

*The Blank Slate, the Noble Savage, and the
Ghost in the Machine*

STEVEN PINKER

THE TANNER LECTURES ON HUMAN VALUES

Delivered at

Yale University
April 20 and 21, 1999

STEVEN PINKER is professor in the department of brain and cognitive sciences and director of the McConnell-Pew Center for Cognitive Neuroscience at the Massachusetts Institute of Technology. He was educated at McGill University and at Harvard University, where he received his Ph.D. He was recently elected to the American Academy of Arts and Sciences, and is a fellow of the American Association for the Advancement of Science, the American Psychological Association, and the American Psychological Society. His publications include *Learnability and Cognition: The Acquisition of Argument Structure* (1989) and *The Language Instinct* (1994), which was named one of the ten best books of 1994 by the *New York Times*, the *London Times*, and the *Boston Globe*. His most recent book, *How the Mind Works* (1997), won the *Los Angeles Times* Book Prize in Science and the William James Book Prize from the APA, and was a finalist for the Pulitzer Prize and the National Book Critics Circle Award.

These are extraordinary times in the history of human knowledge. For hundreds of years the progress of science has been a story of increasing unification and coherence, which the biologist E. O. Wilson has recently termed *consilience*, literally “jumping together.”¹

In 1755, Samuel Johnson wrote that his *Dictionary* should not be expected to “change sublunary nature, and clear the world at once from folly, vanity, and affectation.” Few people today understand his use of the word “sublunary,” literally “below the moon.” It was an allusion to the ancient belief that there was a strict division between the pristine, lawful, unchanging cosmos above and our grubby, chaotic earth below. The division was already obsolete when Johnson wrote; Newton had shown that a single set of laws described the forces pulling the apple toward the ground and keeping the moon in its orbit around the earth.

The collapse of the wall between the terrestrial and the celestial was followed by a collapse of the once equally firm (and now equally forgotten) wall between the creative past and the static present. Charles Lyell showed that today’s earth was sculpted by everyday erosion, earthquakes, and volcanos acting in the past over immense spans of time. The living and nonliving, too, no longer occupy different realms. William Harvey showed that the human body is a machine that runs by hydraulics and other mechanical principles. Friedrich Wöhler showed that the stuff of life is not a magical, quivering gel but ordinary compounds following the laws of chemistry. Darwin showed how the astonishing diversity of life and its ubiquitous signs of good design could arise from the physical process of natural selection among replicators. Mendel,

Preparation of this paper was supported by NIH grant HD 18381.

¹ E. O. Wilson, *Consilience* (New York: Knopf, 1998). See also J. Tooby and L. Cosmides, “Psychological Foundations of Culture,” in *The Adapted Mind*, ed. J. Barkow, L. Cosmides, and J. Tooby (New York: Oxford University Press, 1992).

and then Watson and Crick, showed how replication itself could be understood in physical terms.

But one enormous chasm remains in the landscape of human knowledge. Biology versus culture, nature versus society, matter versus mind, and the sciences versus the arts and humanities survive as respectable dichotomies long after the other walls dividing human understanding have tumbled down.

But perhaps not for long. Four new fields are laying a bridge between nature and society in the form of a scientific understanding of mind and human nature.

The first is cognitive science. Many thinkers believe there is a fundamental divide between human behavior and other physical events. Whereas physical behavior has *causes*, they say, human behavior has *reasons*. Consider how we explain an everyday act of behavior, such as Bill getting on a bus. No one would invoke some physical push or pull like magnetism or a gust of wind, nor would anyone need to put Bill's head in a brain scanner or test his blood or DNA. The most perspicuous explanation of Bill's behavior appeals instead to his beliefs and desires, such as that Bill wanted to visit his grandmother and that he knew the bus would take him there. No explanation has as much predictive power as that one. If Bill hated the sight of his grandmother, or if he knew the route had changed, his body would not be on that bus.

For centuries the gap between physical events, on the one hand, and meaning, content, ideas, reasons, or goals, on the other, has been seen as a boundary line between two fundamentally different kinds of explanation. But in the 1950s, the "cognitive revolution" unified psychology, linguistics, computer science, and philosophy of mind with the help of a powerful new idea: that mental life could be explained in physical terms via the notions of information, computation, and feedback. To put it crudely: Beliefs and memories are information, residing in patterns of activity and structure in the brain. Thinking and planning are sequences

of transformations of these patterns. Wanting and trying are goal states that govern the transformations via feedback from the world about the discrepancy between the goal state and the current situation, which the transformations are designed to reduce.² This general idea, which may be called the computational theory of mind, also explains how *intelligence* and *rationality* can arise from a mere physical process. If the transformations mirror laws of logic, probability, or cause and effect in the world, they will generate correct predictions from valid information in pursuit of goals, which is a pretty good definition of the term “intelligence.”

The second science bridging mind and matter is neuroscience, especially cognitive neuroscience, the study of the neural bases of thinking, perception, and emotion. Our traditional and most familiar conception of the mind is based on the soul: an immaterial entity that enters the fertilized egg at conception, reads the instrument panels of the senses and pushes the buttons of behavior, and leaks out at death. Neuroscience is replacing that conception with what Francis Crick has called the astonishing hypothesis: that all aspects of human thought and feeling are manifestations of the physiological activity of the brain. In other words, the mind is what the brain does, in particular, the information-processing that it does.³

Astonishing though the hypothesis may be, the evidence is now overwhelming that it is true. Many cause-and-effect linkages have a physical event on one side and a mental event on the other. If an electrical current is sent into the brain by a surgeon, the brain's owner is caused to have a vivid, lifelike experience. A host of

² S. Pinker, *How the Mind Works* (New York: W. W. Norton, 1997); H. Gardner, *The Mind's New Science: A History of the Cognitive Revolution* (New York: Basic Books, 1987); J. A. Fodor, *The Elm and the Expert* (Cambridge, Mass.: MIT Press, 1994).

³ F. Crick, *The Astonishing Hypothesis: The Scientific Search for the Soul* (New York: Simon & Schuster, 1994); M. S. Gazzaniga, ed., *The New Cognitive Neurosciences* (Cambridge, Mass.: MIT Press, in press); M. S. Gazzaniga, R. B. Ivry, and G. R. Mangun, *Cognitive Neuroscience: The Biology of the Mind* (New York: W. W. Norton, 1998).

chemicals can find their way to the brain from the stomach, lungs, or veins and change a person's perception, mood, personality, and thoughts. When a patch of brain tissue dies because of trauma, poisoning, infection, or lack of oxygen, a part of the person is gone: he or she may think, feel, or act so differently as to become quite literally "a different person." Every form of mental activity—every emotion, every thought, every perception—gives off electrical, magnetic, or metabolic signals that are being read with increasing precision and sensitivity by new technologies such as positron emission tomography, functional magnetic resonance imaging, electroencephalography, and magnetoencephalography. When a surgeon takes a knife and cuts the corpus callosum (which joins the two cerebral hemispheres), the mind is split in two and in some sense the body is inhabited by two selves. Under the microscope, the tissues of the brain show a breathtaking degree of complexity—perhaps a hundred trillion synapses—that is fully commensurate with the breathtaking complexity of human thought and experience. And when the brain dies, the person goes out of existence. It is a significant empirical discovery that no one has found a way to communicate with the dead.

The third bridging discipline is behavioral genetics. All the potential for complex learning and feeling that distinguishes humans from other animals lies in the genetic material of the fertilized ovum. We are coming to appreciate that the species-wide design of the human intellect and personality and many of the details that distinguish one person from another have important genetic roots. Studies show that monozygotic (identical) twins separated at birth, who share their genes but not their family or community environments, are remarkably alike in their intelligence, personality traits, attitudes toward a variety of subjects (such as the death penalty and modern music), and personal quirks such as dipping buttered toast in coffee or wading into the ocean backward. Similar conclusions come from the discovery that

monozygotic twins are far more similar than dizygotic (fraternal) twins, who share only half their genes, and from the discovery that biological siblings of any kind are far more similar than adoptive siblings. The past few years have also seen the discovery of genetic markers, genes, and sometimes gene products for aspects of intelligence, spatial cognition, the control of speech, and personality traits such as sensation-seeking and excess anxiety.⁴

The fourth bridging science is evolutionary psychology, the study of the phylogenetic history and adaptive functions of the mind. Evolutionary psychology holds out the hope of understanding the *design* or *purpose* of the mind, not in some mystical or teleological sense, but in the sense of the appearance of design or illusion of engineering that is ubiquitous in the natural world (such as in the eye or the heart) and that Darwin explained by the theory of natural selection.⁵

Though there are many controversies within biology, what is not controversial is that the theory of natural selection is indispensable to make sense of a complex organ such as the eye. The eye's precision engineering for the function of forming an image could not be the result of some massive coincidence in tissue formation like the appearance of a wart or tumor or of the random sampling of genes that can lead to simpler traits. And the human

⁴ T. J. Bouchard, Jr., "Genes, Environment, and Personality," *Science* 264: 1700–1701; D. H. Hamer and P. Copeland, *Living with Our Genes: Why They Matter More Than You Think* (New York: Doubleday, 1998); S. E. Fisher, F. Vargha-Khadem, K. E. Watkins, A. P. Monaco, and M. E. Pembrey, "Localisation of a Gene Implicated in a Severe Speech and Language Disorder," *Nature Genetics* 18: 168–70; J. M. Frangiskakis, A. K. Ewart, A. C. Morris, C. B. Mervis, J. Bertrand, B. F. Robinson, B. P. Klein, G. J. Ensing, L. A. Everett, E. D. Green, C. Proschel, N. J. Gutowski, M. Noble, D. L. Atkinson, S. J. Odelberg, and M. T. Keating, "LIM-Kinase1 Hemizygoty Implicated in Impaired Visuospatial Constructive Cognition," *Cell* 86 (1996): 59–69; R. Plomin, J. C. Defries, G. E. McClearn, and M. Rutter, *Behavioral Genetics* 3d ed. (New York: W. H. Freeman, 1997).

⁵ J. H. Barkow, L. Cosmides, and J. Tooby, eds., *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (New York: Oxford University Press, 1997); S. Pinker, *How the Mind Works* (New York: Norton, 1997); D. Buss, *Evolutionary Psychology: The New Science of the Mind* (New York: Allyn & Bacon, 1999).

eye's similarity to the eyes of other organisms, including many arbitrary and quirky design features, could not be the handiwork of some cosmic designer.⁶

Evolutionary psychology extends this kind of argument to another part of the body. For all its exquisite natural engineering, the eye is useless without the brain. The eye is an organ of information processing; it does not dump its signals into some empty chasm, but connects to complicated neural circuits that extract information about the depths, colors, motions, and shapes of objects and surfaces in the world. All this analysis of the visual world would itself be useless unless it fed into higher circuits for categorization: the ability to make sense of experience, to impute causes to events, and to remember things in terms of useful predictive categories. And in turn, categorization would be useless unless it operated in the service of the person's goals, which are set by motives and emotions such as hunger, fear, love, curiosity, and the pursuit of status. Those are the motives that tend to foster survival and reproduction in the kinds of environments in which our ancestors evolved.

Beginning with the eye, we have a chain of causation that leads to faculties, or modules, or subsystems of mind, each of which can be seen as an adaptation akin to the adaptations in the organs of the body. Recent research has shown that aspects of the psyche that were previously considered mysterious, quirky, and inexplicable, such as fears and phobias, an eye for beauty, family dynamics, romantic love, and a passionate desire for revenge in defense of honor, have a systematic evolutionary logic when analyzed like other biological systems, organs, and tissues.⁷

Cognitive science, neuroscience, behavioral genetics, and evolutionary psychology are doing nothing less than providing a sci-

⁶ G. C. Williams, *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought* (Princeton, N.J.: Princeton University Press, 1966); R. Dawkins, *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe without Design* (New York: Norton, 1986).

⁷ Pinker, *How the Mind Works*.

entific understanding of the mind and human nature. It is important to note that this understanding is not an alternative to more traditional explanations in terms of learning, experience, culture, and socialization. Rather, it aims at an explanation of how those processes are possible to begin with. Culture is not some gas or force field or bacterial swarm that surrounds humans and insidiously seeps into or infects them. Culture has its effects because of mental algorithms that accomplish the feat we call learning. And learning can be powerful and useful only if it is designed to work in certain ways. Both a parrot and a human child can learn something when exposed to speech, but only the child is equipped with an algorithm for learning vocabulary and grammar that can extract words and rules from the speech wave and use them to generate an unlimited number of meaningful new sentences. The search for mechanisms of learning animates each of the four new sciences.

A chief goal of cognitive science is to identify the learning algorithms that underlie language and other cognitive feats.⁸ Similarly, a major goal of neuroscience arises from the realization that all mental activity, including learning, arises from the neurophysiology and neuroanatomy of the brain: when people learn, neural tissue must change in some way as the result of experience. The phenomenon is called neural plasticity, and it is currently being explored intensively within neuroscience. Behavioral genetics, too, is not aimed at documenting an exclusively genetic control of behavior. In most studies, only around half of the variance in intellectual or personality traits has been found to correlate with the genes; the other half comes from environmental or random factors. Behavioral genetics, by allowing us to subtract the resemblances between parents and children that are due to their genetic relatedness, and to partition the remaining causes into those operating within the family (such as the correlations between adoptive siblings reared together) and those outside the family (such as the

⁸ S. Pinker, *Language Learnability and Language Development* (Cambridge, Mass.: Harvard University Press, 1984/1996).

lack of a perfect correlation between identical twins reared together), is essential to our understanding of the nature of the socialization process. Finally, according to evolutionary psychology human beings are not robotic automata or bundles of knee-jerk reflexes. Mental adaptations are what biologists call *facultative* adaptations: a crucial part of their design is to sense environmental variation and adjust to find the optimum behavioral strategy.

How will these new sciences bridge the gaps in human knowledge that I alluded to at the outset, completing the consilience that we have enjoyed so long in the physical sciences? The emerging picture is that our genetic program grows a brain endowed with emotions and with learning abilities that were favored by natural selection. The arts, humanities, and social sciences, then, can be seen as the study of the products of certain faculties of the human brain. These faculties include language, perceptual analyzers and their esthetic reactions, reasoning, a moral sense, love, loyalty, rivalry, status, feelings toward allies and kin, an obsession with themes of life and death, and many others. As human beings share their discoveries and accumulate them over time, and as they institute conventions and rules to coordinate their often conflicting desires, the phenomena we call “culture” arise. Given this continuous causal chain from biology to culture through psychology, a fundamental division between the humanities and sciences has become as obsolete as the division between the sublunary and supralunary spheres.

Does this picture deserve the dreaded academic epithet “reductionism”? Not in the bad, indeed, idiotic sense of trying to explain World War I in terms of subatomic particles. It *is* reductionist in the good sense of aiming for the deep and uniquely satisfying understanding we have enjoyed from the unification of sciences such as biology, chemistry, and physics. The goal is not to eliminate explanations at higher levels of analysis but to connect them lawfully to more fundamental levels. The elementary processes at one

level can be explained in terms of more complicated interactions one level down.

Not everyone, needless to say, is enthralled by the prospect of unifying biology and culture through a science of mind and human nature. There have been furious objections from many quarters, particularly the academic left and the religious and cultural right. When E. O. Wilson and other “sociobiologists” first outlined a vision of a science of human nature in the 1970s and 1980s, critics expressed their reservations by dousing him with ice water at an academic conference, protesting his appearances with pickets, bullhorns, and posters urging people to bring noisemakers to his lectures, and angry manifestoes with accusations of racism, sexism, class oppression, genocide, and the inevitable comparison to the Nazis.⁹ In their popular book *Not in Our Genes*, three prominent scientists, Richard Lewontin, Steven Rose, and Leon Kamin, felt justified in the use of nonstandard forms of scientific argumentation such as doctoring quotations and dropping innuendoes about their opponents’ sex lives. When the psychologist Paul Ekman announced at an anthropology conference his discovery that facial expressions of basic emotions are the same the world over, he was shouted down and called a fascist and racist.¹⁰ Though the worst of the hysteria has died down, ad hominem arguments and smears of racism and sexism are not uncommon in both academic and popular discussions of behavioral genetics and evolutionary psychology.

Alarms have been sounded not just by tenured radicals and commissars of political correctness. In a highly publicized article entitled “Sorry, But Your Soul Just Died,” the left-lamprooning author Tom Wolfe discusses the prospects of the new understanding

⁹ See R. Wright, *The Moral Animal* (New York: Pantheon, 1994); E. O. Wilson, *Naturalist* (Washington, D.C.: Island Press, 1994).

¹⁰ For documentation, see Pinker, *How the Mind Works*, pp. 45, 569n45, 366, 580n366.

of mind, brain, genes, and evolution with a mixture of admiration and dread. He predicts:

. . . in the year 2006 or 2026, some new Nietzsche will step forward to announce . . . “The soul is dead.” He will say that he is merely bringing the news, the news of the greatest event of the millennium: “The soul, that last refuge of values, is dead, because educated people no longer believe it exists.” . . . Unless the assurances of [E. O. Wilson and his allies] also start rippling out, the lurid carnival that will ensue may make [Nietzsche’s] phrase “the total eclipse of all values” seem tame.¹¹

Farther to the right, the journalist Andrew Ferguson, writing in the neoconservative magazine *Weekly Standard*, is far less ambivalent. He reviewed a recent book by Francis Fukuyama, which argued that civility and social institutions always reassert themselves because of aspects of human nature recently revealed by the new sciences. The book “is sure to give you the creeps,” Ferguson wrote, because “Whether [a] behavior is moral, whether it signifies virtue, is a judgment that the new science, and materialism in general, cannot make.”¹² In another book review he characterizes the new sciences as saying that people are nothing but “meat puppets,” in contrast to the traditional Judeo-Christian view in which “human beings were persons from the start, endowed with a soul, created by God, and infinitely precious. And this is the common understanding . . . the new science . . . means to undo.”¹³

Clearly the new sciences of mind are widely seen as threatening, almost in the manner of a religious heresy. For observers such as Ferguson it is literally the religious doctrine of the immaterial

¹¹ *Forbes* magazine, 1996.

¹² A. Ferguson, “The End of Nature and the Next Man,” *Weekly Standard*, 1999.

¹³ A. Ferguson, “How Steven Pinker’s Mind Works,” *Weekly Standard*, January 12, 1998.

soul that he sees as threatened. For others it is a modern secular religion, which John Tooby and Leda Cosmides have called the Standard Social Science Model or SSSM.¹⁴

The ascendancy of the SSSM is a key event in modern intellectual history that began in the first decades of the twentieth century and was firmly entrenched by the 1950s.¹⁵ The model embraces three beliefs, which give me the title of this paper.

The first is John Locke's doctrine of the *tabula rasa*, the Blank Slate: that the human mind is infinitely plastic, with all its structure coming from reinforcement and socialization. Here are two of the twentieth century's earliest and most vociferous defenders of the Blank Slate, the psychologist John B. Watson and the anthropologist Margaret Mead:

Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select—doctor, lawyer, artist, merchant-chief, and yes, even beggar-man and thief, regardless of his talents, penchants, tendencies, abilities, vocations, and race of his ancestors. (John B. Watson, *Behaviorism*, 1925)

We are forced to conclude that human nature is almost unbelievably malleable, responding accurately and contrastingly to contrasting cultural conditions. . . . The members of either or both sexes may, with more or less success in the case of different individuals, be educated to approximate [any temperament]. (Margaret Mead, *Sex and Temperament in Three Primitive Societies*, 1935)

The second belief is Jean-Jacques Rousseau's doctrine of the Noble Savage: that evil comes not from human nature but from

¹⁴ Tooby and Cosmides, "Psychological Foundations."

¹⁵ For an excellent history, see C. N. Degler, *In Search of Human Nature: The Decline and Revival of Darwinism in American Social Thought* (New York: Oxford University Press, 1991).

our social institutions. We find the doctrine today in a particularly pure form in the “Seville Statement” of 1986, in which twenty social scientists, with the endorsement of UNESCO and several academic societies, declared that it is “scientifically incorrect” to say that “we have inherited a tendency to make war from our animal ancestors,” that “war or any other violent behavior is genetically programmed into our human nature,” that humans have a “violent brain,” or that war is caused by “instinct.” We see the doctrine as well in the popular image of native peoples living in peaceful co-existence with the ecosystem and with one another.¹⁶

The third doctrine is what Gilbert Ryle called the Ghost in the Machine: the belief that we are separate from biology, free to choose our actions and define meaning, value, and purpose. As Wolfe puts it,

Meantime, the notion of a self—a self who exercises self-discipline, postpones gratification, curbs the sexual appetite, stops short of aggression and criminal behavior—a self who can become more intelligent and lift itself to the very peaks of life by its own bootstraps through study, practice, perseverance, and refusal to give up in the face of great odds—this old-fashioned notion (what’s a boot strap, for God’s sake?) of success through enterprise and true grit is already slipping away, slipping away . . . slipping away . . .

. . . Where does that leave self-control? Where, indeed, if people believe this ghostly self does not even exist, and brain imaging proves it, once and for all?

Similarly, Ferguson writes that the scientific belief that our minds arise from neural activity

runs counter to the most elemental belief every person has about himself. . . . Beyond this, however, the “scientific belief”

¹⁶ “The Seville Statement on Violence,” *American Psychologist* 45 (1990): 1167–68.

would also appear to be corrosive of any notion of free will, personal responsibility, or universal morality. . . .

The old myth of natural law had a means for making moral judgments, of course. But it took as fundamental the very concepts that the new science wants to render meaningless—that human beings are endowed with souls, for example.

At first it would seem that the Ghost in the Machine would be chained to religious thought; that secular thinkers would have nothing to do with an immaterial soul. Those who deny the existence of human nature would attribute behavior instead to the cumulative effects of socialization and conditioning. But in fact it is common for believers in the Standard Social Science Model to invoke an “I,” a “we,” a “you,” or a “person” that somehow floats free of genetics, neurobiology, or evolution and can act as it pleases, constrained only by current environmental circumstances. Rose, an ardent foe of evolutionary psychology and behavioral genetics, repeatedly declares in a recent book that “we have the ability to construct our own futures, albeit not in circumstances of our own choosing”; the statement is intended as a refutation of “reductionist” biology.¹⁷ But he never explains who the “we” is, if not highly structured neural circuits, which must get that structure in part by genes and evolution.

The Standard Social Science Model arose in part as a legitimate backlash against many deplorable events of the nineteenth and the first half of the twentieth century. These include pseudoscientific doctrines of racial and ethnic inferiority, coercive eugenic policies, the oppression of women, the maltreatment and neglect of children, the theory of “Social Darwinism,” which tried to justify inequality and conquest as part of the wisdom of nature, policies of

¹⁷ S. Rose, *Lifelines: Biology beyond Determinism* (New York: Oxford University Press, 1998).

racial discrimination, and outright genocide.¹⁸ As a result, the precepts of the SSSM, which would appear to undermine any ideology that could permit such horrors, have acquired a moral authority and are felt to be the foundation for political and ethical decency.

The precepts, however, are factual claims, many of which are being refuted. Does this mean that we are forced to return to repugnant doctrines and horrific practices? My aim in this lecture is to convince you that the answer is no: the supposed conflict between the new sciences of mind and human nature and our ethical values is misconstrued. The habit of basing equality, dignity, and human rights on the doctrines of the Blank Slate, the Noble Savage, and the Ghost in the Machine is a product of fuzzy thinking about both ethics and science, and we can bid these doctrines good riddance without compromising human values at all. I will try to allay four fears that have surrounded the prospects of a science of mind and human nature.

The first fear is of the possibility of biological differences. If the mind has an innate structure, the worry goes, different people (or different classes, sexes, and races) could have different innate structures, and that would justify discrimination and oppression. But if there were no innate structure, there could be, by definition, no individual or group differences in innate structure, and thus no basis for discrimination. Therefore, according to this moral argument for the SSSM, there is no human nature. The argument, however, is fallacious both empirically and morally.

The empirical problem is that discoveries about a universal human nature—the bread and butter of cognitive science and evolutionary psychology—do not imply innate differences between individuals, groups, or races. Any page of *Gray's Anatomy* will show a complex design of systems, organs, and tissues that are qualitatively alike in every normal human being (though of course

¹⁸ Degler, *In Search of Human Nature*.

with numerous quantitative differences in size and shape). The same is likely to be true of the mental equivalents of systems, organs, and tissues. This is not because it would be nice if it were true, but because of particular properties of the forces that shaped human nature. Sexual recombination and natural selection (which adapts organisms to an environmental niche by weeding out variants that are less fit) are homogenizing forces, making the members of a species qualitatively alike.¹⁹ And in the case of humans, the racial divisions in the family tree probably opened up only recently and are constantly being bridged by the fact that humans migrate and interbreed with gusto, which has resulted in a steady shuffling of genes across racial groups for tens of thousands of years. It is therefore unlikely that individuals or races differ qualitatively in any mental faculty, and indeed the striking universals in language, emotions, and cognitive categories that emerge from the ethnographic record suggest that in fact the differences are small to nonexistent.²⁰ All this means that research on human nature does not necessarily lead to invidious assertions about the mental traits of specific people or groups.

But of course there *could* be genetic variation, most likely quantitative, among people and races; it would be absurd to declare this outcome impossible a priori just because it would be uncomfortable if true. If such variation were discovered, what would follow? Would discrimination or oppression be justified? Of course not! Discrimination against an individual on the basis of the person's sex, race, ethnicity, or sexual orientation is simply wrong on moral grounds. Now, conceivably someone could argue on grounds of economic efficiency that a rational agent ought to factor in group statistics in making a decision about an individual (say, whether the person should be admitted to a university, or be

¹⁹ J. Tooby and L. Cosmides, "On the Universality of Human Nature and the Uniqueness of the Individual: The Role of Genetics and Adaptation," *Journal of Personality* 58 (1990): 17–67.

²⁰ D. E. Brown, *Human Universals* (New York: McGraw-Hill, 1991).

released on parole), because that is what standard Bayesian inference demands: prior probabilities (such as the probability that people of a given race or sex will succeed in school or commit a crime) should influence the estimate of posterior probabilities (whether a *given* individual will succeed in school or commit a crime). But I think most people would agree that the rights of an individual to be considered on his or her individual merits and accomplishments trump any gain in overall accuracy of decision-making based on the use of the statistics of races, ethnic groups, or genders. It is one of many cases in which we willingly sacrifice a modicum of freedom and economic efficiency for a more general good. (Other examples include laws safeguarding individuals' privacy and those that outlaw the voluntary sale of one's vote, one's organs, or one's freedom.)

Crucially, the moral argument against discrimination can be made *regardless* of the existence or nonexistence of any empirically discovered biological differences among people or groups. And that is surely the way we want it. Is there *any* conceivable finding on group differences in any trait that would undermine our belief in the evil of racial or sexual discrimination against an individual? If not, we should not fear the study of human nature just because it may stumble upon some innate difference.

This is especially important to keep in mind when it comes to possible differences between men and women. When it comes to the sexes, the *Gray's Anatomy* argument breaks down. The difference between male and female anatomy is a vivid illustration that there *can* be important biological differences between members of the human race. Though some writers still insist that all sex differences are products of sexism or socialization practices, the argument requires standards of evidence worthy of the tobacco industry. The more honest of these writers admit that they are motivated by a fear that the discovery of any biologically influenced sex differences will compromise the ideals of feminism or gender equality. (Thus we have the strange situation in which some writ-

ers, under the banner of a dubious form of feminism, argue that women are identical to men in terms of their inherent propensity for promiscuity, infidelity, taste for pornography, and violence.) The assumption appears to be that *fairness* requires *sameness*, and that is absurd. Whether or not males are identical to females in some or all psychological traits, it is intolerable for public institutions to discriminate against individual men and women on the basis of their sex. We can all agree with Gloria Steinem when she said, “There are really not many jobs that actually require a penis or a vagina, and all the other occupations should be open to everyone.”

The second fear is the possibility of evil instincts. The unstated assumption is that if deplorable behavior such as aggression, war, rape, clannishness, exploitation, xenophobia, and the pursuit of status and wealth is innate, that would make them “natural” and hence good. And even if we agree that they are not good, they are “in the genes” and therefore cannot be changed, so attempts at social reform are futile. Aggression is objectionable, and social reform is desirable; therefore, the argument seems to go, *Homo sapiens* must be a bunch of nice guys. Only “society” is at fault.

The lunatic version of this argument is, of course, the Seville Statement, with its fiat that all claims about biological propensities toward dominance, violence, and war are “scientifically incorrect.” The signatories were at least clear about their motives. They alleged that the “incorrect” statements “have been used, even by some in our disciplines, to justify violence and war” (they gave no examples) and concluded that “biology does not condemn humanity to war, and that humanity can be freed from the bondage of biological pessimism and empowered with confidence to undertake the transformative tasks needed in the International Year of Peace and in the years to come.”

The Seville Statement is a textbook example of what the philosopher G. E. Moore called the Naturalistic Fallacy: that whatever is found in nature is morally right. In this case, the fallacy is that if people are prone to violence, that would make it justifiable.

Hence the signatories' decision to legislate empirical claims about people's natural propensities was, in their minds, a tactic to bring about peace. Apparently it was inconceivable to these leading social scientists that there could be selection for violent behavior *and* that violent behavior is morally unjustifiable. Their manifesto is especially egregious because the legislated factual claims are a blatant kind of disinformation.

The notion that the human brain houses no inherent tendency to use violence, and that violence is an artifact of some particular culture at a particular time, has to confront an obvious fact about human history. Winston Churchill wrote, "The story of the human race is war. Except for brief and precarious interludes there has never been peace in the world; and long before history began murderous strife was universal and unending." Or as one biologist put it, "*Homo sapiens* is a nasty business."

For many years intellectuals tried to deny the significance of history with two myths. One is the myth of the peaceful savage, where "savages" or hunter-gatherers are thought to be representative of a human nature uncorrupted by the malign influences of civilization. According to this myth, among preagricultural peoples war is rare, mild, and ritualized, or at least it used to be before contact with Westerners. Recent books by anthropologists, biologists, and historians who have examined the factual record, such as Napoleon Chagnon, Richard Keeley, Jared Diamond, Martin Daly and Margo Wilson, Richard Wrangham, and Michael Ghiglieri, have shown that this is romantic nonsense; war has always been hell.²¹

It is not uncommon among preagricultural peoples for a third

²¹ N. A. Chagnon, "Life Histories, Blood Revenge, and Warfare in a Tribal Population," *Science* 239 (1988): 985–92; J. Diamond, *Guns, Germs, and Steel* (New York: Norton, 1997); L. H. Keeley, *War before Civilization: The Myth of the Peaceful Savage* (New York: Oxford University Press, 1996); M. Daly and M. Wilson, *Homicide* (Hawthorne, N.Y.: Aldine de Gruyter, 1988); R. Wrangham and D. Peterson, *Demonic Males* (n.p., 1996); M. Ghiglieri, *The Dark Side of Man: Tracing the Origins of Violence* (New York: Perseus Books, 1999).

of the men to die at the hands of other men, and for almost half of the men to have killed someone. As compared to modern warfare, in primitive warfare mobilization is more complete, battles are more frequent, casualties are proportionally higher, prisoners are fewer, and the weapons are more damaging. Even in the more peaceable hunter-gatherer societies such as the !Kung San of the Kalahari desert, the murder rate is similar to that found in modern American urban jungles such as Detroit. In his survey of human universals gleaned from the ethnographic record, the anthropologist Donald Brown includes violent conflict, rape, envy, sexual jealousy, and in-group/out-group conflicts as traits documented in all cultures.²²

A related romantic myth is the harmony and wisdom of nature. Many intellectuals still believe that animals kill only for food, that among animals war is unknown, and that, in the words of the Seville Statement, dominance hierarchies are a form of bonding and affiliation that benefits the group. The reality was summed up by Darwin: “What a book a devil’s chaplain might write on the clumsy, wasteful, blundering, low, and horribly cruel works of nature!” The most chilling example is the one closest to home. The primatologists Jane Goodall and Richard Wrangham have documented behavior in our closest relatives, the chimpanzees, that would surely be called genocide if it had been observed in humans. In evolutionary terms, killing a member of one’s own species is hardly an anomalous or puzzling event. As Daly and Wilson have pointed out, “Killing one’s antagonist is the ultimate conflict resolution technique, and our ancestors discovered it long before they were people.”²³

Is any of this a “justification” for war or other violent conflict? Obviously not. As we used to say in the 1960s, war is not healthy for children or other living things. Nothing about the behavior of

²² Brown, *Human Universals*.

²³ Daly and Wilson, *Homicide*, p. ix.

hunger-gatherers or primates could conceivably push us from abhorring war and trying to eliminate it.

But is war nonetheless inevitable, making attempts to prevent it fruitless? Here too the answer is no. The human mind is a complex system with many parts. One may be an urge to neutralize rivals by any means necessary. But another is a calculator that can come to the realization that conflict has terrible costs and that everyone can come out ahead by dividing up the surplus that results from laying down arms. According to Brown's survey, what is *also* universal across human societies is the deploring of conflict, violence, rape, and murder and the use of mechanisms to reduce them, including laws, punishment, redress, and mediation.

And another obvious empirical fact is that the human condition can improve. For all the horrors of the past few centuries, they have seen the disappearance of war, slavery, conquest, blood feuds, despotism, the ownership of women, apartheid, fascism, and Leninism from vast swaths of the earth that had known them for decades, centuries, or millennia. Even at their worst, the homicide rates in American cities were twenty times lower than those measured among many foraging peoples. Modern Britons are twenty times less likely to be murdered than their medieval counterparts.²⁴

There are many reasons that war and aggression can decline despite a constancy of human nature. They include a knowledge of the lessons of history and the use of face-saving measures, mediation, contracts, deterrence, equal opportunity, a court system, enforceable laws, monogamy, and limits on perceived inequality. These are humble, time-tested methods that *acknowledge* human nature and its dark side. They are likely to continue to be more humane and effective than attempts to re-engineer culture and redesign human nature, as we are reminded by the recent history of the Soviet empire, Cambodia, and mainland China. Indeed, the strongest argument against totalitarianism may be recognition of

²⁴ Daly and Wilson, *Homicide*.

a universal human nature: that all humans have innate desires for life, liberty, and the pursuit of happiness. The doctrine of the Blank Slate, which justifies the dismissal of people's stated wants as an artifact of a particular time and place and thereby licenses the top-down redesign of society, is a totalitarian's dream.

The third fear aroused by a science of human nature is the dissolution of free will and the resulting universal abdication of responsibility. If behavior is a physical consequence of ricocheting molecules in the brain shaped in part by genes that were put into place by natural selection, where is the "person" whom we hold responsible for his or her actions? If the rapist is following a biological imperative to spread his genes, the worry goes, it's a short step to saying that it's not his fault. This worry has been stated by voices of both the left and the right. When E. O. Wilson suggested that humans resemble most other mammals in the great male desire for multiple sexual partners, Rose accused him of really saying, "Don't blame your mates for sleeping around, ladies, it's not their fault they are genetically programmed."²⁵ Tom Wolfe writes in a similar vein (though with tongue partly in cheek):

The male of the human species is genetically hardwired to be polygamous, i.e., unfaithful to his legal mate. Any magazine-reading male gets the picture soon enough. (Three million years of evolution made me do it!) Women lust after male celebrities, because they are genetically hardwired to sense that alpha males will take better care of their offspring. (I'm just a lifeguard in the gene pool, honey.) Teenage girls are genetically hardwired to be promiscuous and are as helpless to stop themselves as dogs in the park. (The school provides the condoms.) Most murders are the result of genetically hardwired compulsions. (Convicts can read, too, and they report to the prison psychiatrist: "Something came over me . . . and then the knife went in.")

²⁵ S. Rose, "Pre-Copernican Sociobiology?" *New Scientist* 80 (1978): 45–46.

But this argument is fallacious for two reasons. First, the apparent threat to the traditional notion of free will has nothing to do with genetic, neurobiological, or evolutionary explanations of behavior; it is raised by *any* explanation of behavior. In this century it has been far more common to excuse behavior because of putative *environmental* causes. Remember the gang members in *West Side Story*, who explained, “We’re depraved on accounta we’re deprived”?

Dear kindly Sergeant Krupke,
You gotta understand,
It’s just our bringing up-ke,
That gets us out of hand.
Our mothers all are junkies,
Our fathers all are drunks.
Golly Moses, naturally we’re punks!

Stephen Sondheim’s lyrics lampooned the psychoanalytic and social science explanations of behavior popular in the 1950s and 1960s. Since then we have seen the Twinkie Defense that mitigated the sentence of the mayor-murdering Dan White, the Abuse Excuse that led to a mistrial of the Menendez brothers, the Black Rage Defense offered to the Long Island Railroad gunman Colin Ferguson, and the Pornography-Made-Me-Do-It defense attempted by several attorneys for rapists. Clearly there is nothing specific to brains, genes, or evolutionary history that lends itself to bogus justifications for bad behavior; any explanation can be abused in that way.

And that leads to Rose’s and Wolfe’s second fallacy, the confusion of explanation with exculpation. The difference between them is nicely captured in the old saying “To understand is not to forgive.” We would do well to keep the two separate. If some moral system identifies personal responsibility with a ghost in the machine, we ought to discard that moral system, because the

ghost is being exorcised, but we still need the notion of individual responsibility, if for no other reason than to construct policies of effective deterrence and to satisfy people's sense of justice. A better moral system would separate causation from responsibility as two sets of rules played out over the same entities (humans and their actions). We don't want the morality of killing, raping, lying, and stealing to depend on what comes out of the psychology or neuroscience lab at the other end of town. The autonomous moral agent is an indispensable construct that makes judicial and moral reasoning possible. It allows us to distinguish voluntary from involuntary acts, intended from unintended consequences, and the acts of rational adults from those of children, animals, and the patently deluded. It does not literally require a ghost in the machine as an alternative to a causal explanation in biological terms.

What about the more practical worry that the exorcising of the ghost implies that there is no way to hold people responsible for their behavior, and hence no way to reduce bad behavior? If bad behavior results from biological urges, is it inevitable, no matter how much we may condemn it? The answer is the same as the one to the question of whether urges toward violent conflict imply that war is inevitable. Since the mind has more than one part, one urge can counteract another and prevent it from pressing the buttons of behavior. Together with motives to hurt, lie, philander, and crave status, the human brain houses motives to avoid punishment, condemnation, loss of reputation, loss of self-esteem, and mistrust or abandonment by allies and loved ones. These faculties of social reasoning and emotion are every bit as "biological" as the deadly sins, so an approach to behavior that is consilient with biology does not dissolve hopes of improving standards of individual behavior.

The final fear is that a scientific explanation of mind will lead to a dissolution of meaning and purpose. The worry is that if emotion and feeling are just biochemical events in our brains, and if emotions are just patterns of activity in circuits ultimately designed by

natural selection as a way of propagating our genes, then our deepest ideals would be shams. Life would be a Potemkin village with only a facade of value and worth. For example, if we love our children because the genes for loving children are in the bodies of those children and the genes are thereby benefiting copies of themselves, wouldn't that undermine the inherent goodness of that love and the value of the self-sacrifice that parenting entails? If our empathy and good deeds toward others evolved, as evolutionary psychologists suggest, as ways of obtaining favors in the future, and if our sense of fairness and justice evolved as a way to avoid getting cheated when exchanging favors, wouldn't that imply that there is no such thing as altruism or justice, that deep down we're really selfish?

The worry reminds me of the opening scene of *Annie Hall* in which the young Alvy Singer is taken by his mother to a doctor:

Mother: He's been depressed. All of a sudden, he can't do anything.

Doctor: Why are you depressed, Alvy?

Mother: (Nudging Alvy) Tell Dr. Flicker. (Young Alvy sits, his head down. His mother answers for him.) It's something he read.

Doctor: Something he read, huh?

Alvy: (His head still down) The universe is expanding.

Doctor: The universe is expanding?

Alvy: (Looking up at the doctor) Well, the universe is everything, and if it's expanding, someday it will break apart and that would be the end of everything!

Mother: (Disgusted, she looks at him. Shouting) What is that your business? (She turns back to the doctor.) He stopped doing his homework.

Alvy: What's the point?

Mother: (Excited, gesturing with her hands) What has the universe got to do with it? You're here in Brooklyn! Brooklyn is not expanding!

Alvy's mother has a good point. Brooklyn is not expanding. What may seem depressing at the ultimate level of scientific analysis can be without consequence at the day-to-day scale on which we live our lives. The worry that our motives are "selfish" in an ultimate, evolutionary sense, and that therefore our supposedly selfless motives are really shams, is a confusion, a misreading of Richard Dawkins's metaphor of the selfish gene.

Dawkins pointed out that an excellent way to understand the logic of natural selection is to imagine that genes are agents with selfish motives.²⁶ The metaphor provides insight into complex processes of evolution and has led to countless successful empirical predictions. Unfortunately, the idea easily leads to a confusion, flowing from the assumption that the genes are our deepest hidden self, our essence. If genes are selfish, one might be tempted to think, then deep, deep down *we* must be selfish. The conclusion is a strange hybrid between evolutionary biology and Freud's theory of an unconscious self with ignoble motives.

The fallacy is that the *metaphorical* motives of the *genes* are different from the *real* motives of the *person*. Sometimes, the most selfish thing the genes can do is to help build a thoroughly unselfish person. For example, the love of children, at the psychological

²⁶ R. Dawkins, *The Selfish Gene* (New York: Oxford University Press, 1976/1989).

level of analysis at which we make sense of our own behavior, is obviously pure and heartfelt. It is only at a different level of analysis—the ultimate or evolutionary level at which we seek to explain *why* we have that pure emotion—that “selfishness” comes into the picture. The selfishness at one level does not contradict a selflessness at a different level, any more than the fact that the entire universe is expanding over billions of years undermines the fact that Brooklyn was not expanding in the 1940s.

A more general worry arises from the undisputed fact that experimental psychology has taught us that some of our experiences are figments. For example, the qualitative difference between the color red and the color green does not correspond to any qualitative physical difference in the light producing the sensation of red and green; wavelength, which gives rise to the perception of hue, is a continuous variable. The difference in kind between red and green is a construct of our perceptual system and could be different in an organism with slightly different chemistry or wiring (indeed, such organisms exist: people with red-green colorblindness). The new sciences of mind seem to be implying that the same is true of our perception of the difference between right and wrong—that the attainment of meaning and moral purpose may be just our way to tickle certain pleasure centers in the brain. They would have no more reality than the difference between red and green and could be meaningless to a person with a slightly different constitution.

But the analogy is imperfect. Many of our mental faculties evolved to mesh with real things in the world. Our perception of depth is the product of complicated circuitry in the brain, circuitry that appears to be absent in other species and even in certain impaired people. But that does not mean that there aren't real trees and cliffs out there or that the world is as flat as a cartoon. And this argument can be carried over to more abstract properties of the world. Humans (and many other animals) appear to have an innate sense of number, which can be explained by the utility of reasoning about numerosity in our evolutionary history. That is perfectly compatible with the Platonist theory of number believed

by many mathematicians and philosophers of mathematics, according to which abstract mathematical entities such as numbers have an existence independent of minds. The number three is not a figment like greenness; it has real properties, which are discovered and explored, not invented out of whole cloth. According to this view, the number sense evolved to mesh with real truths in the world that in some sense exist independent of human knowers.

A similar argument can be made for morality. According to the theory of moral realism, right and wrong have an existence and an inherent logic that licenses some chains of argument and not others. If so, our moral sense evolved to mesh with the logic of morality; it did not invent it out of the whole cloth.²⁷ The crucial point is that something can be *both* a product of the mind *and* a genuinely existing entity.

Even if one is uneasy with the admittedly difficult idea that moral truths exist in some abstract Platonic sense, one can preserve the inherent value of our moral judgments in other ways. One could be agnostic about the realism of moral judgments and simply note that our moral sense cannot work unless it *believes* that right and wrong have an external reality. That is, we cannot reason other than by presupposing that our moral judgments have some inherent validity (whether or not one could ever determine that they do). So when we have a moral debate, we would still appeal to external standards, even if moral reasoning is a biological adaptation; we would not merely be comparing idiosyncratic emotional or subjective reactions.

CONCLUSION

As I mentioned at the outset, these are exciting times in the study of the human mind and in the state of human knowledge in general. Thanks to cognitive science, neuroscience, behavioral

²⁷ See R. Nozick, *Philosophical Explanations* (Cambridge, Mass.: Harvard University Press, 1981), pp. 317–62.

genetics, and evolutionary psychology, we are beginning to arrive at an understanding of human nature that will bridge the last remaining chasms of knowledge: between matter and mind, and between biology and culture. This promises to lead to a particularly satisfying depth of understanding of our own kind, fulfilling the ancient injunction to know thyself.

In addition, a better understanding of mind and brain holds out the promise of indispensable practical applications. To take just one example, Alzheimer's disease will surely be one of the leading causes of human misery in the industrial world over the next several decades, as we live longer and stop dying of other causes. The successful treatment of Alzheimer's will come not from treating memory and personality as manifestations of an immaterial soul or of some irreducible, dignified agent. It will come from treating memory and personality as phenomena of biochemistry and physiology.

But the coming of a science of mind consilient with biology is, I recognize, not an innocuous development. It challenges beliefs that are deeply held in modern intellectual life and that are, in the minds of many, saturated with moral import. The most fundamental of these beliefs are the doctrines of the Blank Slate, the Noble Savage, and the Ghost in the Machine.

I have argued that the new developments in the sciences of mind do not have to undermine our moral values. On the contrary, they present opportunities to sharpen our ethical reasoning and put our moral and political values on a firmer foundation. In particular, it is a bad idea to say that discrimination is wrong only because the traits of all humans are identical. It is a bad idea to say that war, violence, rape, and greed are bad because humans are not naturally inclined to them. It is a bad idea to say that people are responsible for their actions because the causes of those actions are mysterious. And it is a bad idea to say that our motives are meaningful in a personal sense only if they are meaningless in a biological sense. These are bad ideas because they imply that either

scientists must be prepared to fudge their data or we must all be prepared to give up our values.

I argue that we do not have to make that choice. With a clearer separation of ethics and science, we can have our values and greet the new understanding of mind, brain, and human nature not with a sense of terror but with a sense of excitement. In the sixteenth century people attached grave moral significance to the question of whether the earth revolved around the sun or vice versa. Today it is hard to understand why people were willing to base moral beliefs on a plainly empirical claim, and we know that morals and values easily survived the claim's demise. I suggest that the same is true of the grave moral significance currently attached to the denial of human nature and of a materialist understanding of the mind.