to anchor itself firmly in evolutionary theory, we must take seriously Dawkins' concept of a meme, and there are reasons to doubt that we should.

2. Cultural Replicators: A Central or Peripheral Phenomenon?

In some neighborhoods, ball bearings outnumber rabbit turds; in other neighborhoods, this imbalance is reversed, and in yet other neighborhoods, no entities of either variety are to be found. These differential production patterns change over time, and there are reasons for them, but they are not, in the main, Darwinian reasons. Not all production is replication, and not all distribution is emigration. Variety and similarity are also found among cultural items, and the question is: are any (or many) of the reasons for patterns in changing "populations" of cultural items Darwinian?

Dan Sperber (this volume) notes that the dictionary definition of "memes" is too bland to be of much interest: "an element of culture that may be considered to be passed on by non-genetic means," while a more radical definition, more faithful to Dawkins' arresting proposal, "cultural replicators propagated through imitation," is far from obvious. Indeed, it is in need of defense against two objections. The "simplest and most serious objection" is that the copying of cultural items is in general too low-fidelity to permit natural selection to get a purchase. Compare memes to viruses. Viruses travel light, and carry no copy-machines, so they reproduce by entering cells and tricking the cell's proprietary copy-machines into making spurious copies of them instead of copying their usual and proper fare, the cell's own DNA. If memes are like viruses, as Dawkins and other would-be memeticists have claimed, it is because they reproduce by exploiting the copy-machinery resident in the brains of human beings. But how well does this parallel hold? How good is that machinery? Not good enough, it seems. We human beings are actually rather bad at the sorts of "mindless" copying that cells excel at.

Following Williams (1966), Sperber notes that although higher selection biases can tolerate lower fidelity, it still appears that "mutation" rates among memes would be so great that any description of the emerging patterns in terms of descent with modification, as Darwin put it, would be lip-service only. "For memetics to be a reasonable research program, it should be the case that copying, and differential success in causing the multiplication of copies, overwhelmingly plays the major role in shaping all or at least most of the contents of culture." If we are not inveterate and talented copyists, we will make poor hosts for our cultural viruses, and Darwinian descent with (relatively rare) modification will seldom occur. We will need to look elsewhere to explain the patterns of culture.

But is it so clear that our copying is too low-fidelity to work? Dawkins (1999) has responded to this objection with his example of the origami model of a Chinese junk, which people learn to make by following a canonical set of simple "self-normalizing" instructions, but Sperber finds this misleading, since "the normalization of the instructions results precisely from the fact that something other than copying is taking place." Sperber lays down three conditions for "true replication":

For B to be a replication of A,

(1) B must be caused by A (together with background conditions).

(2) B must be similar in relevant respects to A.

(3) The process that generates B must obtain the information that makes B similar to A from A.

It is condition (3), Sperber claims, that is seldom met by cultural transmissions. Infectious laughter is his excellent example of a transmission event that meets (1) and (2) and fails (3), and he extends his analysis of this case by the fanciful example of the ten sound-recorders that trigger each other, but whose productions, in one case, do not consist of replications, but rather of *recognitions*, followed by *re-productions*. Triggered reproduction of this sort is distinct from copying or replication in the one way that matters for Darwin, according to Sperber: it does not slavishly copy the original; instead it is inspired by the original to make another of the same sort—but *without* any systematic attempt or disposition to reproduce any idiosyncrasies of the original. It normalizes its production to an independent ideal, discarding or not even noticing any mutations, good or bad, in the original.

Sperber illustrates this point with another fine example, the contrast between the nonsense scribble and the five-pointed star. The nonsense scribble would degenerate quickly in any series of attempted replications because people are not good copiers of such productions, while the five-pointed star would be "copied" with high fidelity, just as Dawkins says, but, Sperber maintains, the succession of stars would not really be copies of their predecessors, since the "copyists" would normalize to the recipe for the drawing procedure, ignoring the details of the individual productions. But is Sperber looking at the right grain level? Dawkins' point is that a finite repertory of such triggered reproductions is not just a good trick for human beings who want to heighten their transmission fidelity; it is a Good Trick already discovered, several billion years ago, by natural selection. Sperber distinguishes copying from "merely triggering the production of a similar effect," but a repertory of such triggers, called an alphabet, is just what makes high-fidelity copying possible, both in cells and in human culture.

Suppose Tommy writes the letters "SePERaTE" on the blackboard, and Billy "copies" it by writing "seperate." Is this copying or triggered reproduction? The normalization to all lower-case letters shows that Billy is not slavishly copying Tommy's chalk-marks but rather being triggered to execute a series of canonical, normalized acts: *make an "s," make an "e,"* etc. It is thanks to these letter-norms that Billy can "copy" Tommy's word at all. And he *does* copy Tommy's spelling error, unlike Molly, who "copies" Tommy by writing "separate," responding to a higher norm, at the level of word spelling. Sally then goes a step higher, "copying" the phrase "separate butt equal"—all words in good standing in the dictionary—as "separate but equal," responding to a recognized norm at the phrase level. Can we go higher? Of course. Anybody who, when "copying" the line in the recipe "Separate three eggs and beat the yolks until they form stiff white cones," would replace "yolks" with "whites" knows enough about cooking to recognize the error and correct it. Above spelling and syntactic norms are a host of semantic norms as well.

DNA has an alphabet, the famous ACGT, and words, the three-letter codons that "spell" the twenty amino acids. In fact, the high-fidelity of genetic transmission depends on the sub-cellular machinery being triggered to "recognize" and "re-produce" a small repertoire of types, whose idiosyncrasies, if any, are ignored, not slavishly copied: "*make a cytosine,*" "*make a guanine*," etc. There are error-correcting enzymes as well, but they don't ascend (so far as we know!) above the level of a spell-checker, correcting "typos" by brute template-matching against the original.[3]

Does the human capacity (and irresistible disposition) to respond to higher, semantic norms—our capacity to correct not only typos but what hackers call *thinkos*—rule out cultural transmission as a candidate for natural selection? Sperber seems to think it does. "Contrary to what Dawkins writes," he claims, "the instructions are not 'self-normalizing.' It is the process of attribution of intentions that normalizes the implicit instructions that participants infer from what they observe." (ms p8) Sperber is partly right: the attribution of intentions is the key difference between this sort of human transmission and genetic replication. The point comes out even more clearly if we mutate Sperber's example slightly, adding a point to his star: Consider the fact that there are two distinct recipes (and many other less obvious ones) for making a regular six-pointed star:

(A) Make a regular hexagon and put equilateral triangle points on each side.

(B) Make an equilateral triangle and then superimpose on it another one, upside-down.

A series of six-pointed-star "replications" might be accomplished by a random alternation between these two recipes with no loss of fidelity. Which recipe did various individual copyists follow? It wouldn't matter, because what is being copied is not the recipe but the result *understood as an intended object having certain features.*[4]

Sperber thinks that this reliance on attribution of intentions on the part of the copyists disqualifies cultural transmission as a Darwinian process of natural selection. He supposes that this invocation of intelligent, semantically sensitive, intention-attributing agents in the purported replication process flies in the face of a fundamental requirement of Darwinian processes: mindless, purposeless mechanicity. He is almost right. To see the force of this interesting objection, imagine a creationist variant on standard neo-Darwinian genetic evolution. It postulates that God watches over each moment of DNA replication, and whenever He sees some copying that is "wrong" (relative to God's great plan), he undoes it. Thus when He chooses, He lets mutations flourish, and when He does not, those mutations get corrected by a gentle miraculous nudge of the error-correcting enzymes. Here Intelligence is playing a guiding role in evolution—just the sort of role that orthodox (devout, "fundamentalist") Darwinians abjure. As Richard Powers (199x) has observed, "Natural selection edits with an eye only toward what the message says, not to what it means." Clever human beings, in contrast, edit with an eye toward meaning. If such clever editors are inserted into the process of cultural transmission and revision, what would be left of a Darwinian theory of culture?

This worry ignores the fact that *Homo sapiens* is not itself a miracle, a skyhook, but something that has evolved by non-miraculous natural selection; its capacity to respond to semantic norms is itself something that has evolved under a regime that could not respond to semantic norms. Before there could be eyes, good for distal perception, there had to be mere photo-sensitive responders to proximal stimulation, out of which eyes could gradually be built. Before there could be minds, good for semantic discrimination, there had to be copying machinery that could only discriminate alphabet letters. Put otherwise, DNA error-correcting enzymes have always responded to semantic norms, but just *local* or *proximal* semantic norms—*make a "G"*—as contrasted to more *distal* semantic norms: *make a codon for asperagine or make some lysosyme or make a protein that blocks seratonin uptake, or even make something that will fight off infection.*

Why shouldn't evolution go right on working once the copying machinery graduates to less myopic norms? Even our lowest-level mindless copying avails itself of correction to a norm; is there a "highest permissible" level of normalization in any Darwinian process? Darwin (and Fisher, and Williams, and others) saw the need for a sufficiently "strong principle of inheritance" to keep evolution going, but nothing has been said about how that fidelity is to be maintained, mechanical-

ly. Let there be copyists that take themselves to be responding to semantic norms; there will still be a suitably long-distance evolutionary perspective from which their copying efforts, for all their editorial work, will appear myopic and unwitting, oblivious to—and hence unresponsive to—the larger scale pattern of differential replication that ensures that a Darwinian process is occurring.

In "Pierre Menard, Author of the *Quixote*" (1962) Jorge Luis Borges tells the fanciful tale of a literary theorist who sets out to *compose* (not copy, not write from memory) Cervantes' great work anew in the 20th century. This will be an act of bizarre self-control, since Menard is a Cervantes scholar who no doubt has at least large portions of the text of *Don Quixote* committed to memory, but Menard is determined to bracket that memory and *create*, with his own authorial intentions, all of Cervantes's sentences anew, like an experienced wheelwright setting out to *re-invent* the wheel! He succeeds (though how can he tell?), and Borges tells us: "Cervantes' text and Menard's are verbally identical, but the second is almost infinitely richer." (p42) In one sense, Menard did not *copy* or *memorize* Cervantes' text, but in another sense, he did, in spite of his virtuoso self-control, his obsessive act of re-creation. He did, because, as Sperber puts it, "(3) The process that generates B must obtain the information that makes B similar to A from A," and surely Menard's prior study of Cervantes' text is an essential part of the scholarship that permits him to "compose" *Don Quixote* anew. Of course Menard has used a lot of other information as well; the surplus is presumably what permits him, unlike an ordinary reader, to claim to have re-composed, not written down from brute memory, the work. But so what? According to Borges, the texts are "verbally identical" so high-fidelity reproduction has occurred. Imagine a world in which Menards abound, devoting their lives to the re-composition of their favorite works. The transmission of texts will proceed just fine in such a world—as fine as if photocopy machines were the underlying machinery.

In fact, of course, a pastel version of that fantasy is just what has happened in the transmission of ancient texts in our world, for seldom if ever have the scribes taking dictation been entirely uncomprehending of the words they were dutifully "copying," and so they have willy-nilly "corrected" whatever they heard in the process of transcribing it. Their corrections have been governed by several levels of norms: orthographic/lexical, syntactic, and finally semantical. The imaginary Menard can be conceived to have "transcribed" the entire poem of Cervantes *modulo* the "semantic norm" of the whole text. Most of us lack that highly sophisticated norm; we tend to fall back on our sense of the gist of such a narrative, or when all else fails, rote memory or parroting (but even "parroting" is not like a parrot's parroting—unless it is, as it very seldom is, a matter of reiterating formulae in a language we don't understand).

When Sperber notes that in cultural transmission "the information provided by the stimulus is complemented with information already in the system" he is right, but the same is true of DNA replication. The main difference, so far as I can see, is that unlike DNA replication, human cultural replication is accomplished by processes of highly variable semantic depth, responding to perceived (and mis-perceived) "copying" errors relative to norms at many levels. The alphabets of written languages provide us with the most vivid and best understood system of such norms of replication, but the phenomenon of semantic norms is not directly tied to language. Musical notation relies on the staff to digitalize the roughly inked spots, so that a musician can see at a glance that a chord is A-C#-E-G even though the A is written almost twice as far beneath the staff as it is "supposed" to be. A sketch of a new sort of axle for a wagon need not make the wheels exactly round; the user of the sketch will recognize those irregular closed curves as representations of wheels, which are to be round, of course. As we move through our various apprenticeships in life, we learn to perceive new families of categories—new alphabets, in an extended sense—from which to construct high-fidelity copies. Only a skilled potter can see at a glance what another potter is doing and copy it, or teach it to others.

Consider a chef demonstrating the making of a sauce to an apprentice. The *description in words* might be "deglaze the pan, reduce the sauce and thicken with cornstarch," but the words aren't really necessary, if the apprentice appreciates the goal of each process. Here is a stack of three analog processes, none of which could be exactly copied (in the manner of Sperber's figure A) by the apprentice; the cook didn't measure the water he sloshed into the pan, didn't time the reduction period, and added cornstarch freehand until the sauce took

on the desired consistency. But the recipe can be transmitted faithfully, all the same, thanks to the shared norms for these analog processes, already inculcated in the apprentice.

But if the error-correction in the case of memes is semantically appreciative, at many levels, doesn't this show that cultural evolution is NOT a mindless algorithm but rather a system that must invoke high-level semantic comprehension at every juncture? The variable depth of semantic norms does guarantee that memeticists will have a problem providing identity conditions for memes that are more severe than the (already severe) problems afflicting the identity conditions for genes. If we consider that the meme ought to be understood to be the smallest unit of *information worth copying*, then we have already accumulated a wealth of understanding of just such problems, arising in patent law, and the law of copyright and "trade secrets." How big is an idea? When is one idea an illicit copy of another idea? We have no single bedrock criterion for answering such questions, but we manage quite well with them in practice, counting on the costs of reinvention to stabilize our sense of what is worth copying in particular cases.[5]

It is undeniable that cultural transmission depends on comprehension at almost every juncture. We human beings are just not in the habit of copying formulae we don't understand and then passing them mindlessly on to our neighbors. This in itself is not a fatal blow to the proposed Darwinian theory of cultural evolution because the intelligent agents active at these junctures are not miraculous. They are themselves products of earlier mindless evolution, cranes not skyhooks. Moreover, the comprehension they exhibit, even in extreme cases, is typically insufficient to account for the cultural patterns their many attempts at copying and transmission eventually yield.

Just as genetic engineers, for all their foresight and insight into the innards of things, are still at the mercy of natural selection when it comes to the fate of their creations (that is why, after all, we are so cautious about letting them release their brainchildren on the outside world), so too the memetic engineer, no matter how sophisticated, still has to contend with the daunting task of winning the replication tournaments in the memosphere. One of the most sophisticated musical memetic engineers of the age, Leonard Bernstein, wryly noted this in a wonderful piece he published in 1955 entitled "Why don't you run upstairs and write a nice Gershwin tune?" (*New York Times*, April 1955; reprinted in *The Joy of Music*, 1959, pp52-62)

Bernstein had credentials and academic honors aplenty in 1955, but no songs on the Hit Parade.

A few weeks ago a serious composer-friend and I . . . got boiling mad about it. Why shouldn't we be able to come up with a hit, we said, if the standard is as low as it seems to be? We decided that all we had to do was to put ourselves into the mental state of an idiot and write a ridiculous hillbilly tune.

They failed—and not for lack of trying. As Bernstein wistfully remarked, "It's just that it would be nice to hear someone accidentally whistle something of mine, somewhere, just once." (p54)

His wish came true, of course, a few years later in 1961, when *West Side Story* burst into the memosphere. Leonard Bernstein was a brilliant, comprehending, foresighted, ambitious creator of musical designs that he hoped would replicate like viruses in brains around the world. He succeeded in a few cases, but so did many musically ignorant, lackadaisical, inadvertent exuders of equally infectious melodies. Other unforgettable melodies have no identifiable composer at all, but have emerged from untold rounds of differential replication. A theory that can encompass, and ultimately explain, all such varieties of cultural production will need to track the differential ability of authorless and authored items to get people to harbor them and pass them on, with or without comprehension.

References

Borges, J. L., "Pierre Menard, Author of the *Quixote,*" (1962)
Boyd, Robert, forthcoming, [Aunger memes conference talk]
Boyd and Richerson, 1985
Dawkins, 1999, Foreword to Susan Blackmore, *The Meme Machine*, Oxford: OUP.
Dennett, D. C. 1971 Intentional Systems
—1987 *The Intentional Stance*
—1989 "Cognitive Ethology: Hunting for Bargains or a Wild Goose Chase?" in A. Montefiore and D. Noble, eds, *Goals, No-Goals and own Goals: A Debate on Goal-Directed And Intentional Behaviour*, London: Unwin Hyman.
— 1995, *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, New York: Simon & Schuster.
Feldman and Cavalli-Sforza, 1981;
McFarland, D., 1989, "xxxx" in Montefiore and Noble.
Powers, R. 19xx, The Gold Bug Variations, XXX
Sperber, D. and Wilson, D., *Relevance:*
Williams, G., 1966

1. The demands of this minimal Darwinism are far from trivial, and the ferocity with which Darwinian accounts of the evolution of language and sociality are attacked by some critics from the humanities and social sciences shows that mere consistency with evolutionary theory is not yet an accepted constraint in many quarters. This is a fact of life that we must deal with: fear of the thin edge of the wedge misleads many who hate the idea of a *strong* Darwinian theory of cultural evolution to resist conceding even consistency with evolutionary theory as the obvious requirement it is.

2. Such organisms need not be deemed to be making conscious decisions, of course, but the rationality, such as it is, of the "decisions" they make is typically anchored to the expected benefit to the individual organism. See, e.g., McFarland's (1989) distinction between an organism's goal function and its cost function, and Dennett, 1989. See also Sperber and Wilson (1998) for good discussions of gene, individual, and group benefits of such decision-making.

3. One might be tempted to treat the tolerance for variant "spellings" for proteins—e.g, there are over a hundred different ways of "spelling" lysozyme—as a *sort* of higher-level norm-correction, but this is not strictly parallel, since the copying at each locus is by local spelling, without ad lib interchange, except for mutations.

4. Following Sperber and Wilson's reasoning in a different domain (1995), we can note that no complicated ("Gricean") reasoning is required by the individual vectors in the series of transmissions, since they need not reconstruct the hidden recipe behind the production but simply use the optimality assumptions built into the intentional stance to home in on the intended production. It may often be difficult to "reverse engineer" the recipes for cultural products (styles of pottery, for instance) (Boyd, forthcoming [Aunger conference]), but it is not typically necessary, since the intended properties of the result can be read off so readily.

5. There is considerable debate among memeticists about whether memes should be defined as brain-structures, or as behaviors, or some other presumably well-anchored *concreta,* but I think the case is still overwhelming for defining memes abstractly, in terms of information worth copying (however embodied) since it is the information that determines how much design work or R and D doesn't have to be re-done. That is why a wagon with spoked wheels carries *the idea* of a wagon with spoked wheels as well as any mind or brain could carry it.

IN DARWIN'S WAKE, WHERE AM I?

Parfois je pense; et parfois, je suis. —Paul Valéry[1]

Valéry's "*Variation sur Descartes*" excellently evokes the vanishing act that has haunted philosophy ever since Darwin overturned the Cartesian tradition. If *my body* is composed of nothing but a team of a few trillion robotic cells, mindlessly interacting to produce all the large-scale patterns that tradition would attribute to the non-mechanical workings of my mind, there seems to be nothing left over to be *me*. Lurking in Darwin's shadow there is a bugbear: the incredible Disappearing Self.[2] One of Darwin's earliest critics saw what was coming and could scarcely contain his outrage:

> In the theory with which we have to deal, Absolute Ignorance is the artificer; so that we may enunciate as the fundamental principle of the whole system, that, IN ORDER TO MAKE A PERFECT AND BEAUTIFUL MACHINE, IT IS NOT REQUISITE TO KNOW HOW TO MAKE IT. This proposition will be found, on careful examination, to express, in condensed form, the essential purport of the Theory, and to express in a few words all Mr. Darwin's meaning; who, by a strange inversion of reasoning, seems to think Absolute Ignorance fully qualified to take the place of Absolute Wisdom in all the achievements of creative skill.[3]

This "strange inversion of reasoning" promises–or threatens—to dissolve the Cartesian *res cogitans* as the wellspring of creativity, and then where will we be? Nowhere, it seems. It *seems* that if creativity gets "reduced" to "mere mechanism" *we* will be shown not to exist at all. Or, we will exist, but we won't be thinkers, we won't manifest genuine "Wisdom in all the achievements of creative skill." The individual as Author of works and deeds will be demoted: a person, it seems, is a barely salient nexus, a mere slub in the fabric of causation.

Whenever we zoom in on the act of creation, it seems we lose sight of it. The genius we thought we could see from a distance gets replaced at the last instant by stupid machinery, an echo of Darwin's shocking substitution of Absolute Ignorance for Absolute Wisdom in the creation of the biosphere. Many people dislike Darwinism in their guts, and of all the ill-lit, murky reasons for antipathy to Darwinism, this one has always struck me as the deepest, but only in the sense of being the most entrenched, the least accessible to rational criticism. There are thoughtful people who scoff at Creationism, dismiss dualism out of hand, pledge allegiance to academic humanism—and then get quite squirrely when somebody proposes a Darwinian theory of creative intelligence. The very idea that all the works of human genius can be understood *in the end* to be mechanistically generated products of a cascade of generate-and-test algorithms arouses deep revulsion in many otherwise quite insightful, open-minded people.

Absolute Ignorance? Fie on anybody who would thus put "A" and "I" together! Serendipity is the wellspring of evolution, so it is fitting that an evolutionist such as I should adapt MacKenzie's happy capitalization for a purpose he could hardly have imagined. His outraged scoffing at the powers of Absolute Ignorance has an uncannily similar echo more than a century later in the equally outraged scoffing at those who believe in what John Searle[4] has called "strong AI," the thesis that *real* intelligence can be made by artifice, that the difference between a mindless mechanism and a mindful one is a difference of design (or *program*—since whatever you can design in hardware you can implement in a virtual machine that has the same competence).[5]

Darwin's "strange inversion of reasoning" turns an ancient idea upside-down. The "top-down" perspective on creative intelligence supposes that it always takes a big, fancy, smart thing to create a lesser thing. No horseshoe has ever made a blacksmith; no pot has fashioned a potter. Hence we—and all the other fancy things we see around us—must have been created by something still fancier, something like us only more so. To many—perhaps most—people, this idea is *just obvious*. Consider this page from a creationist propaganda mailing:

> 1. Do you know of any building that didn't have a builder? YES/NO
> 2. Do you know of any painting that didn't have a painter? YES/NO
> 3. Do you know of any car that didn't have a maker? YES/NO

If you answered "YES" for any of the above give details.

But however strongly the idea appeals to common sense, Darwin shows us how it can be, in a word, false. Darwin shows us that a bottom-up theory of creation is, indeed, not only imaginable but empirically demonstrable. Absolute Ignorance *is* fully qualified to take the place of Absolute Wisdom in all the achievements of creative skill—*all* of them.

John Searle's Chinese Room thought experiment is a variation on the desperate joke of the creationists:

Do you know of any machine that can understand Chinese? YES/NO

If you answered "YES" give details!

While the creationists' rhetorical questions merely gesture towards the presumed embarrassments facing anybody who tries to "give details" of an instance of bottom-up creation, Searle's challenge offers a survey of possible avenues the believers in strong AI might take in their attempts to "give details" and purports to rebut them one and all. The believers in strong AI have been remarkably unmoved by Searle's attempts at refutation, and the comparison of Searle's position with creationism shows why. Biologists who cannot *yet* explain some particular puzzle about the non-miraculous path that led to one marvel of nature or another, who cannot *yet* "give details" to satisfy the particular critic, nevertheless have such a fine track record of success in giving the details, and such a stable and fecund background theory to use in generating and confirming new details, that they simply dismiss the rhetorical implication: "You'll never succeed!" They calmly acknowledge that they may need to develop a few new wrinkles before they can declare victory. Believers in strong AI are similarly content to concede that all AI models to date have been deficient in many respects, orders of magnitude too simple, many of them pursuing particular visions of AI that are simply mistaken. They go on to note that Searle isn't challenging particular details of the attempts to date; he purports to be offering an argument for the *in principle* impossibility of strong AI, a conclusion that he insists is meant to cover all *imaginable* complications of the underlying theoretical framework. They know that their underlying theoretical framework is nothing other than the straightforward extension, into the human brain and all its peripheral devices and interfaces, of the Darwinian program of mindless mechanism doing, in the end, all the work. If Darwinian mechanisms can explain the existence of a skylark, in all its glory, they can surely explain the existence of an ode to a nightingale, too.[6] A poem is a wonderful thing, but not clearly more wonderful than a living, singing skylark.

Unsupportable antipathies often survive thanks to protective coloration: they blend into the background of legitimate objections to overstatements of the view under attack. Since the reach of Darwinian enthusiasm has always exceeded its grasp, there are always good criticisms of Darwinian excesses to hide amongst. Likewise, of course, for the excesses of the ideologues of AI. And so the battle rages, generating as much suspicion as insight. Darwinians who are sure that a properly nuanced, sophisticated Darwinism is proof against all the objections and misgivings—I am one such—should nevertheless recall the fate of the Freudian nags of the 50s and 60s, who insisted on seeing everything through the perspective of their hero's categories, only to discover that by the time you've attenuated your Freudianism to accommodate *everything*, it is Pickwickian Freudianism most of the way. Sometimes a cigar is just a cigar, and sometimes an idea is just an idea—not a meme—and sometimes a bit of mental machinery is not usefully interpreted as an adaptation dating back to our ancestral hunter-gatherer days or long before, even though it is, obviously, descended (with modifications) from some combination or other of such adaptations. We Darwinians will try to remind ourselves of this, hoping our doughty opponents will come to recognize that a Darwinian theory of creativity is not just a promising solution but the only solution in sight to a problem that is everybody's problem: *how can an arrangement of a hundred billion mindless neurons compose a creative mind, an I?*

William Poundstone has put the inescapable challenge succinctly in terms of "the old fantasy of a monkey typing *Hamlet* by accident." He calculates that the chances of this happening are "1 in 50 multiplied by itself 150,000 times."

In view of this, it may seem remarkable that anything as complex as a text of *Hamlet* exists. The observation that *Hamlet* was written by Shakespeare and not some random agency only transfers the problem. Shakespeare, like everything else in the world, must have arisen (ultimately) from a homogeneous early universe. Any way you look at it *Hamlet* is a product of that primeval chaos.[7]

Where does all that design come from? What process-

es could conceivably yield such improbable "achievements of creative skill"? What Darwin saw is that design is always both valuable and costly. It does not fall like manna from heaven, but must be accumulated the hard way, by time-consuming, energy-consuming processes of mindless search through "primeval chaos," automatically preserving happy accidents when they occur. This broadband process of Research and Development is breathtakingly inefficient, but—this is Darwin's great insight—if the costly fruits of R and D can be thriftily conserved, copied, and re-used, they can be accumulated over time to yield "the achievements of creative skill." "This principle of preservation," Darwin says, "I have called, for the sake of brevity, Natural Selection."[8]

There is no requirement in Darwin's vision that these R and D processes run everywhere and always at the same tempo, with the same (in-)efficiency. Consider the unimaginably huge multidimensional space of all *possible* designed things—both natural and artificial. Every imaginable whale and unicorn, every automobile and spaceship and robot, every poem and mathematical proof and symphony finds its place somewhere in this Design Space. If we think of design work or R and D as a sort of *lifting* in Design Space[9] then we can see that the gradualistic, frequently back-sliding, maximally inefficient basic search process can on important occasions yield new conditions that speed up the process, permitting faster, more effective local lifting.[10] Call any such product of earlier R and D a crane, and distinguish it from what Darwinism says does not happen: *skyhooks*.[11] Skyhooks, like manna from heaven, would be miracles, and if we posit a skyhook anywhere in our "explanation" of creativity, we have in fact conceded defeat—'Then a miracle occurs."[12]

What, then, is a mind? The Darwinian answer is straightforward. A mind is a crane, made of cranes, made of cranes, a mechanism of not quite unimaginable complexity that can clamber through Design Space at a giddy—but not miraculously giddy—pace, thanks to all the earlier R and D, from all sources, that it exploits. What is the anti-Darwinian answer? It is perfectly expressed by one of the 20th century's great creative geniuses (though, like MacKenzie, he probably didn't mean by his words what I intend to mean by them).

Je ne cherche pas; je trouve.—Pablo Picasso

Picasso purports to be a genius indeed, someone who does not need to engage in the menial work of trial and error, generate-and-test, R and D; he claims to be able to *leap* to the summits of the peaks—the excellent designs—in the vast reaches of Design Space without having to guide his trajectory (he searches not) by sidelong testing at any waystations. As an inspired bit of bragging, this is *non pareil*, but I don't believe it for a minute. And anyone who has strolled through an exhibit of Picasso drawings (as I recently did in Valencia) looking at literally dozens of variations on a single theme, all signed—and sold—by the artist, will appreciate that whatever Picasso may have meant by his *bon mot,* he could not truly claim that he didn't engage in a time-consuming, energy-consuming exploration of neighborhoods in Design Space. At best he could claim that his own searches were so advanced, so efficient, that it didn't seem—to himself—to be design *work* at all. But then what did he have within him that made him such a great designer? A skyhook, or a superb collection of cranes?[13]

We can now characterize a mutual suspicion between Darwinians and anti-Darwinians that distorts the empirical investigation of creativity. Darwinians suspect their opponents of hankering after a skyhook, a miraculous gift of genius whose powers have no decomposition into mechanical operations, however complex and informed by earlier processes of R and D. Anti-Darwinians suspect their opponents of hankering after an account of creative processes that so diminishes the Finder, the Author, the Creator, that it disappears, at best a mere temporary locus of mindless differential replication. We can make a little progress, I think, by building on Poundstone's example of the *creation of the creator* of *Hamlet.* Consider, then, a little thought experiment.

Suppose Dr. Frankenstein designs and constructs a monster, Spakesheare, that thereupon sits up and writes out a play, *Spamlet*. My question is not about the author of *Waverley* but about the author of *Spamlet*.

Who is the author of *Spamlet*?

First, let's take note of what I claim to be irrelevant in this thought experiment. I haven't said whether Spakesheare is a robot, constructed out of metal and silicon chips, or, like the original Frankenstein's monster, constructed out of human tissues—or cells, or proteins, or amino acids, or carbon atoms. As long as the design work and the construction were carried out by Dr. Frankenstein, it makes no difference to the example what the materials are. It might well turn out that the only way to build a robot small enough and fast

enough and energy-efficient enough to sit on a stool and type out a play is to construct it from artificial cells filled with beautifully crafted motor proteins and other carbon-based nanorobots. That is an interesting technical and scientific question, but not of concern here. For exactly the same reason, if Spakesheare is a metal-and-silicon robot, it may be allowed to be larger than a galaxy, if that's what it takes to get the requisite complication into its program—and we'll just have to repeal the speed limit for light for the sake of our thought experiment. These technical constraints are commonly declared to be off-limits in these thought experiments, so so be it. If Dr. Frankenstein chooses to make his AI robot out of proteins and the like, that's his business. If his robot is cross-fertile with normal human beings and hence capable of creating what is arguably a new species by giving birth to a child, that is fascinating, but what we will be concerned with is Spakesheare's purported brainchild, *Spamlet*. Back to our question:

Who is the author of *Spamlet*?

In order to get a grip on this question, we have to look inside and see what happens in Spakesheare.[14] At one extreme, we find inside a file (if Spakesheare is a robot with a computer memory) or a basically *memorized* version of *Spamlet*, all loaded and ready to run. In such an extreme case, Dr. Frankenstein is surely the author of *Spamlet*,[15] using his intermediate creation, Spakesheare, as a mere storage-and-delivery device, a particularly fancy word processor. All the R and D work was done earlier, and copied to Spakesheare by one means or another.

We can visualize this more clearly by imagining a sub-space of Design Space, which I call the Library of Babel, after Jorge Luis Borges' classic short story by that name.[16] Borges invites us to imagine a warehouse filled with books that appears to its inhabitants to be infinite; they eventually decide that it is not, but it might as well be, for it seems that on its shelves—in no order, alas—lie all the possible books.

Suppose that each book is 500 pages long, and each page consists of 40 lines of 50 spaces, so there are two thousand character-spaces per page. Each space is either blank, or has a character printed on it, chosen from a set of 100 (the upper and lower case letters of English and other European languages, plus the blank and punctuation marks).[17] Somewhere in the Library of Babel is a volume consisting entirely of blank pages, and another volume is all question marks, but the vast majority consist of typographical gibberish; no rules of spelling or grammar, to say nothing of sense, prohibit the inclusion of a volume. Five hundred pages times two thousand characters per page gives a million character-spaces per book, so there are $100^{1,000,000}$ books in the Library of Babel. Since it is estimated[18] that there are only 100^{40} (give or take a few) particles (protons, neutrons and electrons) in the region of the universe we can observe, the Library of Babel is not remotely a physically possible object, but thanks to the strict rules with which Borges constructed it in his imagination, we can think about it clearly.

We need some terms for the quantities involved. The Library of Babel is not infinite, so the chance of finding anything interesting in it is not literally infinitesimal.[19] These words exaggerate in a familiar way, but we should avoid them. Unfortunately, all the standard metaphors—astronomically large, a needle in a haystack, a drop in the ocean—fall comically short. No actual astronomical quantity (such as the number of elementary particles in the universe, or the time since the Big Bang measured in nanoseconds) is even visible against the backdrop of these huge-but-finite numbers. If a readable volume in the Library were as easy to find as a particular drop in the ocean, we'd be in business! Dropped at random into the Library, your chance of ever encountering a volume with so much as a grammatical sentence in it is so vanishingly small that we might do well to capitalize the term—*Vanishingly* small—and give it a mate, *Vastly*, short for Very-much-more-than-astronomically.[20]

It is amusing to reflect on just how large this finite set of possible books is, compared with any actual library. Most of the books are pure gibberish, as noted, so consider the Vanishing subset of books composed entirely of English words, without a single misspelling. It is itself a Vast set, of course, and contained within it, but Vanishingly hard to find, is the Vast subset whose English words are lined up in grammatical sentences. A Vast but Vanishing subset of this subset in turn is the subset of books composed of English sentences that actually make sense. A Vast but Vanishing subset of these are about somebody named John, and a Vast but Vanishing subset of these are about the death of John F. Kennedy. A Vast but Vanishing subset of these are true... and a Vast but Vanishing subset of the possible true books about the death of JFK are written entirely in limericks. There are many orders of magnitude more

possible true books in limerick form about the death of JFK than there are books in the Library of Congress.

Now we are ready to return to that needle-in-a-haystack, *Spamlet*, and consider how the trajectory to this particular place in the Library of Babel was traversed in actual history. If we find that the whole journey was already completed by the time Spakesheare's memory was constructed and filled with information, we know that Spakesheare played no role at all in the search. Working backwards, if we find that Spakesheare's only role was running the stored text through a spell-checker before using it to guide its typing motions, we will be unimpressed by claims of Spakeshearian authorship. This is a measurable, but Vanishing, part of the total R and D. There is a sizable galaxy of near-twin texts of *Spamlet*—roughly a hundred million different minor mutants have but a single uncorrected typo in them, and if we expand our horizon to include one typo per page, we have begun to enter the land of Vast numbers of variations on the theme. Working back a little further, once we graduate from typos to *thinkos*,[21] those arguably mistaken, or suboptimally chosen, words, we have begun to enter the land of serious authorship, as contrasted with mere copy-editing. The relative triviality of copy-editing, and yet its unignorable importance in shaping the final product gets well represented in terms of our metaphor of Design Space, where every little bit of lifting counts for something, and sometimes a little bit of lifting moves you onto a whole new trajectory. As usual, we may quote Ludwig Mies van der Rohe at this juncture: "God is in the details."

Now let's turn the knobs on our thought experiment, as Douglas Hofstadter has recommended,[22] and look at the other extreme, in which Dr. Frankenstein leaves most of the work to Spakesheare. The most realistic scenario would surely be that Spakesheare has been equipped by Dr. Frankenstein with a virtual past, a lifetime stock of pseudo-memories of experiences on which to draw while responding to its Frankenstein-installed obsessive desire to write a play. Among those pseudo-memories, we may suppose, are many evenings at the theater, or reading books, but also some unrequited loves, some shocking close calls, some shameful betrayals and the like. Now what happens? Perhaps some scrap of a "human interest" story on the network news will be the catalyst that spurs Spakesheare into a frenzy of generate-and-test, ransacking its memory for useful tidbits and themes, transforming—transposing, morphing—what it finds, jiggling the pieces into temporary, hopeful structures that compete for completion, most of them dismantled by the corrosive processes of criticism that nevertheless expose useful bits now and then, and so forth, and all of this multi-leveled search would be somewhat guided by multi-level, internally generated evaluations, including evaluation of the evaluation ...of the evaluation functions as a response to evaluation of...the products of the ongoing searches.[23]

Now if the amazing Dr. Frankenstein had actually anticipated all this activity down to its finest grain at the most turbulent and chaotic level, and had hand-designed Spakesheare's virtual past, and all its search machinery, to yield just this product, *Spamlet,* then Dr. Frankenstein would be, once again, the author of *Spamlet*, but also, in a word, God. Such Vast foreknowledge would be simply miraculous. Restoring a smidgen of realism to our fantasy, we can set the knobs at a rather less extreme position and assume that Dr. Frankenstein was unable to foresee all this in detail, but rather delegated to Spakesheare most of the hard work of completing the trajectory in Design Space to *one literary work or another*, something to be determined by later R and D occurring within Spakesheare itself. We have now arrived, by this simple turn of the knob, in the neighborhood of reality itself, for we already have actual examples of impressive artificial Authors that Vastly outstrip the foresight of their own creators. Nobody has yet created an artificial playwright worth serious attention, but an artificial chess player—IBM's Deep Blue—and an artificial composer—David Cope's EMI—have both achieved results that are, *in some respects*, equal to the best that human creative genius can muster.

Who beat Garry Kasparov, the reigning World Chess Champion? Not Murray Campbell or any of his IBM team. Deep Blue beat Kasparov. Deep Blue designs better chess games than any of them can design. None of them can author a winning game against Kasparov. Deep Blue can. Yes, but. Yes, but. I am sure many of you are tempted to insist at this point that when Deep Blue beats Kasparov at chess, its brute force search methods are *entirely* unlike the exploratory processes that Kasparov uses when he conjures up his chess moves. But that is simply not so—or at least it is not so in the only way that could make a difference to the context of

this debate about the universality of the Darwinian perspective on creativity. Kasparov's brain is made of organic materials, and has an architecture importantly unlike that of Deep Blue, but it is still, so far as we know, a massively parallel search engine which has built up, over time, an outstanding array of heuristic pruning techniques that keep it from wasting time on unlikely branches. There is no doubt that the investment in R and D has a different profile in the two cases; Kasparov has methods of extracting good design principles from past games, so that he can recognize, and know enough to ignore, huge portions of the game space that Deep Blue must still patiently canvass *seriatim*. Kasparov's "insight" dramatically changes the shape of the search he engages in, but it does not constitute "an *entirely* different" means of creation. Whenever Deep Blue's exhaustive searches close off a *type* of avenue that it has some means of recognizing (a difficult but not impossible task), it can re-use that R and D whenever it is appropriate, just as Kasparov does. Much of this analytical work has been done for Deep Blue by its designers, and given as an innate endowment, but Kasparov has likewise benefitted from hundreds of thousands of person-years of chess exploration transmitted to him by players, coaches and books. It is interesting in this regard to contemplate the suggestion recently made by Bobby Fischer, who proposes to restore the game of chess to its intended rational purity by requiring that the major pieces be *randomly* placed in the back row at the start of each game (random, but mirror image for black and white). This would instantly render the mountain of memorized openings almost entirely obsolete, for humans and machines alike, since only rarely would any of this lore come into play. One would be thrown back onto a reliance on fundamental principles; one would have to do more of the hard design work in real time—with the clock running. It is far from clear whether this change in rules would benefit human beings more than computers. It all depends on which type of chess player is relying most heavily on what is, in effect, rote memory—reliance *with minimal comprehension* on the R and D of earlier explorers.

The fact is that the search space for chess is too big for even Deep Blue to explore exhaustively in real time, so like Kasparov, it prunes its search trees by taking calculated risks, and like Kasparov, it often gets these risks pre-calculated. Both presumably do massive amounts of "brute force" computation on their very different architectures. After all, what do neurons know about chess? Any work *they* do must be brute force work of one sort or another.

It may seem that I am begging the question in favor of a computational, AI approach by describing the work done by Kasparov's brain in this way, but the work has to be done somehow, and no *other* way of getting the work done has ever been articulated. It won't do to say that Kasparov uses "insight" or "intuition" since that just means that Kasparov himself has no privileged access, no insight, into how the good results come to him. So, since nobody knows how Kasparov's brain does it—least of all Kasparov—there is not yet any evidence at all to support the claim that Kasparov's means are "entirely unlike" the means exploited by Deep Blue. One should remember this when tempted to insist that "of course" Kasparov's methods are hugely different. What on earth could provoke one to go out on a limb like that? Wishful thinking? Fear?

But that's just chess, you say, not art. Chess is *trivial* compared to art (now that the world champion chess player is a computer). This is where David Cope's EMI comes into play.[24] Cope set out to create a mere efficiency-enhancer, a composer's aid to help him over the blockades of composition any creator confronts, a high-tech extension of the traditional search vehicles (the piano, staff paper, the tape recorder, etc.). As EMI grew in competence, it promoted itself into a whole composer, incorporating more and more of the generate-and-test process. When EMI is fed music by Bach, it responds by generating musical compositions in the style of Bach. When given Mozart, or Schubert, or Puccini, or Scott Joplin, it readily analyzes their styles and composes new music in their styles, better pastiches than Cope himself—or almost any human composer—can compose. When fed music by two composers, it can promptly compose pieces that eerily unite their styles, and when fed, all at once (with no clearing of the palate, you might say) all these styles at once, it proceeds to write music based on the totality of its musical experience. The compositions that result can then also be fed back into it, over and over, along with whatever other music comes along in MIDI format, and the result is EMI's own "personal" musical style, a style that candidly reveals its debts to the masters, while being an unquestionably idiosyncratic integration of all this "experience." EMI can now compose not just two-part inventions and art songs but whole symphonies—and has

composed over a thousand, when last I heard. They are good enough to fool experts (composers and professors of music) and I can personally attest to the fact that an EMI-Puccini aria brought a lump to my throat—but then, I'm on a hair trigger when it comes to Puccini, and this was a good enough imitation to fool me. David Cope can no more claim to be the composer of EMI's symphonies and motets and art songs than Murray Campbell can claim to have beaten Kasparov in chess.

To a Darwinian, this new element in the cascade of cranes is simply the latest in a long history, and we should recognize that the boundary between authors and their artifacts should be just as penetrable as all the other boundaries in the cascade. When Richard Dawkins notes[25] that the beaver's dam is as much a part of the beaver phenotype—its extended phenotype—as its teeth and its fur, he sets the stage for the further observation that the boundaries of a human author are exactly as amenable to extension. In fact, of course, we've known this for centuries, and have carpentered various semi-stable conventions for dealing with the products of Rubens, of Rubens' *studio*, of Rubens' various students. Wherever there can be a helping hand, we can raise the question of just who is helping whom, what is creator and what is creation. How should we deal with such questions? To the extent that anti-Darwinians simply want us to preserve some tradition of authorship, to have some *rules of thumb* for determining who or what shall receive the honor (or blame) that attends authorship, their desires can be acknowledged and met, one way or another (which doesn't necessarily mean we should meet them). To the extent that this is not enough for the anti-Darwinians, to the extent that they want to hold out for authors as an objective, metaphysically grounded, "natural kind" (oh, the irony in those essentialist wolf-words in naturalist sheep's clothing!), they are looking for a skyhook.

The renunciation of skyhooks is, I think, the deepest and most important legacy of Darwin in philosophy, and it has a huge domain of influence, extending far beyond the skirmishes of evolutionary epistemology and evolutionary ethics. If we commit ourselves to Darwin's "strange inversion of reasoning," we turn our backs on compelling ideas that have been central to the philosophical tradition for centuries, not just Aristotle's essentialism and irreducible *telos*, but also Descartes's *res cogitans* as a causer outside the mechanistic world, to name the three that had been most irresistible until Darwin came along. The siren songs of these compelling traditions still move many philosophers who have not yet seen fit to execute the inversion, sad to say. Clinging to their pre-Darwinian assumptions, they create problems for themselves that will no doubt occupy many philosophers for years to come.[26] The themes all converge when the topic is creativity and authorship, where the urge is to hunt for an "essence" of creativity, an "intrinsic" source of meaning and purpose, a locus of responsibility somehow insulated from the causal fabric in which it is embedded, so that within its boundaries it can generate, from its *own* genius,[27] its *irreducible* genius, the meaningful words and deeds that distinguish us so sharply from mere mechanisms.

Plato called for us to carve nature at its joints, a wonderful biological image, and Darwin showed us that the salient boundaries in the biosphere are not the crisp set-theoretic boundaries of essentialism, but the emergent effects of historical processes. As one species turns into two, the narrow isthmus of intermediates disappears as time passes, leaving islands, concentrations sharing family resemblances, surrounded by empty space. As Darwin noted (in somewhat different terms), there are feedback processes that enhance separation, actively depopulating this middle ground. We might expect the same sort of effects in the sphere of human mind and culture, cultural habits or practices that favor the isolation of the processes of artistic creation in a single mind. "Are you the author of this?" "Is this all your own work?" The mere fact that these are familiar questions shows that there are cultural pressures encouraging people to *make* the favored answers come true. A small child, crayon in hand, huddled over her drawing, slaps away the helping hand of parent or sibling, because she wants this to be *her* drawing. She already appreciates the norm of pride of authorship, a culturally imbued bias built on the palimpsest of territoriality and biological ownership. The very idea of being an artist shapes her consideration of opportunities on offer, shapes her evaluation of features she discovers in herself. And this in turn will strongly influence the way she conducts her own searches through Design Space, in her largely unconscious emulation of Picasso's ideal, or, if she is of a contrarian spirit, defying it, like Marcel Duchamp:

> Cabanne: What determined your choice of readymades?

> Duchamp: That depended on the object. In general, I had to beware of its "look." It's very difficult to choose an object, because, at the end of fifteen days, you begin to like it or to hate it. You have to approach something with an indifference, as if you had no aesthetic emotion. The choice of ready-mades is always based on visual indifference and, at the same time, on the total absence of good or bad taste....[28]

There is a persistent problem of imagination management in the debates surrounding this issue: people on both sides have a tendency to underestimate the resources of Darwinism, imagining simplistic alternatives that do not exhaust the space of possibilities. Darwinians are notoriously quick to find (or invent) differences in *genetic fitness* to go with every difference they observe, for instance. Meanwhile, anti-Darwinians, noting the huge distance between a beehive and the *St. Matthew Passion* as created objects, are apt to suppose that anybody who proposes to explain both creative processes with a single set of principles must be guilty of one reductionist fantasy or another: "Bach had a gene for writing baroque counterpoint just like the bees' gene for forming wax hexagons" or "Bach was just a mindless trial-and-error mutator and selector of the musical memes that already flourished in his cultural environment." Both of these alternatives are nonsense, of course, but pointing out their flaws does nothing to support the idea that ("therefore") there must be irreducibly *non-Darwinian* principles at work in any account of Bach's creativity. In place of this dimly imagined chasm with "Darwinian phenomena" on one side and "non-Darwinian phenomena" on the other side, we need to learn to see the space between bee and Bach as populated with all manner of mixed cases, differing from their nearest neighbors in barely perceptible ways, replacing the chasm with a traversable gradient of non-minds, protominds, hemi-demi-semi minds, magpie minds, copycat minds, aping minds, clever-pastiche minds, "path-finding" minds, "groundbreaking" minds, and eventually, genius minds. And the individual minds, of each caliber, will themselves be composed of different sorts of parts, including, surely, some special-purpose "modules" adapted to various new tricks and tasks, as well as a cascade of higher-order reflection devices, capable of generating ever more rarefied and delimited searches through pre-selected regions of the Vast space of possible designs.

It is important to recognize that genius is itself a product of natural selection and involves generate-and-test procedures all the way down. Once you have such a product, it is often no longer particularly perspicuous to view it solely as a cascade of generate-and-test processes. It often makes good sense to leap ahead on a *narrative* course, thinking of the agent as a self, with a variety of projects, goals, presuppositions, hopes....In short, it often makes good sense to adopt the intentional stance towards the whole complex product of evolutionary processes. This effectively brackets the largely unknown and unknowable mechanical microprocesses as well as the history that set them up, and puts them out of focus while highlighting the patterns of rational activity that those mechanical microprocesses track so closely. This tactic makes especially good sense to the creator himself or herself, who must learn not to be oppressed by the revelation that on close inspection, even on close *intro*spection, a genius dissolves into a pack rat, which dissolves in turn into a collection of trial-and-error processes over which nobody has ultimate control.

Does this realization amount to a loss—an elimination—of selfhood, of genius, of creativity? Those who are closest to the issue—the artistic and scientific geniuses who have reflected on it—often confront this discovery with equanimity. Mozart is reputed to have said of his best musical ideas: "Whence and how do they come? I don't know and I have nothing to do with it."[29] The painter Philip Guston is equally unperturbed by this evaporation of visible self when the creative juices start flowing:

> When I first come into the studio to work, there is this noisy crowd which follows me there; it includes all of the important painters in history, all of my contemporaries, all the art critics, etc. As I become involved in the work, one by one, they all leave. If I'm lucky, every one of them will disappear. If I'm really lucky, I will too.[30]

In closing, I would like to acknowledge a few of my co-authors: Anonymous, Jorge Luis Borges, David Cope, Charles Darwin, Richard Dawkins, Susan Dennett, René Descartes, Marcel Duchamp, Thomas Edison, Bobby Fischer, Philip Guston, Douglas Hofstadter, Nicholas Humphrey, Robert MacKenzie, Tony Marcel, Victoria McGeer, Ludwig Mies van der Rohe, Pablo Picasso, William Poundstone, John Searle, William Shakespeare, Mary Shelley, Paul Valéry.

ENDNOTES

1. Paul Valéry, *Cahiers*, ed. Judith Robinson, 2 vols. (Paris: Edition de la Pleiade 1973-74). vol. 2 (1974), p. 1388.
2. Dennett, D. 1984, *Elbow Room*, p13.
3. Robert Beverley MacKenzie, 1868, *The Darwinian Theory of the Transmutation of Species Examined* (published anonymously "By a Graduate of the University of Cambridge"). London: Nisbet & Co. Quoted in a review, *Athenaeum*, no 2102, February 8, p217.
4. Searle, John, 1980, "Minds, Brains and Programs," *Behavioral and Brain Sciences*, 3, pp.417-58.
5. This is obviously true of all competences of information-processing or control, but not of productive or transformative processes, such as lactation, which requires the transport and assembly of particular materials. Since Searle purports to distinguish the brain's "control powers" from its "bottom-up causal powers" that "produce intentionality," some have thought Searle imagines intentionality to be a special sort of substance secreted by the brain. Since he denies this, he owes us some other way to distinguish these mysterious causal powers from the control powers that software can implement and an explanation of why they are not implementable in a virtual machine.
6. This perspective helps to explain the visceral appeal to many onlookers of the various *apparent* alternatives to Darwinian mechanism that have flourished over the years. The most prominent recently have been the appeal to "self-organization" "on the edge of chaos" (Stuart Kauffman, Per Bak, and others), and "dynamical systems theory" in both evolution and cognition (Esther Thelen, Walter Freeman, Timothy van Gelder, and others), and, of course, Stephen Jay Gould's insistence that evolution is not, as I have claimed (building on the work of theorists from Darwin to Fisher and Haldane to Williams and Maynard Smith), fundamentally an algorithmic process. After the smoke of battle clears, these ideas can be readily seen to be, at best, interesting complications of the basic Darwinian mechanisms, just as connectionist architectures and embodied cognition models are interesting complications of the basic ideas of AI. These controveries are, at best, constructive disagreements over how to "give the details," not challenges to the basic Darwinian vision.
7. William Poundstone, 1985, *The Recursive Universe: Cosmic Complexity and the Limits of Scientific Knowledge*, New York: Wm Morrow, p23.
8. Charles Darwin, *Origin of Species*, Ch. 4 summary (p.127 in facs. edition)
9. This tactic of mapping evolutionary processes and results onto space is a natural and oft-used metaphor, exploited in models of hill-climbing, and peaks in adaptive landscapes, to name the most obvious and popular applications. Its naturalness does not guarantee its soundness, of course, and may even mask its limitations, but since the basic mapping strategy has proven to be particularly useful in expressing *criticisms* of over-simple evolutionary ideas (e.g., Kauffman's "rugged landscape," Eigen's "quasispecies"), it is not obviously biased in favor of simplistic visions of Darwinism.
10. John Maynard Smith and Eörs Szathmary, 1995, *The Major Transitions in Evolution*, Oxford: Freeman, identify eight occasions (major transitions) when the evolutionary process became more efficient, creating cranes.
11. Daniel C. Dennett, *Darwin's Dangerous Idea*, 1995, "The Tools for R and D: Skyhooks or Cranes?" pp 73-80.
12. See the famous cartoon by Sydney Harris, in which the physicist's blackboard is covered with impressive formulae, except for this bracketed phrase in the middle, which leads the onlooker scientist to say "I think you should be more explicit here in step two" (reprinted in Daniel Dennett, *Consciousness Explained*, 1991, p38).
13. I have been unable to discover the source of Picasso's claim, which is nicely balanced by a better known remark by a more down-to-earth creative genius, Thomas Edison: "Genius is one percent. inspiration and ninety-nine percent. perspiration" (in a newspaper interview. *Life* [1932], ch. 24, according to the *Oxford Dictionary of Quotations*).
14. Yes, I intend the homage to an old favorite of mine, *What Happens in Hamlet*, by J. Dover Wilson (1951, Cambridge Univ. Press).
15. Unless we find there is a Ms. Shelley who is the author of Dr. Frankenstein. . . !
16. Borges, J. L. 1962, *Labyrinths: Selected Stories and Other Writings*, New York: New Directions. [La Biblioteca de Babel, 1941, in El jardin de los senderos que se bifurcan, published with another in *Ficciones*. 1956, Emece Editores, S. A., Buenos Aires.]

17. Borges chose slightly different figures: books 410 pages long, with 40 lines of 80 characters. The total number of characters per book is close enough to mine (1,312,000 versus 1,000,000) to make no difference. I chose my rounder numbers for ease of handling. Borges chose a character set with only 25 members, which is enough for uppercase Spanish (with a blank, a comma and a period as the only punctuation), but not for English. I chose the more commodious 100 to make room without any doubt for the upper and lower case letters and punctuation of all the Roman alphabet languages.

18. Steven Hawking insists on putting it this way: "There are something like ten million million million million million million million million million million million million million million million (1 with eighty zeroes after it) particles in the region of the universe that we can observe." *A Brief History of Time,* New York: Bantam, p.129. Michael Denton (*Evolution: A Theory in Crisis*, London: Burnett Books. 1985) provides the estimate of 10^{70} atoms in the observable universe. Manfred Eigen (*Steps Towards Life*, Oxford University Press 1992, p.10) calculates the volume of the universe as 10^{84} cubic centimeters.

19. The Library of Babel is finite, but curiously enough, it contains all the grammatical sentences of English within its walls. But that's an infinite set, and the library is finite! Still, any sentence of English, of whatever length, can be broken down into 500-page chunks, each of which is somewhere in the library! How is this possible? Some books may get used more than once. The most profligate case is the easiest to understand: since there are volumes which each contain a single character and are otherwise blank, repeated use of these one hundred volumes will create any text of any length. As Quine *(Quiddities: An Intermittently Philosophical Dictionary*, Cambridge, MA: Harvard Univ. Press. 1987) points out, in his informative and amusing essay "Universal Library," if you avail yourself of this strategy of re-using volumes, and translate everything into the ASCII code your word processor uses, you can store the whole Library of Babel in two extremely slender volumes, in one of which is printed a 0 and in the other of which appears a 1! (Quine also points out that Theodor Fechner, the psychologist, propounded the fantasy of the universal library long before Borges.)

20. Quine, *loc.cit.*, coins the term "hyperastronomic" for the same purpose. The previous two paragraphs are drawn, with minor changes, from *Darwin's Dangerous Idea*, pp108-9.

21. For more on this concept, see my "From Typo to Thinko: When Evolution Graduated to Semantic Norms," forthcoming in the Fyssen conference volume on cultural evolution.

22. Douglas R. Hofstadter, 1981, "Reflections" (on Searle) in Hofstadter and Dennett, eds., *The Mind's I: Fantasies and Reflections on Self and Soul*, 1981, New York: Basic Books and Hassocks, Sussex: Harvester.

23. Shakespeare himself was, of course, a tireless exploiter of the design work of others, and may well have been poking fun at his own reputation, quoting a critic, when he had Autolycus describe himself as "a snapper-up of unconsidered trifles" in *A Winter's Tale* (Act IV, scene iii). Thanks to Tony Marcel for drawing this passage to my attention.

24. For the details, see David Cope, ed., *Virtual Music* (forthcoming from MIT Press), including my commentary, "Collision Detection, Muselot, and Scribble: Some Reflections on Creativity."

25. Richard Dawkins, 1982, *The Extended Phenotype*, Oxford and San Francisco: Freeman.

26. Three examples: Jerry Fodor 's series of flawed theories of psychosemantics; John Searle's inability to account for how "intrinsic intentionality" could evolve when it has no "control power" consequences visible to selective pressure; John McDowell's quest for a non-Darwinian alternative to what he calls "bald naturalism," a struggle to secure a variety of normativity that is not the mere *as-if* normativity he finds discernible in evolution. See Dennett, *Darwin's Dangerous Idea*, 1995, and "Granny versus Mother Nature—No Contest," *Mind & Language,* 11 no.3, 1996, pp 263-269, and "Review of John Searle, *The Rediscovery of the Mind,*" *Journal of Philosophy*, 60 (4), 193-205, Apr. 1993, for my analysis of Fodor 's and Searle's difficulties. My discussion of McDowell must be deferred to another occasion.

27. See "Do-It-Yourself Understanding," in *Brainchildren*, Cambridge MA: MIT Press, 1998, pp59-80, for my analysis of this theme in Fred Dretske's search for a privileged place where the understanding happens.

28. *Dialogues with Marcel Duchamp*, Pierre Cabanne (transl Ron Padgett), New York, Viking Press, 1971, p.48. Thanks to Nicholas Humphrey and Victoria McGeer for ideas expressed in the previous paragraph.

29. In an oft-quoted but possibly spurious passage—see *Darwin's Dangerous Idea*, p346-7.
30. I have been unable to locate the source of Guston's quote, but I have found much the same remark attributed to the composer John Cage, a close friend and contemporary of Guston's, who (is said to have) said this about painting:

When you are working, everybody is in your studio—the past, your friends, the art world, and above all, your own ideas—all are there. But as you continue painting, they start leaving, one by one, and you are left completely alone. Then, if you are lucky, even you leave.

Like all other creators, Guston and I like to re-use what we find, adding a few touches from time to time.