Jason Blakely
Excerpts from *We Built Reality: How Social Science Infiltrated Culture, Politics, and Power*

So efforts at descriptive scientific theory, when applied to human beings, can actually produce new identities, practices, and worlds of meaning. This is due to humans' uniquely creative meaning-making capacities and is what philosophers refer to as the "double hermeneutic effect," in which an interpretation of the world shapes the very interpretations that comprise it. Throughout the book I call these "double-H effects" for short. Double-H effects make social science profoundly unlike the natural sciences, where the objects of study exist in a certain splendid seclusion and isolation. When a Ptolemaic astronomer places Earth at the center of the cosmos, the sun and the planets do not suddenly swivel violently, modifying their placement to match the theories on the page. Yet in the social sciences, the equivalent of Ptolemaic and Galilean astronomers change the basic social coordinates and field of objects with great frequency, in ways both intended and unintended by the theorists.

Much of this book is dedicated to identifying, diagnosing, and critically analyzing the social scientific double-H effects that created the world we currently inhabit. Viewed from the perspective of the double-H effect, much of social science (and particularly in its popular, vulgarized forms) is not simply descriptive but also performative. Social science as a genre can be read not in its official guise of neutral efforts at description but as artifacts of culture that participate in enacting and inaugurating certain political realities. I hope readers, once they view social scientific theories through these eyes, will be able to see that they are often slippery, escaping the hands of their creators and turning politically ambiguous, sometimes even menacing. This is to suggest nothing less than a form of power and politics completely unknown to epochs prior to the scientific revolution. Whereas the abuse of, say, religious or familial authority was well known to premoderns, the abuse of power by scientists or rather by scientism did not exist. And yet this frequently unrecognized form of domination pervades our societies.

Few people perceive this pervasive form of power, because scientism offers itself as the public, official, neutral, and objective way of doing things. Indeed, in extreme form scientism even tries to actively ban or eliminate other ways of knowing and experiencing the world as prescientific and illegitimate. The humanities, history, literature, the arts, philosophy, and religion are all disparaged as a kind of soft or even magical thinking. Even in mild forms, a culture of scientism subtly marginalizes the liberal arts and the humanities. Fewer people find it "useful" to study such things in college or fund them in primary schools, let alone name humanities scholars to positions of counsel in government and policy.

In the pages that follow, I uncover the hidden underbelly of a culture of scientism and reveal how what often presents itself as social science is instead culture and power. In doing so I seek to provoke readers to think carefully about the ways they themselves evoke the authority of science in everyday ethical and political life. Modern people must develop a critical sensibility for when science has flipped into a form of meaning-making with political and power dimensions flowing through it. Modern people must become critical readers of their own scientific cultures. This book is an effort to break out of the paradigm that holds us captive and tells us we are only allowed to read social science in one officially sanctioned manner.

Our rightful pride in the natural sciences has created a uniquely modern blind spot. In our quest to turn science into the measure of all things, we have generated a new kind of irrationality. Science becomes irrational when its ideal of knowledge is extended beyond its proper bounds and applied in areas where it does not rightfully hold sway. What is needed is a deeper awareness of the value of the humanities and the sensitive, interpretive intelligence required to grasp human life. Intellectually we still need to recover a profounder sense of our own—and others'—humanity.

...

Something strange happened in science fiction dramas in the 2000s: the robots, increasingly played by actors, were depicted as more human than the humans. Two of the most popular examples of this, *Westworld* and *Ex Machina*, both obsessively played on the visual trick that the robots onscreen might later turn out to be humans and vice versa. In doing so they rejected the obvious mechanical androids of earlier sci-fi films, like Star Wars's C-3PO and R2-D2. Instead, in the opening of *Westworld* a male character asked his female host, "Are you real?" She replied, "Well, if you can't tell, does it matter?"

Such anxieties about machine turned man were not new to the history of film and dated back to Fritz Lang's brilliant 1920s classic Metropolis. But in Metropolis the entrancing Maschinenmensch played by Brigitte Helm was an unambiguously evil foil to her human counterpart. By contrast, *Westworld* and *Ex Machina* implied that robots might surpass humans in their very humanity. In other words, there might be no bright line dividing the humans from the machines anymore. Viewers of these films were being taught to imagine themselves as existing on a spectrum of robots.

No less strange is that *Westworld* and *Ex Machina* both popularized an academic theory of intelligence known as the Turing test. The famous mathematician Alan Turing had invented this test many decades earlier. Turing believed that a machine would count as artificially intelligent (i.e., as "AI") when it could deceive humans into believing it was a fellow human being in a blind conversation. *Ex Machina* went so far as to explicitly explain the Turing test to viewers before launching into a series of plot twists in which the audience's ability to distinguish the AIs from the humans was increasingly confounded. The film climaxed in a bloodbath in which human characters were left for dead in the wilderness while AIs self-emancipated and entered civilization. As the

fictional inventor of AI, a mad genius named Nathan, foretold in the movie's key dialogue: "One day the AIs are going to look back on us the same way we look at fossil skeletons . . . an upright ape living in dust with crude language and tools, all set for extinction."

The popularity of *Ex Machina* and *Westworld* coincides with a heated debate over the role of AI and computing machines in modern societies. The United States in particular is immersed in deep anxieties about mechanization, the effects of computer technologies on social life, the gig economy, and the loss of traditional forms of work to robotics. On one side of the debate are tech doomsayers, typified by business tycoon Elon Musk. Following the Swedish philosopher Nick Bostrom, Musk has warned his followers that AI robots are on the verge of making a great leap forward that will mark them off as an independent line of evolution, and that they will usurp human dominance of the globe. If we are not careful, Homo sapiens will be relegated to inferior species status (AI's "house cats," as Musk put it).

On the other side are tech boosters like Mark Zuckerberg, who publicly denounce Musk's "doomsday scenario" as "irresponsible." Zuckerberg is a longtime promoter of Silicon Valley technologies and their ability to improve human life. For Zuckerberg technologies like his own Facebook and AI are the positive result of human innovation. Zuckerberg and his followers believe AI will increase productivity, improve services, facilitate new job creation, and promote human flourishing. AI and a computerized economy are to be embraced.

From the perspective of interpretive philosophy, however, what makes this debate notable are not the differences between doomsayers and boosters but their unspoken, deep agreement. Specifically, both sides accept the basic premise that computational systems are capable of attaining and even superseding human intelligence. They achieve this tacit consensus, moreover, because of a shared faith in a metaphor crafted by researchers in the social sciences. This is the anthropological metaphor of *Homo machina*, or humans conceived as highly complex machines. Of course the metaphor (like all metaphors) has recognized limits: for instance, engineers intentionally design and create machines, while nature accidentally produces humans out of a process of random molecule mutation. But although not designed, according to this vision the human species is nonetheless a system of mechanics akin to engineered androids.

This anthropological metaphor has deep roots in the history of the social sciences. In fact, the idea that humans might be conceptualized as natural or "wet" machines is nearly as old as the scientific revolution itself and was born with the nascent social sciences. In the early 1600s, Galileo had already formulated the notion of a machine as a system of mechanical parts operated by input energy. At nearly the same time, philosophers who contributed to the founding of the modern social sciences, such as Thomas Hobbes, offered speculative machinist accounts of human nature, as in the opening pages of his masterwork *Leviathan*, in which he asked: "Why may we not say that all *Automata*

(Engines that move themselves by springs and wheels as doth a watch) have an artificial life? For what is the *Heart* but a *Spring*, and the *Nerves*, but so many *Strings*; and the *Joints* but so many *Wheels*, giving motion to the whole Body?"

A century later the French philosopher Julien Offray de La Mettrie repeated the machine metaphor in even blunter form in his influential tract, *Man a Machine*. La Mettrie had become convinced that all higher order human capacities were determined by the mechanics of muscles. Indeed, according to La Mettrie, "the human body is a watch" that "winds its own springs," and "the brain has its muscles for thinking, as the legs have muscles for walking." In this way, Hobbes and Le Mettrie helped inaugurate a speculative tradition in the human sciences in which the latest developments of technology are used to reverse engineer the workings of human behavior.

More sophisticated theorizations of the machine metaphor were devised in the twentieth century. For example, the American psychologist B. F. Skinner led a widely influential research program known as behaviorism, which taught that human beings were a kind of stimulus-response contraption, continually shaped and triggered by the environment around them.3 More recently machinist anthropologists turned to neuroscience and the structures of the brain. Social scientists inspired by these discoveries even attempted to construct a total neuroscience of human behavior, inventing fields such as neuroeconomics, neuropolitics, neuroethics, and neurolaw. Defenders of these research programs believed all the social sciences would eventually become immersed in "the advancing tide of neuroscience."

But perhaps the most popular account of *Homo machina* to date (and certainly the model that forms the background of popular culture phenomena like Ex Machina and Westworld) was born out of a synthesis of cognitive psychology and computer science. Just as Hobbes and La Mettrie noticed that the state of art in technology was the mechanics of a watch and imagined humans accordingly, cognitive scientists working in the late twentieth century recognized that the state of art in technology was computers and conceptualized a computational view of humans. Harvard cognitive psychologist Steven Pinker is the most formidable popularizer of this anthropological metaphor, conferring upon it a broad cultural authority. In his 1997 book *How the Mind Works*, Pinker laid out the basic outlines of what cognitive psychologists called the "computational theory of mind."

According to Pinker, the computer revolution allowed psychologists to envision how mind, consciousness, and intelligence might be a feature of a system composed of "lifeless gumball-machine parts." The key breakthrough was Alan Turing's envisioning of a symbol-processing machine. Turing was the first to picture a machine that would combine the automatic triggers of traditional mechanics with the symbolic relationships of an algorithm. Such "Turing machines" could in principle execute a simple fixed set of steps like a recipe by initiating a basic series of causal mechanisms. As Pinker put it, the computational revolution was inspired by the insight that symbolic calculations could be

carried out by “arrangements of matter” that had “both representational and causal properties . . . that simultaneously carry information about something and take part in a chain of physical events.”

In cognitive psychology this meant that a crucial step in understanding the human mind was mapping the basic conceptual features of computer technology in order to reverse engineer (at least in abstract theoretical terms) a human brain. Much of Pinker’s account of the mechanics of the human brain was an effort to speculatively imagine this organ of soft nerve tissue as a form of symbol processor. At the center of this metaphor was the view that the brain was the “hardware” of chemical processes, while the mind was the “software” of algorithmic steps or informational processing. Indeed, for Pinker the metaphor of a computer clarified the basic disciplinary boundaries for scientifically studying human beings. Neurobiology would focus on investigating the biological “hardware,” while psychology would inquire into the “mental software.”

Pinker declared that with the brain as a neurobiological computer and the mind as its software, ordinary people could grasp the basic metaphors of a universal science of human behavior. The self, society, language, morality, politics and the arts would all be decoded as evolutionary software, determined by neurological hardware. Ultimately cognitive neuroscience would show that “every aspect of our mental lives depends entirely on physiological events in the tissues of the brain.” The notion that human mind and intelligence were essentially algorithmic persisted in later popularizations of computational theory promoted by those computer scientists like Andrew Ng who taught massive online seminars on the engineering behind “neural networks” that allowed machines to execute tasks they were not explicitly programmed to achieve beforehand. This used the technology of machine learning (the basis for self-driving cars, photo and voice recognition, and other cutting-edge computer applications) to imply that human intelligence operated essentially in the same manner.

What is of primary importance from this discussion is not initially whether the computational theory of mind or the other social scientific theories backing *Homo machina* are correct or true to reality. Rather, the point for my analysis is the way that these theories are always at the same time a form of meaning creation: a suggestive, imaginative act that enters the popular realm and shapes the ethical and political practices of contemporary society. In other words, my focus in what follows is to argue via examples that the extended highly complex metaphor of *Homo machina* (like all social science) does not solely exist on the page, as it were, but becomes flesh and creates a world. This repressed feature of the social scientific theories extending from Hobbes to Pinker is a blind spot in their own theorization. What claims to be merely an act of discovery is in fact always also an act of ethical and ideological creation. But to say this is already to point beyond the metaphor of man as simply a computational mechanics.

…

The entire foregoing analysis of double-H effects implies that there is something remarkable about humans that drastically differs from machines and makes the metaphor of *Homo machina* misleading at best. Even the most sophisticated machines engineered to date lack an experience of meaning or purpose as integral to their actions. The philosopher John Searle famously argued that the major distinction between even the most sophisticated computers and human beings was this "semantic" feature. Where a computer processed a formal algorithm (what Searle called "syntax"), humans experienced states of meaning that were the very stuff of consciousness.21 In fact, computers had long been able to outperform the human mind in running algorithmic calculations. But for the computer there was no inherent meaning or semantic content to this process. In other words, the central feature of specifically human intelligence was missing: that things matter to a human being.

By contrast, things do not matter to a computing machine, nor does a symbol processor experience meanings that orient it toward a purpose or goal. A computing (or any other kind of) machine does not experience disappointment, shame, triumph, fellow feeling, or pride when executing its operations. And faster and faster algorithms do nothing to narrow this yawning gap between semantics and syntax. This is because although computer engineers have accomplished vastly impressive feats, they have yet to bring machines a single step closer to the experience of meaning that makes the double-H effect possible in the first place. As the interpretive philosopher Charles Taylor put it, machines are missing a "significance feature," which is crucial to all purposive agents, such that when it comes to the question "what is [that machine] really doing? There is no answer . . . attributions of action-terms to such devices are relative to our interests and purposes."

In other words, when it comes to bridging the gap between syntax and semantics, computers are no closer than the most rudimentary tools from the distant past. After all, the light from a computer screen might be used as an impromptu lantern in a dark room even as the device runs algorithms designed to flash certain symbols on the screen. The computer's process has no more intrinsic meaning than a hammer or any other human tool, which as an instrument can be made to serve different ends.

This suggests that the metaphor of *Homo machina* has the whole relationship between humans and machines backward. We are not computational machines, but in a culture of scientism we poeticize to trick ourselves into making our computing machines appear more like us. Like ancient pagans who found the agency of gods and spirits in rivers and earthquakes, we humanize or anthropomorphize our machines. Again, scientific authority is paradoxically involved in what is often taken to be a centrally prescientific way of thinking. This anthropomorphizing of machinery is a spellbinding act that we experience as if it were a metaphysical reality. Overawed by our own imaginative powers and meaning-making abilities, we see in technology something that is not there: human purposive intelligence. There is indeed an entire metaphorical poetics behind this anthropomorphizing move. Thus we frequently say that the computer is "calculating,"

"working," "learning," and "thinking" when it is in fact, strictly speaking, doing no such thing.

Does this mean AI is impossible? Not necessarily; it simply shifts the goal of genuine AI from symbol processing to a form of agency capable of purposive experiences of meaning. Everything else is highly sophisticated tool making and nothing more. But a future Alan Turing might yet imagine how the gap between syntax and semantics can be bridged. In any case, the refutation of the computational theory of mind is not my main point here. My point is to suggest that the significance feature of human intelligence is what makes much of the social sciences a poeticizing, creative act of meaning and not merely a descriptive science of the world that was already there waiting for us. The condition for the possibility of the double-H effects discussed here is a being who experiences meanings; this is the stuff of human agency, not algorithmic calculation, which is something that even the greatest savants among us do only mediocrely.

In this regard, a far better proposal for evaluating AI than the Turing test is suggested by the science fiction classic *Blade Runner*. This film—based on a novel by Philip K. Dick—opens with a scene depicting an interview in which a human is testing for the presence of AI. The test requires determining whether an android (known in the movie as a "replicant") is capable of empathy. Empathy is a state that involves an awareness of how another person is experiencing a situation: what matters to him or her and what the emotional significance of a set of circumstances might be. This is closer to the criterion for human intelligence that Searle calls "semantics" and Taylor the "significance" factor. A reworked Turing test would need to be able to determine if an agent were experiencing significance or meanings. Such a test would be an interpretive or hermeneutic threshold for intelligence.

Blade Runner also serves as a powerful interpretive fable for the anxieties surrounding technological society. Taking place in a future version of Los Angeles, the plot follows a man named Deckard, whose profession is "blade running," or hunting and destroying rogue replicants. Yet Deckard finds himself increasingly disturbed and alienated not only by his own severe loss of empathy for those around him but also by the atomized social relations of an impersonal, consumer society dominated by distant corporations. In this setting an awakening of empathy comes from a strange place: Deckard falls in love with one of the replicants he has been hired to kill.

At the center of this story is a deeper cultural fear that is the actual, repressed object of anxiety in the contemporary AI debate between doomsayers and boosters. This is a repressed fear of ourselves and what we might become if we go further down the road of the form of selfhood presented by *Homo machina*. That is to say, fear of robots is fear of ourselves without humanity, without empathy. Or perhaps more accurately, fear of AI is fear not of technology but of a new constellation of meanings opened up by technological society. The machine-self is one possible form of identity that humans embody in a culture of scientism.

This in turn might be linked to the distinctively modern cultures of violence—as scientifically planned by military experts and technocrats—so common in societies across the ideological spectrum. Consider in this light Joseph Stalin's conviction that social science had revealed society could be explained "in accordance with the laws of movement of matter." This machine view of society was the prologue to treating people like basic parts, to be replaced with other purportedly better parts. Stalinism was only one extreme version of the propensity of modern societies to conduct "scientific" mass killings. This is the kind of killing carried out remotely and planned by scientific experts. A dark dream that began in the French Revolution with the guillotine has reached an apotheosis with the invention of the concentration camp-laboratory, where violence is perfectly justified because it is perfectly rational. There is no "I" behind the system of violence in the camp-laboratory; neither is there a "you" on the receiving end. In the last analysis, there is only the impersonal mechanics of a machine grinding humanity into cinder and fire.

In *Blade Runner* we are offered a capitalist version of this mechanistic culture of violence and antihumanism. The humans who populate a future, dystopian Los Angeles have become radically more robotic in this way; they are no longer attuned to the experiences of their neighbors and are willing to treat them like mute objects. The streets of this Los Angeles are filled with a babble of tongues, homeless people dig through the trash, and crowds rush through the sidewalks distracted by their own individual market activity. No one speaks to one another, while neon advertisements shout platitudes about enjoying soft drinks or starting a new life on an "off-world" colony in outer space. Deckard at one point remarks that his ex-wife used to call him a "cold fish," but the audience is relentlessly confronted with an entire society of cold fishes. What distinguishes Deckard is that he struggles mightily throughout the film to overcome his hardened willingness to assassinate others as simply part of his job, a mere market transaction. The entire plot is thus absorbed in the problem of the loss of human empathy and its replacement with a roboticized self that sees all relationships—even those of violence—as mechanical and rational. In all these ways, the city and inhabitants depicted in *Blade Runner* are not a portrait of the future at all but a dramatic picture of the present: the world as built by *Homo machina*.