Universal Minds

Explaining the New Science of Evolutionary Psychology

Leda Cosmides & John Tooby
Co-Directors, Center for Evolutionary Psychology
University of California, Santa Barbara

CONTENTS

1. INTRODUCTION

Debauching the mind: Evolutionary psychology's past and present

In the final pages of the *Origin of Species*, after Darwin had presented the theory of evolution by natural selection, he made a bold prediction: "In the distant future I see open fields for far more important researches. Psychology will be based on a new foundation, that of the necessary acquirement of each mental power and capacity by gradation." More than a century later, a group of young scholars began to work out exactly how Darwin's fundamental insights could be used as a foundation on which to build a more systematic, scientifically accurate psychology. The resulting field is called *evolutionary psychology*. The goal of research in evolutionary psychology is to discover, understand, and map in detail the design of the human mind, as well as to explore the implications of these new discoveries for other fields.

Evolutionary psychology is an *approach* to psychology, in which knowledge and principles from evolutionary biology are put to use in research on the structure of the human mind. It is not an area of psychology, like vision, reasoning, or social behaviour. It is a *way of thinking* about psychology that can be applied to any topic within it.

What, then, is the mind? When evolutionary psychologists refer to "the mind", they mean the set of information-processing devices, embodied in the human brain, that are responsible for all conscious and nonconscious mental activity, and that generate all behaviour. What allows evolutionary psychologists to go beyond traditional approaches in studying the mind is that they make active use in their research of an often overlooked fact: That the circuits comprising the human mind were designed by natural selection to solve the adaptive problems faced by our hunter-gatherer ancestors. This way of thinking about the brain, mind, and behaviour is transforming how scientists approach old topics, and is opening up a multitude of new ones. It is also transforming everything from economics and anthropology, to philosophy and the humanities. This book introduces the concepts and arguments that animate evolutionary psychology.

Evolutionary psychology, in its modern incarnation, gained force in the 1980s and 1990s. But evolutionary approaches to human behaviour have erupted from time to time in psychology since its inception. For example, in the 1890s, thirty years after the *Origin of Species*, William James tried to fulfill Darwin's prediction in his seminal book, *Principles of Psychology*. *Principles* was both a founding work of experimental psychology and an attempt to ground this new field in a Darwinian framework. In it, James talked about *instincts*. This term was used to refer to specialized neural circuits that are common to every normal member of a species and that are the product of that species' evolutionary history. Taken together, this universal set of circuits constitutes (for our own species) what one can think of as *human nature*.

It was (and is) common to think that other animals are ruled by "instinct", whereas we humans have lost our instincts and had them replaced with "reason", "intelligence", or "learning". This consensus view is used to explain why we are so much more flexible and intelligent than other animals. William James, however, argued against this common sense view. He maintained that human behaviour is more flexibly intelligent than that of other animals because we have *more* instincts than they do, not fewer. If instincts are like tools in a toolbox, then the larger the number that the mind is endowed with, the more abilities it has. James's view fits presciently with work in modern computer science, in which each additional program expands the computer's ability to solve problems.

As a species, we have been blind to the existence of these instincts – not because we lack them, but precisely because they work so well. Because they process information so effortlessly and automatically, their operation disappears unnoticed into the background. The outputs they produce -- such as color, beauty, or the warmth of a smile -- seem to be part of the external world rather than products of the internal world. As William James argued, these instincts structure our thought so powerfully that it can be difficult to imagine how things could be otherwise. As a result, we take normal behaviour for granted: We do not realize that normal behaviour needs to be explained at all. This instinct blindness makes the study of psychology difficult, because our attention is misdirected away from the major features of our minds, into quirks and individual differences. To overcome this problem, James suggested that we try to make the "natural seem strange":

> "It takes...a mind debauched by learning to carry the process of making the natural seem strange, so far as to ask for the why of any instinctive human act. To the metaphysician alone can such questions occur as: Why do we smile, when pleased, and not scowl? Why are we unable to talk to a crowd as we talk to a single friend? Why does a particular maiden turn our wits so upside-down? The common man can only say, Of course we smile, of course our heart palpitates at the sight of the crowd, of course we love the maiden, that beautiful soul clad in that perfect form, so palpably and flagrantly made for all eternity to be loved! And so, probably, does each animal feel about the particular things it tends to do in the presence of particular objects. ... To the lion it is the lioness which is made to be loved; to the bear, the she-bear. To the broody hen the notion would probably seem monstrous that there should be a creature in the world to whom a nestful of eggs was not the utterly fascinating and precious and never-to-be-too-much-sat-upon object which it is to her. " (William James, 1890)

Making the natural seem strange is unnatural and effortful. It requires comparing our minds to those of other species, and to artificial minds, and it requires persistently asking why our minds are built the way they are, and not in other, logically possible ways. In short, it requires the twisted outlook of a Gary Larson cartoon. This change in perspective is a critical first step necessary to discovering and mapping the immense but hidden landscape of the human mind. Historically, many psychologists have avoided the study of our natural competences, thinking that there is little there to be explained. As a result, social psychologists, for example, are disappointed unless they find a phenomenon that would, as they say, surprise their grandmothers, and cognitive psychologists spend more time studying how we solve problems that, as a species, we are bad at, such as learning calculus or playing chess, rather than ones we are good at, such as reading emotional expressions or making friends. But our natural competences -- our abilities to see, to speak, to find someone beautiful, to reciprocate a favour, to fear disease, to fall in love, to initiate an attack, to experience moral outrage, to navigate a landscape, and myriad others -- are possible only because there is a vast and varied array of complex computational machinery supporting and regulating these activities. Because psychologists (whatever you may have heard to the contrary) are also members of the human species, and hence also suffer from instinct blindness, we have, from the dawn of science, neglected to study most of the interesting machinery in the human mind.

An evolutionary approach provides powerful lenses that correct for instinct blindness. It allows researchers to recognize what natural competences exist, it indicates that the mind is a heterogeneous collection of these competences and, most importantly, it provides explicit, schematic theories of their designs. These theories are far from complete, and require experimentation and research to flesh out, but they are far better guides than anything scientists have ever had in the past. Einstein once commented "It is the theory which decides what we can observe." He meant by this that, to take the actions necessary to detect something – to design the right experiment, look at the right part of the sky, or invent an essential piece of equipment – you must have a well-developed idea of what you are looking for. Evolutionary theories have made possible a flood of new observations and discoveries because they provide condensed descriptions of what is likely to be found in the unknown territories of the mind. An evolutionary focus is invaluable for psychologists, who are studying a biological system of fantastic complexity, because it can make the intricate outlines of the mind's design clearly visible. Theories of the adaptive problems our ancestors faced can guide the search for the cognitive programs that evolved to solve them; knowing what cognitive programs exist can, in turn, guide the search for their neural basis. A new route has been found into previously remote areas of the mind, and the prospects are galvanizing research.

The Standard Social Science Model

One of our colleagues, Don Symons, often remarks that you cannot understand what a person is saying until you know who they are arguing with. And so, to follow the evolutionary debates, you need to understand what both sides are saying. Applying evolutionary biology to the study of the mind has brought evolutionary psychology into conflict with a traditional view of its structure, which arose long before Darwin – a view we have elsewhere called the Standard Social Science Model.* This view is no historical relic: institutionally, it remains the dominant paradigm more than a century after Darwin and William James wrote, and its assumptions permeate our thinking.

Both before and after Darwin, a common view among philosophers and scientists has been that the human mind resembles a blank slate, virtually free of content until written on by the hand of experience. According to the 13th-century philosopher Aquinas, there is "nothing in the intellect that was not previously in the senses". Working within this framework, the 17th- and 18th-century British Empiricists and their successors produced elaborate theories about how experience, refracted through a small handful of innate mental procedures, inscribed content onto the mental slate.

Over the years, the technological metaphor used to describe the structure of the human mind has been consistently updated, from blank slate to switchboard to general-purpose computer. But the central tenet of these Empiricist views has remained the same. Indeed, it has become the reigning orthodoxy in mainstream anthropology, sociology and most areas of psychology. According to this orthodoxy, all of the specific content of the human mind originally derives from the "outside" -- from the environment and the social world -- and the evolved architecture of the mind consists solely or predominantly of a small number of general-purpose mechanisms that are content-independent. These hypothetical entities float, unverified, through our intellectual life under names such as "learning", "induction", "intelligence", "imitation", "rationality", "the capacity for culture", or simply "culture".

According to this view, the same mechanisms are thought to govern how one acquires a language, learns to recognize emotional expressions, thinks about incest, responds to an attack or a sonata, or adopts ideas about friendship and reciprocity – indeed everything but perception, which is often accepted as being innate. This is because the mechanisms that govern reasoning, learning, and memory are assumed to operate uniformly, according to unchanging principles, regardless of the content they are operating on or the larger category or domain involved. (For this reason, they are described as *content-independent* or *domain-general*.) Such mechanisms, by definition, have no pre-existing content built in to their procedures; they are not designed to construct certain contents more readily than others; and they have no features specialized for processing particular kinds of content. Because these hypothetical mental mechanisms have no content to impart, it follows that all the particulars of what we think and feel are derived externally, from the physical and social world. The social world organizes and injects meaning into individual minds, but our universal human psychological architecture has, supposedly, no distinctive structure that organizes the social world or imbues it with characteristic meanings. According to the Standard Social Science Model, the contents of human minds are primarily (or entirely) free social constructions, and so the social sciences are autonomous and disconnected from any evolutionary or psychological foundation. In this view, the evolutionary process explains the evolution of the human body, and the capacity for culture, but the blank slate nature of the human mind interposes a barrier between biology and human mental content that renders evolution irrelevant to human affairs. Unlike other animals, our evolution washed us clean of instincts and innate mental organization.

Three decades of progress and convergence in cognitive psychology, evolutionary biology and neuroscience have shown that this plausible and persuasive view of the human mind is radically defective. Evolutionary psychology provides an alternative framework that is beginning to replace this earlier view. According to this new perspective, all normal human minds reliably develop a standard collection of reasoning, emotional, and motivational circuits or programs. These programs are individually tailored to the demands of particular evolutionary functions, and often come equipped with what philosophers would once have called "innate ideas". There are far more of them than anyone had suspected, and they respond far more sensitively to the particulars of human life than anyone had imagined. Humans appear to have evolved circuits specialized for the domains of friendship, incest avoidance, coalitions, landscape preference, status, number, aggression, mating, language, intuiting what others are thinking, judging personality, and hundreds of other functions. These circuits organize the way we interpret our experiences, inject certain recurrent concepts and motivations into our mental life, give us our passions, and provide cross-culturally universal frames of meaning that allow us to understand the actions and intentions of others. Beneath the level of surface variability, all humans share certain views and assumptions about the nature of the world and human action by virtue of these universal circuits. Evolutionary psychologists have now embarked on the exploration of what until now has been largely unknown territory, the evolved programs that make up human nature. If this research continues, after several decades humanity will have an accurate, detailed, and empirically well-tested circuit diagram of human nature, rather than the traditional mishmash of vague impressions, contradictory claims, and endless particularistic descriptions. A scientifically validated model of human nature will transform the entire intellectual landscape we have been living in, providing a systematic foundation not only for psychology, but also for economics, anthropology, psychiatry, sociology, political science, philosophy, literary criticism, cultural studies, or any other field which implicitly or explicitly depends on assumptions about human nature.

2. BACK TO BASICS

How did evolutionary psychologists arrive at this new view? When rethinking a field, it is necessary to go back to first principles, to ask basic questions such as "What is behaviour?" "What do we mean by 'mind'?" "How can something as intangible as a 'mind' have evolved?" and "What is the mind's relation to the brain?". The answers to such questions provide the framework within which evolutionary psychologists operate. We shall try to summarize some of them here.

Psychology is that branch of biology that studies (1) brains, (2) how brains process information, and (3) how the brain's information-processing programs generate behaviour. Once one realizes that psychology is a branch of biology, inferential tools developed in biology -- its theories, principles, and observations -- can be used to understand psychology. Once one realizes that the brain is a computational system, it becomes critical to integrate computer science and cognitive science with evolutionary biology. And once one realizes that the brain evolved to solve computational problems faced by our hunter-gatherer ancestors, hunter-gatherer studies and paleoanthropology become necessary to provide the functional context within which to understand human psychology. Evolutionary psychology was founded on interlocking contributions from evolutionary biology, cognitive science, and anthropology, as well as many other fields, and cannot be understood without thinking through, from first principles, how these fields relate to each other and to the study of the mind.

Here are five basic principles, drawn from the intersection of these three fields, that evolutionary psychologists apply in attempting to understand the design of the human mind. These five principles can be applied to any topic in psychology. They organize observations in a way that allows one to see connections between areas as seemingly diverse as vision, reasoning, and sexuality.

Principle 1. The brain is a physical system. It functions as a computer. Its circuits are designed to generate behaviour that is adaptive given your environmental circumstances.

The brain is a physical system whose operation is governed solely by the laws of chemistry and physics. What does this mean? It means that all of your thoughts and passions and impulses are produced by chemical reactions going on in your head -- a materialist commonplace that, when deeply appreciated, is quite enough to send a jolt of yawning terror through your intimate chemistry. Moreover, each organ in your body evolved to serve a function, and the brain's evolved function is to produce behaviour that is sensitively contingent upon *information* from the environment. In other words, it is a physical system that processes information, and uses that information to guide behavior.

Realizing that the function of the brain is to process information has allowed cognitive scientists to resolve (at least one version of) the ancient mind/body problem. For cognitive scientists, brain and mind are terms that refer to the same system, described in two complementary ways -- either in terms of its physical properties (the brain) or in terms of its information-processing operation (the mind). Your desktop computer, too, can be described in these two complementary ways: in terms of electrons, magnetic fields, voltages, polarity, and so on, or in terms of its logical architecture and the programs running on it. Engineers designed your desktop computer and its programs. They selected and layed out the physical properties of the system so that these properties brought about certain information-processing relationships.

Evolution designed your brain: its physical organization was (naturally) selected because that physical organization brought about certain information-processing relationships: the ones that were adaptive for your ancestors.

So, the brain is a computer – a wet one, made of long-chained carbon polymers rather than silicon chips. The brain is composed of cells, primarily neurons and their supporting structures. Neurons are cells that evolved modifications that make them particularly well-suited for the transmission, storage, and processing of information. Electrochemical reactions cause neurons to fire.

Neurons are connected to one another in a highly organized way. One can think of these connections as circuits -- just as a computer has circuits. These circuits determine how the brain processes information, just as the circuits in your computer determine how it processes information. Neural circuits in your brain are connected to sets of neurons that run throughout your body. Some of these neurons are connected to sensory receptors, such as the retina of your eye. Others are connected to your muscles. Sensory receptors are cells that are specialized for gathering information from the outer world and from other parts of the body. (You can feel your stomach churn because evolution equipped it with sensory receptors, but you cannot feel your spleen, which lacks them.) Sensory receptors are connected to neurons that transmit this information to your brain. Other neurons send information from your brain to motor neurons. Motor neurons are connected to your muscles; they cause your muscles to move. Reduced to this most basic level of description, this muscular movement is what everyone calls *behaviour*. It is the need to regulate behaviour, often in response to externally monitored conditions, that favored the evolution of brains and these other exacting arrangements.

In contrast, organisms that don't move, don't need – and so don't have – brains. Trees don't have brains, bushes don't have brains, flowers don't have brains. In fact, there are some animals that don't move during certain stages of their lives -- and, during those stages, they don't have brains. The sea squirt, for example, is an aquatic animal that inhabits oceans. During the early stage of its life cycle, the sea squirt swims around looking for a good place to attach itself permanently. Once it finds the right rock, and attaches itself to it, it doesn't need its brain anymore because it will never need to move again. So it eats (resorbs) most of its brain. After all, why waste energy on a now useless organ? (The alert reader will have noted the obvious biological parallels between the sea squirt and the professor: only during the job-searching portion of the life stage does the scholar actually need a brain; once a tenured post has been found, and the gut has attached to the payroll, the brain is no longer necessary.)

In short, brains are not inevitable. They are wet computers: biological machines that were created by natural selection. The brain's programs are embodied in neural tissue, and these programs were designed to collectively compute and execute the answer to the question, *What behaviour is most likely to be adaptive for this organism, given the information available to it about its situation?*

Of course, to say that the function of your brain is to generate behaviour that is likely to be adaptive, given your situation, is not saying much, unless you have a well-justifed method for spelling out what *adaptive* means. What counts as adaptive behaviour? As we will see, the theory of natural selection allows us answer this question with precision.

Principle 2. The brain's circuits were designed by natural selection to solve problems that our ancestors regularly faced during our species' evolutionary history.

Other minds. Each species' ancestors faced different environments and different sets of adaptive problems. As a result, different species have evolved different sets of neural programs. Contrasting the behaviour of various species can, therefore, tell us something profound about the causes of behaviour.

Consider the differences between you and a dung fly, for example. You have sensory receptors that are stimulated by the sight and smell of feces -- to put it more directly, you can see and smell dung. So can a dung fly. But on detecting the presence of feces in the environment, what counts as adaptive behaviour for you differs from what is adaptive for the dung fly. On smelling feces, adaptive behaviour for a female dung fly is to move toward the feces, land on them, and lay her eggs. Feces are food for a dung fly larva -- therefore, adaptive behaviour for a dung fly larva is to eat dung. And, because female dung flies frequent piles of dung, adaptive behaviour for a male dung fly is to buzz around these piles, trying to mate; for a male dung fly, these piles are alluring locales for erotic adventures.

But for you, feces are potentially a source of contagious diseases. For you, they are not food, they are not a good place to raise your children, and they do not promise sexual encounters. So, adaptive behaviour for you is to minimize direct contact, moving away from the source of the smell. Perhaps your facial muscles will form the cross-culturally universal disgust expression as well, in which your nose wrinkles and yourtongue protrudes slightly, as it would were you ejecting something from your mouth. If someone were to ask a human why she avoided the feces, we would consider it self-explanatory if she answered "Because it is disgusting!". But for a female dung fly, looking for a good neighbourhood and a gracious home for raising her children, that pile of dung is an Eden. So which is it really, disgusting or welcoming? In any discussion between human and fly – or in any scientific discussion – we cannot take as self-explanatory our species' reactions. We must look further to find their causes and meaning.

Comparisons across species shake us out of our instinct blindness, force us to question the seemingly self-evident causes of our behaviour, and show us how our behaviour could be organized very differently. More importantly, it drives home the pivotal point that environments do not and cannot, in and of themselves, cause behaviour. The same environmental stimulus – feces – leads to completely different behaviours in different species. The neural circuits in the minds of each species imbue the same elements of the world with radically different meanings and promptings. For this reason, the environment cannot be used, by itself, as an explanation for behaviour. Although we might wish to say "My environment made me do it!", and many social scientists have founded their theories on psychology-free notions of environmental causation, no modern scientist can responsibly think in such a fashion any more.

In principle, a computer or neural circuit could be designed so that any given stimulus in the environment (e.g., feces) could cause any kind of resulting behaviour (approach, avoid, dance, meditate, declaim, sculpt, etc.). Which behaviour a stimulus gives rise to is a function of the neural circuitry of the organism. This means that if you were a superhuman designer of brains, you could have engineered the human brain to respond in any way that you wanted, to link any environmental input to any behavioural output. You could have made a human being who scowls when pleased, is erotically transported by cuniform writing, or builds a beaver dam, coos, and devotedly incubates chicken eggs with her body heat whenever the days grow short enough. To explain behaviour, therefore, we need a theory of brain organization that describes how circuits are designed to respond to environmental inputs throughout the lifecycle, and why they have the form they do. We will call this organization the *design of the mind*. Because how the brain is organized to respond to the environment, prior to experience, cannot itself be supplied by the environment, it is easy to see that the design of the mind – including its learning circuits – must be built in to the developing brain. This means that the design was created by evolution.

If superhuman brain designers could have manufactured any relationship at all between the world and the behaviour our minds produce, what did the actual designer of the human brain do, and why? Why do we find fruit sweet and dung disgusting, given that these objects in the world do not, by themselves, explain our responses to them? In other words, how did we get the circuits that we have, rather than those that the dung fly, the beaver, or the sea cucumber has?

When we are talking about a personal computer, the answer to this question is simple. Its circuits were designed by engineers, and the engineers designed them one way rather than another so they would successfully solve problems that the engineer wanted them to solve: problems such as adding or subtracting, redrawing a screen rapidly, or accessing a particular address in the computer's memory. Your neural circuits were also designed to solve problems. But they were designed by the evolutionary process instead of by a human (or superhuman) engineer. Indeed, we can be more specific about the identity of the designer. Though many chance evolutionary forces have shaped us in various ways, natural selection is the component of the evolutionary process that is capable of engineering complexly organized functional machinery into organisms. So, whenever we are looking at complex functionality in our brains or our bodies, natural selection was the designer or engineer.

Natural selection builds *adaptations* – that is, problem-solving machinery – to solve evolutionarily long-standing *adaptive problems*. Accordingly, we can understand and explain the design of each program our neural circuits create by understanding (1) how natural selection works, and (2) what adaptive problems our adaptations were designed to solve.

How does natural selection work? The notion that species evolve – that their design changes over time – had been proposed and hotly debated before Darwin was born. But the early evolutionists lacked a clear and convincing account of how or why this happens. That is what Darwin and Wallace provided. They discovered a materialist mechanism – natural selection – that explains how organisms acquire their design, as well as why that design changes over time. The revolution that ensued bears Darwin's name because he is the one who elaborated the theory, provided the most extensive evidence for it, and was willing to pursue its implications wherever they led – even when they led to the human mind.

Many breakthroughs in science happen not because of new data, but because of a new way of looking at things. This was true for Darwin. Everyone already knew that organisms reproduce, and that when they do, they give rise to similar organisms: rabbits give birth to rabbits, not to ducks. They also knew that, while offspring closely resemble their parents, they are not perfect replicas of them. They vary a bit, and some of these variants are able to perform certain tasks, such as producing milk, better than others. This was common knowledge based on centuries of animal husbandry in which people selectively bred individual animals with special abilities – cows that produced more milk, sheep that grew softer wool. And Darwin, like Descartes, Harvey, and many others before him, knew that an organism can be thought of as a machine: a system whose parts are designed to perform certain functions.

All of these facts fall into place, Darwin realized, if you think of an organism as a *self-reproducing machine*. What distinguishes living from nonliving machines is reproduction: the presence in a machine of devices (organized components) that cause it to produce new and similarly reproducing machines. Given a population of living machines, this property – self-reproduction – will drive a system of positive and negative feedback that can explain the remarkable fit between organisms and their environment.

In contrast to human-made machines, which are designed by inventors, living machines acquire their intricate functional design over deep time, as a downstream *consequence* of the fact that they reproduce themselves. Indeed, modern Darwinism has an elegant deductive structure that logically follows from Darwin's initial insight that reproduction is the defining property of life, the driving force that causes species to change over time. That logic is as follows:

When an organism reproduces, replicas of its design features are introduced into its offspring. But the replication of the design of the parental machine is not always error-free. As a result, randomly modified designs (i.e., mutants) are introduced into populations of reproducers. Because living machines are already exactingly organized so that they cause the otherwise improbable outcome of constructing offspring machines, random modifications will usually introduce disruptions into the complex sequence of actions necessary for self-reproduction. Consequently, most newly modified but now defective designs will remove themselves from the population – a case of negative feedback.

However, a small number of these random design modifications will, by chance, improve the system's machinery for causing its own reproduction. Such improved designs (by definition) cause their own increasing frequency in the population – a case of positive feedback.

This increase continues until (usually) such modified designs outreproduce and thereby replace all alternative designs in the population, leading to a new species-standard design. After such an event, the population of reproducing machines is different from the ancestral population: The population- or species-standard design has taken a step "uphill" toward a greater degree of functional organization for reproduction than it had previously. Over the long run, down chains of descent, this feedback cycle pushes designs through state-space toward increasingly well-engineered – and otherwise improbable – functional arrangements. These arrangements are *functional* in a specific sense: the elements are well-organized to cause their own reproduction in the environment in which the species evolved.

For example, if a more sensitive retina, which appeared in one or a few individuals by chance mutation, causes predators to be detected more quickly, individuals who have the more sensitive retina will produce offspring at a higher rate than those who lack it. Those of their offspring that inherit that more sensitive retina will also evade predators better and therefore produce offspring at a higher rate, and so on and so on, down the generations. By promoting the reproduction of its bearers, the more sensitive retina thereby promotes its own spread over the generations, until it eventually replaces the earlier-model retina and becomes a universal feature of that species' design. This spontaneous feedback process – natural selection – causes functional organization to emerge naturally and inevitably, without the intervention of an intelligent designer or supernatural forces.

Genes are simply the means by which design features replicate themselves from parent to offspring. They can be thought of as particles of design: elements that can be transmitted from parent to offspring, and that, together with an environment, cause the organism to develop some design features and not others. Because design features are embodied in individual organisms, there are ususally only two ways they can propagate themselves: by solving problems that increase the probability that offspring will be produced either by the organism in which they are situated, or by that organism's kin.

An individual's relatives, by virtue of having received some of the same genes from a recent common ancestor, have an increased likelihood of having the same design feature as compared to other conspecifics. This means that a design feature in an individual that causes an increase in the reproductive rate of that individual's kin will, by so doing, tend to increase its own frequency in the population. A circuit that motivates an individual to help feed her sisters and brothers, if they are starving, is an example of a program that increases kin reproduction. When the individual's siblings reproduce, they might pass on this same circuit to their children. Hence, design features that promote both direct reproduction and kin reproduction, and that make efficient trade-offs between the two, will replace those that do not.

How well a design feature systematically promotes direct and kin reproduction is the bizarre but real engineering criterion determining whether a specific design feature will be added to or discarded from a species' design. Therefore, we can understand why our brains are constructed in the way they are, rather than in other perfectly possible ways, when we see how its circuits were designed to cause behaviour that, in the world of our ancestors, led to direct reproduction or kin reproduction.

We can now define the concept of *adaptive behaviour* with precision. Adaptive behaviour, in the evolutionary sense, is behaviour that tends to promote the reproduction of the individual or that individual's blood relatives. By promoting the replication of the genes that built them, circuits that – systematically, and over many generations – cause adaptive behaviour become incorporated into a species' neural design. In contrast, behaviour that undermines the reproduction of the individual or his or her blood relatives removes the circuits causing those behaviours from the species, by removing the genes that built those circuits. Such behaviour is *maladaptive*, in the evolutionary sense.

So, evolutionists continually analyze how design features are organized to contribute to lifetime reproduction not because of a warped and biasing obsession with sexuality, but because reproduction was the final causal pathway through which a functionally improved design feature caused itself to become more numerous with each passing generation, until it became standard equipment in all ordinary members of the species.

Adaptive problems create adaptations. Enduring conditions in the world that create reproductive opportunities or obstacles, such as the presence of predators, the need to share food, or the vulnerability of infants, constitute *adaptive problems.* Adaptive problems have two defining characteristics. First, they are conditions or cause-and-effect relationships that many or most individual ancestors encountered, reappearing again and again during the evolutionary history of the species. Second, they are that subset of enduring relationships that could, in principle, be exploited by some property of an organism to increase its reproduction or the reproduction of its relatives. Enduring relationships of this kind constitute reproductive opportunities or obstacles in the following sense: *if* the organism had a property that interacted with these conditions in just the right way, *then* this property would cause an increase in its own reproductive rate.

You can think of these reproductive opportunities and obstacles as *problems*. A property is a *solution* to such a problem when it allows organisms with this property to take advantage of prevailing conditions, where "advantage" means a reproductive advantage. If a bird would realize a reproductive advantage by being able to travel at night, and stars are prevailing conditions that – given the right brain mechanism – would make this possible, then a brain mechanism that uses stars for navigation would be a solution to the problem of traveling at night. Egg-eating predators pose an obstacle to a bird's reproduction. A property that circumvents this obstacle – e.g., a program that causes the bird to remove from its nest broken eggs shells whose bright white interiors are easily spotted by predators – would be a solution to this problem. A property is a solution to an adaptive problem if it had the systematic effect, over generations, of increasing the reproduction of individual organisms or their relatives. The causal chain linking that property to reproduction may be indirect, and the effect on the organism's own offspring or the offspring of kin may be relatively small. As long as its consequences on relative reproduction are the cause of its spreading through the population, that property is a solution to an adaptive problem. All solutions are, of course, temporary and subject to improvement over time. But each modification that spread because it improved reproduction – however stop-gap or impermanent it may turn out to have been – counts as a solution to an adaptive problem.

Most adaptive problems have to do with relatively mundane aspects of how an organism lives from day to day: what it eats, what eats it, who it mates with, who it socializes with, how it communicates, and so on. Adaptive problems for our hunter-gatherer ancestors included such recurrent tasks as giving birth, winning social support from band members, remembering the locations of edible plants, hitting game animals with projectiles, breast-feeding, breathing, identifying objects, recognizing emotional expressions, protecting family members, maintaining mating relationships, self-defense, heart regulation, assessing the character of self and others, causing impregnation, acquiring language, maintaining friendships, thwarting antagonists, and so on.

An enduring adaptive problem constantly selects for design features that promote the solution to the problem. Over evolutionary time, more and more design features accumulate that fit together to form an integrated structure or device that is well-engineered to solve its particular adaptive problem. Such a structure or device is called an *adaptation.* Indeed, an organism can be thought of as largely a collection of adaptations, such as the functional subcomponents of the eye, liver, hand, uterus, or circulatory system. Each of these adaptations exists in the human design now because it contributed ancestrally to the process of self and kin reproduction.

The adaptive problems that required information-processing for their solutions selected for neural adaptations organized to compute these solutions. Hence, the brain is largely a collection of computational adaptations – sometimes called adaptive specializations, modules, programs, and circuits, as well as cognitive, information-processing, neural, psychological, or mental adaptations.

Over evolutionary time, our neural circuits were cumulatively added to the design of the human brain because they reasoned or processed information in a way that enhanced the adaptive regulation of behaviour and physiology for these enduring adaptive problems. Such cognitive adaptations include emotion programs, such as fear of falling, fear of snakes, or parental love; motivational programs, such as sexual attraction or revenge; learning programs such as the language acquistion device or the food aversion system; and reasoning instincts such as bluff detection, mechanical inference, and the cheater detection circuits that we will discuss at the end of this book. Even "learned" behaviours, such as speaking English, are the product of evolved learning programs. Consequently, the mental and neural organization that results from learning is simply another example of the operation of our evolved adaptations, not an exception.

It is vital to understand that the only kind of problems that natural selection designed circuits for solving are adaptive problems. Therefore, the only functional machinery we can expect the human mind to include is machinery that was specifically designed to solve some ancestral human problem. As we will discuss, the structure of the evolved problem-solver necessarily reflects the structure of the adaptive problem. So, we study the nature of the adaptive problems our ancestors faced in order to get clues into the structure of the mental components that evolved to solve them. If you can identify an adaptive problem that was important for our ancestors, you have taken a first step toward discovering a new aspect of the human mind.

Reciprocally, it is important to realize that our circuits weren't designed haphazardly to solve just any kind of problem, from taking square roots to speaking in rhyme. Problems our ancestors did not face (such as taking square roots) played no role in constructing our minds. Moreover, any problems that our ancestors did face, but that had no impact on lifetime reproduction or kin reproduction also played no role in shaping our minds. Instead, our circuits were retained by selection solely to the extent that they solved the problems that our ancestors actually faced, in proportion to their reproductive importance and frequency.

Novel problems.   Obviously, we are *capable* of solving many problems that no hunter-gatherer ever had to solve -- we can learn to drive cars, slam dance, speak in rhyme, call science an oppressive myth, and play video poker.   Moreover, no one knows what humans may be able to achieve, through explicit instruction, elaborate tools, computers, technology, and evolutionarily novel social and cultural arrangements. But the fact that our evolved programs allow us to solve many evolutionarily novel problems does not mean that our brains consist of a blank slate plus the magical ability to do anything (see Principle 4).

Indeed, our minds are incapable of solving certain problems on their own – that is, without computers and other tools to do the computations and store the information for us. Yet, by seeing what other animals can do, we know with certainty that these same problems can easily be solved by a single brain – just not by a human brain. Unlike bats, we cannot echolocate flying insects; unlike bees, we do not automatically know and use the solar ephemeris; unlike food-caching birds, we do not automatically remember, for months on end, the whereabouts of thousands of food items that were buried in arbitrary, unmarked locations.

When our native (rather than educated) abilities do allow us to solve certain new kinds of problems, this is an accidental, hit-or-miss byproduct of circuits that were designed to solve an earlier, more specific set of adaptive problems. For example, when our ancestors became bipedal -- walking on two legs instead of four -- they had to develop an excellent sense of balance. As a result, we evolved very intricate mechanisms in our inner ear and elsewhere that allow us to balance superbly. These specializations open the door to far more than walking or running. Because of them, we can jitterbug, skateboard, ski, surf, skate, or even ride the winds, sky-boarding. These abilities do not derive from our hunter-gatherer ancestors shooting the curl in the primordial soup. Nor are they evidence for a blank slate mind. They are instead byproducts of adaptations designed for arboreal locomotion and a subsequent history of balancing while walking or running on two legs.   To discover the circuit architectures of our various abilities, and to understand scientifically why our minds have the design that they do, we need to cut through all the dazzlingly diverse things the human mind is capable of doing, and focus in on the hunter-gatherer activities it was specifically engineered to perform (see Principle 5). It is these ancestral adaptive problems and activities that selected for the mind's adaptations, and so studying these problems allows us to reverse engineer the adaptations that evolved to solve them.

What natural selection is not. Despite its clear formal structure, scientists – including many biologists – harbor many misconceptions about natural selection, and therefore have distorted expectations of what functions minds will have evolved to accomplish. For example, natural selection does not engineer organisms to necessarily act for the good of the species, for species survival, for social harmony, or even for personal happiness, health, or individual survival. It is possible that the evolved properties of organisms may occasionally cause such beneficial outcomes as an incidental byproduct – but they were not actually designed to produce these outcomes for their own sake.

For example, in aggregate, any species would reproduce more effectively without having to put energy into aggression or weaponry designed for competing with other members of the species. Indeed, individuals who avoided aggression usually would have lived longer and more pain-free lives. However, individuals with mutations that gave them more effective aggressive weaponry or strategies gained more mates and resources, and so produced more descendants, who had these same characteristics. This is why members of virtually all social animal species – including humans – evolved the ability to engage in aggressive behaviours, along with neural programs to compute when, how, or if to act aggressively. By understanding what conditions these programs monitor to decide whether a situation calls for an aggressive response, and by mapping the adaptive logic these programs use to make such decisions, evolutionary psychologists might eventually be able to increase the welfare of our own species by minimizing or even eliminating violence in human societies. But we can't count on natural selection itself to help our species in this way: The welfare of the species as a whole is irrelevant to how natural selection drives the displacement of some traits by others within a species. In engineering organisms, selection only multiplies design features that cause themselves to replicate more effectively, usually by increasing the number of offspring produced by the specific organism they are embodied in, or its relatives.

Similarly, all organisms die sooner or later, so survival per se is not what drives natural selection. Organisms have evolved many features that sacrifice survival for reproduction, because the engine that drives natural selection is the spread of design features that systematically cause more effective reproduction. Consider the fate of a circuit that had the effect, on average, of enhancing the net lifetime reproduction of the organisms that sported it, but shortened their average lifespan in so doing. If this effect persisted across generations, then its frequency in the population would increase until (virtually) all members of the species carried the trait. Indeed, this is a common evolutionary outcome rather than simply an idle thought experiment: In many species, males engage in dangerous combat or travel long distances through predator patrolled areas for the opportunity to mate, even though nearly every male would live far longer by avoiding such risks. Mutations of all types continuously enter species: among baboons and elephant seals, pacifist males surely appear periodically. The point is that, although the nonpacifist, combative male baboon may live a shorter life, he will leave more descendants and so, over the generations, the species comes to be composed of combative, rather than pacifist, males.

Many clinical psychologists, psychotherapists, anthropologists, and sociologists call a behaviour "adaptive" when it leads to personal happiness or well-being. But this bears no relation to the evolutionary meaning of the term. Any circuit whose average effect was to lower its own rate of reproduction into the next generation would eventually disappear from the population, no matter what effect it had on personal happiness, survival, or anything else. For example, it is perfectly possible to engineer a brain that lives in Buddha-like bliss, bathed in endogenous opiates, indifferent to worldly action or concern. Such an individual, liberated from the wheel of attachment and desire, is unlikely to leave many descendents. Because of this, natural selection has created organisms whose circuits usually enchain them to the desire for worldly rather than spiritual success. That is, evolutionarily well-designed organisms are built so that their desires track the kinds of things that would normally have increased lifetime reproduction for self and family in the environments in which they evolved, such as food, matings, warmth, safety, and social status. As Gautama Buddha himself observed, desire is nearly inexhaustible – a fact that becomes comprehensible in light of how natural selection works. Because reproduction could usually have been increased through an improvement in situation, selection should have shaped human desire to scan the world for opportunities to hill-climb to better situations, no matter how well-off the individual already is. Natural selection has not designed our brains to guarantee happiness or to be easily contented, and it is no accident that sainthood, the cessation of desire, and Nirvana are notoriously difficult for most humans to achieve. (Of course, no one knows what may be possible as the arts and sciences advance, and we become more knowledgeable about the engineering details of our evolutionary design.)

One virtue of the theory of natural selection is that it is not mysterious or occult or esoteric. It can be clearly understood by anyone who is willing to sweep aside their preconceptions and instead think about nothing but simple physical causation, applied in a clear, linear step by step fashion, to sets of individual organisms over time. Do ostriches stick their heads in holes in the ground when menaced by predators? Assume some individual ostriches had such inherited, defective design features. What would happen? They would be eaten by predators far more often than the ostriches whose design features caused them to flee predators. Physically, the action of predators removes defective ostrich designs from the world, and so those ostriches do not become parents as often. Physically, what is reproduced into the next generation are the design features that cause greater lifetime reproduction. Physically, the composition of the population changes, generation by generation, so that designs that cause greater reproduction become more numerous while designs that are less well-engineered become rarer or disappear entirely. Simply by knowing the theory of natural selection, one can deduce that ostriches who stick their heads in the ground to avoid seeing scary things, or lemmings who commit suicide under the irresistible pressure of social conformity, are useful social metaphors, but not biological realities. These examples are obvious. But applied carefully, and in a sustained way, the theory of natural selection can lead us to discover, and help us to understand, very subtle aspects of human and nonhuman design.

Principle 3. Consciousness is just the tip of the iceberg; most of what goes on in your mind is hidden from you. As a result, your conscious experience can mislead you into thinking that our circuitry is far simpler that it really is. Most problems that you experience as easy to solve are actually very difficult to solve -- they require very complicated neural circuitry.

You are not, and cannot become, consciously aware of most of your brain's ongoing activities. Think of the brain as the entire U.S. federal government, and of your consciousness as the President of the United States. Now think of yourself -- the self that you consciously experience as "you" -- as the President. If you were President, how would you know what is going on in the world? Officials, such as the Secretary of Defense, would come and tell you things -- for example, that the Pakistani military is preparing to intervene in Kashmir. How do members of the Cabinet know things like this? Because of bureaucrats in the State Department, CIA analysts pouring over satellite data, National Security Agency specialists listening to electronic interceptions, intelligence operatives in the Indian subcontinent and in other parts of the world, military attaches and advisors stationed in the Middle East and elsewhere, as well as journalists and reporters – perhaps tens of thousands of individuals, investigating, interpreting, communicating, and evaluating enormous amounts of information from all over the world. But you, as President, do not -- and in fact, could not possibly -- know what each of these thousands of individuals were doing when gathering this information over the last few months: what each of them saw, what each of them read, who each of them talked to, what satellite imaging techniques were used (or their limitations), what conversations were clandestinely taped, whose offices were bugged. All that you, as President, know is the final conclusion that the Secretary of Defense came to based on the information that was passed on to him. And all he knows is what other high level officials passed on to him, and so on. In fact, no single individual knows all of the facts about the situation, because these facts are distributed among thousands of people. Moreover, each of the thousands of individuals involved knows all kinds of details about the situation that they decided were not important enough to pass on to higher levels.

So it is with your conscious experience. The only things you become aware of are a few high level conclusions (I realize this, I am angry at that, I want something else) passed on by thousands and thousands of specialized mechanisms. Some are gathering sensory information from the world, others are analyzing and evaluating that information, checking for inconsistencies, and filling in the blanks; still others are using these cues to select an emotion program to activate, deciding what it all means, or feeding the situation through the motivational system, to get direction on what to do.

It is important for any scientist who is studying the human mind to keep this in mind. In investigating how the mind works, your conscious experience of yourself and the world can suggest some valuable hypotheses. But these same intuitions can seriously mislead you as well. They can fool you into thinking that our neural circuitry is much, much simpler and more transparent than it really is.

Consider vision. Your conscious experience tells you that seeing is simple: You open your eyes, light hits your retina, and – voila! – you see. It is effortless, automatic, reliable, fast, unconscious and requires no explicit instruction – no one has to go to school to learn how to see. But this apparent simplicity is deceptive. Your retina is a two-dimensional sheet of light sensitive cells covering the inside back of your eyeball. Figuring out what three-dimensional objects exist in the world based only on the light-dependent chemical reactions occurring in this two dimensional array of cells poses enormously complex problems – so complex, in fact, that no computer programmer has yet been able to create a robot that can see nearly as well as we do. You see with your brain, not just your eyes, and your brain contains a vast array of dedicated, special-purpose circuits, each set specialized for solving a different part of the problem. What is involved in seeing your mother walk? You have circuits that are specialized for (1) parsing surfaces into discrete objects; (2) analyzing their shapes; (3) detecting the presence of motion; (4) detecting the direction of motion; (5) judging distance; (6) analyzing colour; (7) identifying an object as human; (8) recognizing whose face it is (Mother vs. everyone else), as well as hundreds of other tasks and subtasks. Each individual circuit is shouting its information to higher level circuits, which check the "facts" generated by one circuit against the "facts" generated by the others, resolving contradictions, blocking discrepant information, and so on, until they settle on a high level interpretation. Then these conclusions are handed over to even higher level circuits, which piece them all together and hand the final report to the President – your consciousness. But all the "president" ever becomes aware of is the sight of Mother walking, gin bottle in hand. Although each circuit is specialized for solving a delimited task, they work together to produce a coordinated functional outcome – in this case, your conscious experience of the visual world. Seeing seems effortless, automatic, reliable, and fast to us precisely because it sits on top of an immense, unseen pyramid of intricate, functionally dedicated machinery.

So our intuitions deceive us. We experience an activity as easy when evolution has provided us with specialized machinery to perform it, regardless of how computationally complex it is in reality. And evolution built machinery to solve tasks that were *natural* – that is, that were important and common adaptive problems for our distant ancestors. In contrast, we experience certain activities as difficult, such as playing chess or finding the mean of a string of numbers, when we have no machinery specifically designed for solving them, even if they are computationally simple. As a result, mental tasks are experienced as difficult if they are evolutionarily novel, that is, not similar to any ancestral activity.

Thus our conscious experience of an activity as "easy" or "natural" leads us to grossly underestimate the complexity of the circuits that make it possible. Doing what comes naturally, effortlessly or automatically is rarely simple from an engineering point of view. To find someone beautiful, to fall in love, to feel jealous – all can seem as simple and automatic and effortless as opening your eyes and seeing, so simple that it seems as if there is nothing much to explain. But these activities feel effortless only because there is a vast array of complex, evolved, unnoticed neural circuitry supporting and regulating them. Now that we know they are there, as scientists we can proceed to explore them. This was and is not possible for researchers who are unaware of their existence.

Principle 4. Different neural circuits are specialized for solving different adaptive problems.

Functional specialization. Despite a century's investment in the Standard Social Science Model, with its hypothesis that the mind is a general learning system connected to a blank slate, researchers are coming to recognize that our minds consist of a very large number of distinct circuits that are functionally specialized.

Why should this be true? A basic engineering principle is that the same device is rarely capable of solving two different problems equally well. We have both screwdrivers and saws because each solves a particular problem better than the other. It would be futile to cut planks of wood with a screwdriver or to turn screws with a saw.

For exactly the same reason, natural selection has divided our body into organs such as the heart and the liver. Pumping blood throughout the body and detoxifying poisons are two very different problems. Consequently, your body has evolved a different machine for solving each of them. The design of the heart is specialized for pumping blood; the design of the liver is specialized for detoxifying poisons. Your liver cannot function as a pump, and your heart cannot detoxify poisons.

The same principle applies to the mind. When carefully considered, it leads to the conclusion that the mind has many independent evolved programs. One way of seeing this is to put yourself, once again, in the position of a superhuman engineer. Your task is to design an organism like ourselves – one that has values and uses them to make choices. What would your organism be like if you gave it only one set of choice criteria?

Assume your science project is to design a human female, and first, you intend to build in the ability to choose more nutritious foods over less nutritious foods or non-foods. Natural selection has engineered into humans an elaborate set of neural circuits organized to choose nutritious food on the basis of taste, smell, and digestive consequences. Knowing this, you decide to implant your science project with the same programs. But if this is the only set of choice criteria she has, what kind of *mate* would she end up choosing? A goat cheese pizza or a giant chocolate bar? Although superior to a bad date, they won't be adequate as a parent to her children. To solve the adaptive problem of finding the right mate, her mental machinery would have to be guided by qualitatively different standards and values than when she is choosing the right food, or the right word, or the right path to get home. And that's just what natural selection did: There is a large and vigorous literature in evolutionary psychology showing that the criteria that guide human mate choice are specialized for this task (and not nearly as capricious as you might think!)*

We humans solve many different adaptive problems well. To accomplish these feats, *there must be at least as many independent evolved mental programs as there are adaptive domains in which the standards for successful behaviour are qualitatively different.* There are hundreds or thousands of these domains, ranging from food choice, mate choice, friendship maintenance, predator-escape, contagion avoidance, thermoregulation, and fluid intake, to social status, grammar acquisition, child survival, deception detection, aggressive defense and offense, foraging decision-making, incest avoidance, navigation, and coalition formation.

Environments themselves cannot provide organisms with definitions of problem-solving success – remember that avoiding dung or seeking dung are both adaptive solutions for different species. Therefore many mental problem-solvers must be built in to the structure of the brain by evolution. For this and many other reasons, the brain must be composed of a large collection of evolved circuits, with different circuits specialized for solving different problems. You can think of each of these specialized circuits as a computer that is dedicated to solving one kind of problem. What we are coming to understand is that the brain is a diverse collection of thousands of dedicated computers networked together. As a part of this system there must, of course, be circuits whose design is specialized for integrating the output of all these dedicated computers, in order to produce behaviour. But this only gives us the illusion that the mind and the self are unitary – it does not change the underlying reality.

Functional specialization can take many forms. For choice behavior, knowledge of the appropriate criteria must somehow be embodied in the program, either as a database, or implicitly, within procedures or in the nature of the cues to which the procedures that cause attraction, repulsion or disinterest respond. But information about proximal goals is not the only kind of functional specialization that one sees in the mind. For example, specialized circuits can often be designed to make inferences and produce knowledge far more efficiently than general-purpose circuits can. Biological machines are tailored to the structure of the environments in which they evolved, and information about the stably recurring properties of these ancestral worlds can be embodied in the very way their procedures work. For example, one function of vision is object recognition, and this is easier if the same object – e.g., a banana – appears to have the same color – yellow – from one situation to the next, regardless of changes in the wavelengths of the light illuminating it. This is called *colour constancy*, and our visual system does it very well. Natural selection has created colour constancy circuits that automatically compensate for the wild changes in illumination that occur on the surface of the Earth as the sun traverses the sky, and under variations in cloud cover and forest canopy. As a result, the banana looks yellow to us at high noon and at sunset, even though, objectively speaking, it is swamped by red light at sunset, causing far more red than yellow light to reach our eyes from its surface. Natural – that is, ancestrally recurrent – changes in terrestrial illumination pose no problems for these circuits, because they are calibrated to them: their procedures were shaped by them and embody knowledge about them. But these circuits cannot compensate for evolutionarily novel changes in illumination, such as the unearthly spectrum cast by the sodium vapor lights that illuminate many parking lots at night. The cars that we think of as red and green and blue all look a muddy brown when they are illuminated by these golden lights, because our colour constancy mechanisms were not shaped by, and embody no knowledge of, the spectral properties of sodium.*

Evolved crib sheets. This principle applies not just to perception, but to learning and reasoning as well. Many evolved problem-solvers are equipped with crib sheets: they come to a problem already "knowing" a lot about it. This allows them to be far more intelligent than they otherwise would be if they embodied no innate knowledge. For example, a newborn's brain has response systems that expect faces to be present in the environment; babies less than 10 minutes old turn their eyes and head in response to face-like patterns, but not to scrambled versions of the same pattern. Neural maturation brings other evolved circuits on line subsequently. Infants have strong assumptions, deriving from the evolutionary past, about how the world works and what kinds of things it contains, even at 2 ½ months (the point at which they can see well enough to be tested). They assume, for example, that the world will contain rigid objects that are continuous in space and time, and they have preferred ways of dividing the world into separate objects. Indeed, an infant's mind is designed to *privilege* some hypotheses about what counts as an object over others. Ignoring shape, colour, and texture (all of which they can see), they treat any surface that is cohesive, bounded, and moves as a unit as a single object. Another privileged hypothesis is that solid objects are impenetrable. So when one solid object appears to pass through another, these infants are surprised, just as you or I would be.*

A baby with a completely open mind – one lacking any privileged hypotheses – would be undisturbed by such displays. Why shouldn't a toy train travel smoothly through a solid block of wood? If the superhuman engineer were to remove these privileged hypotheses from the baby's mind, the baby would be left without informative guidance in the world in which we actually live. By definition, a blank-slate system must entertain all possible hypotheses: that it was born into a world in which objects are like mercury droplets, no one has a face, and surfaces that move together are physically unconnected to each other. These are properties of imaginable universes, but not of the one in which we evolved. There is nothing in our evolutionary past that would cause our brains to be organized in such a futile way.

So babies have innate ideas built into them about the universe and niche in which they actually evolved, instead of being prepared to deal with all worlds, whether they exist or not. In watching objects interact, babies less than a year old distinguish causal events from non-causal ones that have similar spatio-temporal properties; they distinguish objects that move only when acted upon from ones that are capable of self-generated motion (making the inanimate/animate distinction)*; and they assume that the self-propelled movement of animate objects is caused by invisible internal states – goals and intentions. Toddlers have a well-developed mind-reading system (that is, a system for intuiting what is on others' minds), which uses eye direction and movement to infer what other people want, know, and believe. This system is *domain-specific*: it is designed only for understanding the behavior of animate beings. It is *content-dependent*: it is activated by stimuli that have properties ancestrally associated with animate beings, such as eyes or self-propelled motion (seeing a boulder rarely excites curiosity about its hopes, ambitions, or beliefs). And it is *functionally-specialized*: it is designed to compute beliefs, desires, and intentions, not color, trajectory, or weight. Indeed, the mind-reading system is so functionally-specialized that it can be *selectively* impaired, that is, impaired while other cognitive abilities are intact. This can be clearly seen in certain people with autism.

Crib sheets expand our abilities. Autism graphically illustrates what happens when an evolved crib sheet is missing. A person with autism may have a normal IQ, be better than normal at many cognitive tasks, such as Where's Waldo?, and be able to make sophisticated inferences about machines. Yet this same person cannot make simple inferences about other people's beliefs and desires. If a normal three-year-old sees a character, Charlie, looking at one of four candies, and is asked "Which candy does Charlie want?", she will point to the one Charlie's eyes are trained on. But an autistic person will answer randomly, even though he can tell you exactly which candy Charlie is looking at.* The autistic person can detect eye direction but, unlike you or me, he cannot use it to infer what someone wants. This shows that the mental toolkit that comes with having a normal IQ and normal abilities to reason about the physical world is not sufficient for reasoning about the mental world. Because the mind of an autistic person is missing an evolved crib sheet, he does not know that eye direction can indicate desire. Similarly, having an intact mind-reading system is insufficient for reasoning about the physical world: Children with Williams syndrome are good at inferring other people's mental states, yet they are profoundly retarded and have difficulty learning even very simple spatial tasks.

Domain-specialized inferential tools and knowledge bases are found not just in the learning systems of infants and toddlers, but in those of adults as well. For example, it is now well-established (if not universally assented to) that the learning mechanisms that govern the acquisition of a language are different from those that govern the acquisition of food aversions, and both of these are different from the learning mechanisms that govern the acquisition of snake phobias. Each program has knowledge of its particular domain built into its structure, which allows it to perform its function far more efficiently than any blank-slate system could. The language acquisition device knows, for example, that the names of objects are nouns. The snake phobia system knows what snakes look like, knows what fear looks like on others' faces, and has a procedure specialized for using fear on other's faces to change the intensity of fear you feel in the presence of snakes. The food aversion system knows that nausea is usually caused by foods recently ingested, that it is more likely to be caused by novel foods than by familiar foods, and uses the contingency between food ingestion and nausea to regulate the subsequent attractiveness of food items. How did these systems get these specialized procedures and knowledge? Those mutations that, for example, built in the knowledge of what snakes looked like and what a fear-face looked like, increased the efficiency with which one learns which snakes should be avoided; hence, they were selected for.

So, the mind is not packed with specialized programs merely because they afford small differences in efficiency. Not only do they usually afford enormous increases in efficiency, but more critically, they make many things possible that would otherwise be completely impossible. Different problems often *require* different crib sheets for their solution. Reciprocally, this means that the implicit knowledge present in one specialized program will be distorting or useless if applied outside of its evolved domain. For example, knowledge about beliefs and desires, which is often indispensible for inferring the behaviour of other people, will be misleading if it is applied to rocks and lakes. Knowing that concrete objects are nouns will not allow you to avoid venomous snakes. Two machines are better than one when the crib sheet that helps solve problems in one domain is misleading – or useless – in another. This is why many evolved computational mechanisms are engineered to be domain-specific: To be useful, they must be activated in some domains but remain inactive in others.

The more crib sheets a system has, the more problems it can solve. A brain equipped with a multiplicity of specialized inference engines will be able to generate sophisticated behaviour that is sensitively responsive to many more aspects of its environment. In this view, the flexibility and power often attributed to blank slates and content-independent algorithms is illusory. All else being equal, a content-rich system will be able to infer far more than a content-poor one.

Why content-rich is better than content-poor. This view of the mind is radically at variance with the Standard Social Science Model. Its advocates attributed everything – from hopscotch to romance to rivalry – to the operation of "learning", "reasoning", and "decision-making". Regrettably, these are simply names for mysterious hypothetical processes, not scientific theories of how things actually happen in cause-and-effect terms inside the head. To fill this gap, later scientists proposed that the mind comes factory-equipped with general purpose computational circuits that are jacks-of-all-trades. Prime candidates were so-called "rational" algorithms: programs that implement formal methods for inductive and deductive reasoning, such as the laws of probability, mathematics, or formal logic. Others proposed simpler, associationist systems, that compute correlations or contingencies. These methods were inviting precisely because they are content-free. After all, the mind was assumed to be a blank slate, initially free of all content, so that all of its content could be derived from experience alone.

What do we mean by a content-free program? Consider *modus ponens* and *modus tollens*, two domain-general rules of logic. (They form the background for some experiments on domain-specific reasoning that we will describe at book's end). Whenever "If *P* then *Q*" is true and "*P*" is true, *modus ponens* allows you to validly conclude that "*Q*" is also true. *Modus tollens* licenses a different inference: when "If *P* then *Q*" is true, but "Q" is false, it allows you to conclude that "*P*" is also false. These rules are *content-independent*: they allow you (or an automaton, such as a computer or a neural circuit) to deduce true conclusions from true premises, no matter what is substituted in for "*P*" and "*Q*". Let's say that *P* = "you snooze" and *Q* = "you lose". If it is true that "*If you snooze, you lose*" then you can conclude that anyone who snoozed lost (*modus ponens*), and anyone who won didn't snooze (*modus tollens*). They will produce new knowledge whenever a true premise is combined with a true if-then statement – anything from *If it rains, the ground gets wet* to *If you can keep your head while all those around you are losing theirs, then you'll be a man, my son*. Bayes's rule, a widely used equation for computing the probability that a hypothesis is true given data about that hypothesis, is also content-independent. It can be applied indiscriminately to medical diagnosis, deciding whether Paul McCartney was dead before *Abbey Road* was recorded, playing Baccarat against James Bond, or any other subject matter.

In addition to the fact that they lack any content, and can therefore be applied across domains, rational methods like these were inviting models for the mind for another reason. They are, after all, methods that scientists believe they should use to learn about the world. If they are good enough for science, aren't they good enough for everyday learning?*

No. First, note that although they are content-free, they are not actually general purpose. Their purpose is to add to knowledge about the way the world is; they contain no values or choice criteria that could guide behavior (equipped only with these, our science project woman might learn that men exist, but wouldn't have the desire to choose one for a mate – or, indeed, to mate at all). Brains evolved to produce behaviour, and merely knowing what is true will not tell you what to do. We're not arguing that it is useless to know what is true: Indeed, one reason content-specialized learning mechanisms are superior is that they are better at truth-discovery, and indeed already embody knowledge that is true – or at least useful. And natural selection may in fact have built a few rational algorithms, such as modens ponens, into our brains to supplement our many crib sheeted ones (we ourselves have done experiments showing that people are much better at Bayesian inference than many psychologists thought).

Nevertheless, these rational algorithms are not *sufficient* to solve the adaptive problems our hunter-gatherer ancestors faced. This is because rational algorithms are defined by what they lack: content. Machines limited to executing Bayes's rule, modus ponens and other rational procedures derived from mathematics or logic are computationally weak compared with an evolved system of dedicated, content-specialized inference mechanisms. The theories of rationality embodied by "rational" procedures are environment-free; they were designed to produce valid inferences in *all* domains, in all environments, even in imaginable but non-existent worlds. They can be applied to a wide variety of domains, however, *only because they have been stripped of all information and all procedures that would be helpful in one domain but counter-productive in another*. Having no crib sheets, there is little they can deduce about a domain; having no privileged hypotheses, there is little they can induce before their operation is hijacked by combinatorial explosion. These jacks of all trades are, necessarily, masters of none. They achieve generality only at the price of general ineptitude. Domain-specific algorithms do not need to make the same trade-off: Each can be master of a different domain. The difference between domain-specific methods and domain-independent ones is akin to the difference between experts and novices: experts can solve problems faster and more efficiently than novices because they already know a lot about the problem domain.

The flexibility of human intelligence – that is, our ability to solve many different kinds of problems – was once thought to be conclusive evidence that the circuits that generate it are general-purpose and hence content-free. *Homo sapiens* was thought of as the rational animal, a species whose instincts, rendered unnecessary by culture, were erased by evolution. William James's counter-argument was ignored for much of the 20^{th} century. But his view of a human mind that is flexibly intelligent exactly because it is packed with a wealth of instincts is being vindicated today.

Indeed, instincts should not be thought of as the polar opposite of reasoning and learning. The computational systems we described above – the ones that govern how we reason and learn about faces, objects, language, snakes, mind-reading, nausea and so on – have the following five properties: (1) they are complexly structured for solving a specific type of adaptive problem, (2) they reliably develop in all normal human beings, (3) they develop without any conscious effort and in the absence of any formal instruction, (4) they are applied without any conscious awareness of their underlying logic, and (5) they are distinct from more general abilities to process information or behave intelligently. In other words, they have all the hallmarks of what one usually thinks of as an instinct.* In fact, one can think of these functionally-specialized, content-rich computational systems as reasoning instincts and learning instincts. They make certain kinds of inferences just as easy, effortless, and natural to us as humans, as spinning a web is to a spider or dead-reckoning is to a desert ant.

*Principle 5. Our modern skulls house a stone age mind.**

Good design takes time. Natural selection, the process that designed our brain, operates very slowly. This is because it often takes thousands of generations for reproductive advantage to propel a new mutation, often appearing originally in only one individual, upwards in frequency across the generations until everyone in the species carries it. This process may need to be repeated cyclically many times in order to piece together a complex, new program, lengthening the required time still further. This is because a complex program will require many component parts, and assembling these parts will require increasing the frequencies of many underlying mutant genes in order to construct them. Like a stone being sculpted by wind-blown sand, the time it takes to evolve complex circuits that are tailored to a given environment is slow compared not only to the lifetime of a single human, but also slow compared to the rise and fall of civilizations or even of all of recorded human history.

For this reason, the world that our minds were engineered to operate in is the hunter-gatherer world of our distant ancestors, and not the agricultural or industrial world that we live in now. Our programs were sculpted by the millions of years our ancestors spent as foragers. In contrast, the period humans have lived by agriculture – a transition that first began ten thousand years ago, and is only now being completed – is a very brief time. In fact, it is far too brief an interval to have caused natural selection to build complex new species-typical adaptations suited to these dramatically changed conditions. For this reason, our modern skulls house a stone age mind. Our minds have not been biologically re-engineered by the last few thousand years to solve the novel problems of the modern world.

Evolutionary high technology. In saying that our modern skulls house a stone age mind, we do not mean to imply that our minds are unsophisticated. Quite the contrary: They are full of computer technology of a very high order – something closer to what saucer-flying extraterrestrials might have biologically engineered than what Intel or Cambridge researchers are likely to come up with any time soon. Our brains are many quantum leaps ahead of the most advanced of modern computers, robots, or any other cutting edge technology existing today, and we cannot yet duplicate most of their computational powers or feats. Modern computers can outperform the human mind only on problems that are evolutionarily novel, not problems our minds evolved to solve. We leave supercomputers in the dust on any problem we evolved to solve. Humans may be bad spreadsheets, but equally, Microsoft Excel would be an abysmal failure as a hunter or a lover or a parent. In popular culture, "stone age" and "cave man" are treated as synonyms for *crude* or *brutish,* but in reality the challenges and tasks of this vanished way of life were at least as subtle and as exacting as anything we deal with today. Any party of modern sophisticates would find it quite intellectually challenging to be dropped naked in a wilderness to survive for a lifetime, while delivering and raising their children. We cannot build robots that could do this at all. However, the evolved circuits in our ancestors' minds were elegantly designed by natural selection to solve these challenging problems. We inherited those sophisticated neural programs from our remote hunter-gatherer ancestors without any significant changes, and so our minds are based on the same hardware blueprints. Our mental worlds differ from the worlds of our ancestors only in the incidental specifics of our cultures, not in the minds we experience them through.

The EEA. All devices are designed to operate in certain environments rather than in others, and their designs assume the presence of certain background conditions. The clipper ship was designed to operate in the oceans and the commercial climate of the early 19^{th} century, and the Panzer tank was designed to operate in the battlefields of World War II. The key to understanding how the modern mind works is to realize that its circuits were not designed to solve the day-to-day problems of the modern member of industrial society; they were designed to solve the day-to-day problems of our hunter-gatherer ancestors, and to operate successfully in those environments, under the conditions they faced. Researchers call these sets of conditions the *environment of evolutionary adaptedness,* or EEA.

We need to appreciate just how different the ancestral way of life was from our modern way of life. To understand the functions of the various design features of our minds, first we need to conceptually remove ourselves from modernity, from which our ordinary common sense notions of usefulness and sensible action are derived, and place ourselves instead in this ancient world. Imagine being on a camping trip that lasts a lifetime. Absent from our predecessors' lives were the familiar props of industrial society that we rely on now – stoves, furnaces, refrigerators, hospitals, doctors, police, roads, schools, grocery stores, laws, maps, antibiotics, factories, fabrics, farms, armies and nation-states. Instead, our forebears lived in small, nomadic bands of (typically) a few dozen individuals who got all of their food by gathering plants or by hunting animals; who made all of their own clothes, tools, and shelter; who relied on themselves and each other for the treatment of disease and injury, for provisioning during incapacitation, for protection against the attack of predators, for protection against competing groups and extortionate individuals, and so on. Over millions of years, those mutant circuits that were better designed for solving these problems left more children, and we are descended from them.

The fact that our skulls house a stone age mind explains why it is easier for us to learn to fear snakes or spiders than electric sockets or guns, even though sockets and guns now kill far more people than snakes and spiders do. In many cases, our brains are better at solving the kinds of problems our ancestors faced on the African savannahs than they are at solving the familiar tasks we face in a college classroom or a modern city.

The past explains the present. For these reasons, evolutionary psychology is relentlessly past-oriented. Indeed, evolutionary psychologists reject the notion that one has "explained" a behaviour pattern by showing that it spreads genes under modern conditions. Instead, a necessary (though not sufficient) component of any explanation of behaviour – modern or otherwise – is a description of the design of the computational machinery that generates it. Behaviour in the present is generated by information-processing mechanisms in our brains that were inherited unchanged from the past. They are designed to make the choices that would have been adaptive for our ancestors, in the environments in which the human line evolved. These may, or may not, be adaptive in the industrial world, depending on the details of the differences between the EEA and the industrial world. So, for example, we do not need to puzzle over why modern humans self-destructively indulge in fast food: A taste for salt, fat, and sugar – all limiting nutrients for our often deprived ancestors – was functional for them, even though it is destructive for those modern humans living in abundant nations. For our ancestors, physical beauty tracked fertility, health, and remaining years of effective parenting. Men and women still sexually pursue what their circuits perceive as beauty, even though beauty no longer predicts health and fertility the way it did among our ancestors. Women still appreciate male athleticism, even though the hunting success that male athleticism once predicted is now economically valueless. Our minds still traffick in ingroup favoritism and outgroup hostility, even though we live in a world where the circuits that cause these thoughts and feelings are self-destructive and reproductively pointless. We have evolved adaptations that make sex pleasurable to us now, regardless of whether the individuals involved are using contraception, because sex once reliably led to conception, in a world in which there was no contraception. We have created an industrial world in which we have reshaped modern activities, sometimes poorly, to fit into the constraints of hunter-gatherer mental programs. We have not changed our neural architecture to fit the modern world.

These points separate evolutionary psychology from those approaches that view humans as fitness-maximizers, or posit that they do what they do in order to spread their genes. Humans do what they do because their evolved neural programs execute their evolved program structures, generating behaviour. These adaptations acquired their program designs because, under ancestral conditions, they had the *net effect* of spreading genes, not the *goal* of spreading genes. They simply follow out the dictates of their various designs, which in the modern world may lead to outcomes very distant from gene-propagation. So modern humans may feel motivated to exact revenge, even though this may be disasterous in a modern world packed with lawyers, lawsuits, police, and courts, simply because on average among our self-reliant ancestors it was an adaptive thing to do. Some of the multiplicity of evolved neural programs have their own distinct goals built into them (e.g., sex, child protection, revenge, way-finding), while some operate without represented goals (e.g., face recognition, language acquisition, coughing, vision). However, there is no motivational program in the human mind with the imperative "to spread genes" and evolutionary psychologists have no expectation that humans will necessarily be maximizing fitness under modern conditions, although sometimes they may. *Humans are adaptation executers, not fitness-maximizers*.

Can we know the past? Although the hominid line is thought to have originated on the African savannahs, the environment of evolutionary adaptedness, or EEA, is not a particular place or time. After all, different groups of ancestors faced different conditions at various times and places. Instead, it is the statistical composite of all of selection pressures or cause-and-effect relationships that pushed the genes underlying our adaptations upwards in frequency until they became species-typical (which most adaptations are). The EEA is the net sum of all the forces of all of these past conditions on the designs of our adaptations. One can think of it as a set of enduring conditions, a set of selective events, or a set of adaptive problems. Because adaptations evolved and assumed their modern form at different times, the EEA for one adaptation may be somewhat different from that for another. Conditions of terrestrial illumination, which form (part of) the EEA for the vertebrate eye, remained relatively constant for hundreds of millions of years. In contrast, the EEA that selected for neural programs that cause human males, under certain conditions, to provision and care for their offspring may be less than two million years old.

It is often argued we can know nothing about the past that is relevant to psychology because behaviour doesn't fossilize, and so the whole field rests on uncertain speculation or conjecture. But through the work of researchers in many disciplines, we know with certainty thousands of important things about our ancestors, many of which can be useful in guiding psychological (or medical) research. Our ancestors nursed, had two sexes, hunted, gathered, chose mates, used tools, had color vision, bled when wounded, were predated upon, were subject to viral infections, were incapacitated from injuries, had deleterious recessives and so were subject to inbreeding depression if they mated with siblings, fought with each other, cooperated with each other, lived in a biotic environment with predatory cats, venomous snakes, and plant toxins. They were omnivorous, ground-living primates, and mammals with helpless infants, long periods of biparental investment in offspring, and an extended period of physiologically obligatory female investment in pregnancy and lactation. It is a certainty that our ancestors lived in a world in which certain principles of physics governed the motions of objects: facts that allowed Roger Shepard* to develop his theories about the evolutionary foundations of motion perception and cognition. It is equally certain that hominids had eyes, looked at what interested them, and absorbed information about what they were looking at, making eye-gaze direction informative to on-lookers: facts that led Simon Baron-Cohen* to create a subtle and far-reaching research program on mindreading, or inferring the mental states of others. Although we are certain about a huge inventory of facts about the ancestral world that have not yet been harnessed to guide psychological research, certainty about the past is not necessary to build better hypotheses. One can derive valuable experimental hypotheses from possible rather than certain features of the ancestral world: The worst that will happen is that the hypothesis will be experimentally falsified – something that routinely happens to the traditional researchers who have no principled source from which to derive their hypotheses. There are, of course, many features of the ancestral world about which we are completely ignorant: These simply do not form the basis for experiments.

The five principles above are tools for thinking about psychology and human behaviour that can be applied to any topic: sex and sexuality, how and why people cooperate, whether people are rational, how babies see the world, intergroup relations, when and why people conform, aggression, hearing, vision, sleeping, eating, depression, schizophrenia, deception, and on and on. The framework that they provide links areas of study that formerly seemed unconnected, and saves researchers from drowning in particularity. Whenever you try to understand some aspect of human behaviour, they encourage you to ask the following fundamental questions: 1. Where in the brain are the relevant circuits and how, physically, do they work? 2. What kind of information is being processed by these circuits? 3. What is the specific series of information processing steps that these circuits carry out? and 4. What were these circuits designed to accomplish (when they functioned in a hunter-gatherer context)?

Using these tools to reverse engineer the human mind, evolutionary psychologists hope they can turn the mystery of human nature into a set of blueprints, exploring the human psychological architecture circuit by circuit, specialization by specialization, function by function.

3. HOW TO DO EVOLUTIONARY PSYCHOLOGY

Most psychological research is based on theories that are stand-alone: they are not tied to any other psychological theory, or deduced from a set of fundamental principles, or linked to any theories outside of psychology, or derived from functional considerations (at least not when *functional* is used in the evolutionary sense). Indeed, if by *hypothesis* one means a proposition derived from a theory, then an uncomfortable amount of research in psychology is atheoretical: the testing of hunches and folk beliefs rather than hypotheses.

Research in evolutionary psychology is different. Evolutionary psychologists use theories of adaptive function to guide their investigations of the structure of the human mind and body. For example, they can go from theories of how our ancestors cooperated and competed in groups to theories about the design of evolved mental programs specialized for group interaction; or from an understanding of the ups and downs of daily foraging success to theories about the structure of the programs that guide the impulse to share with others who are deprived. These models of mental programs can then be tested and refined experimentally. This research strategy is enormously helpful, but to use it correctly you need to understand why it works. You need to understand how the evolutionary process designs organisms.

Common descent and adaptation

The goal of Darwin's theory was to explain the designs of organisms: Darwin asked, for example, Why do the beaks of finches differ from one species to the next, and have the forms that they do? Why do animals expend energy attracting mates that could be spent on survival? Why are human facial expressions of emotion similar to those found in other primates?

One of the most important evolutionary principles accounting for the characteristics of organisms is common descent. An increasing body of evidence indicates that all organisms alive today are the descendents of a single originating organism. Over the course of evolutionary time, new species originate because one breeding population sometimes becomes subdivided into two or more populations, and stops interbreeding. Although they start out with the same set of genes, they subsequently can evolve independently because the different populations no longer exchange genes through matings. For example, Darwin recognized that the different finch species he encountered in the Galapagos Islands were all descended from the same ancestral mainland species, and shared many characteristics in common as a result. This process of species splitting gives a hierarchical tree structure of similarity to all species on Earth. Offspring inherit their parents' genes and design features, which stay the same across the generations unless selection or chance modifies genes. So, the more recently two species were descended from the same ancestral species, the more design features they will share in common. Because we are related to all other species by virtue of common descent, we expect to find many similarities between humans and our closest primate relatives. Paul Ekman has found, for example, that humans and chimpanzees both use the same muscles in making parallel facial expressions.

This is the *phylogenetic* approach, and it consists of the search for features (called *homologous features*) that are similar because both species inherited them from the species that was their common ancestor. This approach has a long and productive history in psychology. But as valuable as it is for many questions, this approach cannot adequately address features that evolved uniquely in only one lineage, because there are then no similarities to compare. Because there is the widespread misimpression that evolutionary psychology consists solely or primarily of applying a phylogenetic approach, many take it as a given that evolutionary psychology cannot address the large set of properties that make us uniquely human. They think it is limited to the study of characteristics that we share with other animal species.

However, the second principle that accounts for the characteristics of organisms is that natural selection builds adaptations into their designs. Indeed, natural selection can cause the designs of different species to diverge from one another, sometimes producing characteristics that are unique to a given species, such as the elephant's trunk or the cognitive mechanisms that allow humans to learn language. *Adaptationism* is the name for the research program that gives a central role to exploring how natural selection functionally organizes the designs of organisms. It can be applied to analyze features that are unique to humans, because the theory of natural selection illuminates equally well features that are shared in common with many species, and features that are limited to one species.

George Williams' 1966 book, *Adaptation and Natural Selection*, clarified the logic of adaptationism. In so doing, this work laid the foundations of modern evolutionary psychology. Although evolutionary psychologists certainly appreciate and invoke phylogenetic explanations where they are appropriate (as well as other relevant theories and analytic tools), it is the application of adaptationist logic that has provided the brightest illumination to formerly murky issues in human psychology, physiology, and the social sciences.

Engineering and reverse engineering. Those who study species from an adaptationist perspective adopt the stance of an engineer. In discussing sonar in bats, for example, Richard Dawkins proceeds as follows: "...I shall begin by posing a problem that the living machine faces, then I shall consider possible solutions to the problem that a sensible engineer might consider; I shall finally come to the solution that nature has actually adopted"*. At first glance, this may seem strange: Why would a biologist, a student of the warm, wet, and living, find it useful to think like an engineer?

Engineers start with a problem, and then design machines that are capable of solving that problem in an efficient manner. As a result, the machine's structure reflects its function: it has certain properties and components rather than others *because* those structures solve a problem better than alternative ones.

Engineers can also work in reverse: Given a mysterious machine, they can figure out what its design features are – i.e., which of its components are functional and how their arrangement accomplishes the machine's function. Doing this is relatively simple if you know what problem the mysterious machine was designed to solve, because you can then look for structures capable of accomplishing that function. But – as any engineer will tell you – reverse engineering is crushingly difficult when you have no idea what the machine was designed to do. Without a theory of *function*, how do you determine which parts are function*al*?

Psychologists are like engineers working in reverse: the brain is a mysterious machine, and we are supposed to figure out how it works, i.e., which of its components are functional and how their arrangement accomplishes various functions. Doing this is difficult, however, without knowing what problems this mysterious organic machine was designed to solve.

That is why evolutionary biology is so helpful. Evolutionary biologists figure out what adaptive problems a given species encountered during its evolutionary history, and then ask themselves, "What would a machine capable of solving these problems well under ancestral conditions look like?" Against this background, they empirically explore the design features of the evolved machines that, taken together, comprise an organism.

Despite all the obvious differences between living beings and human-made machines, this engineering strategy works because the two resemble each other in one crucial respect. Like human-made machines, organisms are comprised of structures that reflect their function. This is an inexorable result of how natural selection works, and fundamental to the logic of evolutionary psychology. Let's see why.

Why does structure reflect function in living organisms?

Organisms, like watches or automobile engines, exhibit a multitude of parts and subassemblies that are arranged in precise and highly ordered ways so that they operate to achieve the functional ends they were designed to perform. The eyes, immune system, umbilical cord, cell nucleus, and lungs, to pick a handful of examples, all display a very advanced technology, built out of organic molecules. The more that chance events, such as accidents or violence, act to randomize this internal order, the more the watch, automobile, or organism is damaged. In a world where everything is bombarded by chance forces, where did all this functional order in animals and plants come from?

The evolutionary process has only two components, chance and natural selection, that govern how the genes in a species change over time. Chance processes act to randomize relationships within the organism, and so cannot account for the accumulation of the highly ordered arrangements of functional parts that permeate organisms. For this reason, modern researchers now understand that natural selection is the only component of the evolutionary process that can build complex functional organization into a species' structure.* This means that all complex functional design in organisms was created by natural selection. Consequently, we know that all functional organization in humans must be built in a way that is consistent with the principles of natural selection. This recognition is what makes evolutionary biology the foundation of psychology and neuroscience, not to mention anatomy, physiology, the medical sciences, and the social sciences. Our functional order originally comes from evolution.

To be sure, there is much that is not functional in organisms as well, introduced by chance evolutionary and non-evolutionary processes. But the functional architecture of organisms is central to their organization, and they would not exist without it. Indeed, in evolved systems there is a sense in which *function determines structure*, and that is the key that unlocks our understanding of the human cognitive architecture.

To see this, consider our prior discussion of how natural selection works. Any organism can be described as a collection of design features – micro-machines, such as the functional components of the eye, the limbic system, the milk ducts, or the musculature. Individuals die, but their design features live on in their descendants – if they have any. As mutations arise, alternative design features are either added to or discarded from the species' design because of their *relative* consequences on their own reproduction. More particularly, a design feature can cause its own spread over generations if it has the consequence of *improving* the ability of those who have it to solve an adaptive problem: e.g., if it improves the quality of the milk a mother produces for her infant, or increases a child's resistance to a disease, or makes a man more alert to cues that a rival is sexually interested in his mate.

To accomplish each of these functions, the relevant design features must mesh with the matrix of cause-and-effect relationships that actually exist in the world. Consider milk production, the adaptation that defines us as mammals. Only certain design modifications can increase the fat content of milk; others will decrease it, or increase the protein content by too much (cow's milk, for example, has too much protein for a human neonate to digest), or make the milk too watery. A different set of modifications are necessary to make the system more energy efficient by coordinating milk production with the infant's needs: to create sensory receptors that respond to a feature of the world – an infant's sucking rate – plus a hormonal feedback system capable of adjusting the quantity of milk produced in response to the infant's sucking. Another feature of the world is that human infants cry when hungry. To take advantage of this cause-and-effect relationship in the world, still other modifications – including circuits connecting the mother's auditory cortex to her mammary system – are needed to improve the timing of milk delivery: to orchestrate the let-down reflex so that the milk flows when a mother hears her infant cry. Random processes continually churn out design modifications of all kinds: some of these will improve the milk production-and-delivery system, some will degrade it, some will have no effect on it. But the feedback process of natural selection chooses among these alternative designs *on the basis of how well they function*. As a result, the ones that improve milk production and delivery persist in the population, while the ones that degrade it disappear from our design. In this way, function determines structure.

Moreover, natural selection is not a "good-enough" principle: It contains no mechanism for stopping once a certain threshold level of functionality has been reached. It is a hill-climbing process, in which a design feature that solves an adaptive problem well can be outcompeted by a new design feature that solves it better. So every time mutations produce a design feature that improves performance – that accomplishes an adaptive function, such as nourishing a dependent infant, better – that design feature is in contention to replace the previous, alternative model. In this way, natural selection tinkers with a species' components over immense expanses of time, choosing among enormous numbers of alternative designs, and cumulatively saving most improvements that appear.

As a result of this vast Research & Development enterprise, natural selection constructs devices that are very well-engineered for accomplishing particular functions. Evolved systems are not optimal or perfect, but they are very good by human engineering standards. We can say this with confidence because human engineers have not been able to match the quality of what evolution produces. Natural selection has produced exquisitely engineered biological machines – the vertebrate eye, photosynthetic pigments, efficient foraging algorithms, colour constancy systems – whose performance is unrivaled by any machine yet designed by humans. Indeed, scientists continually underestimate our adaptations. The human lactation system is a case in point. The more scientists study lactation, the more remarkable they find it is. Lactation not only produces dramatic changes in the physiology of a woman's breast, but in how that physiology responds to information deriving from the infant. The system is designed to deliver antibodies to the infant; to shut off ovulation so that energetic resources will not be diverted to a new pregnancy until the existing infant is robust enough to survive with less milk; to link nursing patterns to circadian rhythms, such that the breasts fill with milk at the times of day that the infant habitually nurses; even the composition of the milk itself changes over time, in ways that match the changing nutritional needs of the developing infant.

A causal process does not need the human properties of foresight and intention to be capable of designing something. The selection of parts on the basis of their functional consequences is the crux of the concept of design (e.g., we say a thermocouple has been designed because the two different metals, each with different heat conducting properties, did not come together by chance; they were selected for the thermocouple *because* this has functional consequences if the goal is to regulate something's temperature). From this perspective, it does not matter whether the causal system that does the selection is a volitional agent or a feedback process. A system can be said to be designed whenever the *cause* of its having the parts and properties that it has – rather than others – is that they have functional consequences, i.e., that they solve a problem of some kind. By this criterion, natural selection designs organisms – no scare quotes needed around "design".

Chance events, such as mutations, cause alternative parts (design features) to be introduced into a population of organisms, but natural selection is not a chance process. Natural selection is a systematic feedback process that retains or discards parts because of their consequences on the functional performance of the system. By selecting designs on the basis of how well they solve adaptive problems, natural selection creates structures that reflect their function.

Using function to discover the structure of mental programs. Because naturally selected outcomes are not crude but very well-engineered, there is a very tight fit between function and structure, making selection a very informative guide to evolved design.

Examining the properties of a lock can tell a locksmith a great deal about the structure of the key that unlocks it. Indeed, it allows the locksmith to deduce enough about the design of a lost key to make a duplicate. In exactly the same way, by identifying the function – the adaptive problem – to be solved, researchers can determine a great deal about what the structure must be like that solves it. This can guide research much more effectively than the present practice of studying the mind without knowing the functions it evolved to serve. The study of hunter-gatherers, primate behaviour and ecology, human origins, and the theory of natural selection are indispensible to the future of psychology because they provide a window onto the adaptive problems that shaped the human mind.

Because the brain evolved to solve adaptive problems of a computational nature, its structure can be described in two different but complementary ways: as a set of neural elements or as a set of programs. Because function determines structure at all levels of organization, at some point a field called *evolutionary neuroscience* will surely exist. But, much as a lock places more constraints on the shape of a key than on the metal from which it is made, theories about adaptive computational problems place more constraints on the structure of the programs that solve them than on the structure of the neural tissue that implements them. This becomes clear when we consider what it means for a problem to be computational in nature.

Organisms are composed of many parts, but only some of these parts are computational. By *computational* we mean that they are designed to (1) monitor the environment for specific changes, and (2) regulate the operation of other parts of the system functionally on the basis of the changes detected. For example, the diaphragm muscle, which causes the lungs to contract and expand, is not computational. But the system that measures carbon dioxide in the blood and regulates the contraction and extension of the diaphragm muscle is. The plastic cover on a thermostat is not computational, nor are the parts of a furnace that generate heat. But the thermocouple that responds to ambient temperature by toggling the switch on the furnace, and the connections between them form a computational system. Muscles are not computational, but the visual system that detects the presence of a hungry-looking lion, the inference mechanisms that judge whether that lion has seen you or not, and the circuits that cause your muscles to either run to a nearby tree (if the lion has seen you) or freeze (if it hasn't seen you) do compose a computational system. The language of information-processing can be used to express the same distinction: One can identify the computational components of a system by isolating those aspects that were designed to regulate the operation of other parts of the system on the basis of *information* from the internal and external environment.

By "monitoring the environment for specific changes", we mean the system is designed to detect a change in the world. That change can be internal to the organism (such as fluctuations in carbon dioxide levels in the blood or the activation of a memory trace) or external to the organism (such as the onset of a rainstorm or the arrival of a potential mate). Changes in the world become *information* when (1) they interact with a physical device that is designed to change its state in response to variations in the world (i.e., a transducer), and (2) the changes that are registered then participate in a causal chain that was designed to regulate the operation of other parts of the system. A photon, for example, does not become information until it causes a chemical reaction in a retinal cell, which was designed for this purpose and is part of a causal system that was itself designed to regulate an organism's behavior on the basis of inferences about what objects exist in the world and where they are.

A set of features is not computational unless they were *designed* to exhibit these properties. For example, the outer cells of a dead tree stump expand in the rain, and as this happens, the inner portions of the stump might become compressed. But these dead cells were not designed for detecting changes in weather. More importantly, although their swelling does cause a change in the inner part of the stump, it is not *regulating* the operation of the stump. Regulation means more than merely influencing or changing something. It means systematically modifying the operation of a system so that a *functional* outcome is achieved. In the case of a thermostat, that function was determined by the intentions of the engineer who designed it. In the case of an organism, that function was determined by natural selection, which acted to organize the properties of the organism.

Neurons do not perform any significant metabolic function for an organism. They exist because of the computational relationships they create. Natural selection retains neural mechanisms on the basis of their ability to create functionally organized relationships between information and behaviour (e.g., the sight of a predator activates inference procedures that cause the organism to hide or flee) or between information and physiology (e.g., the sight of a predator increases the organism's heart rate in preparation for flight). Each neural program was selected for because it created the correct information-behaviour or information-physiology relationship and, so long as a physical implementation produces this relationship, its particular form is free to vary according to other factors. (Indeed, when people recover function after brain damage, repair processes restore the original information-processing relationship – but using a different set of physical connections.)

In other words, the brain was designed by natural selection to be an information processing device. The brain has the physical structure that it does *because* this structure embodies a particular set of programs, and each program has the computational structure that it does *because* that structure solved a particular problem in the past. This is the causal chain that licenses inferences from function to program structure to physical structure. If you know what problems our ancestors faced, then you can make educated guesses about what programs evolved to solve them, including what computational procedures they would have required. Once you have experimentally confirmed whether these programs exist, you can search for their neural basis. Having a theory of adaptive function is useful to psychologists and neuroscientists because it allows us to see beyond our instinct blindness, to look for programs we would not otherwise look for: If, as Einstein said, what we can observe depends on the theory we are using, evolutionary biology provides the theory that allows us to observe the functional architecture of the human mind.

Example 1: Spatial cogniton. For many years, spatial cognition was considered to be a unitary phenomenon, which could be measured by tests assessing one's ability to imagine rotations of objects in space. As a group, men score higher on these tests than women, and this average difference in performance was taken as evidence for a *general* male advantage in spatial cognition. Some speculated, post-hoc, that this difference may reflect spatial tasks that arise in hunting (e.g., aiming projectiles at moving targets; dead-reckoning to navigate home through unfamiliar territory).

In contrast, Irwin Silverman and Marion Eals reasoned from function to structure. They asked what kind of spatial cognition an ancestral woman would have needed to solve the problems of gathering. Unlike animals, plants do not wander around. On the other hand, a small vine may develop an edible root in a couple of weeks, and the fig tree you saw today might bear fruit three months from now. To be an efficient forager, a woman would need to note and remember the spatial location of wild plants within a complex array so she can return to each plant later, when it is ready to be harvested. So Silverman and Eals developed tests of this form of spatial cognition, and found that, as a group, women score higher on them than men do. Moreover, this difference is most pronounced when learning is incidental (that is, when no one tells you to remember the locations of things.) Their research is interesting not only because of the sex difference favoring females, but also because it indicates that spatial cognition is better thought of as a collection of different computational processes, each designed for solving functionally different spatial tasks.

Without a theory of what the mind was designed to do, you have to rely on your intuitions for hypotheses. But if most scientists are men, those intuitions will be male intuitions. 100 years of intuition-based ressearch on spatial cognition never turned up a female advantage. But the *first time* someone asked what kind of spatial cognition a gatherer would need, they found a female advantage. A functional theory corrects for instinct blindness.

Example 2. Parental care. Although parental care is one of the defining properties of mammals, you will be hard pressed to find an introductory textbook in psychology that devotes a chapter – or even an index reference – to the psychology of parenting in humans. That parents care for children is one of those facts of human life that is so obvious that it does not seem to be in need of explanation, and has therefore not received the study it deserves.

There is, however, a huge literature in evolutionary biology about the conditions under which selection will favor design features in adults that motivate them to care for children. Based on this theory, Martin Daly and Margo Wilson were able to identify a major risk factor for child abuse.

Into every life, some stress does fall. Things go badly at work, you haven't had enough sleep, the baby won't stop crying, and your older children are fighting…everyone who has been a parent has had days like this. Sometimes parents do yell at or even strike their children on days like this, but usually they do not. As anyone with children will tell you, what carries parents through such moments is the overwhelming love they feel when they look at these small people, who are – to a parent – the most enchantingly lovely beings to have ever graced the earth. Put in scientific terms, this overwhelming parental love is a computational system designed by natural selection that motivates parents to keep protecting and provisioning their children, even under adversity, and that buffers their reactions when the child is – unwittingly or not – adding to their stress.

Daly and Wilson have argued that this system will not be activated equally for every child who is under one's care: natural children will benefit more from its buffering effects than step-children.

In humans (unlike most mammals) both parents cooperate to invest in children. During the Pleistocene – as today – parents sometimes acquired a new mate after having lost a spouse due to death or desertion, so there would have been step-families among ancestral hunter-gatherers, just as there are now. This means many adults would have been "at risk" of investing in children who were not their own. We say "at risk" because a clear prediction of parental investment theory is that programs that motivated an individual to invest in children indiscriminately would quickly be selected out, in favor of ones that motivated an individual to invest selectively in his or her own children. The programs that cause the constellation of behaviors that we call "parental love" will be more strongly activated in response to own children than step-children. Not only will this create a difference in how parents react to stress with natural children versus step-children, but this inclination to invest more in one's natural children can itself create conflict within a step-family: a step-mother would be motivated to invest more in her own children than her step-children; the father would be motivated to have her invest more equally; the step-children would feel that their step-mother was treating them unfairly; and so on. This conflict will sometimes erupt into violence.

For all these reasons, Daly and Wilson predicted that, all else equal, there would be more child abuse in families with a step-parent than in those with two natural parents. This is not because step-parents will be motivated to specifically harm or kill step-children (as is found in some species, such as langurs), but as a byproduct of the parental investment mechanisms not kicking in for step-children. And indeed, although levels of child abuse are low for both kinds of families, Daly and Wilson found that the probability that a child will be beaten or killed is 10 to 100 times higher if that child lives in a step-family than with two natural parents.

Before Daly and Wilson's research, those who work in shelters and programs for the prevention of child abuse were unaware that step-families were so much more dramatically at risk. Everyday observation is not sufficient to notice the relationship. Consider: if the probability of abuse in a step-family is 10 times higher, but only 1 out of 10 families is a step-family, then you would see just as many abused children at a shelter from families with two natural parents as from step-families. Once again, a functional theory led to the discovery of a previously unknown phenomenon: as Einstein said, "It is the theory which decides what we can observe."

Seeing organization

[Focusing in on the important and relevant is a difficult achievement in science. Humans engage in an endless variety of actions, and have an endless number of properties. How should they be studied? An evolutionary and functionalist perspective allows researchers to carve nature at the joints, without becoming lost in a sea of particularity and irrelevance.] While the poetic imagination is free to describe organisms in any way, only narrowly defined aspects of organisms fit together to form functional systems: most ways of dividing an organism into parts and relating them to one another will not capture its functional properties.

Within an evolutionary functional framework, all of an organism's properties can be partitioned into one of three kinds: (1) Adaptations, which are present because their functional consequences caused them to be selected for; (2) byproducts, which are present in the architecture because they are causally coupled to traits that were selected for; and (3) evolutionary noise, which are features that were injected into the architecture by the chance components of evolution, such as mutation and drift. So, for example, the neural programs that allow humans to acquire and use language are adaptations; the fact that humans can learn to read and write are byproducts of the language system; and the gene variants that in some individuals cause dyslexia – difficulties with learning to read – are random evolutionary noise.

Unfortunately, some have misrepresented the well-supported claim that selection creates functional organization as the obviously false claim that all traits of organisms are functional -- something no sensible evolutionary biologist would ever maintain. Similarly, it is important to realize that not every beneficial property is an adaptation. Suppose, for example, that a novelist were to become wealthy through writing, and used that wealth to sire many children. This would not make writing – let alone novel-writing – an adaptation. Adaptations are present because of a prior history of selection; they are not defined as any ability or trait – however rare or however modern – that happens to allow a particular individual to have more children.

Design evidence

One can distinguish between adaptations, byproducts, and noise, and test hypotheses about them, by using *design evidence*. Of these three kinds of properties, adaptations are the most important and illuminating, because they explain why the systems have certain parts, why these participate in certain cause-and-effect relationships with one another, and why they interact with the world in the way that they do. Adaptations are problem-solving machines, and can be identified using the same standards of evidence that one would use to test hypotheses about the function of any human-made machine. For example, one can identify a machine as a television rather than a stove by finding evidence that it is extremely well-designed to serve the functions that televisions serve, but poorly designed for cooking food. To do this, one would demonstrate that it has many coordinated design features (an antenna, a cathode ray tube, circuits for isolating channels, and so on) that are complexly specialized for transducing television signals and transforming them into a colour bit map displayed on an external (and, therefore, visible) screen. Such an unusual assemblage of elements is highly unlikely to have arisen by chance alone. It is only likely to come into existence if the system's suitability for fulfilling this function played a role in its construction – something that only happens with natural selection, or intentional human design. Throwing parts together randomly (in engineering, or in evolution) is unlikely to produce anything that is intricately organized to serve a function that is otherwise difficult to achieve. For example, a television has virtually no design features that would make it good for frying eggs, taking baths, cutting paper, brushing teeth, thickening gravy, or paving roads.

One can identify an aspect of an organism' physical or psychological structure -- its phenotype -- as an adaptation by showing that (1) it has many design features that are improbably well-suited to solving an ancestral adaptive problem, (2) these phenotypic properties are unlikely to have arisen by chance alone, and (3) they are not better explained as the byproduct of mechanisms designed to solve some alternative adaptive problem, or some more inclusive class of adaptive problem. Finding that an architectural element solves an adaptive problem with reliability, precision, efficiency, and economy* is *prima facie* evidence that one has located an adaptation. This is like showing that an oddly shaped piece of metal opens the lock on your front door. To show something is a byproduct, one must first establish that something else is an adaptation (e.g, blood as an oxygen transport system), and then show how the feature is a side-effect of the adaptation (e.g., the redness of blood is a side-effect of the oxygen-carrying iron in hemoglobin). Features that are uncoordinated with functional demands are evolutionary noise (e.g., the locations of freckles).

Design evidence is important not only for explaining why known properties of organisms have the form that they do (i.e., why is the lens of the eye transparent, rather than opaque). It also is a tool for discovering mechanisms that no one had previously thought to look for. Thus, evolutionary psychologists use theories of adaptive function heuristically, to guide the mapping of the unknown precincts of the mind. In our own research with our students and colleagues, we have used models of adaptive problems to design studies to detect new mental adaptations. Our studies have produced evidence for cognitive adaptations for reciprocity, bluff detection, risk-sensitive decision-making, statistical reasoning, coalition identification and participation, alliance detection, hazard management, female evaluation of male parenting circuits, modeling predator and prey mental states, understanding death, and incest avoidance, among others.

Definitions of adaptive problems do not, of course, uniquely specify the design of the mechanisms that solve them. Because there are often multiple ways of achieving any solution, empirical studies are needed to decide which solution "nature has actually adopted." But the more precisely one can define an adaptive information-processing problem, the more clearly one can see what a mechanism capable of producing that solution would have to look like. Later, we will illustrate this practice by showing how we have applied this strategy to the study of social reasoning.

Nature and nurture: An adaptationist perspective

To fully understand the concept of design evidence, we need to consider how modern adaptationists think about nature and nurture. Debates about the relative contribution (as it is misleadingly put) of genes and environment during development have been among the most contentious in psychology and in the discussion of human nature. The premises that underlie these debates are flawed, yet they are so deeply entrenched that many people – scientists and nonscientists alike -- have difficulty seeing that there are better ways to think about these issues.

Rather than there being one nature-nurture issue, there are many independent issues but, unfortunately, they have become so tangled together that most discussions in psychology and the social sciences are hopelessly confused. Over the next several sections, we will try to pull the major questions apart, and look at them one by one. Some of them are conceptual confusions, whereas others are genuine scientific questions whose resolution will depend on research, rather than on clear thinking alone.

Despite widespread belief to the contrary, evolutionary psychology is not another swing of the nature/nurture pendulum. It shatters the traditional framework and the old categories entirely, rather than siding with any position within the old debate. Indeed, a defining characteristic of the field is the explicit rejection of the usual nature/nurture dichotomies -- instinct versus reasoning, innate versus learned, biological versus cultural, nativist vs. environmentalist, socially determined vs. genetically determined, and so on -- because they do not correspond to the actual distinctions that need to be made in the real world.

"Innate" is not the opposite of "learned"

To begin with, Everyone is a nativist, whether they know it or not. Even the most extreme advocates of the role of the environment in shaping human behaviour, from Foucault to

Skinner, are making – explicitly or implicitly – nativist claims about the "innate" structure of the evolved neural machinery that learns or responds to the environment. The only difference is whether they make the nature of their claims about this machinery explicit, or allow them to remain implicit, forcing the reader to deduce them from their arguments about why people act as they do.

Imagine, once again, that you are a superhuman engineer, and this time your science project is to create a brain that can learn. To be able to learn, this brain would have to have a certain kind of structure – after all, 3 lb. bowls of oatmeal don't learn, but 3 lb. brains do. To get your brain to learn, you would have to arrange the neurons in particular ways: you would have to create circuits that cause learning to occur. In short, you would have to equip your brain with programs that *cause* it to learn. The same is true when natural selection is the engineer.

Even if a program that causes a particular kind of learning was itself learned, there had to be a prior program that caused that learning to occur, and so on. Logic forces us to conclude that there had to be, at some point in the causal chain, a program that caused learning, but that was itself unlearned. These unlearned programs got into the brain by virtue of being part of its evolved architecture: they are programs that reliably develop across the ancestrally normal range of human environments.

This is a point on which both environmentalists and nativists must agree: Pavlov, Skinner, and Chomsky alike. They may disagree strongly about the computational structure of the evolved programs that cause learning, but not about whether evolved learning programs exist. Consider classical and operant conditioning. These are widely viewed as the simplest and most general forms of learning in humans and other animals. Yet even operant conditioning presumes the existence of evolved mechanisms that change the probability of a behavior by a certain amount, as a function of its consequences (and according to very precise equations). It also presumes that a handful of consequences – food, water, pain – are innately reinforcing (that is, that these consequences are capable of changing the probability of a subsequent behavior is a design feature of the brain). Classical conditioning presumes the existence of a great deal of innate equipment: in addition to the programs that compute contingencies, the animal is filled with unconditioned – that is, *unlearned* – responses, such as salivating in response to meat. Salivating in response to meat is considered to be part of the dog's evolved architecture, and what the evolved learning program does is calculate when an arbitrary stimulus – such as a bell – predicts the appearance of the meat.* So even in classical conditioning, the learned behavior – salivating to the sound of the bell – is caused by an innate, evolved learning program, which takes as input both innate stimulus-response pairs (meat and salivation) and information from the external environment (the contingency between the sound of the bell and the appearance of meat). The only substantive disagreement between a Skinner and a Chomsky is about the structure of the evolved programs that cause learning.

Consequently, any learned behaviour is the joint product of innate equipment interacting with environmental factors, and so cannot be solely attributed to the action of the environment on the organism (see Principle 2). Thus "innate" cannot be the opposite of "learned". It is also mistaken is to think of "evolved" as the opposite of "learned." This is because our evolved learning programs were organized by evolution to learn some things and not others.

To say behaviour is learned in no way undermines the claim that the behaviour was organized by evolution, because the behaviour was learned through the agency of evolved mechanisms that determine what to extract from the input, and how to behave as a response. If natural selection had built a different set of learning mechanisms into an organism, that organism would learn a different set of behaviours in response to the very same environment (remember the dung fly). It is these evolved mechanisms that organize the relationship between the environmental input and behavioural output, and thereby pattern the behaviour. For this reason, *learning is not an alternative explanation to the claim that natural selection shaped the behaviour*, although many researchers unreflectively assume that it is. The same goes for culture. Given that cultural ideas are absorbed via learning – which is caused by evolved learning programs – a behaviour can be, at one and the same time, "cultural", "learned" and "evolved".

Moreover, there does not appear to be a single program that causes learning in all domains (consider food aversions, snake phobias and grammar acquisition, as discussed in Principle 4). Evidence strongly substantiates the view that learning is caused by a multiplicity of programs. Without specifying which program is the cause, one has explained little, if anything, by invoking "learning" as an explanation for a behaviour. Labeling something learning does not remove the requirement to spell out the evolved machinery involved, it only makes the weak claim that interaction with the environment participated in the process – something that is almost always the case, anyway. In short, learning is a phenomenon that itself requires explanation. A coherent explanation for how people learn about a given domain must include (a) a description of what the evolved learning program looks like, (b) why it came to have that structure, both developmentally and over evolutionary time, and (c) what information is available to the organism that is executing that evolved program.

Of course, everyone is also an environmentalist, whether they know it or not. Even the most die-hard nativist understands that organisms learn – or, even more broadly, that an organism's evolved mechanisms extract information from the environment and process it to regulate behaviour. Hence the environment regulates behaviour – and it is the presence of evolved mechanisms that makes this possible.

So evolved programs – instincts, if you will – are not the opposite of learning. They are the engines or programs through which learning takes place. We learn only through instincts – learning and reasoning instincts. There are instincts in song birds for learning songs, instincts in geese for learning which individual is one's mother, instincts in desert ants for learning how to return home, and instincts in humans for learning a language. In the case of language, we have, as Darwin put it, "an instinct for acquiring an art".

Specialized or general-purpose?

If the "innate versus learned" controversy is meaningless, there is a genuine and illuminating question to be answered. It is *What is the precise structure of these evolved learning and regulatory programs?* Are there many, or just a few? Which use evolved crib sheets, and what do these crib sheets say? Which embody knowledge about enduring aspects of the world, what knowledge do their procedures reflect? To what extent is a program – whether it governs learning or not – functionally specialized to produce the outcome that you have observed?

What effect a given environmental factor will have on an organism depends critically on the details of the designs of its evolved cognitive programs. So the discovery of their structure is a pivotal question. Indeed, one of the few genuine nature-nurture issues concerns the extent to which each evolved program is specialized for producing a given outcome. Most nature/nurture issues disappear when one understands more about evolution, cognitive science, and developmental biology, but this one does not.

So, the important question for any particular behaviour is not "Is it learned" but "What kind of circuits produced it?" More specifically, "What is the nature of the universal, species-typical evolved cognitive programs through which we learned this particular type of behaviour or acquired this kind of knowledge?"

For any given outcome, there are three alternative possibilities: (1) It is the product of general-purpose programs (if such exist); (2) it is the product of cognitive programs that are specialized for producing that outcome; or (3) it is a byproduct of specialized cognitive programs that evolved to solve a different problem. (See Principle 4.)

The debate about language acquisition, which began in 1959 when Noam Chomsky reviewed – and panned – B. F. Skinner's book, *Verbal Behavior*, brings this issue into sharp focus, because Chomsky and Skinner disagreed about precisely these issues. Both sides in the ensuing controversy admit, as they must, that the human mind contains innate learning programs built by evolution. But the two camps differ in their answer to the question: Does a single set of general-purpose cognitive programs cause children to learn everything, with language as one incidental example? Or is language learning caused, in part or in whole, by programs that are specialized for performing this task – by what Chomsky called a *language acquisition device*?

Questions about functional specialization cannot be answered *a priori*, by theory or logic alone. Each proposal needs to be evaluated on its own merits: The theoretical tools and empirical studies necessary will differ, depending on whether the proposal is about language learning, mind-reading, the acquisition of gender roles, friendship, jealousy, or something else. For language, thirty years of research strongly supports the hypothesis that humans have evolved programs specialized for various aspects of language acquisition, although the debate remains heated.* With the emergence of evolutionary psychology, and under the weight of discoveries in many biological fields, the debate over adaptive specializations has now widened to include all human competences. Is how people reason or feel about groups the product of general-purpose learning, or are there specializations for coalitional psychology? Are there specializations that govern how people learn and reason about exchange, threats, family relations, friendship, incest avoidance, and so on, or are these processes governed solely by general-purpose learning?

A key adaptationist insight offered by evolutionary psychology is that these learning and reasoning devices were themselves built by natural selection. They are systems of parts that fit together over our evolutionary history because they functioned to solve the actual set of adaptive problems that our ancestors faced. Like other adaptations, their detailed structures should reflect their evolved functions, and so knowing the adaptive problems they evolved to solve catapults us ahead in the enterprise of discovering their structures. The study of what hunter-gatherers had to learn about and reason about should help us reverse engineer the battery of human cognitive devices.

A central goal of evolutionary psychology is to inventory all of the learning programs – specialized or not – in the human mind, and decode their structures to figure out how they work. If it turns out that at least some of these programs are specialized, with their own built-in content, then this will be a turning point in the history of the human sciences, because it will falsify assumptions, such as the belief in the blank slate, on which the modern social and behavioural sciences are built.

Present at birth?

Sometimes people think that to show that a behaviour or program is part of our evolved architecture, one needs to show that it is present from birth. Otherwise, the behavior is "learned" (by which they implicitly mean, learn through general-purpose processes). But this assumes that all of the evolved programs that cause development operate before birth, and none after birth.

This is clearly wrong. Teeth, breasts, and beards are all standard parts of our evolved architecture, but they develop after birth – 10 or 15 years after, in the case of breasts and beards. Newborns lack teeth; but does this mean babies acquire them through learning? Or that the design of their teeth was not organized by evolution?

Organs and design features can mature at any point of the life-cycle, and this applies to the cognitive programs in our brains just as much as it does to the features of our bodies. For this reason, the fact that a behaviour emerges after birth tells one very little about how it was acquired or why it has a certain organization. Organs can be disassembled on schedule as well, as we saw in the case of the brain-eating sea squirt, not to mention the placenta, umbilical cord, fetal haemoglobin, and a suite of infant reflexes that disappear as the frontal lobes mature. Evolutionists expect – and the evidence appears to bear them out – that many mechanisms will appear and disappear on a time-table based on when they would have been needed, under ancestral conditions, to solve the challenges of that life-stage. Infants need the sucking reflex; adolescents need sexual desires.

Presence at birth is only a function of what is needed at birth, not an indicator of whether something is or is not part of our evolved architecture. Accordingly, much of what is present in adult minds may have been put there by evolution, and activated through neural maturation, without depending on the accidents of personal experience. For example, infants who cannot crawl do not need a fear of heights, whereas infants who can crawl do. But experiments have demonstrated that a fear of heights is not learned by trial-and-error; rather, it is an evolved competence that is triggered when the baby starts to self-locomote, even if you contrive the situation such that the baby never experiences a fall.

Appreciating these facts can eliminate a number of misguided arguments. One common error is mistaking cultural transmission as an inescapable explanation for a behavioural pattern, when it is only a possible explanation. For example, baby boys cry, showing that the reluctance to cry is not present at birth. Moreover, boys could be learning that crying is unmasculine from television, where they may see representations of men having trouble crying, or of children making fun of a boy who cries. Taken together, these two facts are treated as proof that the culture responsible for causing men to cry less than women. But which is cause and which is effect? Does the fact that men don't cry on television teach boys not to cry, or does it merely reflect a fact about the reliably developing evolved architecture of boys and men? It could be either. In the absence of research on this particular topic, there is no way of knowing, and no basis for preferring one explanation to the other. To see this, just think how easy it would be to make an analogous argument: that girls acquire breasts through learning. Not only do baby girls lack breasts, but there are strong cultural forces that encourage the development of breasts – the peer pressure during adolescence; the celebration of famous beauties in the media. The culture reinforces the idea that women should have breasts, therefore, *obviously*, breasts are an arbitrary cultural construction no more part of women's design than apron-wearing is.…

Of course, the early presence of features is not completely irrelevant when evaluating alternative hypotheses about our evolved design. For example, the early emergence of a competence, before the social world could plausibly have acted, may falsify or undermine a particular social constructionist hypothesis. But the early absence of a competence does nothing to undermine the claim that it is part of our evolved design.

Co-determinism and the twin fallacies of biological determinism and environmental determinism

Traditional researchers hold a series of beliefs that are widely accepted and that sound eminently reasonable, but that are based on a series of fallacies about how the world works. The first belief is that some behaviour is genetically determined, while other behaviour is environmentally determined. This is wrong for many reasons, of which the most basic is that no behaviour is actually genetically determined, and no behaviour is environmentally determined. *All behaviour is co-determined by the interaction of both genes and environment.*

The next step in this fallacious series is the belief that evolutionary psychology only deals with – and can only deal with – behaviour that is genetically determined, not the much larger set of behaviours that are environmentally determined. Of course, this is wrong because all behaviour – learned or not – is jointly co-determined by genes and environment. However, an equally fundamental error committed here is the implicit assumption that evolutionary organization only emerges in the design of the organism through the action of genes, and environmental causation is non-evolutionary.

In reality, natural selection, acting over evolutionary time, has organized and coordinated the interaction between a species' genes and its environment, so that when they meet in the individual organism the two mesh togther to reliably construct the evolved architecture of the organism. Thus, the evolved organization to be found in organisms flows inward, from the enduring structure of the environment, just as much as it flows outward, from the genes. Evolution stores information useful for the construction of the species-typical design in the environment, just as much as it is stored in the genes.

For example, geese unsurprisingly prefer to mate with other geese when they become sexually mature. How do they acquire this preference? They have a set of specialized learning mechanisms that cause the newly hatched bird to identify the nearby animate entity as its mother, even if the experimenter substitutes a non-goose-like human instead. Evolutionarily, this environmentally presented stimulus would almost always have been the mother goose. Given this enduring environmental regularity, natural selection built a mate-attractiveness system out of this mother-identification circuit. The choice of what an attractive mate looks like to a goose is determined by an interaction between the genes (which use whoever is there when the gosling hatches to subsequently construct a template for mate-attractiveness), and an enduring environmental regularity (this is usually the gosling's mother, and its mother is a member of its species who is likely to have a normal appearance for a goose). Hence, the reliably developing feature of geese -- *mate with a member of your own goose species!* -- flows inward from the environment, and outward from the genes, in a dance in which natural selection has choreographed the roles played by the environment as well as the genes. Parallel accounts could be given for human language acquisition, how maternal feedback stabilizes an infant's tendency to smile, and so on.

How does natural selection choreograph the interaction between genes and environment to cause evolved design? How could natural selection orchestrate the role the environment plays? Developmental programs, by their design, make some parts of the world relevant to development, and other parts irrelevant. Natural selection, acting over evolutionary time, tinkers with the design of the developmental programs, adding and discarding features over time that hook onto or buffer against various parts of the environment. Over many generations, it tests out the utility of using various parts of the environment in building the organism. Genes underlying adaptations are selected so that, in development, genes and specific, stable aspects of the world interact to cause the reliable development of a well-designed adaptation. This means that the information and other structures necessary for the proper development of an adaptation may be stored in the world as well as in the genome, and that selection will shape developmental programs to exploit enduring features of the world. As a result, the organism is constructed out of two evolutionary inheritances: (1) its genes and (2) a set of enduring environmental regularities.

Natural selection adaptively coordinates the interaction of genes and environment to cause the reliable development of functional machinery *when and where it is needed*. For example, the developmental mechanisms of many organisms were designed by natural selection to produce different phenotypes when facing different social environments. Consider the blue-headed wrasse, a Carribean reef-living fish that lives in social groups consisting of one large male and many smaller females. If the male dies, the largest female – but not the smaller ones – turns into a male. The wrasse are designed to undergo a spontaneous sex change in response to two simultaneous social cues: the absence of a male and being the largest female. Without selection for genes that regulate and structure this process, it would not happen, and without species-typical environmental regularities, it also would not happen. This example illustrates why it is meaningless to attempt to distinguish socially or environmentally determined traits from genetically determined traits. The wrasse's genes render the fish's social environment relevant to determining its sex. In contrast, the human genetic endowment renders our social environment irrelevant to determining our sex: The largest woman in a household does not grow testes and produce sperm when the resident men move out.

The closest that the world actually comes to the fallacious distinction between biologically or genetically determined traits versus environmentally or socially determined traits is in the following real distinction: Some neural programs were designed by natural selection to take in substantial amounts of environmental input (e.g., the language acquisition device) whereas others were designed to take in less information (e.g., the reflex that causes the eye to blink in response to a looming figure). But in all cases, there is an underlying neural program designed by natural selection, and a set of environmental regularities necessary for that program's reliable development.

More nature means more nurture

Controversies over "nature versus nurture" – i.e., over how much of an individual's behavior is determined by genes versus environment (or instinct versus learning) – assumed that there is a zero-sum relationship between the two: that the larger the role of instincts and genes, the smaller the role of learning and environment. But it should be clear from the preceding discussion that this belief is entirely wrong.

In fact, evolutionary psychologists argue that, in the case of human evolution, there has been a positive sum relationship. That is, the more instincts or innate neural programs humans evolved, the more intelligently, richly, and sensitively humans could respond to added dimensions of information in the environment. When instincts were equated with inflexible, simple imperatives, this claim seemed paradoxical. It no longer seems so now that instincts are understood to be complex information-processing programs, like one finds on a computer. Given widespread familiarity with computers, researchers and nonacademics alike understand that each new application program installed on the computer's hard disk gives the computer a suite of new abilities and powers that it would not otherwise have. Each new plug-in on a web-browser allows the browser to understand new aspects of the world – new graphics formats, new sound formats, new video formats, or whatever. Similarly, each new evolved circuit in the human brain allows humans to see and understand some additional domain that is in the world – language, facial expressions, eroticism, the way people categorize living things into natural kinds, the humour of babies. In particular, the richer the innate architecture of learning mechanisms, the more an organism will be capable of learning. Toddlers can learn English while elephants (who are large-brained) and the family dog (who lives in an English-speaking household) cannot because the cognitive architecture of humans contains mechanisms that are not present in elephants or dogs. In short, despite widespread belief to the contrary, evolutionary psychology is a strongly environmentalist discipline, if the word *environmentalist* retains any valid meaning. Evolutionary psychology studies how information, extracted from the environment, is processed to sensitively regulate behaviour. Yet this necessarily involves researching and characterizing our evolved architecture – a nativist enterprise.

Fear of genes

Many people, and perhaps most academics, are disquieted by talk of genes, fear the implications for human life that any truthful examination of genes might bring, and dislike fields that introduce discussion of genes into polite conversation. They grudgingly acknowledge that genes may be real, and do important if obscure things, but on the whole they wish that the scientists who study them would go away or keep quiet. Yet obviously genes will not go away. Their effects organize brains and development, ramify through every human act, and they continue to operate in every cell in every human at every moment. The taboo on biology that hovers over the social sciences is worse than weak-minded – it has had a ruinous human cost. Not only has it stalled progress in the social sciences for decades, but it causes exactly the harm that it is supposed to prevent. This motivated ignorance allows harmful myths and superstitions about genes and biology to grow in a hothouse atmosphere, leading large segments of educated people to privately hold unwarranted conclusions – conclusions that could easily be dispelled by relaxed, accurate, and open dialogues about the nature of genes and their effects on minds, bodies, and behaviour.

Ironically, virtually all the fears that well-intentioned people harbor have turned out to be scientifically baseless, but the institutionalized avoidance of modern biology practiced by most social scientists has left them too unfamiliar with current findings to know this. While it would take a course in genetics to correct all of these misimpressions, we can address the chief anxieties and misunderstandings. The principal focus of dread is the well-established fact that individual humans are genetically different from each other. This fact appears disturbing, but only because of three widespread but erroneous beliefs to which it is connected. These are (1) genes have an irresistable causal power to predestine or determine outcomes, beyond the power of anything else to contravene; (2) genes are essences, that encapsulate a true, unchangeable underlying reality about the person; and (3) the genes of different racial groups might, if they were scrutinized, reveal immutable differences that would predestine social outcomes.

In reality, of course, neither genes nor environment determine anything in isolation from the other. Accordingly, genes cannot possibly determine anything all by themselves, but always act jointly with environmental factors to cause any developmental outcome. Hence, by changing the environmental conditions within which a gene acts, one changes the effect of the gene. Moreover, the idea that if something has a genetic cause, it is unchangeable is completely wrong – a throwback to the hidden essences believed in by medieval scholastics and young children. Genes are not supernatural agents operating outside of the world of causality, capable of foreordaining some outcome, regardless of countermeasures. After all, genes (like everything else in the body) are only fragile molecules, regulating the construction of other molecules in a dense network of chemical pathways. Such a network offers innumerable points where interventions could potentially be made to alter an undesired outcome. There is no aspect of the phenotype that cannot be influenced by some environmental manipulation. It just depends on how ingenious humans, as developmental psychologists and biologists, become. Genes have no special immunities from neutralization as causal agents, and so anything that might influence the organism, from Buddhist scripture to diet to wallpaper to reggae, could potentially change an outcome. Popular belief to the contrary notwithstanding, a genetic effect can be countered by an environmental manipulation, if you know which intervention to make. To take a trivial example, both authors have a genetic predisposition to near-sightedness, and became quite nearsighted as children. But we see quite well, thank you: we wear glasses and contact lenses – inventions that were possible because scientists were not afraid to learn about optics and how the eye works.

In short, *biology is destiny only to the extent that humans remain ignorant of it*. This is why the head-in-the-sand attitude of many social scientists has been disasterous, while the pursuit and integration of biological knowledge – resisted by so many – is emancipating. Indeed, rapid advances in molecular biology will soon render the magical belief in the immutability of genetic effects absurd to everyone but the most cloistered of social scientists.

As for racial differences, molecular genetics brings equally welcome news: Despite the differences in skin pigmentation and hair on the surfaces on bodies that make people appear different, geneticists can now travel inward past the skin to look directly at DNA. As they map the genome, they have been unable to find active genes that all normal members of one race have that other races lack, that would genetically demarcate races. At the genetic level, it turns out that human groups do not correspond to genetic types, and there is no objective way of fractionating humanity into "races" at all.   While each individual human embodies a unique combination of genes, making each individual genetically different from every other, the genes that vary between individuals are found widely distributed all over the planet, mixed into all races. As we will see, encompassing these minor differences is a species-typical, universal genetic architecture that endows us all with the same underlying human nature.

Evolutionary psychology is not behaviour genetics

Because of this widespread unfamiliarity with how genes play their roles in human life, many people think that all forms of research that relate genes to behaviour are the same. In reality, evolutionary psychology and behaviour genetics are two very different fields, animated by two radically different questions. Behaviour geneticists ask: "To what extent can *differences* between individuals be accounted for by *differences* in their genes?", whereas evolutionary psychologists ask: "What is the universal, evolved psychological and neural architecture that we all share by virtue of being human?" In behaviour genetics, the focus is on genes as the cause of individual differences; in evolutionary psychology, the focus is on genes as the cause of underlying inter-individual uniformity – indeed of universal human design.

This leads to two different perspectives on genetic differences. Behaviour geneticists are interested in genetically caused differences for their own sake, and because of their medical and psychiatric importance. Indeed, in the next several decades, we may owe the cure of schizophrenia, bipolar depression, autism, and many other disorders to the findings of behaviour geneticists. Behaviour geneticists formalize this investigation through measuring the heritability of behavioural traits, such as schizophrenia or shyness. They do this by computing a heritability coefficient for a trait, which can can vary from 0 to1. Heritability estimates the proportion of variation between individuals (in a given population) in a trait that is attributable to differences in their genes – roughly, variance due to genes divided by variance due to all sources (genes, environment, and their interaction). The data from which they compute this coefficient comes from studying how similar or different the studied trait is in different genetic relatives, such as parents and children, or identical twins, as well as the effect on the trait of being raised in similar or different environments.

In contrast, evolutionary psychologists are less interested in human characteristics that vary due to genetic differences, because they recognize that these are unlikely to be evolved adaptations central to human nature. Of the three kinds of characteristics that are found in organisms – adaptations, byproducts, and noise – traits caused by genetic variants are predominantly evolutionary noise, with little functional significance, while complex adaptations are likely to be universal in the species. Why is uniformity associated with functionality, and variability associated with lack of function? Alternative genes at the same locus (the same place in the human genome) are in a zero-sum competition for relative frequency in the species: The more common one allele is, the less common the others are. Natural selection tends to eliminate genetic differences whenever two alternative alleles (genes) differ in their ability to promote reproduction. Usually, the better functioning gene increases in frequency, squeezing out the less functional gene variant, until it disappears from the species. When this happens, there is no longer genetic variability at that locus – natural selection has produced genetic uniformity, instead. The more important the function, the more natural selection tends to enforce genetic uniformity. This is one reason why complex, important functional machinery tends to be universal in the species. In contrast, whenever a mutation fails to make a functional difference, selection will not act on it, and such minor variants can build up at the locus until there is a lot of genetic variability for the trait. For this reason, genetic variability tends to be predominantly nonadaptive or maladaptive evolutionary noise – neutral variants, negative mutations on their way to being eliminated, and so on. Genetic differences may be of great personal or medical importance – making one person shy, or another more prone to depression or breast cancer. But they are rarely evolved adaptations. If something is highly functional, selection usually acts to spread it to the entire species.

There is a second reason why evolutionary psychologists expect to find universality in our evolved neural programs. The fact that humans reproduce sexually as opposed to asexually constrains the design of genetic systems in a way that ensures that the genetic basis of any complex adaptation (such as the eye, the heart, or a cognitive mechanism) will be universal and species-typical. A complex adaptation is like any other intricate machine, whose parts must all be present and fit together precisely if it is to work. Each new human being is put together sexually: a *randomly* selected half of the mother's set of genes is recombined with a *randomly* selected half of the father's set of genes. If the suite of genes coding for a complex adaptation were in one parent but not the other, the offspring would receive only some of that adaptation's component parts, and it would fail to develop properly.* The only way to prevent the destructive scrambling of our complex adaptations every generation is for all of the *genes* necessary for coding for each complex adaptation to be universal, and hence reliably supplied by each parent. (This is why, for example, men have all the genes necessary to code for ovaries and a uterus; these genes are present in males, but not expressed.) In other words, *functional* aspects of the architecture will tend to be universal at the genetic level, even though their expression may be limited to a particular sex or age, or be contingent on the presence of an eliciting cue in the environment. Humans are free to vary genetically in their superficial, nonfunctional traits, but are constrained by natural selection to share a universal genetic design for their complex, evolved functional architecture.

The claim that the genetic basis for the human cognitive architecture must be universal is not a pious liberal falsehood – it is a profoundly important fact, derivable from adaptationist principles. This is why evolutionary psychologists, as part of their research program, often empirically test their models of cognitive adaptations cross-culturally. If something is truly a complex, evolved adaptation, it should be detectable in the Amazon as well as in London.

There is, nonetheless, a great deal of genetic variability within species, which is in tension with the functional advantages of genetic uniformity. Aside from mutations and neutral variants, there is a third reason for this genetic diversity. Genetic variability, such as the ABO blood group system, is retained in the species because genetically-based biochemical individuality interferes with the transmission of infectious diseases from host to host. Diseases that use or depend on a protein found in their present host are thwarted when the next individual they jump to has a different protein instead. Hence, natural selection sifts for genetic variants which supply approximately the same functional properties to the adaptations they participate in, but which "taste different" from the point of view of disease organisms. Because we catch diseases from those we have contact with – such as our family, neighbors, and other locals -- selection favors maximizing genetically based protein diversity locally – which requires pulling into every local population as many of the genetic variants found anywhere in the species as possible. This is why individuals are genetically different from each other, but different populations tend to be so surprisingly genetically similar. These genetic differences introduce minor perturbations into our universal designs. The result is that each normal human expresses the universal human design but, simultaneously, each human is slightly different from every other in personality, temperment, and appearance. These differences tend to be quantitative in nature – a little more of this, a little less of that – while the overall architecture remains the same. This is why the study of universal design by evolutionary psychologists is so different an enterprise from the behaviour genetics study of differences.

Consider sexual jealousy. Behaviour geneticists would approach its investigation by measuring it as a dimensional trait – more jealous to less jealous – along which individuals vary. They would then attempt to determine if genetic differences between individuals accounted for some of the observed variation in this trait – i.e., whether differences in jealousy are heritable, and which alleles cause differences. Evolutionary psychologists take a quite different approach. Instead of viewing jealousy as a dimensional trait, they view it as a cognitive program with a complex functional design. Their goal is mapping the architecture of the program: What environmental cues activate the program? What conditions is the program designed to conceptually represent? (E.g., partner fidelity vs. infidelity; sexual rivals; opportunities for unmonitored sexual activity by the partner; partner trust.) What behaviour does it motivate?

(E.g., seeking information about potential infidelity; violence against the rival; deterrence of the partner; ending the relationship.) What functional outcomes is the program designed to accomplish? They expect its complex design to be human-universal, and expect that whatever genetic differences show up in its components will be minor perturbations in the species-typical design. Thus, the threshold for activation of the jealousy program may vary between individuals because of the random distribution of minor genetic variants – some individuals might become jealous more easily than others. But all humans should be designed by natural selection to have circuits that cause them to view sex between their mate and a rival negatively, and be motivated to prevent it (despite anthropological myths to the contrary). From the diversity of observed behaviour, the goal of evolutionary psychology is to abstract out the design of an underlying human-universal program that generates the diversity.

What does this mean? Individuals will differ from each other because of two reasons: different genes and different environmental histories. Environmental inputs (e.g., has aggression by the individual been successful in the past?) will be taken as inputs by the brain's evolved programs, which have been designed by evolution to process them and generate adaptively appropriate behavioural output. Hence, individual differences due to different environmental inputs should be patterned according to a functional adaptive logic. Minor genetic perturbations in this design should lead to an additional layer of individual differences that are not adaptively patterned, and hence not reflective of our evolved design. For this reason, evolutionary psychologists focus on the search for cross-culturally universal functional patterns between environmental inputs and behavioural outputs – in this sense, evolutionary psychology is largely an environmentalist discipline.

Reciprocally, evolutionary psychologists are unsurprised by behaviour geneticists' findings that genetic differences play a major role in causing stable lifelong personality differences between individuals. If someone is, for example, angrier than others across situations and across the lifespan, even outside of conditions where this is appropriate, then this is unlikely to be functional – that is, the well-calibrated output of an adaptation. Anger in the well-engineered brain should be activated when it is functional, and disappear when it is not, according to evolved rules that selection has made universal. If someone's brain is stuck at an angry (or depressed or shy) setting regardless of environmental inputs and across the lifespan, then a common reason will be that the chemical products of the genetic variants that she carries bias the neural architecture in that direction. This part of personality formation is just an accident of random gene shuffling.

Genetics can be illuminating, however, not just about individual differences, but also about universal architecture. The field of cognitive neuroscience uses studies of individuals with brain damage to throw light on normal species-typical neural architecture. If a patient can lose one ability (such as the ability to reason about social exchange) while retaining others (such as the ability to reason well about other topics) then one knows that reasoning about social exchange is carried out, at least in part, by a separate program. Similarly, the study of individuals whose rare mutations have unusual cognitive effects have the same potential to reveal the outlines of separate mechanisms. Individuals with Williams' Sydrome, who suffer from profound retardation, often have normal or even exceptional language skills, indicating that the language acquisition system is separate from general learning and reasoning abilities. SLI, or selective language impairment, causes certain grammatical disturbances, while leaving most other cognitive functions intact. Each of these results tells us something important about our evolved cognitive architecture, and the promise in this method has hardly been tapped.

What heritability does not mean

Studies of heritability are often feared because they are imbued with meanings that, as a scientific matter, they simply do not have. For one thing, many people believe that heritability studies rank traits in terms of how much they are genetically fixed or open to change. Because IQ, whatever it may be, appears to be moderately to highly heritable (between 50% - 80%) people think that this establishes that IQ is resistant to change (for example, that it could only be improved under the best possible conditions by 20%). This is a complete misinterpretation of what heritability means: A high heritability says nothing about whether there exist, potentially or actually, manipulations that could, for example, triple (or halve) IQ. Heritability simply does not measure resistance to environmental manipulation – it only measures, for a specific population with a specific distribution of genes and a specific distribution of environmental treatments, what proportion of the *differences* between people *happened* to have been caused by genetic differences as opposed to the environmental treatments that have actually been experienced. Create a new environmental manipulation, or change the relative frequencies or targets of existing environmental treatments, and the heritability coefficient might change by any magnitude. Even more importantly, the trait as expressed in an individual or a targeted population might also change by any magnitude (e.g., for environmental reasons that remain unclear, the average age of menarche has been dropping dramatically in industrial populations). For similar reasons, as a scientific matter it is impossible to tell from heritability studies in humans what proportion of intergroup differences in a trait are due to genetic differences between groups. Despite decades of controversy, heritability studies cannot be made to produce meaningful public policy implications about human groups. In contrast, for disorders in which genetic variants play a contributory role, it is vitally important to identify the mechanism of action, and we may be the first generation to see the prevention or cure of schizophrenia, bipolar depression, autism, obsessive-compulsive disorder, and scores of other harmful conditions.

A second common belief is that high heritabilities signal traits that are more genetically influenced, and hence more evolutionarily significant, while low heritabilities signal traits that are less genetically influenced, more attributable to environmental action, and less evolutionarily sculpted. This is also completely erroneous. All traits are jointly co-determined by the action of genes and environment, and so genes participate equally in the construction of *all* traits, whether they are heritable or not. The difference between a heritable trait and a nonheritable trait is that in the heritable trait, *different genes* situated in different individuals participate with environmental factors to co-determine the outcomes, while in a nonheritable trait, the *same genes*, uniformly present in all individuals, participate with environmental factors to co-determine the outcomes. In both cases, genes inescapably co-determine the outcome. When natural selection eliminates less functional genes from a population, leaving genetic uniformity at a locus, the same trait that exhibited heritable differences in preceding generations now has heritability coefficient of zero. Why? Because at that point, none of the remaining differences between individuals in the trait are caused by genetic differences.

Thus, our evolved, universal human design – having two eyes, two arms, two legs, a brain – has everything to do with our genes, but has low or zero heritability. This is because *differences* between individuals in eye number, leg number, arm number, and so on, are usually caused by environmental factors – accidents, such as car crashes and war injuries – rather than by genetic differences. The *uniformity* in design among all individuals is caused by a species-wide genetic uniformity; by functionally equivalent effects even where genes are biochemically different between individuals; and by species-wide environmental regularities. Because their genetic basis is universal and species-typical, our complex cognitive adaptations, such as a language acquisition device, our emotional adaptations, or our reasoning instincts, are inherited, but not heritable.

Evolutionary psychology focuses on universal architecture

At a certain level of abstraction, every species has a universal, species-typical evolved architecture, which in the case of humans is called human nature. This is not an accident, a trick of definition, nor a piety. As we outlined in the preceding sections, natural selection and sexual reproduction combine to impose uniformity across the species in our complex adaptations, whether neural or not. So, for example, one can open any page of the medical textbook, *Gray's Anatomy*, at random, select the most minute anatomical feature in an illustration, and find that it accurately characterizes humans in Uruguay, Irkutsk, Mauritania, Edinburgh, and Samarkand. Not only do we all have a heart, two lungs, a stomach, intestines and so on, but this functional uniformity extends down to microscopic functional structures such as flagella and mitochondria all the way to enzymatic pathways, such as the Krebs cycle. Of course, no two stomachs are exactly alike – they vary a bit in quantitative and nonfunctional properties, such as size, shape and how much acid they produce. But all humans have stomachs, and their stomachs all have the same basic functional design: each is attached at one end to an esophagus and at the other to the small intestine, each secretes the same chemicals necessary for digestion, and so on. The same is true of the brain and, hence, of the evolved architecture of our cognitive programs -- of the information-processing mechanisms that generate behaviour. The programs underlying vision, romantic love, friendship, jealousy, revenge, and so on, can be expected to be human-universal. Evolutionary psychology seeks to characterize the universal, species-typical architecture of these mechanisms, precisely because complex functionality is associated with species-typicality.

Because the functional operation of complex machinery depends upon its component parts meshing, the developmental process that builds each complex adaptation, over the life of the individual, must have built in safeguards that protect the developmental process from disruption. These safeguards are themselves adaptations, that have evolved ancestrally, and so their designs are only designed to compensate for the types of disruptions present in the ancestral world. So, the development our cognitive or neural architecture is buffered against various genetic and environmental insults that individuals may suffer, such that it reliably develops across the (ancestrally) normal range of human environments.

Finally, evolutionary psychologists focus on our evolved architecture – the constellation of neural or cognitive adaptations -- rather than on behaviour per se. This is because (1) functional design will be recognizable in the structure of the programs, even though behaviour under modern conditions may not be functional; and (2) although our neural adaptations will be universal, there may be few universal behaviours. After all, our programs evolved to generate behavior on the basis of information that they extract from the world, and since the informational input will vary between individuals, behavior will vary between individuals as well. When people with the same cognitive programs are exposed to different information, this will cause their programs to generate different representations about the world – different knowledge and values – and this will lead them to behave in different ways. For this reason, one should expect few behaviours to be universal across cultures, even though, all people have basically the same cognitive programs.

Ultimate versus proximate explanation

As we discussed, natural selection's feedback process creates structures that reflect adaptive functions. To capture this causal relationship between function and structure, biologists had to develop a theoretical vocabulary that distinguishes between them. In evolutionary biology, explanations that appeal to the structure of a device are called *proximate* explanations. When applied to psychology, these would include explanations that focus on genetic, biochemical, physiological, developmental, cognitive, and all other immediate causes of behaviour, including the information that evolved programs take as input. In contrast, explanations that appeal to the adaptive function of a device are called *ultimate* explanations. Ultimate explanations refer to the causes that led, over evolutionary time, to the proximate structure being organized in the way that it is rather than in some other physically possible way. Why did the mother's breast release milk just now? Proximate explanations would include "Because she heard her baby cry", "Because prolactin was released", "Because her breasts were engorged, since the baby habitually nurses at this time of day", and so on. Ultimate explanations – at different levels of specificity – would include, "Because mothers whose circuit designs caused them to deploy energy more efficiently had more offspring", "Because natural selection created circuits whose function was to save the mother's energy by having milk produced when the baby is most likely to need it", "Because under ancestral conditions, certain cues were correlated with the baby needing milk which, given the right circuitry, could be monitored and used to regulate milk production", "Because natural selection solved the problem of being more energy efficient by creating information-processing circuits that (a) cause letdown in response to crying, (b) produce milk as a function of sucking rate, (c) entrain production to the baby's circadian rhythms, as indicated by time of day that nursing occurs, (d) …".

All these levels and types of explanation are complementary and mutually compatible, rather than mutually exclusive. Unfortunately, many of the debates that surround evolutionary psychology are simple misunderstandings, caused by a failure to separate these different kinds of explanations. So, for example, language is the result of an evolved competence *and* language is learned from others in a social community with a unique history: There is no contradiction whatsoever between these two statements.

Is evolutionary psychology politically suspect?

By far the most powerful force that keeps scientists and the public from open-mindedly learning about evolutionary psychology has been a set of academic urban legends about it. The central contention is that Darwinism invites or abets immoral political views. What follows from this is the widespread conviction that any doctrine that undermines the relevance of Darwinism to humans is to be welcomed (no matter how weak-minded or illogical), while any research that illustrates its relevance ought to be fought, dismissed, deflated, or ignored by all ethical people. Of course, many academics have made heavy investments in existing ways of thinking. When new discoveries break existing habits of thought, many traditionalists routinely protect their investments with accusations of immorality. But such accusations are a rhetorical strategy, rather than a genuine moral response based on an accurate characterization of their opponents.

Because most academics are currently on the left, the most effective obfuscation is the accusation that evolutionary psychology is a right-wing movement, just as in other times and places new ideas have been called anti-Christian, anti-German, anti-Soviet, un-American, pro-Lutheran, and so on. Polemicists come forward to play to the market created by such anxieties, and the fast-spreading ad hominem assertions soon acquire the status of "fact."

Among the most imaginatively unfounded of these attacks is the claim that evolutionary psychology was constructed to support systems of social and racial inequality by claiming that ethnic, racial, and class subordination is biologically determined and deserved. This accusation is obviously disingenuous because, by its very nature, the study of those aspects of the human psychological architecture that are human-universal cannot possibly play a role in the argument that some ethnic groups or classes are biologically destined to be unequal. Indeed, some of the most powerful arguments about why all human populations must share the same set of complex psychological adaptations have come out of recent discoveries in evolutionary psychology.

More generally, reality is rich, subtle, textured, and complex, while ideologies are simple-minded and reductive. It is inevitable that any set of scientific innovations will violate some aspects of any and all pre-packaged ideologies. Indeed, humans are such absurdly factional animals that Renaissance Europeans were perfectly able to politicize theories about the behaviour of immensely remote astronomical bodies, forever distant from Earth – theories as irrelevant to the everyday practicalities of human society as anything could possibly be. Evolutionary psychology is not about distant chunks of rock, gas and ice, but about something far closer to home – ourselves – and so has inevitably been attacked by doctrinaire partisans of the right, left, and center, by crank and traditionalist alike. Conversely, imaginative advocates from virtually every political persuasion – from Black Panthers and neo-conservatives to liberals, former communists, New Labourites, anarchists, feminists, libertarians, and democratic socialists – have professed to find support for their views in its discoveries.

Our own view, for what it is worth, is that it is too early by decades to draw political lessons from this science, because the field is only in its infancy. One would need a complete model of human nature (plus some nonscientific source of value judgments) to construct a mature political theory. But the scientific community is only starting to map the human mind – we are not anywhere close to being finished, and no one can predict what the final product will look like. Moreover, human interpretive freedom and value differences would likely lead to a broad range of imaginative political derivations, even then.

In any case, to obstruct the scientific mapping of the human mind is to sentence us to a continued ignorance of our natures. Over the centuries, this ignorance has condemned hundreds of millions to their deaths, and countless more to unbearable suffering. We think there is nothing more destructive or brutal than its continuation, and that nothing would be more emancipating than to work open-mindedly to replace it with genuine scientific knowledge. This will involve the labor of thousands, working over decades. We expect that such knowledge will transform our view of ourselves and the world far more radically and profoundly than the revolutions in the physical and biological sciences which have preceded it.

4. REASONING INSTINCTS: AN EXAMPLE FROM SOCIAL EXCHANGE

To give a brief taste of what a concrete project in evolutionary psychology looks like, we will use an example from our own research. We have been exploring the hypothesis that the human cognitive architecture contains circuits specialized for reasoning about adaptive problems posed by the social world. In categorizing social interactions, there are two basic effects that humans can have on each other: helping or hurting, bestowing benefits or inflicting costs. These acts can be delivered either conditionally or unconditionally. For example, a mother nurses her infant unconditionally – i.e., without asking it for a favour in return. But humans often put conditions on the performance of a social act: for example, “I will help you *if* you help me”, or “I will hurt you *if* you do not do something for me”. Two major categories of social conditionals are social exchanges (benefit for benefit) and threats (produce a benefit to avoid a harm) – acts carried out by individuals or groups on individuals or groups. We believe that the structure of these social interactions among our hunter-gatherer ancestors selected for specialized cognitive programs that could detect, understand, and draw inferences about social conditionals reliably, precisely, and economically. According to this view, such reasoning specializations have endowed humans with the ability to cooperate for mutual benefit, and with the ability to navigate through the dangers of responding to threats and punishments. The evolutionary logic built into these neural programs underpin our legal and economic systems, and choreograph much of the social dance of our daily lives.

In our research, we focused initially on social exchange. We selected this topic for several reasons. First, many aspects of the evolutionary theory of social exchange (sometimes called cooperation, reciprocal altruism, or reciprocation) were well-developed. Hence, they allowed us to derive clear prior hypotheses about the structure of the information-processing procedures necessary for engaging in this activity.

Second, complex adaptations are constructed in response to evolutionarily long-enduring problems, not recent human activities. Evidence from primatology and paleoanthropology suggested that our ancestors had engaged in social exchange for at least several million years,more than long enough to build good adaptations in response.

Third, complex adaptations are built in response to important adaptive problems, and social exchange appears to be an ancient, pervasive and central part of human social life. Moreover, although the universality of a behavior is not, by itself, a sufficient reason for believing that a cognitive adaptation evolved to produce it, it is suggestive. Social exchange is nearly as ubiquitous as the human heartbeat. The heartbeat is universal because the organ that generates it is everywhere the same. A parsimonious hypothesis for explaining the universality of social exchange would be that the neural program that generates it is everywhere the same. Like the heart, its development does not seem to require environmental conditions (social or otherwise) that are idiosyncratic or culturally contingent.

Fourth, theories about reasoning and rationality have played a central role in cognitive science, the social sciences, philosophy, and economics. Virtually every past approach to rationality has assumed that human reasoning consists of nothing but general-purpose, content-independent machinery. Research into the design of human rationality can, as a result, serve as a powerful test of the utility of applying domain-specific adaptationist principles to new areas of psychology, and hence as a test of the central assumption of the Standard Social Science Model: that the evolved architecture of the mind consists solely or predominantly of a small number of content-independent, general-purpose mechanisms. Although innumerable theories invoke human rationality, researchers have yet to provide a circuit diagram of it. Hence, tracing out pieces of this diagram seemed worthwhile.

The evolutionary analysis of social exchange parallels the economist's concept of trade. Social exchange is an "I'll scratch your back if you scratch mine" principle. Economists and evolutionary biologists had already explored constraints on the emergence or evolution of social exchange using game theory.* One important conclusion was that social exchange cannot evolve in a species or be stably sustained in a social group unless cheaters are excluded from future interactions in which they would otherwise continue to exploit cooperators. For the ancestral conditions that humans faced, this required the evolution of cognitive machinery that, among other things, allows a potential cooperator to identify potential exchange interactions, identify the conditions on the exchange that each party imposes on the other and, especially, to detect individuals who cheat, so that they can be excluded. In this context, a cheater is defined as an individual (or group) who accepts a benefit without satisfying the requirement set by the person (or group) providing the benefit.

So, our research strategy has been to (1) model how natural selection operated in ancestral hunter-gatherer contexts in order to (2) construct models of the specific computational problems our ancestors would have had to solve; (3) from these, construct hypotheses about the design features that our evolved neural programs would have to have to solve those problems; (4) construct experiments to test for the presence of the predicted design features; (5) test these findings cross-culturally to see if they are universal; (6) use cognitive neuroscience techniques to determine the neural basis for the cognitive adaptation.

The way we test the structure of hypothesized reasoning specializations is to probe subjects with reasoning problems of varying contents and structures, to see which types they find difficult, which they find intuitive, obvious, and effortless, and what specialized principles they supply from their own minds that are not present in the problems themselves. In particular, we predicted spikes of high performance associated with certain problem contents (such as social exchange, threats, and hazard avoidance) that that do not occur for problems that are logically equivalent, and of equal or greater familiarity. By adding and subtracting properties and contents from the problems, we can see exactly which elements trigger which adaptively organized responses.

The evolved cognitive programs that are cause social exchange have a large number of functionally specialized design features, but the one that we will discuss involves the prediction that humans have inference procedures that are specialized for detecting cheaters. To test this hypothesis, we used an experimental paradigm called the Wason selection task. For about 35 years, psychologists had been using this paradigm (which was originally developed as a test of logical reasoning) to probe the structure of human reasoning mechanisms. In this task, the subject is asked to look for violations of a conditional rule of the form *If P then Q*. Consider the Wason selection task presented in Figure 1 (*P*s and *Q*s are for the reader’s convenience, and do not appear on the actual task).

Figure 1

Ebbinghaus disease was recently identified and is not yet well understood. So an international committee of physicians who have experience with this disease were assembled. Their goal was to characterize the symptoms, and develop surefire ways of diagnosing it.

Patients afflicted with Ebbinghaus disease have many different symptoms: nose bleeds, headaches, ringing in the ears, and others. Diagnosing it is difficult because a patient may have the disease, yet not manifest all of the symptoms. Dr. Buchner, an expert on the disease, said that the following rule holds:

"If a person has Ebbinghaus disease, then that person will be forgetful."
If P then Q

Dr. Buchner may be wrong, however. You are interested in seeing whether there are any patients whose symptoms violate this rule.

The cards below represent four patients in your hospital. Each card represents one patient. One side of the card tells whether or not the patient has Ebbinghaus disease, and the other side tells whether or not that patient is forgetful.

Which of the following card(s) would you definitely need to turn over to see if any of these cases violate Dr. Buchner's rule: "If a person has Ebbinghaus disease, then that person will be forgetful." Don't turn over any more cards than are absolutely necessary.

has Ebbinghaus disease	does not have Ebbinghaus disease	is forgetful	is not forgetful
P	*not-P*	*Q*	*not-Q*

From a logical point of view, the rule has been violated whenever someone has Ebbinghaus disease but is not forgetful. Hence the logically correct answer is to turn over the "has Ebbinghaus disease" card (to see whether this person is not forgetful) and the "is not forgetful" card (to see whether this person has Ebbinghaus disease). More generally, for a rule of the form *If P then Q*, one should turn over the cards that represent the values *P* and *not-Q*. If the human mind develops reasoning procedures specialized for detecting logical violations of conditional rules, this would be intuitively obvious. But it is not. In general, fewer than 25% of subjects spontaneously make this response. Moreover, even formal training in logical reasoning does little to boost performance on descriptive rules of this kind. Indeed, there is a large literature that shows that people are not very good at detecting logical violations of if-then rules in Wason selection tasks, even when these rules deal with familiar content drawn from everyday life.

The Wason selection task provided an ideal tool for testing hypotheses about reasoning specializations designed to operate on social conditionals, such as social exchanges, threats, permissions, obligations and so on because (1) it tests reasoning about conditional rules; (2) the task structure remains constant while the content of the rule is changed; (3) it is easy to change performance by altering the content of the tests; and (4) there was already a body of existing experimental results against which performance on new content domains could be compared.

For example, showing that people who ordinarily cannot detect violations of conditional rules can do so when that violation represents cheating on a social contract would constitute support for the view that people have cognitive adaptations specialized for detecting cheaters in situations of social exchange: it would provide some evidence of special design. Finding that violations of conditional rules are spontaneously detected when they represent bluffing on a threat would, for similar reasons, support the view that people have reasoning procedures specialized for analyzing threats. The inability of subjects to detect violations of conditional rules in 'neutral' contexts acts as a baseline against which to look for the performance-boosting effects of reasoning specializations. By seeing which types of content produce high performance on the tasks, the boundaries of different reasoning specializations can be ascertained.

The results of these investigations were striking. People who ordinarily cannot detect violations of if-then rules can do so easily and accurately when that violation represents cheating in a situation of social exchange. This is a situation in which one is entitled to a benefit only if one has fulfilled a requirement. For example, "If you borrow my car, then you must fill the tank with gas"; "If a man eats cassava root, then he must have a tattoo on his chest"; or, more generally, "If you take benefit B, then you must satisfy requirement R". Cheating is accepting the benefit specified without satisfying the condition that provision of that benefit was made contingent upon (for example, borrowing the car without filling the tank).

When asked to look for violations of social contracts of this kind, the adaptively correct answer is immediately obvious to almost all subjects, who commonly experience a "pop out" effect. No formal training is needed. Whenever the content of a problem asks subjects to look for cheaters in a social exchange – even when the situation described is culturally unfamiliar and downright bizarre – subjects experience the problem as simple to solve, and their performance jumps dramatically. In general, 65-80% of subjects get it right, the highest performance ever found for a Wason selection task. They choose the "benefit accepted" card (for example, "ate cassava root") and the "requirement not met" card (for example, "no tattoo"), for any social conditional that can be interpreted as a social contract, and in which looking for violations can be interpreted as looking for cheaters.

From a domain-general, formal view, investigating men eating cassava root and men without tattoos is logically equivalent to investigating people with Ebbinghaus disease who are not forgetful. The rules of logic, after all, are content-free. But in cross-cultural investigations, wherever it has been tested – including adults in the USA, UK, Germany, Italy, France, and Hong Kong, schoolchildren in Ecuador, and Shiwiar hunter-horticulturalists in the Ecuadorian Amazon – people do not treat social exchange problems as equivalent to other kinds of reasoning problems. Their minds distinguish social exchange contents, and reason as if they were translating these situations into an evolved mental language built into a specialized neural program, equipped with what philosophers would once have called innate ideas – ideas such as "benefit", "requirement", "obligation", "entitlement", "cheater", "intention", and "agent". Indeed, the relevant inference procedures are not activated unless the subject has represented the situation as one in which people are entitled to a benefit only if they have satisfied a requirement.

Do people choose the *requirement not met* card (*not-Q*) merely because social exchange activates a normally dormant rule of logic, the (content-free) modus tollens rule that we met earlier? No. The procedures activated by social contract rules do not behave as if they were designed to detect logical violations per se. Instead, they prompt choices that track an adaptive "logic" that is useful for detecting cheaters, whether or not this results in an answer that is logically correct. For example, by switching the order of requirement and benefit within the if-then structure of the rule, one can elicit responses that are adaptively correct from the point of view of a cheater detection procedure, but that are logically incorrect (see Figure 2). Subjects choose the “benefit accepted” card and the "requirement not met" card -- the adaptively correct response if one is looking for cheaters -- no matter what logical category these cards correspond to. Hence, we have evidence not for a general rationality based on adherence to formal logic, but instead for a diverse series of functionally specialized, domain-specific adaptive logics, each designed to master a particular kind of content (exchanges, threats, dangers, and so on).

Figure 2. The Wason selection task: Generic Structure of a Social Contract Problem

The following rule holds:

“If you take the ***benefit***, then you satisfy the ***requirement***.” (*standard form*)

(If P then Q)

“If you satisfy the ***requirement***, then you take the ***benefit***.” (*switched form*)

(If P then Q)

The cards below have information about four people. Each card represents one person. One side of a card tells whether a person accepted the benefit, and the other side of the card tells whether that person satisfied the requirement. Indicate only those card(s) you definitely need to turn over to see if any of these people are violating the rule.

✔ benefit accepted	benefit not accepted	requirement satisfied	✔ requirement not satisfied	
P	*not-P*	*Q*	*not-Q*	*(standard form)*
Q	*not-Q*	*P*	*not-P*	*(switched form)*

To show that an aspect of the phenotype is an adaptation, one needs to demonstrate a close fit between form and function: one needs design evidence. There are now a large number of experiments comparing performance on Wason selection tasks in which the conditional rule did or did not express a social contract. These experiments have provided evidence for a series of specialized subcircuits predicted by our analysis of the adaptive problems that arise in social exchange. Social contracts activate content-dependent rules of inference that appear to be complexly specialized for processing information about this domain. Indeed, they include subroutines that are specialized for cheater detection, for distinguishing innocent mistakes from intentional cheating, for distinguishing social exchanges from other conditionals, for taking the perspective of each party to the exchange, for taking the personal history of the participants into account, and so on. The programs involved do not operate so as to detect potential altruists (individuals who pay costs but do not take benefits). Nor are they activated in social contract situations in which errors would correspond to innocent mistakes rather than intentional cheating: situations in which the violator breaks the rule accidentally, and does not obtain the benefit specified in the rule by so doing. Nor are they designed to solve problems drawn from domains other than social exchange; for example, they will not allow one to detect bluffs and double crosses in situations of threat, or allow one to detect when a safety rule has been violated. The experiments have also empirically eliminated a number of alternative, byproduct hypotheses. Social contract effects cannot be explained by familiarity with the task, the existence of permission schemas (which are more content-general), the use of a content-independent logic, or any other counter-hypothesis that has been advanced so far in the scientific literature. The pattern of results elicited by social exchange content is so distinctive that we believe reasoning in this domain is governed by computational units that are domain-specific and functionally distinct: what we have called social contract algorithms.* More recently, we and our colleagues have found neural evidence for the existence of a social exchange specialization: Brain damage can selectively impair an individual's ability to detect cheaters in social exchange, while having no detectible effect on that individual's ability to solve logically equivalent problems from other domains. This would not be possible if reasoning were carried out only by general-purpose circuits – damage would degrade performance on all topics, not knock out performance narrowly on one.

There is, in other words, strong design evidence that show that the properties of specialized social exchange algorithms match the demands of the adaptive problem of social exchange. The programs that cause reasoning in this domain have many coordinated features that are complexly specialized in precisely the ways one would expect if they had been designed by a computer engineer to make inferences about social exchange reliably and efficiently: configurations that are highly unlikely to have arisen by chance, and so are likely to be the result of natural selection operating in ancestral environments. By starting with an important adaptive problem that hunter-gatherers had to be able to solve, we were able to discover evidence for a previously unknown mental program. This program makes possible a certain category of human interaction – cooperation – that would not be possible in its absence, and that underlies everything from human economic life to the exchange of favors. Of course, research into social exchange was undertaken as a test case – one demonstration of the usefulness of the paradigm. Evolutionary psychologists have discovered design features of many other evolved programs, but because the human mind is so vast, this is only the beginning: There are thousands of programs left to investigate.

Table 1. Computational machinery that governs reasoning about social contracts (Based on evidence reviewed in Cosmides & Tooby, 1992)

Design features:

1. It includes inference procedures specialized for detecting cheaters.
2. The cheater detection procedures cannot detect violations that do not correspond to cheating (e.g., mistakes where no one profits from the violation).
3. The machinery operates even in situations that are unfamiliar and culturally alien.
4. The definition of cheating it embodies varies lawfully as a function of one's perspective.
5. The machinery is just as good at computing the cost-benefit representation of a social contract from the perspective of one party as from the perspective of another.
6. It cannot detect cheaters unless the rule has been assigned the cost-benefit representation of a social contract.
7. It translates the surface content of situations involving the contingent provision of benefits into representational primitives such as "benefit", "cost", "obligation", "entitlement", "intentional", and "agent."
8. It imports these conceptual primitives, even when they are absent from the surface content.
9. It derives the implications specified by the computational theory, even when these are not valid inferences of the propositional calculus (e.g., "If you take the benefit, then you are obligated to pay the cost" implies "If you paid the cost, then you are entitled to take the benefit.").
10. It does not include procedures specialized for detecting altruists (individuals who have paid costs but refused to accept the benefits to which they are therefore entitled).
11. It cannot solve problems drawn from other domains; e.g., it will not allow one to detect bluffs and double crosses in situations of threat.
12. It appears to be neurologically isolable from more general reasoning abilities (e.g., it is unimpaired in schizophrenic patients who show other reasoning deficits; Maljkovic, 1987).
13. It appears to operate across a wide variety of cultures (including an indigenous population of hunter-horticulturalists in the Ecuadorian Amazon; Sugiyama, Tooby & Cosmides, 1995).

Alternative (byproduct) hypotheses eliminated:

1. That familiarity can explain the social contract effect.
2. That social contract content merely activates the rules of inference of the propositional calculus.
3. That social contract content merely promotes (for whatever reason) "clear thinking".
4. That permission schema theory can explain the social contract effect.
5. That any problem involving payoffs will elicit the detection of violations.
6. That a content-independent deontic logic can explain the effect.

CONCLUSIONS

The view that the mind is a teeming confederation of ancient, expert programs is a dramatic and strange departure from the traditional way scientists and nonscientists had thought about the mind. If it proves true, it will revolutionize how social scientists see the relationship between human nature and human life, and will require changes in nearly every field of knowledge that relates to humans. It matters deeply whether evolved specializations exist.

The old model of human nature as general-purpose learning circuits, connected to a blank slate, was a very bland characterization of humanity – one that supplied few specific insights into human social life or culture. Indeed, it denied the reality of any richer characterization of human nature, and it appealed to many scholars precisely because it provided an intellectual justification for disconnecting human nature from human culture.

The model emerging from evolutionary psychology is anything but bland. Its model of human nature has character, definition, and bite, because each formal model of an evolved mental program provides specific implications about its particular area of human life. According to this new view, there are specific circuits in the mind for cooperation, love, sex, parenting, friendship, language, status, ingroup identification, families, deception, remembering and representing the identities of others, incest, food, number, disgust and contamination, jealousy, revenge, aesthetics, violence, play, imitation, extortion, kindness, trade, foraging, dangers, causality, beauty, coalitions, free-riding, morality, natural landscapes, geometry, predators, weapons, gender, and so on. These programs are not the products of culture – culture is the product of these programs.*

This emerging model of human nature promises to be rich, precise, specific, and detailed. It will be the work of many lifetimes to produce high-resolution maps of all the circuits and features of human nature, but when this enterprise is finished, the product will be an exacting scientific theory that will rival anything in physics, geology or chemistry. Such detailed and reliable information about the program circuitry of the human mind will inevitably provide the foundation on which a revised set of social sciences will be built, because the decision-rules built into these programs are relevant to all aspects of how humans interact socially, what causes them to adopt some cultural elements over others, and what economic choices they make.
The denial of human nature, and the insistence on the blank slate, has been the central fact of modern thought. The dominance of this view has stalled scientific investigation into human nature for most of the 20th century. Its end will change everything about the modern intellectual world.

Acknowledgements:

We should like to thank Donald Brown, David Buss, Martin Daly, Irven DeVore, Steven Pinker, Roger Shepard, Dan Sperber, Donald Symons, and Margo Wilson for many fruitful discussions of these issuesWe are grateful to the James S. McDonnell Foundation Grant to John Tooby, NSF Grant BNS9157-499 to John Tooby, and a Research Across Disciplines grant from Office of Research, the University of California, Santa Barbara, for their financial support during the preparation of this book.

Futher reading:

Barkow, J., Cosmides, L. and Tooby, J. 1992. *The Adapted Mind: Evolutionary psychology and the generation of culture*. NY: Oxford University Press.
Buss, D. (1994). *The evolution of desire*. NY: Academic Press.
Buss, D. (Ed.) (2005). *The Handbook of Evolutionary Psychology*. Hoboken, NJ: Wiley.
Daly, M.& M. Wilson 1988. *Homicide*. Aldine de Gruyter.
Dawkins, R. 1986. *The Blind Watchmaker*. NY: Norton.
Gallistel, C. R. (1990). *The Organization of Learning*. Cambridge: MIT Press.
Pinker, S. 1994. *The Language Instinct*. NY: Morrow.
Pinker, S. 1997. *How the Mind Works*. NY: Norton.
Symons, D. 1979. The Evolution of Human Sexuality. NY: Oxford.
Tooby, J. & Cosmides, L. (in press) *Evolutionary psychology: Foundational Papers*. with a forward by Steven Pinker. MIT Press: Cambridge
Williams, G. 1966. *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.

Footnotes

On the Standard Social Science Model, see "Psychological Foundations of Culture", by Tooby and Cosmides, in *The Adapted Mind.* For a start into the large literature on human mate preferences, see Symons, 1979; Buss, 1994; and Sugiyama, 2005 in *The Handbook of Evolutionary Psychology.*. On evolution and color vision, see R. Shepard (1987) in *The Latest on the Best* (Cambridge, MA: MIT Press). On domain-specific adaptations in infants and toddlers, see L. Hirschfeld & S. Gelman, Eds., (1994), *Mapping the Mind: Domain-specificity in cognition and culture* (NY: Cambridge University Press). On mindreading, see S. Baron-Cohen (1995), *Mindblindness*, Cambridge, MA: MIT Press. On the pros and cons of "rational methods" of reasoning, see G. Gigerenzer, P. Todd, and the ABC group (1999), *Simple Heuristics that Make Us Smart* (NY: Oxford). On hallmarks of instincts, see Pinker, 1994. On "Our modern skulls house a stone age mind": We thank William Allman for suggesting this phrase, which is a very apt summary of our position. On natural selection as the only source of functional organization, see R. Dawkins (1986), *The Blind Watchmaker*, and G. Williams (1966). On reliability, efficiency, and economy as design criteria for adaptations, see G. Williams (1966). Engineering bats: Dawkins, *The Blind Watchmaker*. On the programs that cause classical and operant conditioning, and the inadequacy of associationist theories, see Gallistel, 1990, *The Organization of Learning*. On sexual recombination and universality of genetic basis for complex adaptations: Tooby & Cosmides (1990), "On the universality of human nature and the uniqueness of the individual: The role of genetics and adaptation." *Journal of Personality, 58*: 17-67. Game theorists model social exchange as a repeated Prisoners' Dilemma; e.g., R. Axelrod & W. D. Hamilton, (1981) *Science, 211*, 1390-1396; R. Boyd (1988) *Ethology and Sociobiology, 9*, 211-222; R. Trivers (1971) *Quarterly Review of Biology, 46*, 35-57. On review of evidence for social contract algorithms, see Cosmides & Tooby (1992) in *The Adapted Mind* and Cosmides & Tooby (1997) in : *Characterizing human psychological adaptations* (Ciba Foundation Symposium #208), Chichester: Wiley. On the complex relationships between evolved psycholgy and culture, see "Psychological Foundations of Culture", by Tooby and Cosmides, in *The Adapted Mind.*