1

Toward a Science of Morality

> Science has nothing to be ashamed of even in the ruins of Nagasaki. The shame is theirs who appeal to other values than the human imaginative values which science has evolved. The shame is ours if we do not make science part of our world.… For this is the lesson of science, that the concept is more profound than its laws.
>
> —Jacob Bronowski, *Science and Human Values*, 1956[1]

The metaphor of the bending moral arc symbolizes what may be the most important and least appreciated trend in human history—moral progress—and its primary cause is one of the most understated sources: scientific rationalism.

By *progress* I accept the *Oxford English Dictionary*'s historical usage as "advancement to a further or higher stage; growth; development, usually to a better state or condition; improvement." By *moral* I mean "manner, character, proper behavior" (as from the Latin *moralitas*), in terms of intentions and actions that are right or wrong with regard to another moral agent.[2] Morality involves how we think and act toward other moral agents in terms of whether our thoughts and actions are right or wrong with regard to their *survival and flourishing*. By *survival* I mean the instinct to live, and by *flourishing* I mean having adequate sustenance, safety, shelter, bonding, and social relations for physical and mental health. Any organism subject to natural selection—which includes all organisms on this planet and most likely on any other planet as well—will by necessity have this drive to survive and flourish, for if they didn't they would not live long enough to reproduce and would therefore no longer be subject to natural selection.

Because I include animals (and, perhaps one day, extraterrestrial life-forms) in our sphere of moral consideration, by moral agent I mean *sentient beings*. By *sentient* I mean *emotive, perceptive, sensitive, responsive, conscious*, and therefore able to feel and to suffer. In addition to using criteria such as intelligence, language, tool use, reasoning power, and other cognitive skills, I am reaching deeper into our evolved brains toward more basic emotive capacities. Our moral consideration should be based not primarily on what sentient beings are *thinking,* but on what they are *feeling*. There is sound science behind this proposition. According to the Cambridge Declaration on Consciousness—a statement issued in 2012 by an international group of prominent cognitive neuroscientists, neuropharmacologists, neuroanatomists, and computational neuroscientists—there is a convergence of evidence to show the continuity between humans and nonhuman animals, and that *sentience* is the common characteristic across species.

The neural pathways of emotions, for example, are not confined to higher-level cortical structures in the brain, but are found in evolutionarily older subcortical regions. Artificially stimulating the same regions in human and nonhuman animals produces the same emotional reactions in both.[3]

Further, attentiveness, sleep, and decision making are found across the branches of the evolutionary tree of life, including mammals, birds, and even some invertebrates, such as octopodes. In assessing all the evidence for sentience, these scientists declared, "Convergent evidence indicates that non-human animals have the neuroanatomical, neurochemical, and neurophysiological substrates of conscious states along with the capacity to exhibit intentional behaviors. Consequently, the weight of evidence indicates that humans are not unique in possessing the neurological substrates that generate consciousness."[4] Whether nonhuman animals are "conscious" depends on how one defines consciousness, but for my purposes the more narrowly restricted emotional capacity to *feel* and *suffer* is what brings many nonhuman animals into our moral sphere.

Given these reasons and this evidence, *the survival and flourishing of sentient beings* is my starting point, and the fundamental principle of this system of morality.[5] It is a system based on science and reason, and is grounded in principles that are themselves based on nature's laws and on human nature—principles that can be tested in both the laboratory and in the real world. Thus I take *moral progress* to mean *the improvement in the survival and flourishing of sentient beings*.

Here I am specifically referring to *individual* beings. It is the *individual* who is the primary moral agent—not the group, tribe, race, gender, state, nation, empire, society, or any other collective—because it is the *individual* who survives and flourishes, or who suffers and dies. It is individual sentient beings who perceive, emote, respond, love, feel, and suffer—not populations, races, genders, groups, or nations. Historically, immoral abuses have been most rampant, and body counts have run the highest, when the individual is sacrificed for the good of the group. It happens when people are judged by the color of their skin—or by their X/Y chromosomes, or by whom they prefer to sleep with, or by what accent they speak with, or by which political or religious group they belong to, or by any other distinguishing trait our species has identified to differentiate among members—instead of by the content of their *individual* character. The Rights Revolutions of the past three centuries have focused almost entirely on the freedom and autonomy of individuals, not collectives—on the rights of *persons*, not groups. Individuals vote, not races or genders. Individuals want to be treated equally, not races. Rights protect individuals, not groups; in fact, most rights (such as those enumerated in the Bill of Rights of the US Constitution) protect individuals from being discriminated against as members of a group, such as by race, creed, color, gender, and—soon—sexual orientation and gender preference.

The singular and separate organism is to biology and society what the atom is to physics—a fundamental unit of nature. (Here I am not including social insects such as worker drone bees, whose members are genetically nearly identical.) Thus the first principle of *the survival and flourishing of sentient beings* is grounded in the biological fact that the discrete organism is the principal target of natural selection and social evolution, not the group.[6] We are a social species—we need and enjoy the presence of others in groupings such as families, friends, and assorted social consortia—but we are first and foremost *individuals* within social groups and therefore ought not to be subservient to the collective.[7] Making sacrifices for one's social group is not the same as being sacrificed for the group.

This drive to survive is part of our essence, and therefore the freedom to pursue the fulfillment of that essence is a *natural right*, by which I mean it is universal and inalienable and thus not contingent only upon the laws and customs of a particular culture or government. Natural rights theory arose during the Enlightenment to counter the belief in the divine right of kings, and became the basis of the social contract that gave rise to democracy, a superior system for the protection of human rights. This is what the English philosopher John Locke had in mind in his 1690 *Second Treatise of Government*

(written to rebut Sir Robert Filmer's 1680 *Patriarcha*, which defended the divine right of kings[8]) when he wrote: "The state of nature has a law to govern it, which obliges every one: and reason, which is that law, teaches all mankind, who will but consult it, that being all equal and independent, no one ought to harm another in his life, health, liberty, or possessions."[9] The social contract entered into freely, Locke argued, is the best way to ensure our natural rights.[10]

In rights language, the individual is imbued with *personal autonomy*. As a natural right, the personal autonomy of the individual gives us criteria by which we can judge actions as right or wrong: *do they increase or decrease the survival and flourishing of individual sentient beings?* Morality is not arbitrary, relative, or completely culture-bound. Morality is universal. We are all born with a moral sense, with moral emotions that guide us in our interactions with other people, and that are influenced by local culture, customs, and upbringing. Nature endowed us with the capacity to feel guilt for the violation of promises and social obligations, for example, but nurture can tweak the guilt dial up or down. Thus morality is real, discoverable, "out there" in nature, and "in here" as part of our human nature. From these facts we can build a science of morality—a means of determining the best conditions to expand the moral sphere and increase moral progress through the tools of reason and science.

SCIENCE, REASON, AND THE MORAL ARC

Understanding the nature of things and the causes of effects is what science is designed to do, and ever since the Scientific Revolution there has been a systematic effort by thinkers in all fields to apply the methods of science—which include the philosophical tools of reason and critical thinking—to understanding ourselves and the world in which we live, including and especially the social, political, and economic worlds, with an end toward the betterment of humanity. This effort has produced a worldview known as Enlightenment Humanism (or secular humanism, or simply humanism), which, unlike most other worldviews, is more a method than an ideology; it is a means of solving problems more than it is a set of doctrinaire beliefs. Humanism, as its name implies, is—*and ought to be*—concerned with the survival and flourishing of humans, and its methods of reason and science are directed at figuring out how best to do that. Thus the goal of a science of morality is—*and ought to be*—to determine the conditions under which humans and, by extension, other sentient beings best prosper. To that end I need to define what I mean by science and reason.

Science

Science is a set of methods that describes and interprets observed or inferred phenomena, past or present, and is aimed at testing hypotheses and building theories. By *set of methods* I mean to emphasize that science is more of a procedure than it is a set of facts, and to *describe and interpret* them means that the facts do not just speak for themselves. *Observed or inferred phenomena* means that there are some things in nature that we can see, such as elephants and stars, but other things that we must infer, such as the evolution of elephants and stars. *Past or present* means that the tools of science can be used to understand not only phenomena occurring in the present, but in the past as well. (The historical sciences include cosmology, paleontology, geology, archaeology, and history, including and especially human history.) *Testing hypotheses* means that for something to be truly

scientifically sound it must be testable, such that we can confirm it as probably true or disconfirm it as probably false.[11] *Building theories* means that the aim of science is to explain the world by constructing comprehensive explanations from numerous tested hypotheses.

Defining the *scientific method* is not so easy. The process involves making observations and forming hypotheses from them, then making specific predictions based on those hypotheses, then making additional observations to test those predictions to confirm, disprove, or falsify the initial hypotheses. The process is a constant interaction of making observations, drawing conclusions, making predictions, and checking them against the evidence. But note that data-gathering observations are not made in a vacuum. The hypotheses shape what sort of observations a scientist will make, and these hypotheses are themselves shaped by education, culture, and the particular biases of the observer. Observation is key. The British astronomer Sir Arthur Stanley Eddington employed a legal metaphor to capture the sentiment: "For the truth of the conclusions of physical science, observation is the supreme court of appeal."[12] All facts in science are provisional and subject to challenge and change, therefore science is not a "thing" per se; rather it is a *method* of discovery that leads to *provisional* conclusions.

Reason

Reason is the cognitive capacity to establish and verify facts through the application of logic and rationality, and to make judgments and form beliefs based on those facts. *Rationality* is the application of reason to form beliefs based on facts and evidence, instead of guesswork, opinions, and feelings. That is to say, the rational thinker wants to know what is *really* true and not just what he or she would *like* to be true.[13]

However, as several decades of research in cognitive psychology has shown, we are not the rationally calculating beings we'd like to think we are, but are instead very much driven by our passions, blinded by our biases, and (for better or worse) moved by our moral emotions. The confirmation bias, the hindsight bias, the self-justification bias, the sunk-cost bias, the status-quo bias, anchoring effects, and the fundamental attribution error are just a few of the many ways that our brains work to convince us that what we *want* to be true *is* true—regardless of the evidence—in a general process called "motivated reasoning."[14] Nevertheless, the capacity for reason and rationality is within us as a feature of our brains that evolved to form patterns and make connections (it's called *learning*) in the service of survival and flourishing in the environment of our evolutionary ancestry. Reason is part of our cognitive makeup, and once it is in place it can be put to use in analyzing problems it did not originally evolve to consider. Pinker calls this an open-ended combinatorial reasoning system that "even if it evolved for mundane problems like preparing food and securing alliances, you can't keep it from entertaining propositions that are consequences of other propositions." This ability matters for morality because "if the members of species have the power to reason with one another, and enough opportunities to exercise that power, sooner or later they will stumble upon the mutual benefits of nonviolence and other forms of reciprocal consideration, and apply them more and more broadly."[15]

Drawing inferences about the movement of animals from their tracks—as hunter-gatherer trackers do—has obvious survival advantages, and we have been able to apply those inferential skills to everything from driving to the store to flying rockets to the moon. Historian of science and

professional animal tracker Louis Liebenberg has, in fact, argued that our ability to reason scientifically is a by-product of fundamental skills for tracking game animals that our ancestors developed. Liebenberg's analogy between tracking and the *scientific method* is revealing: "As new factual information is gathered in the process of tracking, hypotheses may have to be revised or substituted by better ones. A hypothetical reconstruction of the animal's behaviors may enable trackers to anticipate and predict the animal's movements. These predictions provide ongoing testing of hypotheses."[16] Liebenberg distinguishes between *systematic tracking* ("the systematic gathering of information from signs, until it provides a detailed indication of what the animal was doing and where it was going") and *speculative tracking* ("the creation of a working hypothesis on the basis of initial interpretation of signs, knowledge of the animal's behavior and knowledge of the terrain" that leads to hypotheses that are tested and, if not confirmed, to new hypothetical reconstructions of the animal's whereabouts). Speculative tracking also involves another cognitive process called "theory of mind," or "mind reading," in which trackers put themselves into the mind of the animal they are pursuing and imagine what it might be thinking in order to predict its actions.

Based on archaeological and anthropological evidence Liebenberg estimates that humans have been hunting and using systematic tracking for at least two million years (as far back as *Homo erectus*), and speculative tracking for at least one hundred thousand years.[17] Whenever these cognitive capacities arose, once the neural architecture is in place to deduce, say, that a lion slept here last night, a person can substitute lion with any other animal or object and can swap "here" with "there" and "last night" with "tomorrow night." The objects and time elements of the reasoning process are interchangeable. In a modern example, once you've mastered the multiplication tables and you know that $7 \times 5 = 35$, you can infer that 5×7 is also 35 because 5 and 7 are interchangeable in the equation. This interchangeability is a by-product of neural systems that evolved for basic reasoning abilities such as tracking animals for food.[18]

This is how a brain that evolved for one purpose can be put to other uses, and this cognitive capacity to substitute Xs and Ys in a representational system that encompasses endless combinations and options—from prey to people—is what enables us to adopt the perspective of another moral agent, and is thus the cognitive architecture underlying moral reasoning.

THE EXPANDING MORAL SPHERE AND THE PRINCIPLE OF INTERCHANGEABLE PERSPECTIVES

The *expanding moral sphere* is the metaphor I use to describe what has been pushing the moral arc upward, derived from the *expanding circle* metaphor first evoked in 1869 by the Irish historian William Edward Hartpole Lecky in his massive two-volume survey titled *History of European Morals*: "History tells us that, as civilisation advances, the charity of men becomes at once warmer and more expansive, their habitual conduct both more gentle and more temperate, and their love of truth more sincere." Such moral progress, however, is not built into our biology, Lecky says. "Men come into the world with their benevolent affections very inferior in power to their selfish ones, and the function of morals is to invert this order." After admitting that "The extinction of all selfish feeling is impossible for an individual, and if it were general, it would result in the dissolution of society," Lecky shows that moral progress is an incremental process: "The question of morals must always be a question of proportion or of degree. At one time the benevolent affections embrace merely the

family, soon the circle expanding includes first a class, then a nation, then a coalition of nations, then all humanity, and finally, its influence is felt in the dealings of man with the animal world."[19] Expanding the moral circle to include animals, in nineteenth-century Europe? That was innovative for the time, and it shows what can happen once you start reasoning from basic moral principles.[20]

The philosopher Peter Singer too was ahead of the curve when he published *The Expanding Circle* in 1981, anticipating the developments in the sciences of evolutionary psychology and evolutionary ethics that unfolded in the 1990s and 2000s and out of which a science of morality could be developed. Singer makes the case for reason and science as providing rational arguments for *why* we should value the interests of X as much as we value our own interests, with X being racial minorities, gay people, women, children, and now animals. To explain the expanding circle Singer invokes what he calls "The principle of impartial consideration of interests": "In making ethical decisions I am trying to make decisions which can be defended to others. This requires me to take a perspective from which my own interests count no more, simply because they are my own, than the similar interests of others. Any preference for my own interests must be justified in terms of some broader impartial principle."[21]

Steven Pinker explains the logic like this: "If I appeal to you to do something that affects me then I can't do it in a way that privileges my interests over yours if I want you to take me seriously. I have to state my case in a way that would force me to treat you in kind. I can't act as if my interests are special just because I'm me and you're not, any more than I can persuade you that the spot I am standing on is a special place in the universe just because I happen to be standing on it."[22]

The reasoning process behind the expanding moral sphere (I prefer the three-dimensionality of a sphere instead of the two-dimensionality of a circle because I imagine it encompassing more range of variability within and across time and space and species) might more broadly be called the *principle of interchangeable perspectives*, and it applies not just to individuals within a group, tribe, or nation, but *between* groups, tribes, and nations as well. I cannot reasonably appeal to your nation to privilege my nation simply because it is my nation and not yours. (When I tell my European friends about a certain US conservative radio talk show host's routine refrain about America being "the greatest nation on God's green Earth"[23] they just roll their eyes.) Any preference for my group's interests over yours must be justified by some unbiased, disinterested ethic, which sounds simple, but given that we're dealing with humans, not Vulcans, it's sometimes difficult for two parties to agree on basic principles—especially parties who are unable or unwilling to switch points of view. This is the power of ethical reasoning, which, as Singer notes, "once begun, pushes against our initially limited ethical horizons, leading us always toward a more universal point of view."[24]

Reason and the *principle of interchangeable perspectives* put morals more on a par with scientific discoveries than cultural conventions. Scientists cannot just assert a claim without backing it up with reasoned arguments and empirical data (well, they can, but they'll be unceremoniously dismissed or publicly trounced by their colleagues). There really is a better way for people to live, and in principle we should be able to discover that way through the tools of science and reason. It is often said that you cannot reason someone out of a belief that they didn't reason themselves into in the first place, but when reasons are given for a belief we are entitled to counter those reasons with better reasons if they are available. And if no reasons are in the offing, we are entitled to dismiss them by invoking what I call *Hitchens's Dictum*, after my late friend and colleague Christopher Hitchens's observation, "What can be asserted without evidence can also be dismissed without

evidence."[25]

The ability and, just as importantly, the willingness to change perspectives and points of view are major drivers of the expanding moral sphere, which is pictured in *figure 1-1*. The expansion outward from ourselves shows that our moral concerns are most directly connected to those most closely related to us genetically, from an identical twin, to our siblings, parents, and offspring, to our grandparents, half siblings, aunts and uncles, nieces and nephews, to our great-grandparents, great-aunts and great-uncles, and great-grandchildren, all the way out to great-great-great-grandparents, to friends, acquaintances, and members of our in-group, to members of other groups, tribes, states, and nations, to all members of our species, to members of other mammal species, to all sentient beings, to the biosphere. Reflected in this ability to change perspectives and expand the moral sphere may be our expanding intelligence and abstract reasoning ability.

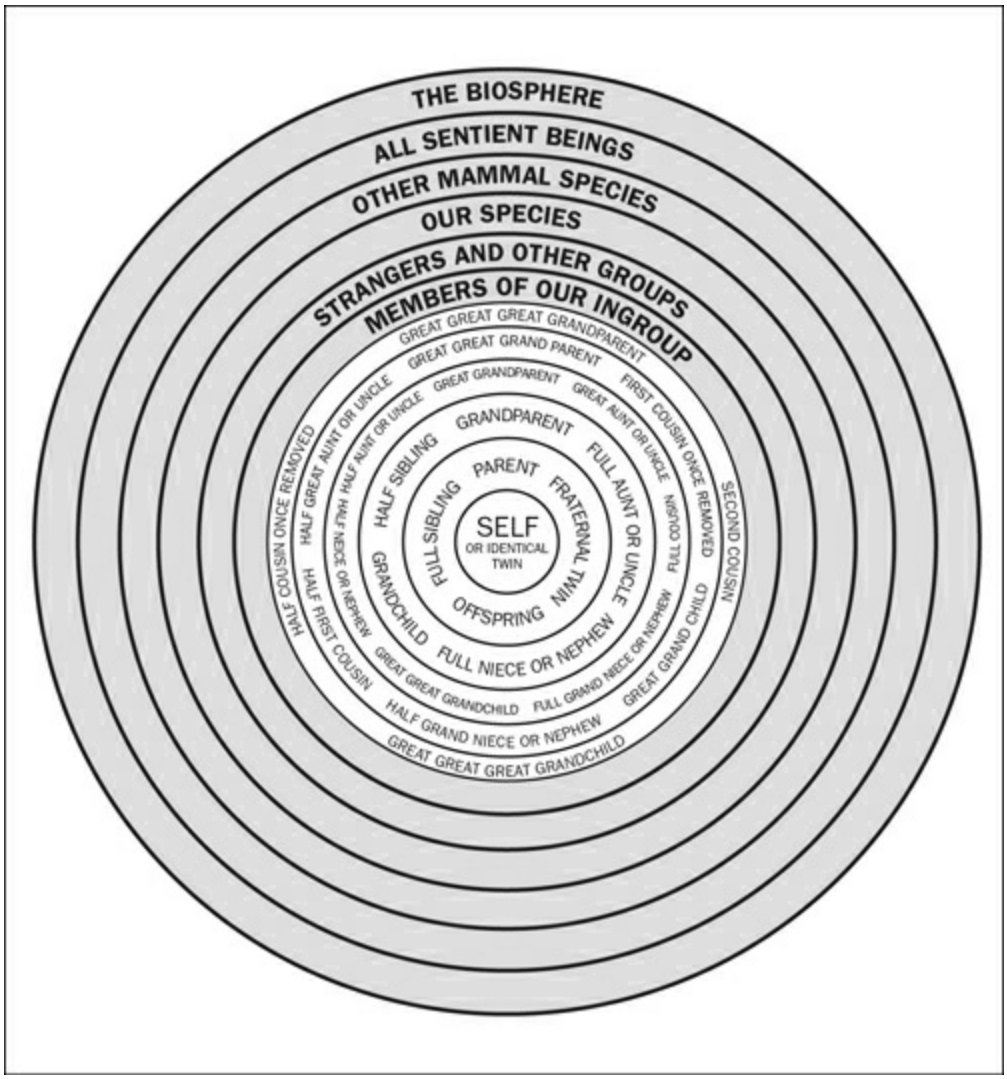

Figure 1-1. The Expanding Moral Sphere

ABSTRACT REASONING AND MORAL INTELLIGENCE

Thinking scientifically requires the ability to reason abstractly, which itself is at the foundation of all morality. Consider the mental rotation required to implement the Golden Rule: Do unto others as you would have them do unto you. This necessitates one to change positions—to become the other—and then to extrapolate what action X would feel like as the receiver instead of the doer (or as the victim instead of the perpetrator). A case can be made that the type of conceptual ratiocination required for

both scientific and moral reasoning not only is linked historically and psychologically, but also that it has been improving over time as we become better at nonconcrete, theoretical reflection.

In the 1980s the social scientist James Flynn discovered that IQ scores have been going up, on average, 3 points every decade for the past century. Now called the *Flynn Effect*, this is an astonishing increase of about 30 IQ points over 100 years. This translates into an improvement of two standard deviations of 15 points each from an "average" IQ of 100 to a "very superior" score of 130. (IQ test scores remain the same, however, as they are regularly "normed" upward to account for the Flynn Effect, which is how Flynn discovered it in the first place.) Were it just the case that we're all getting better at taking tests, then the scores should have been improving across the board. But this is not what has happened. The increases in IQ scores have been almost exclusively in the two subtests that most require abstract reasoning: Similarities and Matrices. The subtests of Information, Arithmetic, and Vocabulary have hardly budged at all.[26] *Figure 1-2* shows the trend lines since the late 1940s.

The subtest called *Similarities* asks questions such as "What do dogs and rabbits have in common?" If you answer, "Both are mammals," says Flynn, you are thinking like a scientist in classifying organisms by type, which is an abstraction. If you said, "You use dogs to hunt rabbits," you are thinking concretely, imagining a tangible use for a dog. According to Flynn, for the past century people have learned to think more abstractly than concretely.

Matrices are abstract figures that require determining a pattern and then deducing the missing piece in the pattern, as in *figure 1-3*.

The cause of the Flynn Effect is controversial. The hypothesis that the rising tide of standardized testing has lifted all boats is contradicted by the fact that the increase in IQ scores preceded the era of standardized test taking and they've continued their rise at a steady rate irrespective of test-taking rates.[29] A more likely explanation is that the improvement is a function of more years in school, more technologies in society, more technical jobs, and a greater need for people to perform conceptual tasks as our economy has shifted from agrarian and industrial to information-based. Instead of manipulating plows, cows, and machinery, many of us are now manipulating words, numbers, and symbols. Even in science classes the trend has been shifting away from the rote memorization of facts about nature to reasoning about nature's laws and processes—content *and* process. And process thinking is a form of abstract reasoning.[30]

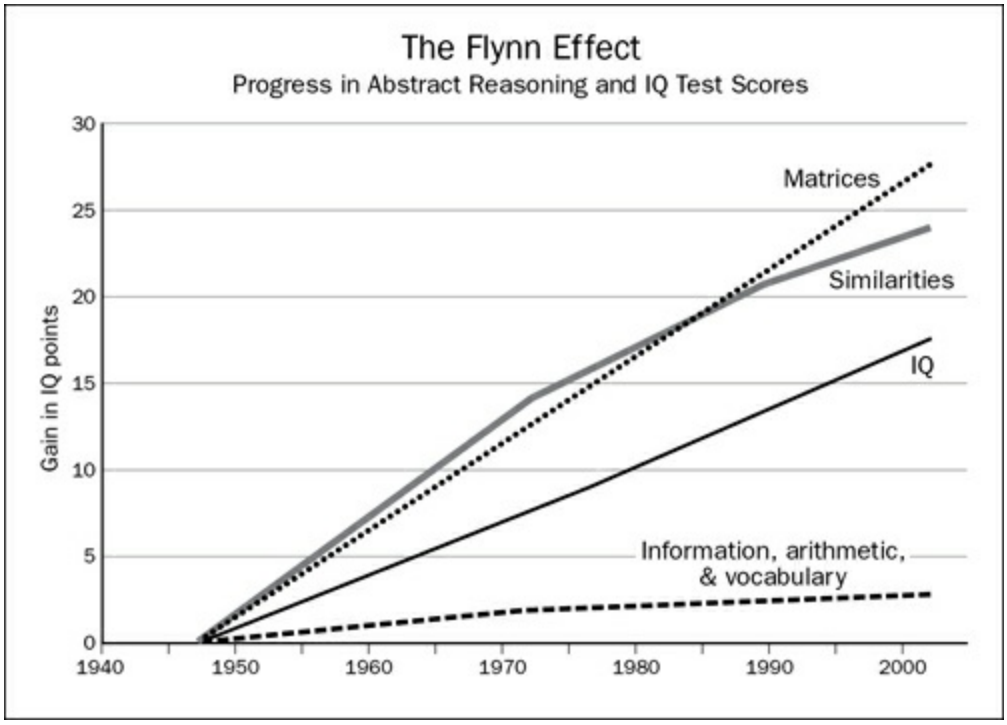

Figure 1-2. The Flynn Effect
The social scientist James Flynn discovered that IQ scores are increasing, on average, 3 points every decade, most noticeably in the two subtests that require abstract reasoning: Similarities and Matrices.[27]

Flynn himself attributes the effect to an accelerating capacity for people to view the world through "scientific spectacles." He contrasts the "prescientific" world of his father with the "postscientific" world of today through a poignant anecdote about how he and his brother tried to mitigate their father's typical prejudice of his generation through a thought experiment: "What if you woke up one morning and discovered your skin had turned black? Would that make you any less of a human being?" The senior Flynn shot back, "Now, that's the stupidest thing you've ever said. Who ever heard of a man's skin turning black overnight?" The Flynn patriarch was intelligent but uneducated, Flynn explained, in attributing the effect to nurture, not nature.[31] The anecdote is symbolic of larger social trends. Each generation is producing not only better *abstract* reasoners, but better *moral* reasoners as well. In an interview in *Skeptic* magazine, Flynn reflected on the early-twentieth-century research by the psychologist Alexander Luria on the reasoning abilities of Russian peasants:

> The illiterate Russian peasants Luria studied were not willing to take the hypothetical seriously. He said, "Imagine that bears come from where there is always snow and imagine that if bears come from where there is always snow they are white. What color would the bears be at the North Pole?" and they would respond something like, "I've only seen brown bears. If an old man came from the North Pole and told me I might believe him." They were not interested in the hypothetical, or abstract categories. They were grounded in concrete reality. "There are no camels in Germany. B is in Germany. Are there camels there?" They said, "Well, it's big enough, there ought to be camels. Or maybe it's too small to have camels." We have wonderful data from the Raven's Progressive Matrices tests from 1950 and 2010 showing that the Raven's games are entirely correlated with freeing your mind of the concrete reference of the symbols in order to take the relationship between the symbols more seriously.[32]

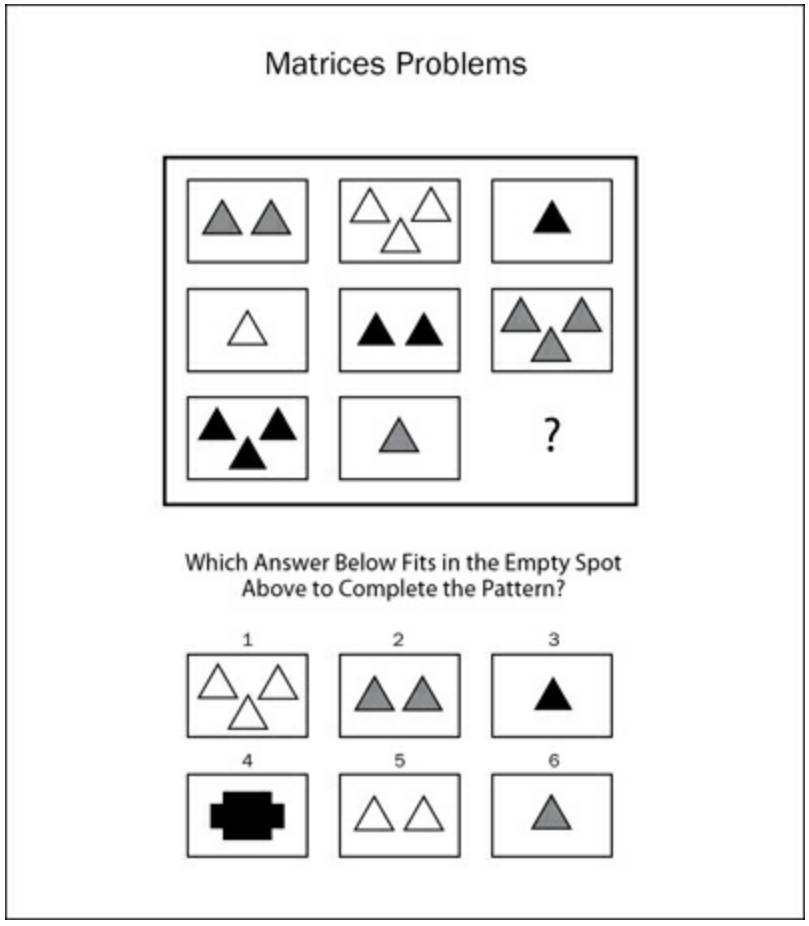

Figure 1-3. Matrices Problems

If you selected answer #5, then you are reasoning abstractly.[28]

Flynn and his colleague William Dickens suggest that the increases in cognitive reasoning may have started centuries ago with the Industrial Revolution, which saw an improvement in both the quantity and quality of education, better nutrition, disease control, and the manipulation of complex machinery. Then, after 1950, "IQ gains show a new and peculiar pattern. They are missing or small on the kind of IQ tests closest to school-taught material like reading and arithmetic. They are huge on tests that emphasize on-the-spot problem-solving, like seeing what verbal abstractions have in common, or finding the missing piece of a Matrices pattern, or making a pattern out of blocks, or arranging pictures to tell a story. Perhaps the Industrial Revolution stopped demanding progress in the basics and started demanding that people take abstract problem-solving more seriously."[33]

Whatever the cause of the Flynn Effect, it isn't genetic or biological, because there hasn't yet been enough time for natural selection to operate, and even though nutrition has improved over time, it more or less stabilized in the mid-twentieth century (and, if anything, may have gotten worse recently due to the prevalence of junk food), and yet IQ gains continue. Steven Johnson, in his intriguing book *Everything Bad Is Good for You*, makes the case for modern pop culture and media—even the "boob tube"—as drivers of abstract reasoning improvements, noting, for example, that the plot lines and character developments of today's television shows are far more complex than those of decades past.[34] Flynn theorizes a suite of cultural factors that came together over the past century:

> Cognitively demanding jobs raise IQ. Look at how much more cognitively demanding a merchant banker's job is today than in 1900, when a banker just had to know who was a good risk for mortgages. Or cognitively demanding leisure: The video games may not be getting you to read good literature, but they do exercise your mind, rather than just playing sports. Leisure, occupation,

and, of course, schooling too make a difference. The schooling of today has introduced into its curriculum much more intellectually challenging material. If you look at the exams given American school children in 1914 in Ohio, it was all about socially valuable material. "What are the capitals of the then 44 states?" Today it would be, "Why is the capital of a state rarely the largest city?"[35]

Being able to rattle off the capitals of the states does not require any abstract reasoning ability, but knowing that rural state legislatures controlled where the capital was put in a state, and that they disliked big cities and so put them in a county seat, leads to a deeper understanding of why Albany is the capital of New York rather than New York City, and why Harrisburg is the capital of Pennsylvania rather than Philadelphia. "So you see they are being asked to make hypotheses about fairly abstract concepts and propositions, and link those hypotheses through logic," Flynn explained. "So the demands conceptually placed on the schools have now changed."[36] And they have changed in the workplace. Flynn points out that in 1900 only 3 percent of Americans had cognitively demanding jobs, whereas that figure was 35 percent in 2000.

A case can be made that our improved ability to reason abstractly is the result of the spread of scientific thinking—that is, science in the broader sense of reason, rationality, empiricism, and skepticism. Thinking like a scientist means employing all our faculties to overcome our emotional, subjective, and instinctual brains to better understand the true nature of not only the physical and biological worlds, but the social world (politics and economics) and the moral world (abstracting how other people should be treated) as well. That is, the moral arc of the universe may be bending, in part, because of something like a *Moral Flynn Effect*, as Pinker calls it.[37] Pinker says "the idea is not crazy," but I would go farther. I claim that our improvement in abstract reasoning generally has translated into a specific improvement in abstract moral reasoning, particularly about other people who are not our immediate kith and kin. Evolution endowed us with a natural tendency to be kind to our genetic relations but to be xenophobic, suspicious, and even aggressive toward people in other tribes. As our brains become better equipped to reason abstractly in such tasks as lumping dogs and rabbits together into the category of "mammal," so too have we improved in our capacity to lump blacks and whites, men and women, straights and gays into the same category of "human." To employ a metaphor from evolutionary theory and the problem of defining a species, we are moving toward becoming "lumpers" instead of "splitters"—seeing similarities instead of differences.

As philosophers and scholars over the pastb two centuries have consciously adopted the methods of science to establish such abstract concepts as rights, liberty, and justice, successive generations have become schooled in thinking of these abstractions as applied to others in a Matrices-like mental rotation. Consider a number of studies and lines of evidence in support of the hypothesis that our moral intelligence is increasing[38]:

- Intelligence and education are negatively correlated with violent crime.[39] As intelligence and education increase, committing violent crimes and falling victim to them decrease, even when controlled for socioeconomic class, age, sex, and race.[40]
- Cognitive style predicts criminal justice attitudes. The psychologist Michael Sargent found a correlation between a high "need for cognition" (enjoying mental challenges such as those employed in intelligence tests) and a low demand for punitive justice, even when such attitudes are controlled for age, sex, race, education, income, and political orientation. In an aptly titled

paper "Less Thought, More Punishment," Sargent's conclusions support the principle that the punishment should fit the crime, a principle that requires grasping the abstract concept of proportionality, a process fundamental to all scientific thought.[41]

- Abstract reasoning ability is positively correlated with cooperation in a game called Prisoner's Dilemma (a classic thought experiment in game theory that proves cooperation creates better outcomes even when a perfectly rational, self-interested actor would not want to do so). The economist Stephen Burks and his colleagues administered a thousand trainee truck drivers the Matrices IQ test and had them participate in a game of Prisoner's Dilemma in which they could either cooperate or defect with a game partner. Those wannabe truck drivers who scored high in ability to solve those Matrices figures were more likely to cooperate on the first move of a Prisoner's Dilemma game, even after the usual intervening variables were controlled for, such as age, race, gender, schooling, and income.[42] The economist Garrett Jones confirmed the connection in a meta-analysis of thirty-six Prisoner's Dilemma experiments conducted between 1959 and 2003 in colleges and universities around the country, finding a positive correlation between a school's mean SAT score and the propensity of its students to respond cooperatively.[43]
- Intelligence predicts classical liberal attitudes toward helping others. An analysis of data from the National Longitudinal Study of Adolescent Health found that among twenty thousand young adults there was a positive correlation between IQ and liberalism, and data from the General Social Survey clarified the link in noting that the correlation was between intelligence and *classical* liberalism of the Enlightenment kind, in which smarter people were less likely to agree that the government should redistribute income from the rich to the poor but more likely to agree that the government should help African Americans to compensate for historical discrimination.[44] In other words, the effect was more in the moral dimension of how people are ethically treated instead of the more concrete dimension of economic adjustment, and thus its import for moral reasoning.
- The psychologist Ian Deary and his colleagues confirmed this link in a paper aptly titled "Bright Children Become Enlightened Adults." Deary found a positive correlation between the IQ of British children at age ten and their endorsement of antiracist, socially liberal, and pro-working-women attitudes at age thirty, holding the usual potentially intervening variables constant. The causal arrow from intelligence to moral abstraction is confirmed by the twenty-year gap between measures.[45] By "enlightened" Deary meant the values that came directly from the Enlightenment, the definition for which he adopted from the *Concise Oxford Dictionary*: "a philosophy emphasizing reason and individualism rather than tradition."
- Intelligence predicts economic attitudes, most notably abstract concepts such as how free trade is a form of positive-sum game that seems counterintuitive to our folk-economic intuitions that most economic exchanges are zero-sum in a fixed pie of wealth. The economists Bryan Caplan and Stephen Miller culled data from the General Social Survey and found a correlation between intelligence and openness to immigration, free markets, and free trade, and a reluctance to endorse government make-work projects, protectionist policies, and business interventionism.[46] Concrete thinking leads us to endorse economic tribalism along with populist and nationalist zero-sum attitudes toward other tribes (nations in the modern world). Abstract reasoning leads us to consider members of other tribes (nations) as potential trading partners to be respected

rather than as potential enemies to be conquered or killed.

- Intelligence predicts democratic tendencies, most notably the rule of law. The psychologist Heiner Rindermann ran correlational studies on a number of datasets from many different countries, examining their average scores on popular intelligence tests and measures of academic achievement from 1960 to 1972, and found that these predicted the level of prosperity, democracy, and the rule of law found in those countries in the subsequent period 1991 to 2003 (and this is even when controlling for the country's prior level of prosperity).[47] In other words, all other things being equal, a country that educates its population in the ability to reason abstractly will be a more prosperous and moral country.
- Most encouraging for those of us who are citizens of the Republic of Letters, evidence is now accumulating that shows a positive correlation between literacy and morality, and most particularly between the reading of fiction and being able to take the perspective of others.[48] Taking the perspective of characters in a novel requires a Matrices-like rotation of relational positions, coupled to an emotional connection of what it would feel like if X happened to you, even though the "you" in this case is a character in the novel. In a 2011 study, for example, the Princeton University neuroscientist Uri Hasson and his team scanned the brain of a woman while she told a story out loud that the scientists recorded and subsequently played back for other subjects while their brains were being scanned as they just listened. Results: when the reader's emotional brain region called the insula lit up during a certain portion of the story, so too did the listeners' insulas; when the woman's frontal cortex became active during a different part of the story, the same region in listeners' brains were also activated.[49] It's almost as if the fictional story synchronized the brains of reader and listeners, enabling a form of mind reading and moral perspective taking. (Such brain synchronicity was found in another study by Hasson and his team when he scanned the brains of subjects while they watched Sergio Leone's classic 1966 film *The Good, the Bad, and the Ugly*, finding that "brain activity was similar across viewers' brains," specifically that about 45 percent of the neocortex across all five of the subjects' brains were lit up in the same areas during the same movie scenes.[50])

A 2013 study published in *Science* titled "Reading Literary Fiction Improves Theory of Mind" reported on the results of studies conducted by the psychologists David Comer Kidd and Emanuele Castano on the causal relationship between reading high-quality literary fiction and the ability to take the perspective of others, as measured by one of several well-tested tools, such as judging the emotions of others and eye-gaze directionality for interpreting what someone is thinking.[51] They found that participants who were assigned to read literary fiction performed significantly better on the Theory of Mind (ToM) tests than did participants assigned to the other experimental groups, who did not differ from one another.

This experiment is important because it nails down the direction of the causal arrow from reading literary fiction to perspective taking, instead of the other direction, in which people who are good at mind reading prefer fiction. That said, this research is in its infancy, and there are reasons to be skeptical about pushing the link between literacy and morality too far. Education in general and literature in particular may have morally salubrious effects for reasons we do not yet fully understand, but I am encouraged by these studies, and others that put theory into practice. For example, in a documentary film titled *The University of Sing Sing*, Tim Skousen documents the work of his parents

—Jo Ann and Mark Skousen—and other teachers at the Sing Sing prison in New York, in which literature is used to broaden the critical thinking skills and expand the moral horizons of the prisoners there.[52]

Working through one of the few college degree-granting programs in the Department of Corrections of New York State, the psychologists interviewed cite statistics showing that the best predictor of success after prison is a college degree. As the psychologist Susan Weiner, who works with the program, noted, "These men and women will come back to the community. How do you want them to come back? This isn't just a gift for them. It's a gift for society. It's the way that we make society a better place." A prisoner named Denis Martinez, for example, explained what getting an education and learning to read deeply into subjects gave him in terms of perspective: "It's given me a new set of glasses. Before I wasn't able to see the things I see now. I was a nineteen-year old knucklehead going around and thinking I knew it all. The more I learned the more I could sense how wrong I was and how many things I didn't know." Inspired by his reading of René Descartes, Martinez reflected, "There are two ways to be in prison—physically and/or mentally. Being in prison mentally is to live in ignorance, closed-mindedness, and pessimism. You can confine me for as long as you want, but my mind will always be free." The title of a painting this prisoner made is revealing: *Cogito Ergo Sum Liber—I Think Therefore I Am Free*. (Now, *there's* a bumper sticker/T-shirt slogan for the modern Enlightenment thinker.)

THE VIRTUE OF CONTINUOUS THINKING

Thinking abstractly is not the only cognitive tool of the scientist that we can apply to moral reasoning. Thinking about concepts both on a *continuous scale* and as *categorical entities* illuminates—and sometimes eliminates—a number of moral problems. In my book *The Science of Good and Evil* I applied the idea of "fuzzy logic"[53] to show that "evil" and "good" are not simply black-and-white categories of reified "things" but are instead fuzzy shades of behavior along a continuous scale. *Good* and *evil* are descriptive terms for behaviors of moral actors that can be assessed along a continuum. Take altruism and selfishness: Like all behaviors, there is a broad range of expression of both types of behavior. Instead of categorizing someone as either altruistic or selfish in a 1 or 0 binary logic system, we might think of this person as 0.2 altruistic and 0.8 nonaltruistic (or selfish), or 0.6 cooperative and 0.4 noncooperative (or competitive).[54]

Most moral problems are better conceived as continuous rather than as categorical. The categorization of the world into cleanly cleaved boxes is a useful cognitive tool for some tasks, but it doesn't always serve us well in understanding social and moral problems. Does democracy decrease the probability of war? If you categorize nations into simple binary boxes of democracy or nondemocracy (1 or 0), you can find lots of exceptions to the democratic peace theory. But if you scale rank countries by degrees of democracy (on a 1–10 scale), and you also scale wars from minor conflicts to world wars, you find a significant negative correlation between the degree of democracy and the probability of fighting (more on this later).

Scientists tend to think about problems on a continuous scale. The evolutionary biologist Richard Dawkins, for example, points out that labeling a fossil as belonging to this or that species is falling prey to "the tyranny of the discontinuous mind," in which "Paleontologists will argue passionately about whether a particular fossil is, say, *Australopithecus* or *Homo*. But any evolutionist knows there

must have existed individuals who were exactly intermediate." Darwin's theory of evolution, in fact, overturned the categorical "essentialism" of how organisms are conceived as fixed entities. "It's essentialist folly to insist on the necessity of shoehorning your fossil into one genus or the other," Dawkins notes. "There never was an *Australopithecus* mother who gave birth to a *Homo* child, for every child ever born belonged to the same species as its mother."[55]

Consider the category of "poverty." The World Bank defines poverty as making less than $1.25 a day, and the Bill and Melinda Gates Foundation has highlighted the fact that since 1990 the percentage of the world's population living in poverty has declined by 50 percent.[56] This is progress, and it should be duly noted what has been done to continue this positive trend, but the categorical thinking that puts people into two boxes of "poverty" and "nonpoverty" obscures the fact that making, say, $2.50 a day is still a serious detriment to one's survival and flourishing. A gradient scale indicating how much people make around the world more accurately depicts their economic well-being (or lack thereof).

Not all moral problems can be so conceived, but throughout this book I will show how thinking about issues on a continuous scale rather than as categorical entities is both instructive and enlightening, and if exceptions to generalizations come to mind it is helpful to consider whether they invalidate the generalization or fall on a continuous scale within the generalization.

FROM *IS* TO *OUGHT* IN MORAL PROGRESS

As we work our way through the evidence for moral progress and the many factors that help to bring it about, keep in mind that establishing the causes of moral progress tells us how to achieve it if that is our goal. But it does not explain *why* we would want to expand the moral sphere in the first place. One could just as well argue that science and reason show us how to *shrink* the moral sphere, and that too would be true. This difference between *how* and *why* gets at a vexing problem that has come to plague the study of morality ever since the philosopher David Hume identified what is known as the *Is-Ought Problem* (sometimes rendered as the *Naturalistic Fallacy,* or *Hume's Guillotine*), or the difference between *descriptive* statements (the way something *is*) and *prescriptive* statements (the way something *ought to be*). Here is how Hume described the problem:

> In every system of morality, which I have hitherto met with, I have always remarked, that the author proceeds for some time in the ordinary ways of reasoning, and establishes the being of a God, or makes observations concerning human affairs; when all of a sudden I am surprised to find, that instead of the usual copulations of propositions, is, and is not, I meet with no proposition that is not connected with an ought, or an ought not. This change is imperceptible; but is however, of the last consequence. For as this ought, or ought not, expresses some new relation or affirmation, 'tis necessary that it should be observed and explained; and at the same time that *a reason should be given*; for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it. But as authors do not commonly use this precaution, I shall presume to recommend it to the readers; and am persuaded, that this small attention would subvert all the vulgar systems of morality, and let us see, that the distinction of vice and virtue is not founded merely on the relations of objects, nor is perceived by reason.[57]

Most people take Hume to mean that there is a wall separating is from ought, and that science can have nothing to say about determining human values and morals. But if morals and values should not be based on the way things are—reality—then on what should they be based? When I use the word "is" I do not just mean what is *natural*—as in biological nature alone—I mean the reality of the "is"

under study.

When we undertake a study of war to understand its causes so that we may lessen its occurrence and attenuate its effects, this is an is-ought transition grounded in the true nature of war—and by nature I do not just mean the biological propensity (or not) of humans to fight. I mean all of the factors that go into the causes of war: biological, anthropological, psychological, sociological, political, economic, and the like.

Philosophers have grappled with the is-ought problem ever since Hume outlined it, and some have proposed solutions, such as John Searle's widely cited 1964 paper "How to Derive 'Ought' from 'Is,'" in which he proposes, for example, that the act of making a promise becomes the "is" and as such it becomes an obligation that one "ought" to fulfill.[58] In any case, note what Hume is actually saying: not that one *cannot* shift (however imperceptively) from "is" to "ought," but that one *should not do so* without providing a *reason*. Fair enough. As with any claim in science, reasons and evidence must be provided in support, or else a claim is little more than an assertion. This appears to be what Hume is recommending to his readers when he suggests "that this small attention [to giving reasons] would subvert all the vulgar systems of morality." To make sure I am not misreading Hume, or reading into him what I think *ought* to be there instead of what *is* there (and thereby be guillotined), I queried one of the world's foremost Hume scholars—the Oxford philosopher Peter Millican—on the matter. He explained:

> It's true that Hume didn't explicitly say the Is/Ought gap was unbridgeable, but his analysis of morals as founded on sentiment did imply that in a sense—moral statements are not matters of fact (or mere relations of ideas, for that matter). I think he would go much the same way as you do—instead of making out that morals can be worked out by logical thinking, he would want to understand it as a natural human phenomenon—which requires scientific understanding, and then have that inform the decisions we make about how to foster it. Those decisions, of course, will be informed by our natural sentiment, so there's an element of circularity here (we're not simply inferring matters of fact from other matters of fact—ethical judgment is playing a role), but if there's enough agreement on basics (like war is bad, trust is good) that doesn't prevent progress.[59]

Not preventing progress is a worthy goal of any project, regardless of its degree of blending is and ought, but we can do better still by applying the knowledge of the causes of moral progress to help bring it about. It is in this sense that this book is *descriptive* in that it describes what has unfolded over time as we have become more moral, and it is *prescriptive* in that it prescribes what we ought to do if we want to continue the trend.

Take homosexuality and same-sex marriage, the latest of the Rights Revolutions that are unfolding in our time. *Descriptively*, science tells us that human beings have an evolved, innate drive to survive and to flourish, and that one of the most necessary and primal requirements among the many preconditions for life, health, and happiness for most people is a loving bond with another human being. *Prescriptively*, we can say that granting only a select group of privileged people the right to fulfill this evolved need—while simultaneously depriving others of the same basic right—is immoral because it robs them of the opportunity to fulfill their essence as evolved sentient beings. This is true even if the case could be made (as it has been by those who oppose same-sex marriage) that such discriminatory practices are better for the group (in a type of utilitarian calculus where the sacrifice of the few is justified if it leads to the greater happiness for the greater number). It is still wrong because the *individual* is the moral agent, not the group. It is the individual who feels the sharp pain of discrimination, the sting of being excluded, and the insult of being treated differently under the law.

Science tells us why they feel this way and reason instructs us what to do about it if we want to continue the moral progress of the Rights Revolutions.

As well, social science shows that humans are naturally tribal and that we tend to exclude people simply because they're not one of us (however "us" is defined). So how can we thwart the mind's natural inclination to pigeonhole people into prejudicial categories that turn them into Others we can exclude, exploit, or kill? Scientific research gives us guidance. For example, studies show that straight people who know gay people as neighbors, friends, or work colleagues are less likely to hold bigoted and prejudicial views of homosexuality, and they are more likely to agree that gay people should be treated equally under the law and should be accorded equal rights, such as the right to marry. A 2009 Gallup study, for example, found that "when controlling for ideology, those who know someone who is gay or lesbian are significantly more supportive of gay marriage than are those of the same political persuasion who do not personally know someone who is gay or lesbian."[60] Thus the existence of plays, films, and television shows that feature LGBTQ actors; of literary and pop cultural references that portray homosexuality and non-normative gender expression in a positive light; of "coming out" campaigns; and of LGBTQ role models in politics, business, and sports—all of these are essential for awakening empathy and understanding, and thus for expanding the moral sphere ever outward.

A PUBLIC HEALTH MODEL OF MORAL SCIENCE

In proffering this rapprochement between the way things are and the way things ought to be—the interface of facts and values—I am doing nothing more than recognizing a trend that has been under way since the Enlightenment in taking the findings of science about the way the world is and applying them to the way we would like the world to be. There's a reason why social scientists—social psychologists, cognitive psychologists, evolutionary psychologists, anthropologists, sociologists, economists, political scientists, and criminologists—along with policymakers and politicians, have been amassing extensive databases and ethnographies, testing hypotheses, and crunching the numbers through models and theories related to violence, aggression, crime, war, terrorism, civil rights violations, and the like: we want to understand causes to effect changes.

This approach may be modeled on *public health*, defined as long ago as 1920 in a *Science* article as "the science and art of preventing disease, prolonging life, and promoting health through the organized efforts and informed choices of society, organizations, public and private, communities and individuals."[61] Public health science includes such fields as epidemiology, biostatistics, behavioral health, health economics, public policy, insurance medicine, occupational health, and others. If you want to know why the average person lives almost twice as long today as a century ago, look to public health. The *maximum life potential* (the age at death of the longest-lived member of the species) has not changed and remains at 120 years. *Life span* (the age at which the average person would die if there were no premature deaths from accidents or disease) has also not changed and remains at about 85 to 95 years. But *life expectancy* (the age at which the average individual would die when accidents and disease have been taken into consideration) has skyrocketed upward from 47 years in 1900 in the United States to 78.9 for all Americans born in 2010, and 85.8 for Asian American women.[62] The cause of this remarkable progress in both the quantity and quality of life is public health science and technology: flush toilets, sewers, and waste disposal technologies, clean

water, hand washing, antiseptic surgery, vaccinations, pasteurization, road traffic safety, occupational safety, family planning, nutrition and diet, and other measures, coupled to the epidemiological study of infectious diseases such as smallpox and yellow fever, chronic diseases such as cancer and heart disease, and disease prevention through these many techniques. The survival and flourishing of humans have progressed in the past century more than in all previous centuries combined. If you agree that it is better that millions of people no longer die of yellow fever and smallpox, cholera and bronchitis, dysentery and diarrhea, consumption and tuberculosis, measles and mumps, gangrene and gastritis, and many other assaults on the human body, then you have offered your assent that the way something *is* (diseases such as yellow fever and smallpox kill people) means we *ought* to prevent them through vaccinations and other medical and public health technologies.

This analogy between social problems and public diseases is not at the level of cause—crime, violence, war, and terrorism are not diseases in the medical sense of being an abnormal condition caused by the equivalent of a virus or bacteria. Instead the analogy is at the level of methodology: how we approach solving the problem by using the best tools of science, technology, and social policy available. Most crimes and acts of violence, along with wars and acts of terrorism, are not abnormal responses in a diseased state; most are normal responses to specific situations and conditions. But the public health model of engaging numerous sciences to change the situations and conditions that cause them is a viable methodology toward the goal of making moral progress.

By way of example, in 2012 I undertook an extensive study of gun violence for a special issue of *Skeptic* magazine in response to the raft of mass public shootings in recent years, such as at the Sandy Hook elementary school.[63] Subsequently I engaged in a series of debates around the country on gun control with the economist John Lott,[64] whose book *More Guns, Less Crime* offers his own prescription in its title.[65] While conducting a literature search I was struck by how much data on gun violence are published in public health and medical journals. For example, Johns Hopkins University houses the Bloomberg School of Public Health, which is one of the leading centers in the country for research on gun violence, and in 2012 they published a scholarly study on the problem titled *Reducing Gun Violence in America*. The subtitle is illustrative of my point: *Informing Policy with Evidence and Analysis*. It showed, for example, that oversight of licensed gun dealers to make sure their sales were to customers without criminal records resulted in a 64 percent decrease in guns diverted to criminals.[66] A 1998 study in the *Journal of Trauma and Acute Care Surgery*, "Injuries and Deaths Due to Firearms in the Home," found that "every time a gun in the home was used in a self-defense or legally justifiable shooting, there were four unintentional shootings, seven criminal assaults or homicides, and 11 attempted or completed suicides." In other words, a gun is twenty-two times more likely to be used in a criminal assault, an accidental death or injury, a suicide attempt, or a homicide than it is for self-defense.[67] A 2009 study published in the *American Journal of Public Health* that was conducted by epidemiologists at the University of Pennsylvania School of Medicine found that on average not only did guns not protect those who possessed them from being shot in an assault, these scientists also determined that people with a gun were 4.5 times more likely to be shot in an assault than those not possessing a gun.[68]

The raw figures for deadly violence in general, and gun violence in particular, are staggering. If ever there were a public health problem that needs solving, this is it. According to the FBI's crime reports, between 2007 and 2011 the United States experienced an annual average of 13,700

homicides, with guns responsible for 67.8 percent of them.[69] That's an annual average of 9,289 people shot dead by a gun, or 774 a month, 178 a week, 25 a day, or a little more than 1 per hour. It's a disquieting thought that just in the United States, every hour of every day someone is shot to death. That fact alone should convince us of the value of understanding the causes that underlie it, but the problem is even worse: according to the National Center for Injury Prevention and Control, in 2010 a total of 19,392 US residents killed themselves with a firearm,[70] another 11,078 were shot to death, and 55,544 were injured by gunfire and treated in emergency rooms.[71]

The public health model can also be applied on a larger scale and time horizon by pulling back to see that, in fact, homicide rates have plummeted over the millennia, from almost 1,000 per 100,000 people per year in prehistoric times and in modern nonstate societies, to about 100 per 100,000 people per annum in Western societies through the Middle Ages, to about 10 per 100,000 each year by the time of the Enlightenment, to less than 1 per 100,000 today in Europe (and a little more than 5 per 100,000 in the United States), an improvement of four orders of magnitude. How do we know this? Science. Archaeologists can estimate the rates of violent deaths in prehistoric bands through skeletal remains (see chapter 2). The ethnographies of anthropologists record the rates of violence among modern prestate peoples from their oral accounts and histories. And historians have used old court and county records to calculate that homicide rates have, to give just one example, dropped from 110 homicides per 100,000 people per year in fourteenth-century Oxford to less than 1 homicide per 100,000 in mid-twentieth-century London. Similar patterns have been documented in Italy, Germany, Switzerland, the Netherlands, and Scandinavia, and by the same order of magnitude: from about 100 per 100,000 to less than 1 between the fourteenth century and the twenty-first century. *Figure 1-4* shows the unmistakable trend in the decline in homicides from the thirteenth century to the twentieth century in multiple countries.[72] The uptick in murders at the end of the long downward trends captures the crime wave of the 1970s and 1980s, but the rate returned to historical lows by the early years of the twenty-first century.

If that isn't moral progress as I have defined it—an increase in the survival and flourishing of sentient beings—I don't know what is. Regardless of your position on gun control, my point is that treating gun violence as a problem to be solved by better science and public policy is now common practice and part of the long-term trend to address moral issues pragmatically. If you agree that the hundreds of millions of lives that have been saved by public health sciences, technologies, and policies over the past two centuries is a moral good, then there is no reason why you might not also concur that applying social sciences to solving problems such as crime and violence is also something we ought to do. Why? Because saving lives is moral. Why is saving lives moral? Because the survival and flourishing of sentient beings is our moral starting point. But why should organisms *want* to survive and flourish in the first place? The answer may be found in the logic of the evolutionary process that created this drive.

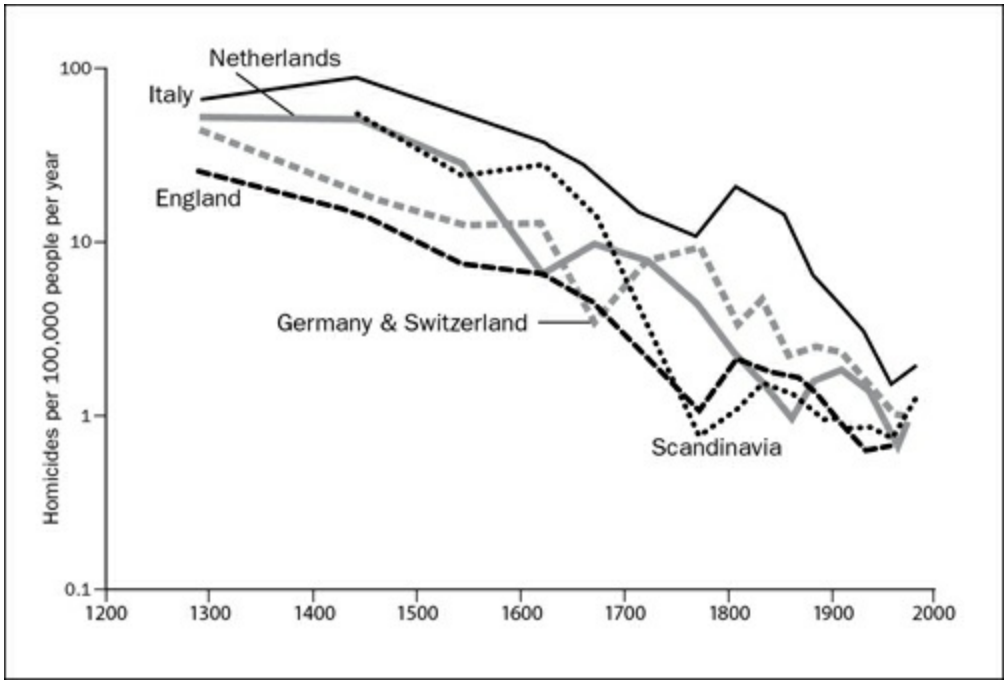

Figure 1-4. The Decline in Homicides
Homicide rates per 100,000 people per annum in five Western European regions from the thirteenth century to the twentieth century compiled by the criminologist Manuel Eisner.[73]

THE EVOLUTIONARY LOGIC OF THE MORAL STARTING POINT

If you were a molecule, what would you do to survive? First you would need to build a substrate on which to generate a replication system inside a cell that contains machinery for energy consumption, maintenance, and repair, and other features that keep the molecule intact long enough to reproduce. Once such molecular machinery is up and running, the replicating molecules become immortal as long as there is energy to feed the system and an ecosystem in which these processes can take place. In time these replicating molecules will outsurvive nonreplicating molecules by virtue of the very process of replication—those that don't, die—thus the cells or bodies in which the replicators are housed are survival machines. In modern jargon, the replicators are called genes and the survival machines are called organisms, and this little thought experiment is what Richard Dawkins means by the "selfish gene" in his book of that title.[74] A cell, or body, or organism—a survival machine—is the gene's way of surviving and perpetuating itself. Genes that code for proteins that build survival machines that live long enough for them to reproduce will win out over genes that do not. Genes that code for proteins and enzymes that protect its survival machine from assaults such as disease help not just the organism to survive, but the genes as well. Survival, reproduction, flourishing: this is what survival machines do by their very nature. It is their—our—essence to strive to survive.

The problem is that survival machines scurrying around in, say, a liquid environment such as an ocean or pond will bump into other survival machines, all of whom are competing for the same limited resources. "To a survival machine, another survival machine (which is not its own child or another close relative) is part of its environment, like a rock or a river or a lump of food," says Dawkins. But there's a difference between a survival machine and a rock. A survival machine "is inclined to hit back" if exploited. "This is because it too is a machine that holds its immortal genes in trust for the future, and it too will stop at nothing to preserve them." Thus, Dawkins concludes, "Natural selection favors genes that control their survival machines in such a way that they make the

best use of their environment. This includes making the best use of other survival machines, both of the same and of different species."[75] Survival machines could evolve to be completely selfish and self-centered, but there is something that keeps their pure selfishness in check, and that is the fact that other survival machines are inclined "to hit back" if attacked, to retaliate if exploited, or to attempt to use or abuse other survival machines first.

So in addition to selfish emotions that drive survival machines to want to hoard all resources for themselves, they evolved two additional pathways to survival in interacting with other survival machines: *kin altruism* ("blood is thicker than water") and *reciprocal altruism* ("I'll scratch your back if you'll scratch mine"). By helping its genetically related kin, and by extending a helping hand to those who will reciprocate its altruistic acts, a survival machine is helping itself. Thus there will be a selection for those who are inclined to be altruistic—to a point. With limited resources, a survival machine can't afford to help all other survival machines, so it must assess whom to help, whom to exploit, and whom to leave alone. It's a balancing act. If you're too selfish, other survival machines will punish you; if you're too selfless, other survival machines will exploit you. Thus, developing positive relationships—social bonds—with other survival machines is an adaptive strategy. If you are there to help your fellow group members when times are tough for them, they are more likely to be there when times are tough for you.

In this way survival machines develop networks and relationships that lead to interactions with one another that in addition to being neutral may be helpful or hurtful. From this we may derive the logic of moral emotions. In a social species such as ours, sometimes the most selfish thing you can do to help yourself is to help others, who will pay you back in kind, not necessarily out of some nebulous notion of "altruism" for its own sake, but because it pays to help others. Our legacy is a moral emotion system that includes the capacity for us to help and hurt other survival machines, depending on what they do. Sometimes it pays to be selfish, but other times it pays to be selfless as long as you're not a milquetoast who lies down and lets others run roughshod over your generosity.

THE EVOLUTIONARY LOGIC OF EMOTIONS

The language I'm using to describe these interactions makes it sound like it is a rational process, a moral calculation conducted by survival machines as they interact with one another. But that is not what is going on. Organisms are driven by their passions more than they are by rational calculations. Natural selection has done the calculating for organisms, who evolved emotions as proxies for those calculations. Let's drill down deeper into the brain to figure out why we evolved emotions in the first place, and then pull back to see how moral emotions work.

Emotions interact with our cognitive thought processes to guide our behaviors toward the goal of survival and reproduction. The neuroscientist Antonio Damasio has shown that at low levels of stimulation, emotions act in an advisory role, carrying additional information to the decision-making process along with input from higher-order cortical regions of the brain. At medium levels of stimulation, conflicts can arise between high-road reason centers and low-road emotion centers. At high levels of stimulation, low-road emotions can so overrun high-road cognitive processes that people can no longer reason their way to a decision and report feeling "out of control" or "acting against their own self-interest."[76]

The emotion of *fear*, for example, directs an organism to steer clear of danger. The anthropologist

Björn Grinde, a rock climber himself, uses the sport as an example of a situation where the positive emotion of the thrill of risk taking can quickly shift to the negative emotion of fear of death when a climber loses his or her grip: "The brain is designed to induce us to take some chances, otherwise we would never have laid down a large prey or ventured into uncharted land; but it is also designed to stop us from causing harm to ourselves, that is, to avoid hazards. The 'adrenaline kick' associated with climbing a mountain or riding a roller coaster may feel good, presumably because it improves the chance of survival if voluntarily encountered dangerous situations induce a positive mood and a high self-esteem. At the moment one loses the grip on the mountain, the unpleasant sensations devoted to harm avoidance kick in."[77] The survival value of fear has obvious adaptive value.

Consider *hunger* as a motivating drive that leads to such emotions as desire, longing, or lust for food. A little bit of hunger may be perceived as mildly pleasant in the anticipation of eating, and we've evolved to understand these small pangs to mean we should seek and find food. If too much time goes by, however, and our body becomes depleted to the point of feeling weak, hunger morphs into an unpleasant feeling. By this example, emotions act as a feedback mechanism to alert the brain when the body is out of balance. This is a *homeostatic* theory of emotions, in which the process operates like an emotional thermostat. When our bodies are low on energy we feel hungry, and that emotion is triggered by a number of internal and external feedback cues—such as shrinking or distension of the stomach, elevated or reduced blood glucose levels, or the sight or smell of food—which act as cues to drive us to raise the calorie thermostat, or bring the body back into homeostasis by eating.

In some cases it is literally like a thermostat, as when our core body temperature deviates above or below the 98.6 degree Fahrenheit set point and certain physiological systems kick in to correct the imbalance, such as sweating to cool the body or shivering to warm it. Departing from the set point of a homeostatic system *feels bad*, and this negative emotion motivates the organism to take action to correct the imbalance. Moving the out-of-balance system back toward homeostasis *feels good*, and behaviors that feel good tend to be repeated, which is the very definition of reinforcement (anything that causes an organism to repeat a behavior).

Thus our need to maintain homeostasis is caused by our emotions that direct us to avoid pain and to pursue pleasure, or to approach a stimulus that is reinforcing and to avoid a stimulus that is punishing. In this sense, in seeking something we call "pleasure," what we are really after is a homeostatic cue, an information signal to tell us what to do next. When those cues are unclear, or in conflict, it can set up a state that psychologists call approach-avoidance behavior. In the case of apparent moral dilemmas—such as what subjects experienced between obedience to authority and the discomfort of harming a fellow human in Stanley Milgram's famous shock experiments—the phenomenon becomes an *approach-avoidance conflict*. (More on this in chapter 9.)

THE EVOLUTIONARY LOGIC OF AGGRESSIVE EMOTIONS

Conflicts among survival machines are inevitable by-products of the evolutionary logic of the essence to survive and flourish and the many different ways there are to fulfill that need in an environment of limited resources. This approach helps us see that there is a certain evolutionary logic to violence and aggression, a taxonomy of which Steven Pinker classified into five types:[78]

1. *Predatory and instrumental*: violence as a means to an end, a way of getting something you want. Theft, for example, can grant the thief more resources necessary for survival and reproduction, and thus there evolved a capacity for cheating, stealing, and free riding (taking without giving in a social system) among some individuals in a group.
2. *Dominance and honor*: violence as a means of gaining status in a hierarchy; power over others; prestige in a group; or glory in sports, gangs, or war. Bullying, for example, can grant individuals higher status in the pecking order of social dominance.[79] A reputation for being aggressive can be a credible deterrent against other aggressors.
3. *Revenge and self-help justice*: violence as a means of punishment, retribution, or moralistic justice. Revenge murders, for example, are an evolved strategy for dealing with cheaters and free riders. *Jealousy* is another type of moralistic emotion that evolved to direct survival machines to mate guard against potential poachers of their sexual partner (and thus, for men, the bearer of their children who carry their genes), which when expressed violently can lead to spousal murders. Even *infanticide* has an evolutionary logic to it, as evidenced by the statistic that infants are fifty times more likely to be murdered by their stepfather than their biological father, an act far more common among species—including our own—than we care to admit.[80]
4. *Sadism*: violence as a means of gaining pleasure at someone else's suffering. Serial killers and rapists, for example, seem at least partially motived by the pain and suffering they cause, especially when it has no other apparent motive (such as instrumental, dominance, or revenge). It is not clear if sadism is adaptive or, more likely, is a by-product of something else in the brain that evolved for some other reason.
5. *Ideology*: violence as a means of attaining some political, social, or religious end that results in a utilitarian calculus whereby killing some for the sake of many is justified. (I deal with this cause of violence in detail in chapter 9.)

THE EVOLUTIONARY LOGIC OF MORAL EMOTIONS

On the platform of a subway station a woman and two men are standing a few feet away from the open track pit, when all of a sudden one of the men reaches forward and shoves the woman by her shoulders. She staggers backwards, loses her footing, and starts to fall toward the pit. The other man reaches out to catch her but he's too late—into the pit she goes. In an instant he reacts. But instead of reaching down to pull the woman to safety before a train arrives to crush her, the would-be rescuer turns on his heels and cold-cocks the perpetrator. It's a magnificent roundhouse blow right out of a Hollywood movie that snaps the perpetrator's head back with the sweet crash of fist against chin. Satisfied with this act of revenge, the rescuer steps back and pauses for a moment, then appears to remember what needs to be done next, dashes over to the pit, and pulls the woman to safety. He says something to her that appears reassuring, and then takes off after the perpetrator, who had beat a hasty retreat through an open door. The entire incident takes twenty seconds, and you can see it yourself in a viral video that includes a number of heroic rescues.[81] In that brief moment—too short a span for rational calculation—a conflict of pure emotional morality unfolds between revenge and rescue, hurting and helping. In a flash two neural networks in the rescuer's brain launched into action—help a fellow human in trouble or punish the person who caused the trouble. What's a morally motivated

primate to do? In this case, the rescuer had time to do both as no train arrived to derail his problematic first choice to avenge the woman's maltreatment. Revenge is sweet and so is rescue. It doesn't always work out so well.

This vignette well illustrates our multifaceted moral nature that evolved to solve several problems at once in our ancestral environment—be nice to those who help us and our kin and kind, punish those who hurt. Evidence that these moral emotions are deeply entrenched in human nature may be found in a series of experiments with babies, succinctly synthesized in the psychologist Paul Bloom's book *Just Babies: The Origins of Good and Evil*.[82] Testing the theory that we have an innate moral sense as proposed by such Enlightenment thinkers as Adam Smith and Thomas Jefferson, Bloom provides experimental evidence that "our natural endowments" include "a moral sense—some capacity to distinguish between kind and cruel actions; empathy and compassion—suffering at the pain of those around us and the wish to make this pain go away; a rudimentary sense of fairness—a tendency to favor equal divisions of resources; a rudimentary sense of justice—a desire to see good actions rewarded and bad actions punished."[83] Consider an experiment conducted in Bloom's lab with a one-year-old baby who watched a puppet show in which one puppet rolls a ball to a second puppet, who passes the ball back to it. The first puppet then rolls the ball to a different puppet, who runs off with the ball. Next, the "nice" and the "naughty" puppets are placed before the baby, along with a treat in front of each; the baby is then given the choice of which puppet to take the treat away from. As Bloom predicted, the infant removed the treat from the naughty puppet—which is what most babies do in this experimental paradigm—but for this little moralist, removing a positive reinforcement (the treat) was not enough. In his inchoate moral mind, punishment was called for, as Bloom recounts: "The boy then leaned over and smacked this puppet on the head."[84]

Numerous permutations on this research paradigm (such as a puppet trying to roll a ball up a ramp, for which another puppet either helps or hinders it) show time and again that the moral sense of right (preferring helping puppets) and wrong (abjuring hurting puppets) emerges as early as three to ten months of age—far too early to attribute to learning and culture.[85] Young children who are exposed in a laboratory to an adult experiencing pain—the experimenter getting her finger caught in a clipboard, say, or the child's mother banging her knee—typically respond by soothing the injured party. Toddlers who see adults struggling to open a door because their arms are full, or to pick up an out-of-reach object, will spontaneously help without any prompting from the adults in question.[86] Another experiment involved three-year-old children who were asked, "Can you hand me the cup so that I can pour the water?" but the cup in question was broken. Remarkably, the youngsters spontaneously went in search of an intact cup to help the experimenter complete the task.[87]

However, children are not always so beneficent, particularly with other children, with whom they clearly show awareness of an unequal distribution of rewards after a shared task (in this case a candy treat), but are not always so eager to unselfishly right the wrong by redistributing the wealth.[88] But as children get older—from three-to-four-year-olds to seven-to-eight-year-olds—they are not only more aware of an unequal and unfair distribution of candy, they are also more likely to give away the extra unearned treat (50 percent of the three-to-four-year-olds did, compared to 80 percent of the seven-to-eight-year-olds), showing that while the moral sense is inborn and instinctive, it is a capacity that can be tuned by learning and culture and brought to bear (or not) in different environments that either encourage or discourage helping or hurting behavior.[89]

As well, research with infants shows how early in life xenophobia takes root. Babies become wary of strangers, or anyone who doesn't look like members of their family on whom they've imprinted, at a very early stage—days, in fact. In one experiment, three-day-old newborns were donned with headphones and special pacifiers that allowed them to pick audio tracks based on how rapidly they sucked on them. These infants not only figured out the connection between sucking and music selections, but also were able to transfer that learned skill to selecting a passage read to them from a Dr. Seuss book by their mother rather than a stranger. For newborns given the option to select among languages being spoken, results showed that "Russian babies prefer Russian, French babies prefer French, and American babies prefer English, and so on," and even more remarkably, says Bloom, "This effect shows up mere minutes after birth, suggesting that babies were becoming familiar with those muffled sounds that they heard in the womb."

This research confirms a classic experiment from the 1960s conducted by a third-grade schoolteacher named Jane Elliott on her class in the small, all-white rural town of Riceville, Iowa. Elliott began her experiment by dividing her class into two groups by eye color—blue and brown—then presented the kids with examples of blue-eyed good people and brown-eyed bad people. In addition, the blue-eyed kids in the class were told that they were superior and were given special privileges, while the brown-eyed kids were called inferior and treated as second-class citizens. Almost immediately a social division followed the physical classification. The blue-eyed kids quit playing with the brown-eyed kids, with some of the students even suggesting to Elliott that school officials should be alerted to the potential criminal behavior of brown-eyed youngsters. When a fight broke out between a blue-eyed and a brown-eyed boy, the latter justified his aggressive actions thusly: "He called me brown-eyes, like being a black person, like a Negro." By the second day of the experiment, the brown-eyed children already started to show signs of poorer performance in class and described themselves as feeling "sad," "bad," "stupid," and "mean."

As a control, the next day Mrs. Elliott reversed the conditions, explaining that she had been mistaken and that, in fact, it was the *brown-eyed* children who were superior, and the blue-eyed kids who were inferior. Just as quickly the self- and other perceptions were reversed, with the "happy," "good," "sweet," and "nice" labels previously used by the blue-eyed children to describe themselves now adopted by the brown-eyed children. "What had been marvelously cooperative, thoughtful children became nasty, vicious, discriminating little third-graders," Mrs. Elliott explained. "It was ghastly!"[90]

Bloom's conclusion about morality from this sizable body of research supports what I saw in the video vignette from the subway: "it entails certain feelings and motivations, such as a desire to help others in need, compassion for those in pain, anger towards the cruel, and guilt and pride about our own shameful and kind actions."[91] Of course, society's laws and customs can turn the moral dials up or down, but nature endowed us with the dials in the first place. It's as Voltaire said: "Man is born without principles, but with the faculty of receiving them. His natural disposition will incline him either to cruelty or kindness; his understanding will in time inform him that the square of twelve is a hundred and forty-four, and that he ought not to do to others what he would not that others should do to him."[92]

THE LOGIC OF MORAL DILEMMAS

The logic of our moral emotions has been worked out by game theorists in the aforementioned Prisoner's Dilemma paradigm. Here's the scenario: You and your partner are arrested for a crime, and you are held incommunicado in separate prison cells. Neither of you wants to confess or rat on the other, but the DA gives each of you the following options:

1. If you confess but the other prisoner does not, you go free and he gets three years in jail.
2. If the other prisoner confesses and you do not, you get three years and he goes free.
3. If you both confess, you each get two years.
4. If you both remain silent, you each get a year.

Figure 1-5, called a game matrix, summarizes the four outcomes.

With those outcomes, the logical choice is to defect and betray your partner. Why? Consider the choices from the first prisoner's point of view. The only thing the first prisoner cannot control about the outcome is the second prisoner's choice. Suppose the second prisoner remains silent. Then the first prisoner earns the "temptation" payoff (zero years in jail) by confessing but gets a year in jail (the "high" payoff) by remaining silent. The better outcome in this case for the first prisoner is to confess. But suppose, instead, that the second prisoner confesses. Then, once again, the first prisoner is better off confessing (the "low" payoff, or two years in jail) than remaining silent (the "sucker" payoff, or three years in jail). Because the circumstances from the second prisoner's point of view are entirely symmetrical to the ones described for the first, each prisoner is better off confessing no matter what the other prisoner decides to do.

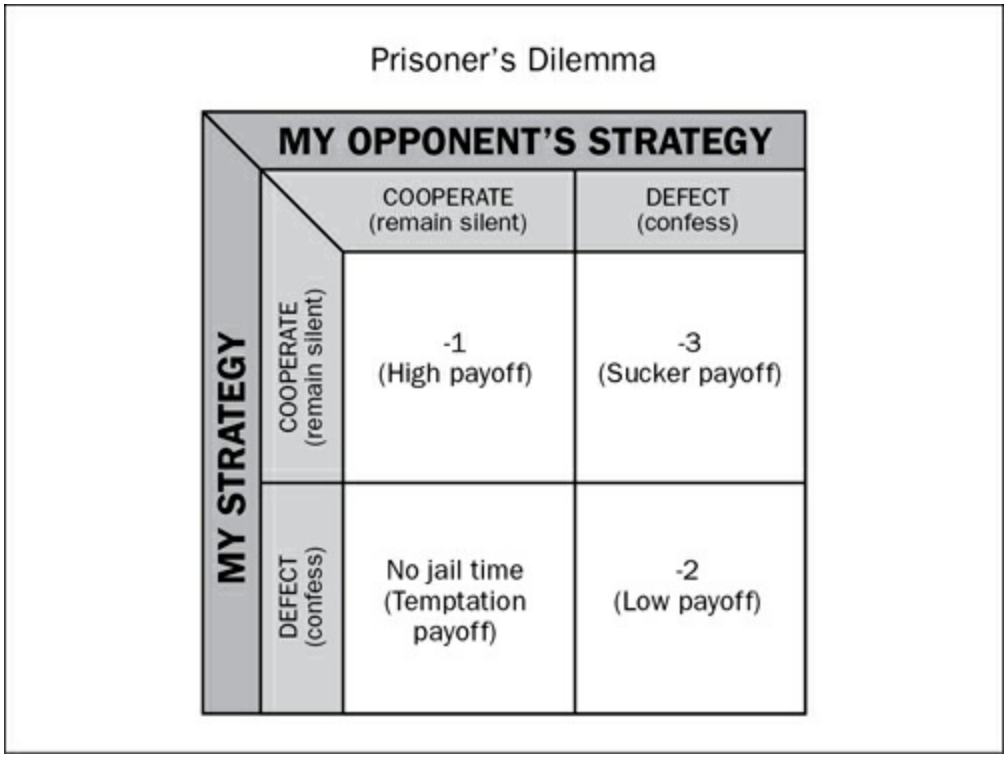

Figure 1-5. Prisoner's Dilemma

Those preferences are not only theoretical. When test subjects play the game just once or for a fixed number of rounds without being allowed to communicate, defection by confessing is the common strategy. But when testers play the game for an unknown number of rounds, the most common

strategy is tit for tat: each begins cooperating with the prior agreement by remaining silent, then mimics whatever the other player does. Even more mutual cooperation can emerge in a many-person prisoner's dilemma, provided the players are allowed to play enough repeated rounds to establish mutual trust. But the research shows that once defection by confessing builds momentum, it cascades throughout the game.

In an analysis for *Scientific American* I worked out the game matrix dynamics of why professional athletes use performance-enhancing drugs, and in particular why cyclists dope.[93] In cycling, as in any sport, the contestants compete according to a set of rules. The rules of cycling clearly prohibit the use of performance-enhancing drugs. But because the drugs are so effective and many of them are so difficult (if not impossible) to detect, and because the payoffs for success are so great, the incentive to use banned substances is powerful. Once a few elite riders defect from the rules by doping to gain an advantage, their rule-abiding competitors feel the need to defect as well—even if they don't want to—leading to a cascade of defection through the ranks. Because of the penalties for breaking the rules, however, a code of silence prevents any open communication about how to reverse the trend and return to abiding by the rules. *Figures 1-6* and *1-7* show the game matrices that favor cheating and that favor playing by the rules.

In game theory, if no player has anything to gain by unilaterally changing strategies, the game is said to be in a Nash equilibrium. The concept was developed by the mathematician John Forbes Nash Jr., who was portrayed in the film *A Beautiful Mind*. To end cheating in sports the doping game must be restructured so that competing clean is in a Nash equilibrium. That is, the governing bodies of each sport must change the payoff values of the expected outcomes identified in the abiding-by-the-rules matrix. First, when other players are playing by the rules, the payoff for doing likewise must be greater than the payoff for cheating. Second, and perhaps more importantly, even when other players are cheating, the payoff for playing fair must be greater than the payoff for cheating. Players must not feel like suckers for following the rules. In the game of Prisoner's Dilemma, lowering the temptation to confess and raising the payoff for keeping silent if the other prisoner confesses increases cooperation. Giving players the chance to communicate before they play the game is the most effective way to increase their cooperation. In sports, that means breaking the code of silence. That will show each player that the payoff for playing fair is greater than the payoff for cheating, no matter what the other players do.

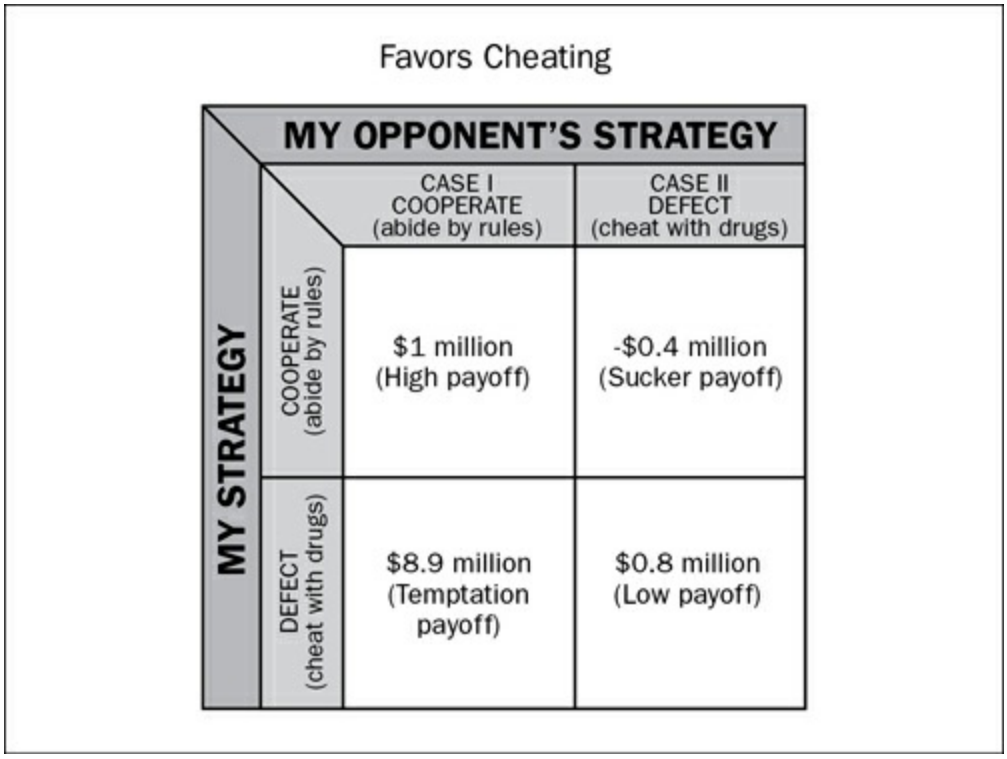

Figure 1-6. Prisoner's Dilemma Matrix for Cheating

The Cheating Matrix assumptions: Value of winning the Tour de France: $10 million. Likelihood that a doping rider will win the Tour de France against nondoping competitors: 100 percent. Value of cycling professionally for a year when the playing field is level: $1 million. Cost of getting caught cheating (penalties and lost income): $1 million. Likelihood of getting caught cheating: 10 percent. Cost of getting cut from a team (forgone earnings and loss of status): $1 million. Likelihood that a nondoping rider will get cut from a team for being noncompetitive: 50 percent. Under these conditions, in case 1 in which my opponent abides by the rules (he "cooperates"), if I also cooperate by not doping, the playing field is level and there is an expected payoff of $1 million. But if I cheat by doping and don't get caught, then I stand to make $8.9 million ($10 million × 90 percent − $0.1 million), which is more than $1 million, so I should cheat. In case 2, in which my opponent cheats by doping, if I play by the rules I'm a sucker and lose $0.4 million, but if I also cheat by doping then I too face the low payoff amount of $0.8 million, so my incentive is once again to cheat.

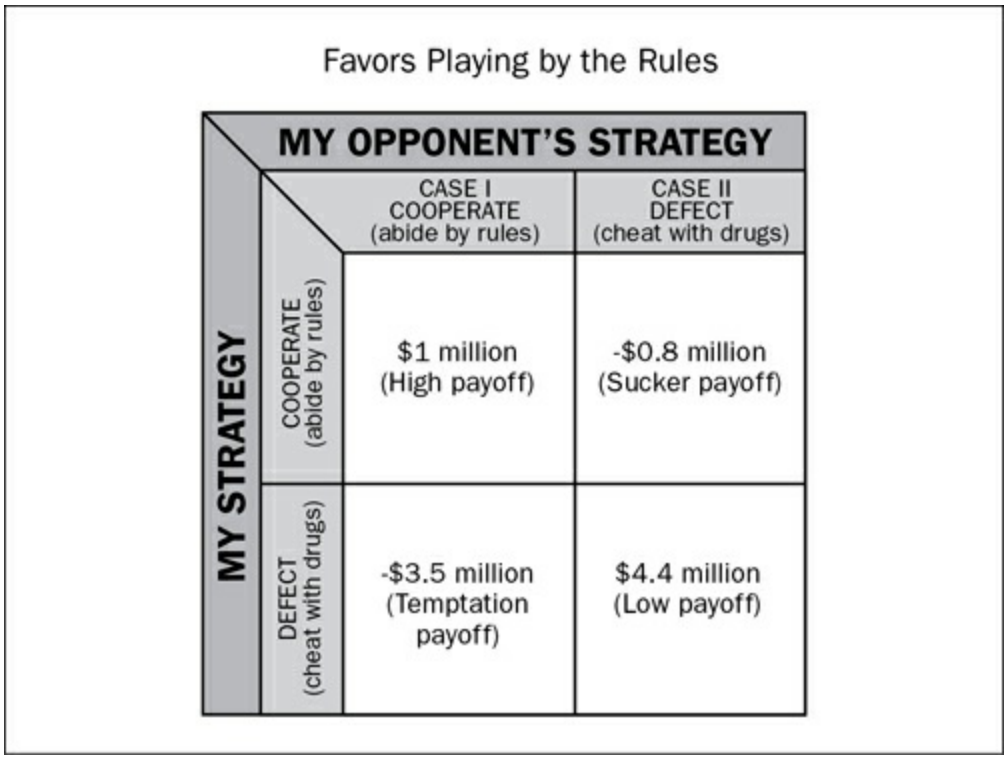

Figure 1-7. Prisoner's Dilemma Matrix for Abiding by the Rules

The Playing by the Rules Matrix assumptions: New, higher cost of getting caught cheating (penalties and lost income): $5 million. New, higher likelihood of getting caught cheating: 90 percent. Consequent new, lower likelihood that a nondoping rider will get cut from a team

for being noncompetitive: 10 percent. Under these conditions, in case 1, in which my opponent abides by the rules (he "cooperates"), if I also cooperate by not doping, the playing field is level and there is an expected payoff of $1 million. But this time, if I cheat by doping there's a 90 percent chance I'll get popped in a drug test, so my expected payoff for cheating is now $1 million minus the expected penalty for cheating of $5 million × 90 percent = −$4.5 million, so I stand to lose $3.5 million, so the incentive is to play by the rules. Even in case 2, in which my opponent dopes and I'm a sucker for cooperating, I still come out on top with a net $0.8 million, compared to my also doping and getting caught and penalized, resulting in a net loss of $4.4 million. Either way, in this matrix, with these conditions, we should all play by the rules.

Whether this has happened in cycling is unclear, but I am encouraged by the startling events of 2012 and 2013 when Tyler Hamilton broke the code of silence in his book *The Secret Race*, and exposed the most sophisticated doping program in the history of sports, orchestrated by his teammate Lance Armstrong, the seven-time Tour de France winner, now stripped of his titles after a thorough investigation by the US Anti-Doping Association.[94] Hamilton revealed how such an elaborate system was maintained through a combination of a code of silence that led everyone to believe that everyone else believed that doping was the norm, which was then reinforced by the threat of punishment for speaking out or not complying. Since then it has been revealed that most athletes who have been caught doping say they didn't want to dope but that they did so out of the belief that *everyone else* was doping, and out of fear of retaliation if they didn't dope, and worse consequences still if they blew the whistle on the system.

As for real prisoners who find themselves in a Prisoner's Dilemma, the abiding-by-the-rules matrix also directs criminologists and policymakers to consider not just the size of the penalty (as most get-tough-on-crime politicians are wont to do) but the probability of getting caught as a factor as well. This harkens back to the eighteenth-century philosopher and reformer Cesare Beccaria, whose 1764 work *On Crimes and Punishments*—a high-water mark of the Italian Enlightenment—launched the movement to apply rational principles to criminal reform, such as adjusting the punishments to fit the crimes (proportionality) instead of, as was the custom of the day, the death penalty for such offenses as poaching, counterfeiting, theft, sodomy, bestiality, adultery, horse theft, being in the company of Gypsies, and two hundred other crimes and misdemeanors. Beccaria opposed the death penalty on two principles: (1) states do not possess the right over life and death, and (2) it doesn't work to deter crime because when would-be criminals are faced with a draconian but improbable penalty, they consider it a risk worth taking—another cost of doing business. To the principles of *proportional* and *probable* punishment, Beccaria added two more Ps: criminal proceedings should be *prompt* and *public*, the latter acting as a signal to other would-be criminals. Beccaria was an early game theorist applying rational principles born of Enlightenment values and observational data from real-world examples—with an end toward tilting the motivational matrices to incentivize the citizenry to commit fewer crimes.[95]

Another Enlightenment thinker we will meet later is Thomas Hobbes, who also proffered a game theoretic model for how people and nations interact. All political theorists—both liberal and conservative—begin with a Hobbesian premise that the state is a necessary evil to protect self-motivated individuals from other self-motivated individuals (think of it as two survival machines in Dawkins's thought experiment). This is sometimes known as the "Hobbesian trap." As Hobbes argued in his classic work in political theory *Leviathan*, we are all motivated to seek pleasure and avoid pain, so inevitably there will be conflict when interests between people overlap. This leads to three forms of "quarrel": competition, diffidence (fear), and glory (honor, status):

> The first maketh men invade for gain; the second, for safety; and the third, for reputation. The first use violence, to make themselves masters of other men's persons, wives, children, and cattle; the second, to defend them; the third, for trifles, as a word, a smile, a different opinion, and any other sign of undervalue, either direct in their persons or by reflection in their kindred, their friends, their nation, their profession, or their name.[96]

As we saw in the Prisoner's Dilemma, for there to be cooperation between competing agents in a game (or nations in the real world), there need to be rules, and the rules must be enforced. With our complex moral nature, people need to be encouraged to do the right thing and discouraged from doing the wrong thing—the proverbial carrots and sticks. The psychology behind this interaction between inner psychological states and external social conditions was explored by the economists Ernst Fehr and Simon Gachter in a study on *moralistic punishment* in which subjects were given the opportunity to punish others who refused to cooperate in a group activity that calls for altruistic giving. They used a "common goods" cooperation game in which the subjects were given money that they then had the option to put into a shared commons that would then be multiplied 1.5 times and divided equally among all the players. Let's say the amount is $10 and there are four players. If everyone puts in the full ten bucks, then $40 × 1.5 = $60, which when split four ways equals $15 each. When this is done anonymously there is a temptation to game the system by putting in less money. Let's say the other three players each put in their full $10, but I put in only $5. The commons now has $35 × 1.5 = $52.5, which divided four ways equals $13.12 each. But I still have my original $5, so I now have $18.12. Sweet! But it doesn't take long for the other players to catch on to the fact that someone is gaming the system, and under these conditions cooperation quickly breaks down and the amount of money put into the commons collapses.

To remedy the free rider problem, in the seventh round Fehr and Gachter introduced a new condition in which contributions to the commons were no longer anonymous and the players were allowed to punish free riders by taking money from them, and this they did with impunity, which immediately triggered a rise in the levels of cooperation and giving by the former free riders.[97] The results, shown in *figure 1-8*, serve as a visual reminder of why people need rules, transparency, and the threat of punishment to be good. This role, the theory goes, is fulfilled by the Leviathan state.

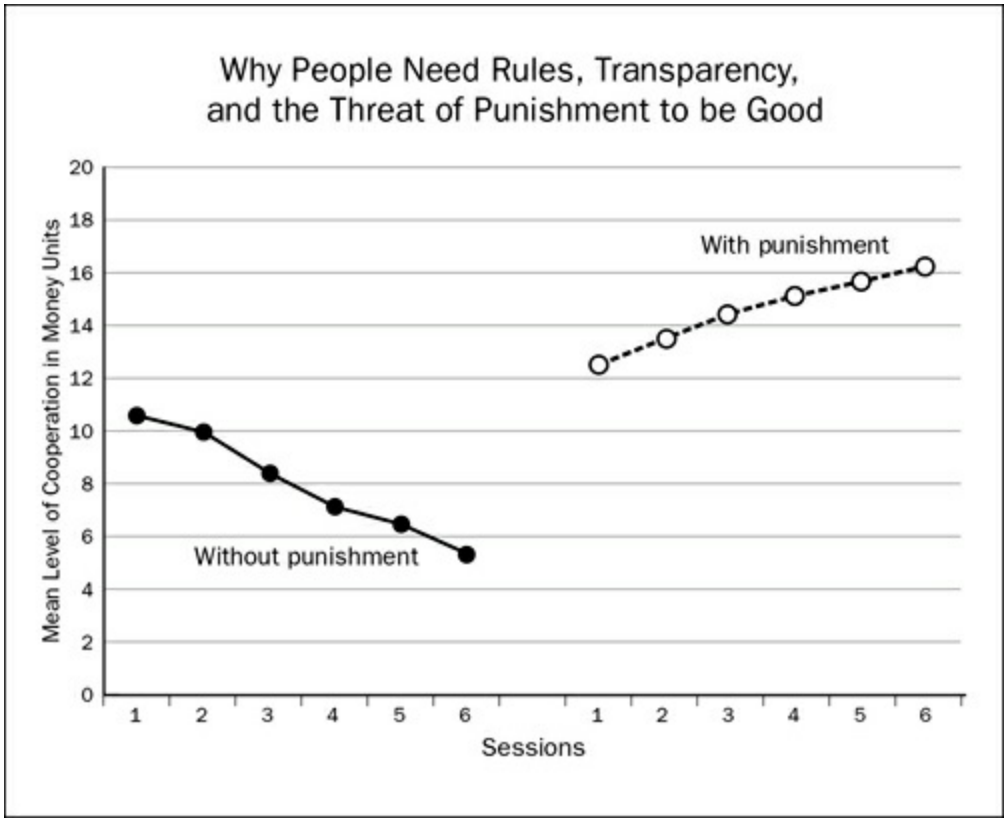

Figure 1-8. Why People Need Rules, Transparency, and the Threat of Punishment

The results of the Fehr and Gachter study on moralistic punishment. In a public goods game players are given a sum of money and have the choice of how much they would like to contribute into a common pool that will then be increased 1.5 times and returned to all the players evenly. A lack of transparency of how much everyone contributes leads to the temptation to reduce the amount given and thereby "free ride" on the others. Since all players face the same temptation, cooperation declines. When transparency is included, plus the opportunity to punish free riders who scrimp on their contributions, cooperation increases. The latter condition is an example of moralistic punishment. It works.

* * *

In this chapter I have outlined the evolutionary origins of morality and the logic of moral interactions, good and bad, and in the next chapter I will show how these principles operate to lessen even the most dangerous threats our species faces—violence, war, and terrorism—and will demonstrate that even here, there has been considerable moral progress.