Moreover, the coordination of this immense mass of esoteric equipment and its operators in the most effective possible way depends upon an extremely highly developed technique both in the employment of each piece of equipment by a specific team of operators and in the management of the program itself. Of course, all large organizations standardize their operating procedures, but it is peculiar to advanced technological systems that their operating procedures embody a very high degree of information drawn from the physical sciences, while their managerial procedures are equally dependent on information drawn from the social sciences. We may describe this situation by saying that advanced technological systems are both "technique intensive" and "management intensive." It should be clear, moreover, even to the most casual observer, that such intensive use of capital, technique, and management spills over into almost every area touched by the technological system in question. An attack program delivering 330,000 tons of munitions more or less selectively to several thousand different targets monthly would be an anomaly if forced to rely on sporadic intelligence data, erratic maintenance systems, or a fluctuating and unpredictable supply of heavy bombs, rockets, jet fuel, and napalm tanks. Thus, it is precisely because the bombing program requires an intensive use of capital, technique, and management that the same properties are normally transferred to the intelligence, maintenance, supply, coordination, and training systems that support it. Accordingly, each of these supporting systems is subject to sharp pressures to improve and rationalize the performance of its machines and men, the reliability of its techniques, and the efficiency and sensitivity of the management controls under which it operates. Within integrated technical systems, higher levels of technology drive out lower, and the normal tendency is to integrate systems.

From this perverse Gresham's Law of Technology follow some of the main social and organizational characteristics of contemporary technological systems: the radical increase in the scale and complexity of operations that they demand and encourage; the rapid and widespread diffusion of technology to new areas; the great diversity of activities that can be directed by central management; an increase in the ambition of management's goals; and, as a corollary, especially to the last, growing resistance to the influence of so-called negative externalities.

Complex technological systems are extraordinarily resistant to intervention by persons or problems operating outside or below their managing groups, and this is so regardless of the "politics" of a given situation. Technology creates its own politics. The point of such advanced systems is to minimize the incidence of personal or social behavior that is erratic or otherwise not easily classified, of tools and equipment with poor performance, of improvisory techniques, and of unresponsiveness to central management....

To define technology so abstractly that it obscures these observable characteristics of contemporary technology — as Mesthene and his school have done — makes no sense. It makes even less sense to claim some magical malleability for something as undefined as "institutional innovation." Technology, in its concrete, empirical meaning, refers fundamentally to systems of rationalized control over large groups of men, events, and machines by small groups of technically skilled men operating through organizational hierarchy. The latent "opportunities" provided by that control and its ability to filter out discordant "negative externalities" are, of course, best illustrated by extreme cases. Hence, the most instructive and accurate example should be of a technology able to suppress the humanity of its rank-and-file and to commit genocide as a by-product of its rationality. The Vietnam bombing program fits technology to a "T."

* * *

IV

Among the conventional explanations for the rise and spread of the democratic ethos in Europe and North America in the seventeenth, eighteenth, and nineteenth centuries, the destruction of the gap in political culture between the mass of the population and that of the ruling classes is extremely important....

Similarly, it is often argued that with the expansion and improvement of road and postal systems, the spread of new tools and techniques, the growth in the number and variety of merchants, the consequent invigoration of town life, and other numerous and familiar related developments, the social experience of larger numbers of people became richer, more varied, and similar in fact to those of the ruling class....

The same period also witnessed a growth in the organized means of popular expression....

This description by no means does justice to the richness and variety of the historical process underlying the rise and spread of what has come to be called the democratic ethos. But it does, I hope, isolate some of the important structural elements and, moreover, it enables us to illuminate some important ways in which the new technology, celebrated by Mesthene and his associates for its potential contributions to democracy, contributes instead to the erosion of that same democratic ethos. For if, in an earlier time, the gap between the political cultures of the higher and lower orders of society was being widely attacked and closed, this no longer appears to be the case. On the contrary, I am persuaded that the direction has been reversed and that we now observe evidence of a growing separation between ruling- and lower-class culture in America, a separation that is particularly enhanced by the rapid growth of technology and the spreading influence of its *laissez innover* ideologues.

Certainly, there has been a decline in popular literacy, that is to say, in those aspects of literacy that bear on an understanding of the political and social character of the new technology. Not one person in a hundred is even aware of, much less understands, the nature of technologically highly advanced systems such as are used in the Vietnam bombing program....

Secondly, the social organization of this new technology, by systematically denying to the general population experiences that are analogous to those of its higher management, contributes very heavily to the growth of social irrationality in our society. For example, modern technological organization defines the roles and values of its members, not vice versa. An engineer or a sociologist is one who does all those things but only those things called for by the "table of organization" and the "job description" used by his employer. Professionals who seek self-realization through creative and autonomous behavior without regard to the defined goals, needs, and channels of their respective departments have no more place in a large corporation or government agency than squeamish soldiers in the army....

However, those at the top of technology's most advanced organizations hardly suffer the same experience. For reasons that are clearly related to the principle of the Altruistic Bureaucracy, the psychology of an individual's fulfillment through work has been incorporated into management ideology. As the pages of *Fortune, Time,* or *Business Week* ... serve to show, the higher levels of business and government are staffed by men and women who spend killing hours looking after the economic welfare and national security of the rest of us. The rewards of this life are said to be very few: The love of money would be demeaning and, anyway, taxes are said to take most of it; its sacrifices are many, for failure brings economic depression to the masses or gains for communism as well as disgrace to the erring managers. Even the essential high-mindedness or altruism of our managers earns no reward, for the public is distracted, fickle, and, on occasion, vengeful.... Hence, for these "real revolutionaries of our time," as Walt Rostow [economist and National Security Advisor to President Lyndon B. Johnson during the Vietnam war – ed.] has called them, self-fulfillment through work and discipline is the only reward. The managerial process is seen as an expression of the vital personalities of our leaders and the right to it an inalienable right of the national elite.

In addition to all this, their lonely and unrewarding eminence in the face of crushing responsibility, etc., tends to create an air of mystification around technology's managers....

It seems fundamental to the social organization of modern technology that the quality of the social experience of the lower orders of society declines as the level of technology grows no less than does their literacy. And, of course, this process feeds on itself, for with the consequent decline in the real effectiveness and usefulness of local and other forms of organization open to easy and direct popular influence, their vitality declines still further, and the cycle is repeated.

The normal life of men and women in the lower and, I think, middle levels of American society now seems cut off from those experiences in which near social means and distant social ends are balanced and rebalanced, adjusted and readjusted. But it is from such widespread experience with effective balancing and adjusting that social rationality derives. To the degree that it is lacking, social irrationality becomes the norm, and social paranoia a recurring phenomenon....

Mesthene himself recognizes that such "negative externalities" are on the increase. His list includes "pollution of the environment, potential damage to the ecology of the planet, occupational and social dislocations, threats to the privacy and political significance of the individual, social and psychological malaise." Minor matters all, however, when compared to the marvelous opportunities *laissez innover* holds out to us: more GNP, continued free-world leadership, supersonic transports, urban renewal on a regional basis, institutional innovation, and the millennial promises of his school.

This brings us finally to the ideologies and doctrines of technology and their relation to what I have argued is a growing gap in political culture between the lower and upper classes in American society. Even more fundamentally than the principles of *laissez innover* and the altruistic bureaucrat, technology in its very definition as the organization of knowledge for practical purposes assumes that the primary and really creative role in the social processes consequent on technological change is reserved for a scientific and technical elite, the elite that presumably discovers and organizes that knowledge. But if the scientific and technical elite and their indispensable managerial cronies are the really creative (and hardworking and altruistic) elements in American society, what is this but to say that the common mass of people are essentially drags on the social weal? This is precisely the implication which is drawn by the *laissez innover* school. Consider the following quotations from an article that appeared in *The New Republic* in December 1967, written by Zbigniew Brzezinski, one of the intellectual leaders of the school.

Brzezinski [who later became National Security Advisor to President Jimmy Carter – ed.] is describing a nightmare that he calls the "technetronic society" (the word, like the concept, is a pastiche of technology and electronics). This society will be characterized, he argues, by the application of "the principle of equal opportunity for all but ... special opportunity for the singularly talented few." It will thus combine "continued *respect* for the popular will with an increasing *role* in the key decision-making institutions of individuals with special intellectual and scientific attainments." [Italics added.] Naturally, "the educational and social systems [will make] it increasingly attractive and easy for those meritocratic few to develop to the fullest of their special potential."

However, while it will be "necessary to require everyone at a sufficiently responsible post to take, say, two years of [scientific and technical] retraining every ten years," the rest of us can develop a new "interest in the cultural and humanistic aspects of life, *in addition to purely hedonistic preoccupations.*" [Italics added.] The latter, he is careful to point out, "would serve as a social valve, reducing tensions and political frustration."

Is it not fair to ask how much respect we carefree pleasure lovers and culture consumers will get from the hardworking bureaucrats, going to night school two years in every ten, while working like beavers in the "key decision-making institutions"? The altruism of our bureaucrats has a heavy load to bear.

Stripped of their euphemisms, these are simply arguments that enhance the social legitimacy of the interests of new technical and scientific elites and detract from the interests of the rest of us....

As has already been made clear, the *laissez innover* school accepts as inevitable and desirable the centralizing tendencies of technology's social organization, and they accept as well the mystification that comes to surround the management process. Thus, equality of opportunity, as they understand it, has precious little to do with creating a more egalitarian society. On the contrary, it functions as an indispensable feature of the highly stratified society they envision for the future. For in their society of meritocratic hierarchy, equality of opportunity assures that talented young meritocrats (the word is no uglier than the social system it refers to) will be able to climb into the "key decision-making" slots reserved for trained talent and thus generate the success of the new society and its cohesion against popular "tensions and political frustration."

The structures that formerly guaranteed the rule of wealth, age, and family will not be destroyed (or at least not totally so). They will be firmed up and rationalized by the perpetual addition of trained (and, of course, acculturated) talent. In technologically advanced societies, equality of opportunity functions as a hierarchical principle, in opposition to the egalitarian social goals it pretends to serve. To the extent that it has already become the kind of "equality" we seek to institute in our society, it is one of the main factors contributing to the widening gap between the cultures of upper- and lower-class America.

V

... *Laissez innover* is now the premier ideology of the technological impulse in American society, which is to say, of the institutions that monopolize and profit from advanced technology and of the social classes that find in the free exploitation of their technology the most likely guarantee of their power, status, and wealth.

This said, it is important to stress both the significance and limitations of what has in fact been said. Here, Mesthene's distinction between the positive opportunities and negative "externalities" inherent in technological change is pivotal; for everything else that I've argued follows inferentially from the actual social meaning of that distinction. As my analysis of the Vietnam bombing program suggested, those technological effects that are sought after as positive opportunities and those that are dismissed as negative externalities are decisively influenced by the fact that this distinction between positive and negative within advanced technological organizations tends to be made among the planners and managers themselves. Within these groups there are, as was pointed out, extremely powerful organizational, hierarchical, doctrinal, and other "technical" factors, which tend by design to filter out "irrational" demands from below, substituting for them the "rational" demands of technology itself. As a result, technological rationality is as socially neutral today as market rationality was a century ago....

This analysis lends some weight (though perhaps no more than that) to a number of wide-ranging and unorthodox conclusions about American society today and the directions in which it is tending....

First, and most important, technology should be considered as an institutional system, not more and certainly not less. Mesthene's definition of the subject is inadequate, for it obscures the systematic and decisive social changes, especially their political and cultural tendencies, that follow the widespread application of advanced technological systems. At the same time, technology is less than a social system per se, though it has many elements of a social system, viz., an elite, a group of linked institutions, an ethos, and so forth. Perhaps the best summary statement of the case resides in an analogy — with all the vagueness and imprecision attendant on such things: today's technology stands in relation to today's capitalism as, a century ago, the latter stood to the free market capitalism of the time....

A second major hypothesis would argue that the most important dimension of advanced technological institutions is the social one, that is, the institutions are agencies of highly centralized and intensive social control. Technology conquers nature, as the saying goes. But to do so it must first conquer man. More precisely, it demands a very high degree of control over the training, mobility, and skills of the work force. The absence (or decline) of direct controls or of coercion should not serve to obscure from our view the reality and intensity of the social controls that are employed (such as the internalized belief in equality of opportunity, indebtedness through credit, advertising, selective service channeling, and so on).

Advanced technology has created a vast increase in occupational specialties, many of them requiring many, many years of highly specialized training. It must motivate this training. It has made ever more complex and "rational" the ways in which these occupational specialties are combined in our economic and social life. It must win passivity and obedience to this complex activity. Formerly, technical rationality had been employed only to organize the production of rather simple physical objects, for example, aerial bombs. Now, technical rationality is increasingly employed to organize all of the processes necessary to the utilization of physical objects, such as bombing systems. For this reason, it seems a mistake to argue that we are in a "postindustrial" age, a concept favored by the *laissez innover* school. On the contrary, the rapid spread of technical rationality into organizational and economic life and, hence, into social life is more aptly described as a second and much more intensive phase of the industrial revolution. One might reasonably suspect that it will create analogous social problems.

Accordingly, a third major hypothesis would argue that there are very profound social antagonisms or contradictions not less sharp or fundamental than those ascribed by Marx to the development of nineteenth-century industrial society. The general form of the contradictions might be described as follows: A society characterized by the employment of advanced technology requires an ever-more socially disciplined population, yet retains an ever-declining capacity to enforce the required discipline....

These are brief and, I believe, barely adequate reviews of extremely complex hypotheses. But, in outline, each of these contradictions appears to bear on roughly the same group of the American population: a technological underclass. If we assume this to be the case, a fourth hypothesis would follow, namely that technology is creating the basis for new and sharp class conflict in our society. That is, technology is creating its own working and managing classes, just as earlier industrialization created its working and owning classes. Perhaps this suggests a return to the kind of class-based politics that characterized the U.S. in the last quarter of the nineteenth century, rather than the somewhat more ambiguous politics that was a feature of the second quarter of this century. I am inclined to think that this is the case, though I confess the evidence for it is as yet inadequate. This leads to a final hypothesis, namely that *laissez innover* should be frankly recognized as a conservative or right-wing ideology....

The point of this final hypothesis is not primarily to reimpress the language of European politics on the American scene. Rather, it is to summarize the fact that many of the forces in American life hostile to the democratic ethos have enrolled under the banner of *laissez innover*. Merely to grasp this is already to take the first step toward a politics of radical reconstruction and against the malaise, irrationality, powerlessness, and official violence that characterize American life today.

ENDNOTE

1. For a more complete statement of the argument that follows, see Suzanne Keller, *Beyond the Ruling Class* (New York: Random House, 1963).

PART III

Debating Technology: Twenty-First-Century Style

The comparison between the 1960s-style debate on technology between Emmanuel Mesthene and John McDermott in Part II of this book and the twenty-first-century debate between Bill Joy and his critics in the next three chapters is a remarkable illustration of both how much and how little has changed during the 35 years that separates these writings.

The Mesthene–McDermott dispute is strongly ideological. McDermott represents the "New Left" of the 1960s and '70s and its attendant counterculture. Although Mesthene would have disagreed strongly with the characterization, McDermott places him squarely in the camp of the reactionary "establishment" seeking to preserve the status quo and all the inequities and injustices it implies. Mesthene, for his part, probably regarded McDermott as a Luddite, heir to the radical anti-technology tradition of those British activists who went about destroying machines in the early days of the Industrial Revolution 200 years ago. The Bill Joy–John Seely Brown/Paul Duguid argument lacks the ideological edge, invective, and bitterness of the 1960s debate. Their arguments are largely devoid of politics and ideology. Both sides in the dispute are technophiles — good friends, in fact — although Joy is a technophile who believes he has seen the future and is very frightened by it. Nonetheless, the opposite sides in both of these debates (Mesthene vs. McDermott and Joy vs. Brown/Duguid) reflect sharply different perspectives on the balance between the benefits and the harmful impacts of technology.

To this mix we have added, as a kind of middle ground, the perspective of Ray Kurzweil, an inventor of extraordinary achievement, winner of the Lemelson-MIT Prize for invention and the National Medal of Technology. Kurzweil is the person Bill Joy says first made him "anxiously aware" of the dangers technology is creating for society. Here Kurzweil presents his side of the story. He shares many of Joy's concerns, but he would deal with them differently. Rather than "relinquishing" the areas of technology that pose threats to humanity, as Joy proposes, he would respond to their threats by more conventional means — regulation and the development of counter-technologies.

The shift back to debating this balance between harm and benefits represents something of a departure during the past few years. In the eighth edition of *Technology and the Future* (published in 2000), the counterpart to this section ("Debating Technology: Turn of the Millennium Style") featured Nicholas Negroponte and Donald Norman in a friendly disagreement over whether the future is going to be "digital" (i.e., shaped by our ability to digitize any kind of information — Negroponte's view) or analog (i.e., whether the headlong rush to digitization is causing us to lose sight of the fact that we are analog creatures and need to design technologies to suit ourselves rather than designing technologies and expecting that we will adapt — Norman's perspective). The technologies about which Bill Joy, John Seely Brown, Paul Duguid, and Ray Kurzweil are arguing — genetic engineering, nanotechnology, and robotics — were barely imagined in the 1960s, and their potential for interaction (which is what concerns Joy) was just coming into focus in 2000. Yet, whether we fear these technologies or look to them as the potential sources of enormous economic, health, and societal benefits (or both), we need to be thinking carefully and constantly about the ways in which society is affected by them and how we can shape their development in positive ways. The articles in this part can help us do this.

9

Why the Future Doesn't Need Us

BILL JOY

The publication in Wired *(April 2000) of Bill Joy's article, "Why the Future Doesn't Need Us," caused quite a stir among the so-called digerati (the elite of the Internet and information technology world). Joy sees the convergence of genetic engineering, nanotechnology, and robotics (GNR) as posing a grave threat to the future of humanity. He argues that these new technologies are qualitatively different from those with which we are familiar. "Accustomed to living with almost routine scientific breakthroughs, we have yet to come to terms with the fact that the most compelling twenty-first-century technologies — robotics, genetic engineering, and nanotechnology — pose a different threat than the technologies that have come before. Specifically, robots, engineered organisms, and nanobots share a dangerous amplifying factor: They can self-replicate. A bomb is blown up only once — but one bot can become many, and quickly get out of control."*

Joy develops his argument in detail, drawing on his knowledge of these technologies, and on conversations with and writings of many contemporary leaders in technology, as well as a number of well-known science fiction writers. His scenario is pessimistic in the extreme. If these technologies are allowed to develop unchecked, Joy fears, they could lead to the extinction of the human species. In response, he calls for the voluntary relinquishment of certain of these technologies by scientists, technologists, and the firms and nations that sponsor their work. Had such a polemic been published by one of the many well-known technological critics (several of whom are represented in this book), it might not have attracted much attention. Bill Joy, however, is one of the leading figures of the high-tech world. At the time he wrote this article, he was chief scientist and *corporate executive officer of Sun Microsystems, a company he cofounded. Joy was the principal designer of the Berkeley version of UNIX, and one of the developers of the Java Programming Language that is now ubiquitous on the Web. He left Sun in 2003 and in 2005 joined the venture capital firm of Kleiner Perkins Caufield & Byers, where he is currently a partner, investing — some might say, ironically — in genetics, nanotechnology, and robotics firms. His lengthy essay, and the rebuttals that follow it, are well worth reading.*

From the moment I became involved in the creation of new technologies, their ethical dimensions have concerned me, but it was only in the autumn of 1998 that I became anxiously aware of how great are the dangers facing us in the twenty-first century. I can date the onset of my unease to the day I met Ray Kurzweil, the deservedly famous inventor of the first reading machine for the blind and many other amazing things.

Ray and I were both speakers at George Gilder's Telecosm conference, and I encountered him by chance in the bar of the hotel after both our sessions were over. I was sitting with John Searle, a Berkeley philosopher who studies consciousness. While we were talking, Ray approached and a conversation began, the subject of which haunts me to this day.

I had missed Ray's talk and the subsequent panel that Ray and John had been on, and they now picked right up where they'd left off, with Ray saying that the rate of improvement of technology was going to accelerate and that we were going to become robots or fuse with robots or something like that, and John countering that this couldn't happen, because the robots couldn't be conscious.

While I had heard such talk before, I had always felt sentient robots were in the realm of science fiction. But now, from someone I respected, I was hearing a strong argument that they were a near-term possibility. I was taken aback, especially given Ray's proven ability to imagine and create the future. I already knew that new technologies like genetic engineering and nanotechnology were giving us the power to remake the world, but a realistic and imminent scenario for intelligent robots surprised me.

It's easy to get jaded about such breakthroughs. We hear in the news almost every day of some kind of technological or scientific advance. Yet this was no ordinary prediction. In the hotel bar, Ray gave me a partial preprint of his then-forthcoming book, *The Age of Spiritual Machines*, which outlined a utopia he foresaw — one in which humans gained near immortality by becoming one with robotic technology. On reading it, my sense of unease only intensified; I felt sure he had to be understating the dangers, understating the probability of a bad outcome along this path.

I found myself most troubled by a passage detailing a dystopian scenario:

> *The New Luddite Challenge*
>
> First, let us postulate that the computer scientists succeed in developing intelligent machines that can do all things better than human beings can

do them. In that case, presumably all work will be done by vast, highly organized systems of machines and no human effort will be necessary. Either of two cases might occur. The machines might be permitted to make all of their own decisions without human oversight, or else human control over the machines might be retained.

If the machines are permitted to make all their own decisions, we can't make any conjectures as to the results, because it is impossible to guess how such machines might behave. We only point out that the fate of the human race would be at the mercy of the machines. It might be argued that the human race would never be foolish enough to hand over all the power to the machines. But we are suggesting neither that the human race would voluntarily turn power over to the machines, nor that the machines would willfully seize power. What we do suggest is that the human race might easily permit itself to drift into a position of such dependence on the machines that it would have no practical choice but to accept all of the machines' decisions. As society and the problems that face it become more and more complex and machines become more and more intelligent, people will let machines make more of their decisions for them, simply because machine-made decisions will bring better results than man-made ones. Eventually, a stage may be reached at which the decisions necessary to keep the system running will be so complex that human beings will be incapable of making them intelligently. At that stage the machines will be in effective control. People won't be able to just turn the machines off, because they will be so dependent on them that turning them off would amount to suicide.

On the other hand, it is possible that human control over the machines may be retained. In that case, the average man may have control over certain private machines of his own, such as his car or his personal computer, but control over large systems of machines will be in the hands of a tiny elite—just as it is today, but with two differences. Due to improved techniques the elite will have greater control over the masses, and, because human work will no longer be necessary, the masses will be superfluous, a useless burden on the system. If the elite is ruthless, they may simply decide to exterminate the mass of humanity. If they are humane, they may use propaganda or other psychological or biological techniques to reduce the birth rate until the mass of humanity becomes extinct, leaving the world to the elite. Or, if the elite consists of soft-hearted liberals, they may decide to play the role of good shepherds to the rest of the human race. They will see to it that everyone's physical needs are satisfied, that all children are raised under psychologically hygienic conditions, that everyone has a wholesome hobby to keep him busy, and that anyone who may become dissatisfied undergoes "treatment" to cure his "problem." Of course, life will be so purposeless that people will have to be biologically or psychologically engineered either to remove their need for the power process or make them "sublimate" their drive for

> power into some harmless hobby. These engineered human beings may be happy in such a society, but they will most certainly not be free. They will have been reduced to the status of domestic animals.[1]

In the book, you don't discover until you turn the page that the author of this passage is Theodore Kaczynski — the Unabomber. I am no apologist for Kaczynski. His bombs killed three people during a seventeen-year terror campaign and wounded many others. One of his bombs gravely injured my friend David Gelernter, one of the most brilliant and visionary computer scientists of our time. Like many of my colleagues, I felt that I could easily have been the Unabomber's next target.

Kaczynski's actions were murderous and, in my view, criminally insane. He is clearly a Luddite, but simply saying this does not dismiss his argument; as difficult as it is for me to acknowledge, I saw some merit in the reasoning in this single passage. I felt compelled to confront it.

Kaczynski's dystopian vision describes unintended consequences, a well-known problem with the design and use of technology and one that is clearly related to Murphy's law — "Anything that can go wrong, will." (Actually, this is Finagle's law, which in itself shows that Finagle was right.) Our overuse of antibiotics has led to what may be the biggest such problem so far: the emergence of antibiotic-resistant and much more dangerous bacteria. Similar things happened when attempts to eliminate malarial mosquitoes using DDT caused them to acquire DDT resistance; malarial parasites likewise acquired multidrug-resistant genes.[2]

The cause of many such surprises seems clear: The systems involved are complex, involving interaction among and feedback between many parts. Any changes to such a system will cascade in ways that are difficult to predict; this is especially true when human actions are involved.

I started showing friends the Kaczynski quote from *The Age of Spiritual Machines*; I would hand them Kurzweil's book, let them read the quote, and then watch their reaction as they discovered who had written it. At around the same time, I found Hans Moravec's book *Robot: Mere Machine to Transcendent Mind*. Moravec is one of the leaders in robotics research and was a founder of the world's largest robotics research program, at Carnegie Mellon University. *Robot* gave me more material to try out on my friends — material surprisingly supportive of Kaczynski's argument. For example:

> *The Short Run (Early 2000s)*
>
> Biological species almost never survive encounters with superior competitors. Ten million years ago, South and North America were separated by a sunken Panama isthmus. South America, like Australia today, was populated by marsupial mammals, including pouched equivalents of rats, deer, and tigers. When the isthmus connecting North and South America rose, it took only a few thousand years for the northern placental species, with slightly more effective metabolisms and reproductive and nervous systems, to displace and eliminate almost all the southern marsupials.

> In a completely free marketplace, superior robots would surely affect humans as North American placentals affected South American marsupials (and as humans have affected countless species). Robotic industries would compete vigorously among themselves for matter, energy, and space, incidentally driving their price beyond human reach. Unable to afford the necessities of life, biological humans would be squeezed out of existence.
>
> There is probably some breathing room, because we do not live in a completely free marketplace. Government coerces nonmarket behavior, especially by collecting taxes. Judiciously applied, governmental coercion could support human populations in high style on the fruits of robot labor, perhaps for a long while.

A textbook dystopia — and Moravec is just getting wound up. He goes on to discuss how our main job in the twenty-first century will be "ensuring continued cooperation from the robot industries" by passing laws decreeing that they be "nice"[3] and to describe how seriously dangerous a human can be "once transformed into an unbounded superintelligent robot." Moravec's view is that the robots will eventually succeed us — that humans clearly face extinction.

I decided it was time to talk to my friend Danny Hillis. Danny became famous as the cofounder of Thinking Machines Corporation, which built a very powerful parallel supercomputer. Despite my current job title of Chief Scientist at Sun Microsystems, I am more a computer architect than a scientist, and I respect Danny's knowledge of the information and physical sciences more than that of any other single person I know. Danny is also a highly regarded futurist who thinks long-term: Four years ago he started the Long Now Foundation, which is building a clock designed to last 10,000 years, in an attempt to draw attention to the pitifully short attention span of our society. (See "Test of Time," *Wired* 8.03, p. 78.)

So I flew to Los Angeles for the express purpose of having dinner with Danny and his wife, Pati. I went through my now-familiar routine, trotting out the ideas and passages that I found so disturbing. Danny's answer — directed specifically at Kurzweil's scenario of humans merging with robots — came swiftly and quite surprised me. He said, simply, that the changes would come gradually and that we would get used to them.

But I guess I wasn't totally surprised. I had seen a quote from Danny in Kurzweil's book in which he said, "I'm as fond of my body as anyone, but if I can be 200 with a body of silicon, I'll take it." It seemed that he was at peace with this process and its attendant risks, while I was not.

While talking and thinking about Kurzweil, Kaczynski, and Moravec, I suddenly remembered a novel I had read almost 20 years ago — *The White Plague*, by Frank Herbert — in which a molecular biologist is driven insane by the senseless murder of his family. To seek revenge, he constructs and disseminates a new and highly contagious plague that kills widely but selectively. (We're lucky Kaczynski was a mathematician, not a molecular biologist.) I was also reminded of the Borg of Star Trek, a hive of partly biological, partly robotic creatures with a strong destructive streak. Borg-like disasters are a staple of science fiction, so why hadn't I been more concerned about such robotic dystopias earlier? Why weren't other people more concerned about these nightmarish scenarios?

Part of the answer certainly lies in our attitude toward the new — in our bias toward instant familiarity and unquestioning acceptance. Accustomed to living with almost routine scientific breakthroughs, we have yet to come to terms with the fact that the most compelling twenty-first-century technologies — robotics, genetic engineering, and nanotechnology — pose a different threat than the technologies that have come before. Specifically, robots, engineered organisms, and nanobots share a dangerous amplifying factor: They can self-replicate. A bomb is blown up only once — but one 'bot can become many and quickly get out of control.

Much of my work over the past twenty-five years has been on computer networking, where the sending and receiving of messages creates the opportunity for out-of-control replication. But while replication in a computer or a computer network can be a nuisance, at worst it disables a machine or takes down a network or network service. Uncontrolled self-replication in these newer technologies runs a much greater risk: a risk of substantial damage in the physical world.

Each of these technologies also offers untold promise: The vision of near immortality that Kurzweil sees in his robot dreams drives us forward; genetic engineering may soon provide treatments, if not outright cures, for most diseases, and nanotechnology and nanomedicine can address yet more ills. Together they could significantly extend our average life span and improve the quality of our lives. Yet, with each of these technologies, a sequence of small, individually sensible advances leads to an accumulation of great power and, concomitantly, great danger.

What was different in the twentieth century? Certainly, the technologies underlying the weapons of mass destruction (WMD) — nuclear, biological, and chemical (NBC) — were powerful and the weapons an enormous threat. But building nuclear weapons required, at least for a time, access to both rare — indeed, effectively unavailable — raw materials and highly protected information; biological and chemical weapons programs also tended to require large-scale activities.

The twenty-first-century technologies — genetics, nanotechnology, and robotics (GNR) — are so powerful that they can spawn whole new classes of accidents and abuses. Most dangerously, for the first time, these accidents and abuses are widely within the reach of individuals or small groups. They will not require large facilities or rare raw materials. Knowledge alone will enable the use of them.

Thus, we have the possibility not just of weapons of mass destruction but of knowledge-enabled mass destruction (KMD), this destructiveness hugely amplified by the power of self-replication.

I think it is no exaggeration to say we are on the cusp of the further perfection of extreme evil, an evil whose possibility spreads well beyond that which weapons of mass destruction bequeathed to the nation-states, on to a surprising and terrible empowerment of extreme individuals.

Nothing about the way I got involved with computers suggested to me that I was going to be facing these kinds of issues.

My life has been driven by a deep need to ask questions and find answers. When I was three, I was already reading, so my father took me to the elementary school, where I sat on the principal's lap and read him a story. I started school early, later skipped a grade, and escaped into books — I was incredibly motivated to learn. I asked lots of questions, often driving adults to distraction.

As a teenager, I was very interested in science and technology. I wanted to be a ham radio operator but didn't have the money to buy the equipment. Ham radio was the Internet of its time: very addictive and quite solitary. Money issues aside, my mother put her foot down — I was not to be a ham; I was antisocial enough already.

I may not have had many close friends, but I was awash in ideas. By high school, I had discovered the great science-fiction writers. I remember especially Heinlein's *Have Spacesuit Will Travel* and Asimov's *I, Robot*, with its Three Laws of Robotics. I was enchanted by the descriptions of space travel and wanted to have a telescope to look at the stars; since I had no money to buy or make one, I checked books on telescope making out of the library and read about making them instead. I soared in my imagination.

Thursday nights my parents went bowling, and we kids stayed home alone. It was the night of Gene Roddenberry's original *Star Trek*, and the program made a big impression on me. I came to accept its notion that humans had a future in space, Western-style, with big heroes and adventures. Roddenberry's vision of the centuries to come was one with strong moral values, embodied in codes like the Prime Directive: to not interfere in the development of less technologically advanced civilizations. This had an incredible appeal to me; ethical humans, not robots, dominated this future, and I took Roddenberry's dream as part of my own.

I excelled in mathematics in high school, and when I went to the University of Michigan as an undergraduate engineering student I took the advanced curriculum of the mathematics majors. Solving math problems was an exciting challenge, but when I discovered computers I found something much more interesting: a machine into which you could put a program that attempted to solve a problem, after which the machine quickly checked the solution. The computer had a clear notion of correct and incorrect, true and false. Were my ideas correct? The machine could tell me. This was very seductive. I was lucky enough to get a job programming early supercomputers and discovered the amazing power of large machines to numerically simulate advanced designs. When I went to graduate school at UC Berkeley in the mid-1970s, I started staying up late, often all night, inventing new worlds inside the machines. Solving problems. Writing the code that argued so strongly to be written.

In *The Agony and the Ecstasy*, Irving Stone's biographical novel of Michelangelo, Stone described vividly how Michelangelo released the statues from the stone, "breaking the marble spell," carving from the images in his mind.[4] In my most ecstatic moments, the software in the computer emerged in the same way. Once I had imagined it in my mind, I felt that it was already there in the machine, waiting to be released. Staying up all night seemed a small price to pay to free it — to give the ideas concrete form.

After a few years at Berkeley, I started to send out some of the software I had written — an instructional Pascal system, Unix utilities, and a text editor called vi (which is still, to my surprise, widely used more than twenty years later) — to others who had similar small PDP-11 and VAX minicomputers. These adventures in software eventually turned into the Berkeley version of the Unix operating system, which became a personal "success disaster" — so many people wanted it that I never finished my Ph.D. Instead, I got a job working for DARPA [Defense Advanced Research Projects Agency] putting Berkeley Unix on the Internet and fixing it to be reliable and to run large research applications well. This was all great fun and very rewarding. And, frankly, I saw no robots here, or anywhere near.

Still, by the early 1980s, I was drowning. The Unix releases were very successful, and my little project of one soon had money and some staff, but the problem at Berkeley was always office space rather than money. There wasn't room for the help the project needed, so when the other founders of Sun Microsystems showed up, I jumped at the chance to join them. At Sun, the long hours continued into the early days of workstations and personal computers, and I have enjoyed participating in the creation of advanced microprocessor technologies and Internet technologies such as Java and Jini.

From all this, I trust it is clear that I am not a Luddite. I have always, rather, had a strong belief in the value of the scientific search for truth and in the ability of great engineering to bring material progress. The Industrial Revolution has immeasurably improved everyone's life over the last couple hundred years, and I always expected my career to involve the building of worthwhile solutions to real problems, one problem at a time.

I have not been disappointed. My work has had more impact than I had ever hoped for and has been more widely used than I could have reasonably expected. I have spent the last twenty years still trying to figure out how to make computers as reliable as I want them to be (they are not nearly there yet) and how to make them simple to use (a goal that has met with even less relative success). Despite some progress, the problems that remain seem even more daunting.

But while I was aware of the moral dilemmas surrounding technology's consequences in fields like weapons research, I did not expect that I would confront such issues in my own field, or at least not so soon.

Perhaps it is always hard to see the bigger impact while you are in the vortex of a change. Failing to understand the consequences of our inventions while we are in the rapture of discovery and innovation seems to be a common fault of scientists and technologists; we have long been driven by the overarching desire to know that is the nature of science's quest, not stopping to notice that the progress to newer and more powerful technologies can take on a life of its own.

I have long realized that the big advances in information technology come not from the work of computer scientists, computer architects, or electrical engineers, but from that of physical scientists. The physicists Stephen Wolfram and Brosl Hasslacher introduced me, in the early 1980s, to chaos theory and nonlinear systems. In the 1990s, I learned about complex systems from conversations with Danny Hillis, the biologist Stuart Kauffman, the Nobel-laureate physicist Murray Gell-Mann, and others. Most recently, Hasslacher and the electrical engineer and device physicist Mark Reed have been giving me insight into the incredible possibilities of molecular electronics.

In my own work, as codesigner of three microprocessor architectures — SPARC, picoJava, and MAJC — and as the designer of several implementations thereof, I've been afforded a deep and firsthand acquaintance with Moore's law. For decades, Moore's law has correctly predicted the exponential rate of improvement of semiconductor technology. Until last year, I believed that the rate of advances predicted by Moore's law might continue only until roughly 2010, when some physical limits would begin to be reached. It was not obvious to me that a new technology would arrive in time to keep performance advancing smoothly.

But because of the recent rapid and radical progress in molecular electronics — where individual atoms and molecules replace lithographically drawn transistors — and related nanoscale technologies, we should be able to meet or exceed the Moore's law rate of progress for another thirty years. By 2030, we are likely to be able to build machines, in quantity, a million times as powerful as the personal computers of today — sufficient to implement the dreams of Kurzweil and Moravec.

As this enormous computing power is combined with the manipulative advances of the physical sciences and the new, deep understandings in genetics, enormous transformative power is being unleashed. These combinations open up the opportunity to completely redesign the world, for better or worse: The replicating and evolving processes that have been confined to the natural world are about to become realms of human endeavor.

In designing software and microprocessors, I have never had the feeling that I was designing an intelligent machine. The software and hardware is so fragile and the capabilities of the machine to "think" so clearly absent that, even as a possibility, this has always seemed very far in the future.

But now, with the prospect of human-level computing power in about thirty years, a new idea suggests itself: that I may be working to create tools which will enable the construction of the technology that may replace our species. How do I feel about this? Very uncomfortable. Having struggled my entire career to build reliable software systems, it seems to me more than likely that this future will not work out as well as some people may imagine. My personal experience suggests we tend to overestimate our design abilities.

Given the incredible power of these new technologies, shouldn't we be asking how we can best coexist with them? And if our own extinction is a likely, or even possible, outcome of our technological development, shouldn't we proceed with great caution?

The dream of robotics is, first, that intelligent machines can do our work for us, allowing us lives of leisure, restoring us to Eden. Yet in his history of such ideas, *Darwin Among the Machines*, George Dyson warns, "In the game of life and evolution there are three players at the table: human beings, nature, and machines. I am firmly on the side of nature. But nature, I suspect, is on the side of the machines." As we have seen, Moravec agrees, believing we may well not survive the encounter with the superior robot species.

How soon could such an intelligent robot be built? The coming advances in computing power seem to make it possible by 2030. And once an intelligent robot exists, it is only a small step to a robot species — to an intelligent robot that can make evolved copies of itself.

A second dream of robotics is that we will gradually replace ourselves with our robotic technology, achieving near immortality by downloading our consciousnesses; it is this process that Danny Hillis thinks we will gradually get used to and that Ray Kurzweil elegantly details in *The Age of Spiritual Machines*. (We are beginning to see intimations of this in the implantation of computer devices into the human body, as illustrated on the cover of *Wired* 8.02.)

But if we are downloaded into our technology, what are the chances that we will thereafter be ourselves or even human? It seems to me far more likely that a robotic existence would not be like a human one in any sense that we understand, that the robots would in no sense be our children, that on this path our humanity may well be lost.

Genetic engineering promises to revolutionize agriculture by increasing crop yields while reducing the use of pesticides; to create tens of thousands of novel species of bacteria, plants, viruses, and animals; to replace reproduction, or supplement it, with cloning; to create cures for many diseases, increasing our life span and our quality of life; and much, much more. We now know with certainty that these profound changes in the biological sciences are imminent and will challenge all our notions of what life is.

Technologies such as human cloning have in particular raised our awareness of the profound ethical and moral issues we face. If, for example, we were to reengineer ourselves into several separate and unequal species using the power of genetic engineering, then we would threaten the notion of equality that is the very cornerstone of our democracy.

Given the incredible power of genetic engineering, it's no surprise that there are significant safety issues in its use. My friend Amory Lovins recently cowrote, along with Hunter Lovins, an editorial that provides an ecological view of some of these dangers. Among their concerns: that "the new botany aligns the development of plants with their economic, not evolutionary, success." Amory's long career has been focused on energy and resource efficiency by taking a whole-system view of human-made systems; such a whole-system view often finds simple, smart solutions to otherwise seemingly difficult problems and is usefully applied here as well.

After reading the Lovins's editorial, I saw an op-ed by Gregg Easterbrook in *The New York Times* (November 19, 1999) about genetically engineered crops, under the headline: "Food for the Future: Someday, rice will have built-in vitamin A. Unless the Luddites win."

Are Amory and Hunter Lovins Luddites? Certainly not. I believe we all would agree that golden rice, with its built-in vitamin A, is probably a good thing, if developed with proper care and respect for the likely dangers in moving genes across species boundaries.

Awareness of the dangers inherent in genetic engineering is beginning to grow, as reflected in the Lovins's editorial. The general public is aware of, and uneasy about, genetically modified foods and seems to be rejecting the notion that such foods should be permitted to be unlabeled.

But genetic engineering technology is already very far along. As the Lovins note, the USDA has already approved about fifty genetically engineered crops for unlimited release; more than half of the world's soybeans and a third of its corn now contain genes spliced in from other forms of life.

While there are many important issues here, my own major concern with genetic engineering is narrower: that it gives the power — whether militarily, accidentally, or in a deliberate terrorist act — to create a White Plague.

The many wonders of nanotechnology were first imagined by the Nobel laureate physicist Richard Feynman in a speech he gave in 1959, subsequently published under the title "There's Plenty of Room at the Bottom." The book that made a big impression on me, in the mid-'80s, was Eric Drexler's *Engines of Creation*, in which he described beautifully how manipulation of matter at the atomic level could create a utopian future of abundance, where just about everything could be made cheaply, and almost any imaginable disease or physical problem could be solved using nanotechnology and artificial intelligences.

A subsequent book, *Unbounding the Future: The Nanotechnology Revolution*, which Drexler cowrote, imagines some of the changes that might take place in a world where we had molecular-level "assemblers." Assemblers could make possible incredibly low-cost solar power, cures for cancer and the common cold by augmentation of the human immune system, essentially complete cleanup of the environment, incredibly inexpensive pocket supercomputers — in fact, any product would be manufacturable by assemblers at a cost no greater than that of wood — spaceflight more accessible than transoceanic travel today, and restoration of extinct species.

I remember feeling good about nanotechnology after reading *Engines of Creation*. As a technologist, it gave me a sense of calm — that is, nanotechnology showed us that incredible progress was possible, and indeed perhaps inevitable. If nanotechnology was our future, then I didn't feel pressed to solve so many problems in the present. I would get to Drexler's utopian future in due time; I might as well enjoy life more in the here and now. It didn't make sense, given his vision, to stay up all night, all the time.

Drexler's vision also led to a lot of good fun. I would occasionally get to describe the wonders of nanotechnology to others who had not heard of it. After teasing them with all the things Drexler described, I would give a homework assignment of my own: "Use nanotechnology to create a vampire; for extra credit create an antidote."

With these wonders came clear dangers, of which I was acutely aware. As I said at a nanotechnology conference in 1989, "We can't simply do our science and not worry about these ethical issues."[5] But my subsequent conversations with physicists convinced me that nanotechnology might not even work — or, at least, it wouldn't work anytime soon. Shortly thereafter I moved to Colorado, to a skunk works I had set up, and the focus of my work shifted to software for the Internet, specifically on ideas that became Java and Jini.

Then, last summer, Brosl Hasslacher told me that nanoscale molecular electronics was now practical. This was new news, at least to me, and I think to many people — and it radically changed my opinion about nanotechnology. It sent me back to *Engines of Creation*. Rereading Drexler's work after more than ten years, I was dismayed to realize how little I had remembered of its lengthy section called "Dangers and Hopes," including a discussion of how nanotechnologies can become "engines of destruction." Indeed, in my rereading of this cautionary material today, I am struck by how naive some of Drexler's safeguard proposals seem and how much greater I judge the dangers to be now than even he seemed to then. (Having anticipated and described many technical and political problems with nanotechnology, Drexler started the Foresight Institute in the late 1980s "to help prepare society for anticipated advanced technologies" — most important, nanotechnology.)

The enabling breakthrough to assemblers seems quite likely within the next twenty years. Molecular electronics — the new subfield of nanotechnology where individual molecules are circuit elements — should mature quickly and become enormously lucrative within this decade, causing a large incremental investment in all nanotechnologies.

Unfortunately, as with nuclear technology, it is far easier to create destructive uses for nanotechnology than constructive ones. Nanotechnology has clear military and terrorist uses, and you need not be suicidal to release a massively destructive nanotechnological device. Such devices can be built to be selectively destructive, affecting, for example, only a certain geographical area or a group of people who are genetically distinct.

An immediate consequence of the Faustian bargain in obtaining the great power of nanotechnology is that we run a grave risk — the risk that we might destroy the biosphere on which all life depends.

As Drexler explained:

> "Plants" with "leaves" no more efficient than today's solar cells could outcompete real plants, crowding the biosphere with an inedible foliage. Tough omnivorous "bacteria" could out-compete real bacteria: They could spread like blowing pollen, replicate swiftly, and reduce the biosphere to dust in a matter of days. Dangerous replicators could easily be too tough, small, and rapidly spreading to stop — at least if we make no preparation. We have trouble enough controlling viruses and fruit flies.
>
> Among the cognoscenti of nanotechnology, this threat has become known as the "gray goo problem." Though masses of uncontrolled replicators need not be gray or gooey, the term "gray goo" emphasizes that replicators able to obliterate life might be less inspiring than a single species of crabgrass. They might be superior in an evolutionary sense, but this need not make them valuable.

> The gray goo threat makes one thing perfectly clear: We cannot afford certain kinds of accidents with replicating assemblers.

Gray goo would surely be a depressing ending to our human adventure on Earth, far worse than mere fire or ice, and one that could stem from a simple laboratory accident.[6] Oops.

It is most of all the power of destructive self-replication in genetics, nanotechnology, and robotics (GNR) that should give us pause. Self-replication is the modus operandi of genetic engineering, which uses the machinery of the cell to replicate its designs, and the prime danger underlying gray goo in nanotechnology. Stories of run-amok robots like the Borg, replicating or mutating to escape from the ethical constraints imposed on them by their creators, are well established in our science fiction books and movies. It is even possible that self-replication may be more fundamental than we thought, and hence harder — or even impossible — to control. A recent article by Stuart Kauffman in *Nature* titled "Self-Replication: Even Peptides Do It" discusses the discovery that a thirty-two-amino-acid peptide can "autocatalyse its own synthesis." We don't know how widespread this ability is, but Kauffman notes that it may hint at "a route to self-reproducing molecular systems on a basis far wider than Watson-Crick base-pairing."[7]

In truth, we have had in hand for years clear warnings of the dangers inherent in widespread knowledge of GNR technologies — of the possibility of knowledge alone enabling mass destruction. But these warnings haven't been widely publicized; the public discussions have been clearly inadequate. There is no profit in publicizing the dangers.

The nuclear, biological, and chemical (NBC) technologies used in twentieth-century weapons of mass destruction were and are largely military, developed in government laboratories. In sharp contrast, the twenty-first-century GNR technologies have clear commercial uses and are being developed almost exclusively by corporate enterprises. In this age of triumphant commercialism, technology — with science as its handmaiden — is delivering a series of almost magical inventions that are the most phenomenally lucrative ever seen. We are aggressively pursuing the promises of these new technologies within the now unchallenged system of global capitalism and its manifold financial incentives and competitive pressures.

This is the first moment in the history of our planet when any species, by its own voluntary actions, has become a danger to itself — as well as to vast numbers of others.

> It might be a familiar progression, transpiring on many worlds — a planet, newly formed, placidly revolves around its star; life slowly forms; a kaleidoscopic procession of creatures evolves; intelligence emerges which, at least up to a point, confers enormous survival value; and then technology is invented. It dawns on them that there are such things as laws of Nature, that these laws can be revealed by experiment, and that knowledge of these laws can be made both to save and to take lives, both on unprecedented scales. Science, they recognize, grants immense powers. In a flash, they create world-altering contrivances. Some planetary civilizations see their way through, place limits on what may and

what must not be done, and safely pass through the time of perils.

Others, not so lucky or so prudent, perish.

That is Carl Sagan, writing in 1994, in *Pale Blue Dot*, a book describing his vision of the human future in space. I am only now realizing how deep his insight was, and how sorely I miss, and will miss, his voice. For all its eloquence, Sagan's contribution was not least that of simple common sense — an attribute that, along with humility, many of the leading advocates of the twenty-first-century technologies seem to lack.

I remember from my childhood that my grandmother was strongly against the overuse of antibiotics. She had worked since before the first World War as a nurse and had a common sense attitude that taking antibiotics, unless they were absolutely necessary, was bad for you.

It is not that she was an enemy of progress. She saw much progress in an almost seventy-year nursing career; my grandfather, a diabetic, benefited greatly from the improved treatments that became available in his lifetime. But she, like many levelheaded people, would probably think it greatly arrogant for us, now, to be designing a robotic "replacement species" when we obviously have so much trouble making relatively simple things work and so much trouble managing — or even understanding — ourselves.

I realize now that she had an awareness of the nature of the order of life and of the necessity of living with and respecting that order. With this respect comes a necessary humility that we, with our early-twenty-first-century chutzpah, lack at our peril. The common sense view, grounded in this respect, is often right, in advance of the scientific evidence. The clear fragility and inefficiencies of the human-made systems we have built should give us all pause; the fragility of the systems I have worked on certainly humbles me.

We should have learned a lesson from the making of the first atomic bomb and the resulting arms race. We didn't do well then, and the parallels to our current situation are troubling.

The effort to build the first atomic bomb was led by the brilliant physicist J. Robert Oppenheimer. Oppenheimer was not naturally interested in politics but became painfully aware of what he perceived as the grave threat to Western civilization from the Third Reich, a threat surely grave because of the possibility that Hitler might obtain nuclear weapons. Energized by this concern, he brought his strong intellect, passion for physics, and charismatic leadership skills to Los Alamos and led a rapid and successful effort by an incredible collection of great minds to quickly invent the bomb.

What is striking is how this effort continued so naturally after the initial impetus was removed. In a meeting shortly after V-E Day with some physicists who felt that perhaps the effort should stop, Oppenheimer argued to continue. His stated reason seems a bit strange: not because of the fear of large casualties from an invasion of Japan, but because the United Nations, which was soon to be formed, should have foreknowledge of atomic weapons. A more likely reason the project continued is the momentum that had built up — the first atomic test, Trinity, was nearly at hand.

We know that in preparing this first atomic test the physicists proceeded despite a large number of possible dangers. They were initially worried, based on a calculation by Edward Teller, that an atomic explosion might set fire to the atmosphere. A revised calculation reduced the danger of destroying the world to a three-in-a-million chance. (Teller says he was later able to dismiss the prospect of atmospheric ignition entirely.) Oppenheimer, though, was sufficiently concerned about the result of Trinity that he arranged for a possible evacuation of the southwest part of the state of New Mexico. And, of course, there was the clear danger of starting a nuclear arms race.

Within a month of that first, successful test, two atomic bombs destroyed Hiroshima and Nagasaki. Some scientists had suggested that the bomb simply be demonstrated, rather than dropped on Japanese cities — saying that this would greatly improve the chances for arms control after the war — but to no avail. With the tragedy of Pearl Harbor still fresh in Americans' minds, it would have been very difficult for President Truman to order a demonstration of the weapons rather than use them as he did — the desire to quickly end the war and save the lives that would have been lost in any invasion of Japan was very strong. Yet the overriding truth was probably very simple: As the physicist Freeman Dyson later said, "The reason that it was dropped was just that nobody had the courage or the foresight to say no."

It's important to realize how shocked the physicists were in the aftermath of the bombing of Hiroshima, on August 6, 1945. They describe a series of waves of emotion: first, a sense of fulfillment that the bomb worked, then horror at all the people that had been killed, and then a convincing feeling that on no account should another bomb be dropped. Yet of course another bomb was dropped, on Nagasaki, only three days after the bombing of Hiroshima.

In November 1945, three months after the atomic bombings, Oppenheimer stood firmly behind the scientific attitude, saying, "It is not possible to be a scientist unless you believe that the knowledge of the world, and the power which this gives, is a thing which is of intrinsic value to humanity, and that you are using it to help in the spread of knowledge and are willing to take the consequences."

Oppenheimer went on to work, with others, on the Acheson-Lilienthal report, which, as Richard Rhodes says in his recent book *Visions of Technology*, "found a way to prevent a clandestine nuclear arms race without resorting to armed world government"; their suggestion was a form of relinquishment of nuclear weapons work by nation-states to an international agency.

This proposal led to the Baruch Plan, which was submitted to the United Nations in June 1946 but never adopted (perhaps because, as Rhodes suggests, Bernard Baruch had "insisted on burdening the plan with conventional sanctions," thereby inevitably dooming it, even though it would "almost certainly have been rejected by Stalinist Russia anyway"). Other efforts to promote sensible steps toward internationalizing nuclear power to prevent an arms race ran afoul either of U.S. politics and internal distrust, or distrust by the Soviets. The opportunity to avoid the arms race was lost and very quickly.

Two years later, in 1948, Oppenheimer seemed to have reached another stage in his thinking, saying, "In some sort of crude sense which no vulgarity, no humor, no overstatement can quite extinguish, the physicists have known sin; and this is a knowledge they cannot lose."

In 1949, the Soviets exploded an atom bomb. By 1955, both the U.S. and the Soviet Union had tested hydrogen bombs suitable for delivery by aircraft. And so the nuclear arms race began.

Nearly twenty years ago, in the documentary *The Day After Trinity*, Freeman Dyson summarized the scientific attitudes that brought us to the nuclear precipice:

> I have felt it myself. The glitter of nuclear weapons. It is irresistible if you come to them as a scientist. To feel it's there in your hands, to release this energy that fuels the stars, to let it do your bidding. To perform these miracles, to lift a million tons of rock into the sky. It is something that gives people an illusion of illimitable power, and it is, in some ways, responsible for all our troubles — this, what you might call technical arrogance, that overcomes people when they see what they can do with their minds.[8]

Now, as then, we are creators of new technologies and stars of the imagined future, driven — this time by great financial rewards and global competition — despite the clear dangers, hardly evaluating what it may be like to try to live in a world that is the realistic outcome of what we are creating and imagining.

In 1947, *The Bulletin of the Atomic Scientists* began putting a Doomsday Clock on its cover. For more than fifty years, it has shown an estimate of the relative nuclear danger we have faced, reflecting the changing international conditions. The hands on the clock have moved fifteen times and today, standing at nine minutes to midnight, reflect continuing and real danger from nuclear weapons. The recent addition of India and Pakistan to the list of nuclear powers has increased the threat of failure of the nonproliferation goal, and this danger was reflected by moving the hands closer to midnight in 1998. [The hands were moved to five minutes before midnight in 2007, then back to six minutes before in January 2010. – Ed.]

In our time, how much danger do we face, not just from nuclear weapons, but from all of these technologies? How high are the extinction risks?

The philosopher John Leslie has studied this question and concluded that the risk of human extinction is at least 30 percent,[9] while Ray Kurzweil believes we have "a better than even chance of making it through," with the caveat that he has "always been accused of being an optimist." Not only are these estimates not encouraging, but they do not include the probability of many horrid outcomes that lie short of extinction.

Faced with such assessments, some serious people are already suggesting that we simply move beyond Earth as quickly as possible. We would colonize the galaxy using von Neumann probes, which hop from star system to star system, replicating as they go. This step will almost certainly be necessary five billion years from now (or sooner if our solar system is disastrously impacted by the impending collision of our galaxy with the Andromeda galaxy within the next three billion years), but if we take Kurzweil and Moravec at their word, it might be necessary by the middle of this century.

What are the moral implications here? If we must move beyond Earth this quickly in order for the species to survive, who accepts the responsibility for the fate of those (most of us, after all) who are left behind? And even if we scatter to the stars, isn't it likely that we may take our problems with us or find, later, that they have followed us? The fate of our species on Earth and our fate in the galaxy seem inextricably linked.

Another idea is to erect a series of shields to defend against each of the dangerous technologies. The Strategic Defense Initiative, proposed by the Reagan administration, was an attempt to design such a shield against the threat of a nuclear attack from the Soviet Union. But as Arthur C. Clarke, who was privy to discussions about the project, observed:

> Though it might be possible, at vast expense, to construct local defense systems that would "only" let through a few percent of ballistic missiles, the much touted idea of a national umbrella was nonsense. Luis Alvarez, perhaps the greatest experimental physicist of this century, remarked to me that the advocates of such schemes were "very bright guys with no common sense."

Clarke continued:

> Looking into my often cloudy crystal ball, I suspect that a total defense might indeed be possible in a century or so. But the technology involved would produce, as a by-product, weapons so terrible that no one would bother with anything as primitive as ballistic missiles.[10]

In *Engines of Creation*, Eric Drexler proposed that we build an active nanotechnological shield — a form of immune system for the biosphere — to defend against dangerous replicators of all kinds that might escape from laboratories or otherwise be maliciously created. But the shield he proposed would itself be extremely dangerous — nothing could prevent it from developing autoimmune problems and attacking the biosphere itself.[11]

Similar difficulties apply to the construction of shields against robotics and genetic engineering. These technologies are too powerful to be shielded against in the time frame of interest; even if it were possible to implement defensive shields, the side effects of their development would be at least as dangerous as the technologies we are trying to protect against.

These possibilities are all thus either undesirable or unachievable or both. The only realistic alternative I see is relinquishment: to limit development of the technologies that are too dangerous, by limiting our pursuit of certain kinds of knowledge.

Yes, I know, knowledge is good, as is the search for new truths. We have been seeking knowledge since ancient times. Aristotle opened his *Metaphysics* with the simple statement, "All men by nature desire to know." We have, as a bedrock value in our society, long agreed on the value of open access to information, and recognize the problems that arise with attempts to restrict access to and development of knowledge. In recent times, we have come to revere scientific knowledge.

But despite the strong historical precedents, if open access to and unlimited development of knowledge henceforth puts us all in clear danger of extinction, then common sense demands that we reexamine even these basic, long-held beliefs.

It was Nietzsche who warned us, at the end of the nineteenth century, not only that God is dead but that "faith in science, which after all exists undeniably, cannot owe its origin to a calculus of utility; it must have originated in spite of the fact that the disutility and dangerousness of the 'will to truth,' of 'truth at any price' is proved to it constantly." It is this further danger that we now fully face — the consequences of our truth seeking. The truth that science seeks can certainly be considered a dangerous substitute for God if it is likely to lead to our extinction.

If we could agree, as a species, what we wanted, where we were headed, and why, then we would make our future much less dangerous — then we might understand what we can and should relinquish. Otherwise, we can easily imagine an arms race developing over GNR technologies, as it did with the NBC technologies in the twentieth century. This is perhaps the greatest risk, for once such a race begins, it's very hard to end it. This time — unlike during the Manhattan Project — we aren't in a war, facing an implacable enemy that is threatening our civilization; we are driven, instead, by our habits, our desires, our economic system, and our competitive need to know.

I believe that we all wish our course could be determined by our collective values, ethics, and morals. If we had gained more collective wisdom over the past few thousand years, then a dialogue to this end would be more practical, and the incredible powers we are about to unleash would not be nearly so troubling.

One would think we might be driven to such a dialogue by our instinct for self-preservation. Individuals clearly have this desire, yet as a species our behavior seems to be not in our favor. In dealing with the nuclear threat, we often spoke dishonestly to ourselves and to each other, thereby greatly increasing the risks. Whether this was politically motivated, or because we chose not to think ahead, or because when faced with such grave threats we acted irrationally out of fear, I do not know, but it does not bode well.

The new Pandora's boxes of genetics, nanotechnology, and robotics are almost open, yet we seem hardly to have noticed. Ideas can't be put back in a box; unlike uranium or plutonium, they don't need to be mined and refined, and they can be freely copied. Once they are out, they are out. Churchill remarked, in a famous left-handed compliment, that the American people and their leaders "invariably do the right thing, after they have examined every other alternative." In this case, however, we must act more presciently, as to do the right thing only at last may be to lose the chance to do it at all.

As Thoreau said, "We do not ride on the railroad; it rides upon us"; and this is what we must fight, in our time. The question is, indeed, Which is to be master? Will we survive our technologies?

We are being propelled into this new century with no plan, no control, no brakes. Have we already gone too far down the path to alter course? I don't believe so, but we aren't trying yet, and the last chance to assert control — the fail-safe point — is rapidly approaching. We have our first pet robots, as well as commercially available genetic engineering techniques, and our nanoscale techniques are advancing rapidly. While the development of these technologies proceeds through a number of steps, it isn't necessarily the case — as happened in the Manhattan Project and the Trinity test — that the last step in proving a technology is large and hard. The breakthrough to wild self-replication in robotics, genetic engineering, or nanotechnology could come suddenly, reprising the surprise we felt when we learned of the cloning of a mammal.

And yet I believe we do have a strong and solid basis for hope. Our attempts to deal with weapons of mass destruction in the last century provide a shining example of relinquishment for us to consider: the unilateral U.S. abandonment, without preconditions, of the development of biological weapons. This relinquishment stemmed from the realization that while it would take an enormous effort to create these terrible weapons, they could from then on easily be duplicated and fall into the hands of rogue nations or terrorist groups.

The clear conclusion was that we would create additional threats to ourselves by pursuing these weapons and that we would be more secure if we did not pursue them. We have embodied our relinquishment of biological and chemical weapons in the 1972 Biological Weapons Convention (BWC) and the 1993 Chemical Weapons Convention (CWC).[12]

As for the continuing sizable threat from nuclear weapons, which we have lived with now for more than fifty years, the U.S. Senate's recent rejection of the Comprehensive Test Ban Treaty makes it clear relinquishing nuclear weapons will not be politically easy. But we have a unique opportunity, with the end of the Cold War, to avert a multipolar arms race. Building on the BWC and CWC relinquishments, successful abolition of nuclear weapons could help us build toward a habit of relinquishing dangerous technologies. (Actually, by getting rid of all but 100 nuclear weapons worldwide — roughly the total destructive power of World War II and a considerably easier task — we could eliminate this extinction threat.[13])

Verifying relinquishment will be a difficult problem, but not an unsolvable one. We are fortunate to have already done a lot of relevant work in the context of the BWC and other treaties. Our major task will be to apply this to technologies that are naturally much more commercial than military. The substantial need here is for transparency, as difficulty of verification is directly proportional to the difficulty of distinguishing relinquished from legitimate activities.

I frankly believe that the situation in 1945 was simpler than the one we now face: The nuclear technologies were reasonably separable into commercial and military uses and monitoring was aided by the nature of atomic tests and the ease with which radioactivity could be measured. Research on military applications could be performed at national laboratories such as Los Alamos, with the results kept secret as long as possible.

The GNR technologies do not divide clearly into commercial and military uses; given their potential in the market, it's hard to imagine pursuing them only in national laboratories. With their widespread commercial pursuit, enforcing relinquishment will require a verification regime similar to that for biological weapons, but on an unprecedented scale. This, inevitably, will raise tensions between our individual privacy and desire for proprietary information and the need for verification to protect us all. We will undoubtedly encounter strong resistance to this loss of privacy and freedom of action.

Verifying the relinquishment of certain GNR technologies will have to occur in cyberspace as well as at physical facilities. The critical issue will be to make the necessary transparency acceptable in a world of proprietary information, presumably by providing new forms of protection for intellectual property.

Verifying compliance will also require that scientists and engineers adopt a strong code of ethical conduct, resembling the Hippocratic oath, and that they have the courage to whistleblow as necessary, even at high personal cost. This would answer the call — fifty years after Hiroshima — by the Nobel laureate Hans Bethe, one of the most senior of the surviving members of the Manhattan Project, that all scientists "cease and desist from work creating, developing, improving, and manufacturing nuclear weapons and other weapons of potential mass destruction."[14] In the twenty-first century, this requires vigilance and personal responsibility by those who would work on both NBC and GNR technologies to avoid implementing weapons of mass destruction and knowledge-enabled mass destruction.

Thoreau also said that we will be "rich in proportion to the number of things which we can afford to let alone." We each seek to be happy, but it would seem worthwhile to question whether we need to take such a high risk of total destruction to gain yet more knowledge and yet more things; common sense says that there is a limit to our material needs — and that certain knowledge is too dangerous and is best forgone.

Neither should we pursue near immortality without considering the costs, without considering the commensurate increase in the risk of extinction. Immortality, while perhaps the original, is certainly not the only possible utopian dream.

I recently had the good fortune to meet the distinguished author and scholar Jacques Attali, whose book *Lignes d'horizons* (*Millennium*, in the English translation) helped inspire the Java and Jini approach to the coming age of pervasive computing, as previously described in [*Wired*]. In his new book *Fraternités*, Attali describes how our dreams of utopia have changed over time:

> At the dawn of societies, men saw their passage on Earth as nothing more than a labyrinth of pain, at the end of which stood a door leading, via their death, to the company of gods and to *Eternity*. With the Hebrews and then the Greeks, some men dared free themselves from theological demands and dream of an ideal City where *Liberty* would flourish. Others, noting the evolution of the market society, understood that the liberty of some would entail the alienation of others, and they sought *Equality*.

Jacques helped me understand how these three different utopian goals exist in tension in our society today. He goes on to describe a fourth utopia, Fraternity, whose foundation is altruism. Fraternity alone associates individual happiness with the happiness of others, affording the promise of self-sustainment.

This crystallized for me my problem with Kurzweil's dream. A technological approach to Eternity — near immortality through robotics — may not be the most desirable utopia, and its pursuit brings clear dangers. Maybe we should rethink our utopian choices.

Where can we look for a new ethical basis to set our course? I have found the ideas in the book *Ethics for the New Millennium*, by the Dalai Lama, to be very helpful. As is perhaps well known but little heeded, the Dalai Lama argues that the most important thing is for us to conduct our lives with love and compassion for others and that our societies need to develop a stronger notion of universal responsibility and of our interdependency; he proposes a standard of positive ethical conduct for individuals and societies that seems consonant with Attali's Fraternity utopia.

The Dalai Lama further argues that we must understand what it is that makes people happy and acknowledge the strong evidence that neither material progress nor the pursuit of the power of knowledge is the key — that there are limits to what science and the scientific pursuit alone can do.

Our Western notion of happiness seems to come from the Greeks, who defined it as "the exercise of vital powers along lines of excellence in a life affording them scope."[15]

Clearly, we need to find meaningful challenges and sufficient scope in our lives if we are to be happy in whatever is to come. But I believe we must find alternative outlets for our creative forces, beyond the culture of perpetual economic growth; this growth has largely been a blessing for several hundred years, but it has not brought us unalloyed happiness, and we must now choose between the pursuit of unrestricted and undirected growth through science and technology and the clear accompanying dangers.

It is now more than a year since my first encounter with Ray Kurzweil and John Searle. I see around me cause for hope in the voices for caution and relinquishment and in those people I have discovered who are as concerned as I am about our current predicament. I feel, too, a deepened sense of personal responsibility — not for the work I have already done, but for the work that I might yet do, at the confluence of the sciences.

But many other people who know about the dangers still seem strangely silent. When pressed, they trot out the "this is nothing new" riposte — as if awareness of what could happen is response enough. They tell me, There are universities filled with bioethicists who study this stuff all day long. They say, All this has been written about before, and by experts. They complain, Your worries and your arguments are already old hat.

I don't know where these people hide their fear. As an architect of complex systems I enter this arena as a generalist. But should this diminish my concerns? I am aware of how much has been written about, talked about, and lectured about so authoritatively. But does this mean it has reached people? Does this mean we can discount the dangers before us?

Knowing is not a rationale for not acting. Can we doubt that knowledge has become a weapon we wield against ourselves?

The experiences of the atomic scientists clearly show the need to take personal responsibility, the danger that things will move too fast, and the way in which a process can take on a life of its own. We can, as they did, create insurmountable problems in almost no time flat. We must do more thinking up front if we are not to be similarly surprised and shocked by the consequences of our inventions.

My continuing professional work is on improving the reliability of software. Software is a tool, and as a toolbuilder I must struggle with the uses to which the tools I make are put. I have always believed that making software more reliable, given its many uses, will make the world a safer and better place; if I were to come to believe the opposite, then I would be morally obligated to stop this work. I can now imagine such a day may come.

This all leaves me not angry but at least a bit melancholic. Henceforth, for me, progress will be somewhat bittersweet.

Do you remember the beautiful penultimate scene in *Manhattan* where Woody Allen is lying on his couch and talking into a tape recorder? He is writing a short story about people who are creating unnecessary, neurotic problems for themselves, because it keeps them from dealing with more unsolvable, terrifying problems about the universe.

He leads himself to the question, "Why is life worth living?" and to consider what makes it worthwhile for him: Groucho Marx, Willie Mays, the second movement of the Jupiter Symphony, Louis Armstrong's recording of "Potato Head Blues," Swedish movies, Flaubert's *Sentimental Education*, Marlon Brando, Frank Sinatra, the apples and pears by Cézanne, the crabs at Sam Wo's, and, finally, the showstopper: his love Tracy's face.

Each of us has our precious things, and as we care for them we locate the essence of our humanity. In the end, it is because of our great capacity for caring that I remain optimistic we will confront the dangerous issues now before us.

My immediate hope is to participate in a much larger discussion of the issues raised here, with people from many different backgrounds, in settings not predisposed to fear or favor technology for its own sake.

As a start, I have twice raised many of these issues at events sponsored by the Aspen Institute and have separately proposed that the American Academy of Arts and Sciences take them up as an extension of its work with the Pugwash Conferences. (These have been held since 1957 to discuss arms control, especially of nuclear weapons, and to formulate workable policies.)

It's unfortunate that the Pugwash meetings started only well after the nuclear genie was out of the bottle — roughly fifteen years too late. We are also getting a belated start on seriously addressing the issues around twenty-first-century technologies — the prevention of knowledge-enabled mass destruction — and further delay seems unacceptable.

So I'm still searching; there are many more things to learn. Whether we are to succeed or fail, to survive or fall victim to these technologies, is not yet decided. I'm up late again — it's almost 6 A.M. I'm trying to imagine some better answers, to break the spell and free them from the stone.

ENDNOTES

1. The passage Kurzweil quotes is from Kaczynski's "Unabomber Manifesto," which was published jointly, under duress, by *The New York Times* and *The Washington Post*

to attempt to bring his campaign of terror to an end. I agree with David Gelernter, who said about their decision:

> It was a tough call for the newspapers. To say yes would be giving in to terrorism, and for all they knew he was lying anyway. On the other hand, to say yes might stop the killing. There was also a chance that someone would read the tract and get a hunch about the author; and that is exactly what happened. The suspect's brother read it, and it rang a bell.
>
> I would have told them not to publish. I'm glad they didn't ask me. I guess. (*Drawing Life: Surviving the Unabomber.* Free Press, 1997:120.)

2. Laurie Garrett, *The Coming Plague: Newly Emerging Diseases in a World Out of Balance* (New York: Penguin, 1994), pp. 47–52, 414, 419, 452.
3. Isaac Asimov described what became the most famous view of ethical rules for robot behavior in his book *I, Robot* in 1950, in his Three Laws of Robotics: 1. A robot may not injure a human being, or, through inaction, allow a human being to come to harm. 2. A robot must obey the orders given it by human beings, except where such orders would conflict with the First Law. 3. A robot must protect its own existence, as long as such protection does not conflict with the First or Second Law.
4. Michelangelo wrote a sonnet that begins:

   Non ha l' ottimo artista alcun concetto
   Ch' un marmo solo in sè non circonscriva
   Col suo soverchio; e solo a quello arriva
   La man che ubbidisce all' intelleto.

   Stone translates this as:

   The best of artists hath no thought to show
   which the rough stone in its superfluous shell
   doth not include; to break the marble spell
   is all the hand that serves the brain can do.

   Stone describes the process: "He was not working from his drawings or clay models; they had all been put away. He was carving from the images in his mind. His eyes and hands knew where every line, curve, mass must emerge, and at what depth in the heart of the stone to create the low relief." *The Agony and the Ecstasy* (New York: Doubleday, 1961), pp. 6, 144.
5. First Foresight Conference on Nanotechnology in October 1989, a talk titled "The Future of Computation." Published in B.C. Crandall, and James Lewis, eds., *Nanotechnology: Research and Perspectives* (Boston; MIT Press, 1992), p. 269. See also <www.foresight.org/Conferences/MNT01/Nano1.html>.
6. In his 1963 novel *Cat's Cradle*, Kurt Vonnegut imagined a gray-goo-like accident where a form of ice called ice-nine, which becomes solid at a much higher temperature, freezes the oceans.
7. Stuart Kauffman, "Self-replication: Even Peptides Do It." *Nature*, 382 (August 8, 1996), p. 496. See <www.santafe.edu/sfi/People/kauffman/sak-peptides.html>.
8. Jon Else, *The Day After Trinity: J. Robert Oppenheimer and The Atomic Bomb* (available at <www.pyramiddirect.com>).
9. This estimate is in Leslie's book *The End of the World: The Science and Ethics of Human Extinction*, where he notes that the probability of extinction is substantially higher if we accept Brandon Carter's Doomsday Argument, which is, briefly, that "we ought to have some reluctance to believe that we are very exceptionally early, for instance in the earliest 0.001 percent, among all humans who will ever have lived. This

would be some reason for thinking that humankind will not survive for many more centuries, let alone colonize the galaxy. Carter's doomsday argument doesn't generate any risk estimates just by itself. It is an argument for revising the estimates which we generate when we consider various possible dangers." (Routledge, 1996, pp. 1, 3, 145.)

10. Arthur C. Clarke, "Presidents, Experts, and Asteroids." *Science*, June 5, 1998. Reprinted as "Science and Society" in *Greetings, Carbon-Based Bipeds! Collected Essays, 1934–1998* (New York: St. Martin's Press, 1999), p. 526.
11. And, as David Forrest suggests in his paper "Regulating Nanotechnology Development," available at <www.foresight.org/NanoRev/Forrest1989.html>, "If we used strict liability as an alternative to regulation it would be impossible for any developer to internalize the cost of the risk (destruction of the biosphere), so theoretically the activity of developing nanotechnology should never be undertaken." Forrest's analysis leaves us with only government regulation to protect us — not a comforting thought.
12. Matthew Meselson, "The Problem of Biological Weapons." Presentation to the 1,818th Stated Meeting of the American Academy of Arts and Sciences, January 13, 1999 (available at <minerva.amacad.org/archive/bulletin4.htm>).
13. Paul Doty, "The Forgotten Menace: Nuclear Weapons Stockpiles Still Represent the Biggest Threat to Civilization." *Nature*, 402 (December 9, 1999), p. 583.
14. See also Hans Bethe's 1997 letter to President Clinton, at <www.fas.org/bethecr.htm>.
15. Edith Hamilton, *The Greek Way* (New York: W. W. Norton & Co., 1942), p. 35.

10

A Response to Bill Joy and the Doom-and-Gloom Technofuturists

JOHN SEELY BROWN AND PAUL DUGUID

Bill Joy's article drew many responses, some concurring with Joy's argument, most not. Among the most interesting and articulate rebuttals of the Joy thesis was one published not long after Joy's in The Industry Standard, *a magazine of e-commerce, technology, media, and politics that was launched in early 1998 and became an unfortunate casualty of the dot.com crash in September 2001. The authors of that piece, colleagues of Joy in high-tech industry, take him to task for basing his pessimistic vision on an overly simplistic way of viewing the influence of technology on society. "Technology and society are constantly forming and reforming new dynamic equilibriums with far-reaching implications," they write. "The challenge for futurology (and for all of us) is to see beyond the hype and past the oversimplifications to the full import of these new sociotechnical formations." In essence, they are restating the outlook articulated by Robert Pool in Part I of this book.*

John Seely Brown is a visiting scholar and advisor to the provost at University of Southern California and the Independent co-chairman of the Deloitte Center for the Edge, a part of Deloitte that "helps senior executives make sense of and profit from emerging opportunities on the edge of business and technology." Brown was formerly chief scientist of the Xerox Corporation from 1992 to 2002, and director of the Xerox Palo Alto Research Center (PARC) from 1990 to

2000. He is also a cofounder of the Institute for Research on Learning, a non-profit institute for addressing the problems of lifelong learning. Paul Duguid, a historian and social theorist, is an adjunct full professor at the School of Information, at the University of California, Berkeley, and a professorial research fellow at Queen Mary, University of London. He was formerly a member of the Institute for Research on Learning, and is coauthor, with John Seely Brown, of The Social Life of Information *(2000).*

If you lived through the 1950s, you might remember President Eisenhower, orderly suburban housing tracts, backyard bomb shelters — and dreams of a nuclear power plant in every home. Plans for industrial nuclear generators had barely left the drawing board before futurists predicted that every house would have a miniature version. From there, technoenthusiasts predicted the end of power monopolies, the emergence of the "electronic cottage," the death of the city and the decline of the corporation.

... Pessimists and Luddites, of course, envisioned nuclear apocalypse. Each side waited for nirvana, or Armageddon, so it could triumphantly tell the other, "I told you so."...With "Why the Future Doesn't Need Us" in the April [2000] issue of *Wired*, Bill Joy invokes those years gone by. No Luddite, Joy is an awe-inspiring technologist — as cofounder and chief scientist of Sun Microsystems, he coauthored, among other things, the Java programming language. So when his article describes a technological juggernaut thundering toward society — bringing with it mutant genes, molecular-level nanotechnology machines, and superintelligent robots — all need to listen. Like the nuclear prognosticators, Joy can see the juggernaut clearly. What he can't see — which is precisely what makes his vision so scary — are any controls.

But it doesn't follow that the juggernaut is uncontrollable. To understand why not, readers should note the publication in which this article appeared. For the better part of a decade, *Wired* has been a cheerleader for the digital age. Until now, *Wired* has rarely been a venue to which people have looked for a way to put a brake on innovation. Therefore, its shift with Joy's article from cheering to warning marks an important and surprising moment in the digital zeitgeist. In an effort to locate some controls, let's go back to the nuclear age. Innovation, the argument went back in the 1950s, would make nuclear power plants smaller and cheaper. They would enter mass production and quickly become available to all.

Even today, the argument might appear inescapable until you notice what's missing: The tight focus of this vision makes it almost impossible to see forces other than technology at work. In the case of nuclear development, a host of forces worked to dismantle the dream of a peaceful atom, including the environmental movement, antinuclear protests, concerned scientists, worried neighbors of Chernobyl and Three Mile Island, government regulators and antiproliferation treaties. Cumulatively, these forces slowed the nuclear juggernaut to a crawl.

Similar social forces are at work on technologies today. But because the digerati, like technoenthusiasts before them, look to the future with technological tunnel vision, they too have trouble bringing other forces into view.

THE TUNNEL AHEAD

In Joy's vision, as in the nuclear one, there's a recognizable tunnel vision that leaves people out of the picture and focuses on technology in splendid isolation. This vision leads not only to doom-and-gloom scenarios, but also to tunnel design: the design of "simple" technologies that are actually difficult to use.

To escape both trite scenarios and bad design, we have to widen our horizons and bring into view not only technological systems, but also social systems. Good designs look beyond the dazzling potential of the technology to social factors, such as the limited patience of most users.

Paying attention to the latter has, for example, allowed the Palm Pilot and Nintendo Game Boy to sweep aside more complex rivals. Their elegant simplicity has made them readily usable. And their usability has in turn created an important social support system. The devices are so widely used that anyone having trouble with a Pilot or Game Boy rarely has to look far for advice from a more experienced user.

As this small example suggests, technological and social systems shape each other. The same is true on a larger scale. Technologies — such as gunpowder, the printing press, the railroad, the telegraph, and the Internet — can shape society in profound ways. But, on the other hand, social systems — in the form of governments, the courts, formal and informal organizations, social movements, professional networks, local communities, market institutions, and so forth — shape, moderate, and redirect the raw power of technologies.

Given the crisp edges of technology and the fuzzy outlines of society, it certainly isn't easy to use these two worldviews simultaneously. But if you want to see where we are going, or design the means to get there, you need to grasp both.

This perspective allows a more sanguine look at Joy's central concerns: genetic engineering, nanotechnology, and robotics. Undoubtedly, each deserves serious thought. But each should be viewed in the context of the social system in which it is inevitably embedded.

Genetic engineering presents the clearest example. Barely a year ago, the technology seemed to be an unstoppable force. Major chemical and agricultural interests were barreling down an open highway. In the past year, however, road conditions changed dramatically for the worse: Cargill faced Third World protests against its patents; Monsanto suspended research on sterile seeds; and champions of genetically modified foods, who once saw an unproblematic and lucrative future, are scurrying to counter consumer boycotts of their products.

Almost certainly, those who support genetic modification will have to look beyond the technology if they want to advance it. They need to address society directly — not just by putting labels on modified foods, but by educating people about the costs and the benefits of these new agricultural products. Having ignored social concerns, however, proponents have made the people they need to educate profoundly suspicious and hostile.

Nanotechnology offers a rather different example of how the future can frighten us. Because the technology involves engineering at a molecular level, both the promise and the threat seem immeasurable. But they are immeasurable for a good reason: The technology is still almost wholly on the drawing board.

Two of nanotechnology's main proponents, Ralph Merkle and Eric Drexler, worked with us at the Xerox Palo Alto Research Center in Palo Alto, California. The two built powerful nano-CAD tools and then ran simulations of the resulting molecular-level designs. These experiments showed definitively that nano devices are theoretically feasible. No one, however, has laid out a route from lab-based simulation to practical systems in any detail.

In the absence of a plan, it's important to ask the right questions: Can nanotechnology fulfill its great potential in tasks ranging from data storage to pollution control, all without spiraling out of control? If the lesson of genetic engineering is any guide, planners would do well to consult and educate the public early on, even though useful nano systems are probably decades away.

Worries about robotics appear premature, as well. Internet "bots" that search, communicate, and negotiate for their human masters may appear to behave like Homo sapiens, but, in fact, bots are often quite inept at functions that humans do well — functions that call for judgment, discretion, initiative, or tacit understanding. They are good (and useful) for those tasks that humans do poorly. So they are better thought of as complementary systems, not rivals to humanity. Although bots will undoubtedly get better at what they do, such development will not necessarily make them more human.

Are more conventional clanking robots — the villains of science fiction — any great threat to society? We doubt it. Xerox PARC research on self-aware, reconfigurable "polybots" has pushed the boundaries of what robots can do, pointing the way to "morphing robots" that are able to move and change shape.

Nonetheless, for all their cutting-edge agility, these robots are a long way from making good dance partners. The chattiness of *Star Wars*' C-3PO still lies well beyond real-world machines. Indeed, what talk robots or computers achieve, though it may appear similar, is quite different from human talk. Talking machines travel routes designed specifically to avoid the full complexities of human language.

Robots may seem intelligent, but such intelligence is profoundly hampered by their inability to learn in any significant way. (This failing has apparently led Toyota, after heavy investment in robotics, to consider replacing robots with humans on many production lines.) And, without learning, simple common sense will lie beyond robots for a long time to come.

Indeed, despite years of startling advances and innumerable successes like the chess-playing Big Blue, computer science is still about as far as it ever was from building a machine with the learning abilities, linguistic competence, common sense or social skills of a five-year-old child.

As with Internet bots, real-world robots will no doubt become increasingly useful. But they will probably also become increasingly frustrating to use as a result of tunnel design. In that regard, they may indeed seem antisocial, but not in the way of *Terminator*-like fantasies of robot armies that lay waste ... human society.

Indeed, the thing that handicaps robots most is their lack of a social existence. For it is our social existence as humans that shapes how we speak, learn, think, and develop common sense. All forms of artificial life (whether bugs or bots) will remain primarily a metaphor for — rather than a threat to — society, at least until they manage to enter a debate, sing in a choir, take a class, survive a committee meeting, join a union, pass a law, engineer a cartel, or summon a constitutional convention.

These critical social mechanisms allow society to shape its future. It is through planned, collective action that society forestalls expected consequences (such as Y2K) and responds to unexpected events (such as epidemics).

THE FAILURE OF A "6-D" VISION

Why does the threat of cunning, replicating robot society look so close from one perspective, yet so distant from another? The difference lies in the well-known tendency of futurologists to count "1, 2, 3 … a million." That is, once the first step on a path is taken, it's very easy to assume that all subsequent steps are trivial.

Several of the steps Joy asks us to take — the leap from genetic engineering to a "white plague," from simulations to out-of-control nanotechnology, from replicating peptides to a "robot species" — are extremely large. And they are certainly not steps that will be taken without diversions, regulations, or controls.

One of the lessons of Joy's article, then, is that the path to the future can look simple (and sometimes downright terrifying) if you look at it through what we call "6-D lenses." We coined this phrase having so often in our research hit up against … such "de-" or "di-" words as demassification, decentralization, disintermediation, despacialization, disaggregation, and demarketization in the canon of futurology.

If you take any one of these words in isolation, it's easy to follow their relentless logic to its evident conclusion. Because firms are getting smaller, for example, it's easy to assume that companies and other intermediaries are simply disintegrating into markets. And because communication is growing cheaper and more powerful, it's easy to believe in the "death of distance."

But things rarely work in such linear fashion. Other forces are often at work, such as those driving firms into larger and larger mergers to take advantage of social, rather than merely technological, networks. Similarly, even though communications technology has killed distance, people curiously can't stay away from the social hotbed of modern communications technology, Silicon Valley.

Importantly, these d-words indicate that the old ties that once bound communities, organizations, and institutions are being picked apart by technologies. A simple, linear reading, then, suggests that these communities, organizations, and institutions will now simply fall apart. A more complex reading, taking into account the multiple forces at work, offers another picture.

While many powerful national corporations have grown insignificant, some have transformed into more powerful transnational firms. While some forms of community may be dying, others, bolstered by technology, are growing stronger.

Technology and society are constantly forming and reforming new dynamic equilibriums with far-reaching implications. The challenge for futurology (and for all of us) is to see beyond the hype and past the oversimplifications to the full import of these new sociotechnical formations. Two hundred years ago, Thomas Malthus, assuming that human society and agricultural technology developed on separate paths, predicted that society was growing so fast that it would starve itself to death, the so-called Malthusian trap.

A hundred years later, H. G. Wells similarly assumed that society and technology were developing independently. Like many people today, Wells saw the advance of technology outstripping the evolution of society, leading him to predict that technology's relentless juggernaut would unfeelingly crush society. Like Joy, both Malthus and Wells issued important warnings, alerting society to the dangers it faced. But by their actions, Malthus and Wells helped prevent the very future they were so certain would come about.

These self-*un*fulfilling prophecies failed to see that, once warned, society could galvanize itself into action. Of course, this social action in the face of threats showed that Malthus and Wells were most at fault in their initial assumption. Social and technological systems do not develop independently; the two evolve together in complex feedback loops, wherein each drives, restrains, and accelerates change in the other. Malthus and Wells — and now Joy — are, indeed, critical parts of these complex loops. Each knew when and how to sound the alarm. But each thought little about how to respond to that alarm.

Once the social system is factored back into the equation like this, the road ahead becomes harder to navigate. Ultimately we should be grateful to Joy for saying, at the least, that there could be trouble ahead when so many of his fellow digerati will only tell us complacently that the road is clear.

11

Promise and Peril

RAY KURZWEIL

Following the publication of Bill Joy's essay, "Why the Future Doesn't Need Us," Ray Kurzweil, another remarkably prolific inventor and technologist, and the author of The Age of Spiritual Machines *(1999), frequently found himself invited to debate Joy in conferences and forums, as a counterpoint, taking (as he does in his book) a more positive view of future technology. Indeed, at the beginning of his essay, Joy attributes his concern about the dangers of emerging technologies to a conversation with Kurzweil in 1998. In this essay, first presented as a paper at a conference on technology at Columbia University entitled "Living with the Genie," Kurzweil presents his side of what he calls a long-term debate with Joy, professing a much more nuanced view than he is sometimes given credit for.*

Mindful of their dangers, he describes the opportunities that the technologies of genetic engineering, nanotechnology, and robotics can offer humankind. The strongest part of his counterpoint, however, is in discussing Joy's recommendation that we relinquish some or all of these technological fields (i.e., give up their pursuit). Some form of "fine-grained" relinquishment, he says, is possible and should be part of technologists' professional ethics. Overall relinquishment, though, is not the answer. Not only will it not work, but there are other, better strategies to protect against the hazards that concern Joy. Technology, Kurzweil says, "will remain a double-edged sword.... We have no choice but to work hard to apply these quickening technologies to advance our human values, despite what often appears to be a lack of consensus on what those values should be."

Ray Kurzweil is perhaps best known as the inventor of the first omni-font optical character recognition device and the first print-to-speech reading machine for the blind. He has founded and developed nine businesses in fields related to optical character recognition, speech and music synthesis, and other areas of *artificial intelligence. He was inducted into the National Inventors Hall of Fame in 2002. He is also the recipient of the National Medal of Technology and the MIT-Lemelson Prize for invention and innovation.*

Consider these articles we'd rather not see on the web:

- Impress Your Enemies: How to Build Your Own Atomic Bomb from Readily Available Materials[1]
- How to Modify the Influenza Virus in Your College Laboratory to Release Snake Venom
- Ten Easy Modifications to the *E. coli* Virus
- How to Modify Smallpox to Counteract the Smallpox Vaccine
- How to Build a Self-Guiding, Low-Flying Airplane Using an Inexpensive Aircraft, GPS, and a Notebook Computer

Or how about the following:

- The Genomes of Ten Leading Pathogens
- The Floor Plans of Leading Skyscrapers
- The Layout of U.S. Nuclear Reactors
- Personal Health Information on 100 Million Americans
- The Customer Lists of Top Pornography Sites

Anyone posting the first item above is almost certain to get a quick visit from the FBI, as did Nate Ciccolo, a fifteen-year-old high school student, in March 2000. For a school science project, he built a papier-mâché model of an atomic bomb that turned out to be disturbingly accurate. In the ensuing media storm, Nate told ABC News, "Someone just sort of mentioned, you know, you can go on the Internet now and get information. And I, sort of, wasn't exactly up to date on things. Try it. I went on there and a couple of clicks and I was right there."[2]

Of course, Nate didn't possess the key ingredient, plutonium, nor did he have any intention of acquiring it, but the report created shock waves in the media, not to mention among the authorities who worry about nuclear proliferation. Nate had reported finding 563 web pages on atomic bomb designs. The publicity resulted in an urgent effort to remove this information. Unfortunately, trying to get rid of information on the Internet is akin to trying to sweep back the ocean with a broom. The information continues to be easily available today. I won't provide any URLs in this essay, but they are not hard to find.

Although the actual article titles above are fictitious, one can find extensive information on the Internet about all of these topics.[3] The web is an extraordinary research tool. In my own experience as a technologist and author, I've found that research which used to require half a day at the library can now typically be accomplished in a couple of minutes. This has enormous and obvious benefits for advancing beneficial technologies, but is also empowering those whose values are inimical to the mainstream of society.

My urgent concern with this issue goes back at least a couple of decades. When I wrote my first book, *The Age of Intelligent Machines*, in the mid-1980s, I was deeply concerned with the ability of genetic engineering, then an emerging technology, to allow those skilled in the art and with access to fairly widely available equipment to modify bacterial and viral pathogens to create new diseases.[4] In malevolent or merely careless hands, these engineered pathogens could potentially combine a high degree of communicability and destructiveness.

In the 1980s this was not easy to do, but was nonetheless feasible. We now know that bioweapons programs in the Soviet Union and elsewhere were doing exactly this.[5] At the time, I made a conscious decision to not talk about this specter in my book, feeling that I did not want to give the wrong people a destructive idea. I had disturbing visions of a future disaster, with the perpetrators saying that they got the idea from Ray Kurzweil.

Partly as a result of this decision, there was some reasonable criticism that the book emphasized the benefits of future technology while ignoring the downside. Thus, when I wrote *The Age of Spiritual Machines* in ... late 1997–98, I attempted to cover both promise and peril.[6] There had been sufficient public attention to the perils by that time (for example, the 1995 movie *Outbreak*, which portrays the terror and panic that follow the release of a new viral pathogen) that I felt comfortable in beginning to address the issue publicly.

It was at that time, in September 1998, with a just-finished manuscript, that I ran into Bill Joy, an esteemed and longtime colleague in the high-technology world, in a bar in Lake Tahoe. I had long admired Bill for his pioneering work in creating the leading software language for interactive web systems (Java) and having cofounded Sun Microsystems. But my focus at this brief get-together was not on Bill but on the third person sitting in our small booth: John Searle, an eminent philosopher from the University of California, Berkeley. John had built a career out of defending the deep mysteries of human consciousness from apparent attack by materialists like me (though I deny the characterization).[7] John and I had just finished debating the issue of whether a machine could be conscious; we'd been part of the closing panel of George Gilder's "Telecosm" conference, which was devoted to a discussion of the philosophical implications of *The Age of Spiritual Machines*.[8]

I gave Bill a preliminary manuscript of the book and tried to bring him up to speed on the debate about consciousness that John and I were having. As it turned out, Bill focused on a completely different issue, specifically the impending dangers to human civilization from three emerging technologies I had presented in the book: genetics, nanotechnology, and robotics, or GNR for short. My discussion of the downsides of future technology alarmed Bill, as he relays in his now-famous cover story for *Wired* magazine, "Why the Future Doesn't Need Us."[9] In the article, Bill describes how he asked his friends in the scientific and technology community whether the projections I was making were credible and was dismayed to discover how close these capabilities were to realization.

Needless to say, Bill's article focused entirely on the downside scenarios, and created a firestorm. Here was one of the technology world's leading figures addressing new and dire emerging dangers from future technology. It was reminiscent of the attention that George Soros, the currency arbitrager and archcapitalist, received when he made vaguely critical comments about the excess of unrestrained capitalism, although the "Bill Joy" controversy became far more intense. The *New York Times* cited about 10,000 articles commenting on and discussing Bill's article, more than any other article in the history of commentary on technology issues.

My attempt to relax in a Lake Tahoe lounge ended up fostering two long-term debates. My dialogue with John Searle has continued to this day, and my debate with Bill has taken on a life of its own. Perhaps this is one reason I now avoid hanging out in bars.

Despite Bill's concerns, my reputation as a technology optimist has remained intact, and Bill and I have been invited to a variety of forums to debate the peril and promise, respectively, of future technologies. Although I am expected to take up the "promise" side of the debate, I often end up spending most of my time at these forums defending the feasibility of the dangers. I recall one event at Harvard during which a Nobel Prize–winning biologist dismissed the "N" (nanotechnology) danger by stating that he did not expect to see self-replicating nano-engineered entities for a hundred years.

I replied that indeed, a hundred years matched my own estimate of the amount of progress required — at *today's* rate of progress. However, since my models show that we are doubling the paradigm shift rate (the rate of technological progress) every ten years, we can expect to make a hundred years of progress — at today's rate — in about twenty-five calendar years, which matches the consensus view within the nanotechnology community. Thus are both promise and peril much closer at hand.

My view is that technology has always been a double-edged sword, bringing us longer and healthier life spans, freedom from physical and mental drudgery, and many new creative possibilities on the one hand, while introducing new and salient dangers on the other. Technology empowers both our creative and our destructive natures. Stalin's tanks and Hitler's trains used technology. We benefit from nuclear power, but live today with sufficient nuclear weapons (not all of which appear to be well accounted for) to end all mammalian life on the planet.

Bioengineering holds the promise of making enormous strides in reversing disease and aging processes. However, the means and knowledge it has created, which began to exist in the 1980s, will soon enable an ordinary college bioengineering lab to create unfriendly pathogens more dangerous than nuclear weapons.[10]

As technology accelerates toward the full realization of G (genetic engineering, also known as biotechnology), followed by N (nanotechnology) and ultimately R (robotics, also referred to as "strong" AI — artificial intelligence at human levels and beyond), we will see the same intertwined potentials: a feast of creativity resulting from human intelligence expanded many-fold, combined with many grave new dangers.

Consider the manner in which extremely small robots, or nanobots, are likely to develop. Nanobot technology requires billions or trillions of such intelligent devices to be useful. The most cost-effective way to scale up to such levels is through self-replication, essentially the same approach used in the biological world. And in the same way that biological self-replication gone awry (i.e., cancer) results in biological destruction, a defect in the mechanism curtailing nanobot self-replication would endanger all physical entities, biological or otherwise. Later in this chapter I suggest steps we can take to address this grave risk, but we cannot have complete assurance in any strategy that we devise today.

Other primary concerns include who controls the nanobots and who the nanobots are talking to. Organizations (e.g., governments or extremist groups) or just a clever individual could put trillions of undetectable nanobots in the water or food supply of an individual or of an entire population. These "spy" nanobots could then monitor, influence, and even control our thoughts and actions. In addition to physical spy nanobots, existing nanobots could be influenced through software viruses and other software "hacking" techniques. When there is software running in our brains, issues of privacy and security will take on a new urgency.

My own expectation is that the creative and constructive applications of this technology will dominate, as I believe they do today. I believe we need to vastly increase our investment in developing specific defensive technologies, however. We are at the critical stage today for biotechnology, and we will reach the stage where we need to directly implement defensive technologies for nanotechnology during the late teen years of this century.

THE INEVITABILITY OF A TRANSFORMED FUTURE

The diverse GNR technologies are progressing on many fronts and comprise hundreds of small steps forward, each benign in itself. An examination of the underlying trends, which I have studied for the past quarter-century, shows that full-blown GNR is inevitable.

The motivation for this study came from my interest in inventing. As an inventor in the 1970s, I came to realize that my inventions needed to make sense in terms of the enabling technologies and market forces that would exist when the invention was introduced, which would represent a very different world than when it was conceived. I began to develop models of how distinct technologies — electronics, communications, computer processors, memory, magnetic storage, and the key feature sizes in a range of technologies — developed and how these changes rippled through markets and ultimately our social institutions. I realized that most inventions fail not because they never work, but because their timing is wrong. Inventing is a lot like surfing; you have to anticipate and catch the wave at just the right moment.

In the 1980s my interest in technology trends and implications took on a life of its own, and I began to use my models of these trends to project and anticipate the technologies of the future. This enabled me to invent with the capabilities of the future in mind. I wrote *The Age of Intelligent Machines*, which ended with the specter of machine intelligence becoming indistinguishable from its human progenitors. The book included hundreds of predictions about the 1990s and early 2000s, and my track record of prediction has held up well.

During the 1990s, I gathered empirical data on the apparent acceleration of all information-related technologies and sought to refine the mathematical models underlying these observations. In *The Age of Spiritual Machines*, I introduced improved models of technology and a theory I called "the law of accelerating returns," which explained why technology evolves in an exponential fashion.

THE INTUITIVE LINEAR VIEW VERSUS THE HISTORICAL EXPONENTIAL VIEW

The most important trend this study has revealed concerns the overall pace of technological progress itself. The future is widely misunderstood. Our forebears expected the future to be pretty much like their present, which had been pretty much like their past. Although exponential trends did exist a thousand years ago, they were at that very early stage where an exponential trend is so flat and so slow that it looks like no trend at all. So their expectation of stasis was largely fulfilled. Today, in accordance with the common wisdom, everyone expects continuous technological progress and the social repercussions that follow. But the future will nonetheless be far more surprising than most observers realize because few have truly internalized the implications of the fact that the rate of change itself is accelerating.

Most long-range forecasts of technical feasibility in future time periods dramatically underestimate the power of future developments because they are based on what I call the "intuitive linear view" of history rather than the "historical exponential view." To express this another way, it is not the case that we will experience 100 years of progress in the twenty-first century; rather, we will witness on the order of 20,000 years of progress (again, at *today*'s rate of progress).

When people think of a future period, they intuitively assume that the current rate of progress will continue into that future. Even for those who have been around long enough to experience how the pace increases over time, an unexamined intuition nonetheless provides the impression that progress happens at the rate we have experienced recently. From the mathematician's perspective, a primary reason for this is that an exponential curve approximates a straight line when viewed for a brief duration. It is typical therefore, for even sophisticated commentators, when considering the future, to extrapolate the current pace of change over the next ten years or hundred years to determine their expectations. This is why I call this way of looking at the future the "intuitive linear view."

But a serious assessment of the history of technology shows that technological change is exponential. In exponential growth, a key measurement such as computational power is multiplied by a constant factor for each unit of time (e.g., doubling every year) rather than just increased incrementally. Exponential growth is a feature of any evolutionary process, of which technology is a prime example. One can examine the data in different ways, on different time scales, and for a wide variety of technologies ranging from electronic to biological, as well as for social implications ranging from the size of the economy to human life span, and the acceleration of progress and growth applies to all.

Indeed, we find not just simple exponential growth, but "double" exponential growth, meaning that the rate of exponential growth is itself growing exponentially. These observations do not rely merely on an assumption of the continuation of Moore's Law (i.e., the exponential shrinking of transistor sizes on an integrated circuit), which I discuss a bit later, but are based on a rich model of diverse technological processes. What the model clearly shows is that technology, particularly the pace of technological change, advances (at least) exponentially, not linearly, and has been doing so since the advent of technology, indeed since the advent of evolution on Earth.

Many scientists and engineers have what my colleague Lucas Hendrich calls "engineer's pessimism." Often engineers or scientists who are so immersed in the difficulties and intricate details of a contemporary challenge fail to appreciate the ultimate long-term implications of their own work and, in particular, the larger field of work that they operate in. Consider the biochemists in 1985 who were skeptical of the announced goal of transcribing the entire genome in a mere fifteen years. These scientists had just spent an entire year transcribing a mere one tenthousandth of the genome, so even with reasonable anticipated advances, it seemed to them as though it would be hundreds of years, if not longer, before the entire genome could be sequenced.

Or consider the skepticism expressed in the mid-1980s that the Internet would ever be a significant phenomenon, given that it included only tens of thousands of nodes. The fact that the number of nodes was doubling every year, so tens of millions of nodes were likely to exist ten years later, was not appreciated by those struggling with a limited "state of the art" technology in 1985 that permitted adding only a few thousand nodes throughout the world in a year.[11]

I emphasize this point because it is the most important failure that would-be prognosticators make in considering future trends. The vast majority of technology forecasts and forecasters ignore altogether this "historical exponential view" of technological progress. Indeed, almost everyone I meet has a linear view of the future. That is why people tend to overestimate what can be achieved in the short term (because we tend to leave out necessary details) but underestimate what can be achieved in the long term (because the exponential growth is ignored).

THE LAW OF ACCELERATING RETURNS

The ongoing acceleration of technology is the implication and the inevitable result of what I call the "law of accelerating returns," which describes the acceleration of the pace and the exponential growth of the products of an evolutionary process. This process includes information-bearing technologies such as computation as well as the accelerating trend toward miniaturization — all the prerequisites for the full realization of GNR.

A wide range of technologies are subject to the law of accelerating returns. The exponential trend that has gained the greatest public recognition has become known as Moore's Law. Gordon Moore, one of the inventors of integrated circuits and then chairman of Intel, noted in the mid-1970s that we could squeeze twice as many transistors on an integrated circuit every twenty-four months.[12] Given that the electrons have a shorter distance to travel, the transistors change states more quickly. This, along with other techniques, allows the circuits to run faster, providing an overall quadrupling of computational power.

The exponential growth of computing is much broader than Moore's Law, however. If we plot the speed (in instructions per second) per $1,000 (in constant dollars) of forty-nine famous calculators and computers spanning the entire twentieth century, we note that there were four completely different paradigms providing exponential growth in the price-performance of computing before integrated circuits were invented. Therefore, Moore's Law was not the first but the fifth paradigm of exponential growth of computational power.[13] And it won't be the last. When Moore's Law approaches its limit, now expected before 2020, the exponential growth will continue with three-dimensional molecular computing, a prime example of the application of nanotechnology, which will constitute the sixth paradigm.

When in 1999 I suggested that three-dimensional molecular computing, particularly an approach based on using carbon nanotubes, would become the dominant computing hardware technology in the teen years of this century, that was considered a radical notion. But so much progress has been accomplished in the past four years, with literally dozens of major milestones having been achieved, that this expectation is now a mainstream view.[14]

The exponential growth of computing is a marvelous quantitative example of the exponentially growing returns from an evolutionary process. We can express the exponential growth of computing in terms of an accelerating pace: it took ninety years to achieve the first million instructions per second (MIPS) per $1,000; now we add one MIPS per $1,000 every day. The human brain uses a very inefficient electrochemical digital-controlled analog computational process. The bulk of the calculations are done in the interneuronal connections at a speed of only about two hundred calculations per second (in each connection), which is about ten million times slower than contemporary electronic circuits. But the brain gains its prodigious powers from its extremely parallel organization *in three dimensions*. There are many technologies in the wings that build circuitry in three dimensions.

Nanotubes (tubes formed from graphite sheets, consisting of hexagonal arrays of carbon atoms) are good conductors and can be used to build compact circuits; these are already working in laboratories. One cubic inch of nanotube circuitry would theoretically be a million times more powerful than the human brain. There are more than enough new computing technologies now being researched, including three-dimensional silicon chips, spin computing, crystalline computing, DNA computing, and quantum computing, to keep the law of accelerating returns as applied to computation going for a long time.

It is important to distinguish between the S curve (an *S* stretched to the right, comprising very slow, virtually unnoticeable growth — followed by very rapid growth — followed by a flattening out as the process approaches an asymptote, or limit) that is characteristic of any specific technological paradigm, and the continuing exponential growth that is characteristic of the ongoing evolutionary process of technology. Specific paradigms, such as Moore's Law, do ultimately reach levels at which exponential growth is no longer feasible. That is why Moore's Law is an S curve.

But the growth of computation is an ongoing exponential pattern (at least until we "saturate" the universe with the intelligence of our human-machine civilization, but that is not likely to happen in this century). In accordance with the law of accelerating returns, paradigm shift (also called innovation) turns the S curve of any specific paradigm into a continuing exponential pattern. A new paradigm (e.g., three-dimensional circuits) takes over when the old paradigm approaches its natural limit, which has already happened at least four times in the history of computation. This difference also distinguishes the tool making of nonhuman species, in which the mastery of a tool-making (or tool-using) skill by each animal is characterized by an abruptly ending *S*-shaped learning curve, from human-created technology, which has followed an exponential pattern of growth and acceleration since its inception.

This "law of accelerating returns" applies to all of technology, indeed to any true evolutionary process, and can be measured with remarkable precision in information-based technologies. There are a great many examples of the exponential growth implied by the law of accelerating returns in technologies, as varied as electronics of all kinds, DNA sequencing, communication speeds, brain scanning, reverse engineering of the brain, the size and scope of human knowledge, and the rapidly shrinking size of technology, which is directly relevant to emergence of nanotechnology.[15]

The future GNR age results not from the exponential explosion of computation alone, but rather from the interplay and myriad synergies that will result from intertwined technological revolutions. Keep in mind that every point on the exponential growth curves underlying this panoply of technologies represents an intense human drama of innovation and competition. It is remarkable, therefore, that these chaotic processes result in such smooth and predictable exponential trends.

ECONOMIC IMPERATIVE

It is the economic imperative of a competitive marketplace that is fueling the law of accelerating returns and driving technology forward toward the full realization of GNR. In turn, the law of accelerating returns is transforming economic relationships.

We are moving toward nanoscale machines, as well as more intelligent machines, as the result of a myriad of small advances, each with their own particular economic justification. There is a vital economic imperative to create smaller and more intelligent technology. Machines that can more precisely carry out their missions have enormous value, which is why they are being built. There are tens of thousands of projects advancing the various aspects of the law of accelerating returns in diverse incremental ways.

Regardless of near-term business cycles, the support for "high tech" in the business community, and in particular for software advancement, has grown enormously. When I started my optical character recognition (OCR) and speech synthesis company (Kurzweil Computer Products) in 1974, high-tech venture deals totaled approximately $10 million. Even during today's high-tech recession, the figure is a hundred times greater. We would have to repeal capitalism and every visage of economic competition to stop this progression.

The economy (viewed either in total or per capita) has been growing exponentially throughout this century. This underlying exponential growth is a far more powerful force than periodic recessions.[16] Even the Great Depression represents only a minor blip on this pattern of growth. Most importantly, recessions and depressions represent only temporary deviations from the underlying curve. In each case, the economy ends up exactly where it would have been had the recession or depression never occurred.

In addition to GDP, other improvements include productivity (economic output per worker), quality and features of products and services (for example, $1,000 of computation today is more powerful, by a factor of more than a thousand, than $1,000 of computation ten years ago), new products and product categories, and the value of existing goods, which has been increasing at 1.5 percent per year for the past twenty years because of qualitative improvements.[17,18]

INTERTWINED BENEFITS...

Significant portions of our species have already experienced substantial alleviation of the poverty, disease, hard labor, and misfortune that have characterized much of human history. Many of us have the opportunity to gain satisfaction and meaning from our work rather than merely toiling to survive. We have ever more powerful tools to express ourselves. We have worldwide sharing of culture, art, and humankind's exponentially expanding knowledge base. Ubiquitous N and R are two to three decades away. A prime example of their application will be the deployment of billions of nanobots the size of human blood cells to travel inside the human bloodstream. This technology will be feasible within twenty-five years, based on miniaturization and cost-reduction trends. In addition to scanning the brain to facilitate reverse engineering (that is, analyzing how the brain works in order to copy its design), nanobots will be able to perform a broad variety of diagnostic and therapeutic functions inside the bloodstream and human body. Robert A. Freitas, for example, has designed robotic replacements for human blood cells that perform hundreds or thousands of times more effectively than their biological counterparts.[19] With Freitas's "respirocytes" (robotic red blood cells), you could do an Olympic sprint for fifteen minutes without taking a breath. His robotic macrophages would be far more effective than our white blood cells at combating pathogens. His DNA repair robot would be able to repair DNA transcription errors and even implement needed DNA changes.

Although realization of Freitas's conceptual designs is two or three decades away, substantial progress has already been achieved on bloodstream-based devices. For example, one scientist has cured type 1 diabetes in rats with a nanoengineered device that incorporates pancreatic islet cells. The device has seven nanometer pores that let insulin out but block the antibodies that destroy these cells. Many innovative projects of this type are already under way.

Clearly, nanobot technology has profound military applications, and any expectation that such uses will be "relinquished" is highly unrealistic. Already, the Department of Defense is developing "smart dust" — tiny robots the size of insects or even smaller.[20] Although not quite nanotechnology, millions of these devices can be dropped into enemy territory to provide highly detailed surveillance. The potential application for even smaller nanotechnology-based devices is even greater. Want to find Saddam Hussein or Osama bin Laden? Need to locate hidden weapons of mass destruction? Billions of essentially invisible spies could monitor every square inch of enemy territory, identify every person and every weapon, and even carry out missions to destroy enemy targets. The only way for an enemy to counteract such a force is, of course, with their own nanotechnology. The point is that nanotechnology-based weapons will one day make larger weapons obsolete.

Nanobots will also expand our experiences and our capabilities by providing fully immersive, totally convincing virtual reality. They will take up positions in close proximity to every interneuronal connection of our sensory organs (e.g., eyes, ears, and skin). We already have the technology that enables electronic devices to communicate with neurons in both directions, and it requires no direct physical contact with the neurons. For example, scientists at the Max Planck Institute have developed "neuron transistors" that can detect the firing of a nearby neuron or, alternatively, can cause a nearby neuron to fire, or suppress it from firing.[21] This amounts to two-way communication between neurons and the electronically based neuron transistors. The institute scientists demonstrated their invention by controlling the movement of a living leech from their computer.

When we want to experience real reality, the nanobots just stay in position (in the capillaries) and do nothing. If we want to enter virtual reality, they suppress all of the inputs coming from the real senses and replace them with the signals that would be appropriate for the virtual environment. You (i.e., your brain) could decide to cause your muscles and limbs to move as you normally would, but the nanobots would intercept these interneuronal signals, suppress your real limbs from moving, and instead cause your virtual limbs to move and provide the appropriate reorientation in the virtual environment.

The primary limitation of nanobot-based virtual reality at this time is only that it's not yet feasible in size and cost. One day, however, the web will provide a panoply of virtual environments to explore. Some will be re-creations of real places; others will be fanciful environments that have no "real" counterpart. Some will be impossible in the physical world, perhaps because they violate the laws of physics. We will be able to "go" to these virtual environments by ourselves, or we will meet other people there, both real people and simulated people. Of course, ultimately there won't be a clear distinction between the two. By 2030, going to a website will mean entering a full-immersion virtual reality environment. In addition to encompassing all of the senses, some of these shared environments will include emotional overlays, as the nanobots will be capable of triggering the neurological correlates of emotions, sexual pleasure, and other derivatives of our sensory experience and mental reactions. In the same way that people today beam their lives from webcams in their bedrooms, "experience beamers" circa 2030 will beam their entire flow of sensory experiences and, if so desired, their emotions and other secondary reactions. We'll be able to plug in (by going to the appropriate website) and experience other people's lives, as in the movie *Being John Malkovich*. We will be able to archive particularly interesting experiences and relive them at any time.

We won't need to wait until 2030 to experience shared virtual-reality environments, though, at least for the visual and auditory senses. Full-immersion visual-auditory environments will be available by the end of this decade, with images written directly onto our retinas by our eyeglasses and contact lenses. All of the electronics for computation, image reconstruction, and a very high bandwidth wireless connection to the Internet will be embedded in our glasses and woven into our clothing, so computers as distinct objects will disappear.

It's not just the virtual world that will benefit from ubiquitous application of nanobots and fully realized nanotechnology. Portable manufacturing systems will be able to produce virtually any physical product from information for pennies a pound, thereby providing for our physical needs at almost no cost. Nanobots will be able to reverse the environmental destruction left by the first industrial revolution. Nanoengineered fuel cells and solar cells will provide clean energy. Nanobots in our physical bodies will destroy pathogens and cancer cells, repair DNA, and reverse the ravages of aging.

These technologies will become so integral to our health and well-being that we will eventually become indistinguishable from our machine support systems. In fact, in my view, the most significant implication of the development of nanotechnology and related advanced technologies of the twenty-first century will be the merger of biological and nonbiological intelligence. First, it is important to point out that well before the end of the twenty-first century, thinking on nonbiological substrates will dominate. Biological thinking is stuck at 1026 calculations per second (for all biological human brains), and that figure will not appreciably change, even with bioengineering changes to our genome.

Nonbiological intelligence, on the other hand, is growing at a double exponential rate and will vastly exceed biological intelligence well before the middle of this century. In my view, however, this nonbiological intelligence should still be considered human, since it is fully derivative of the human-machine civilization. The merger of these two worlds of intelligence is not merely a merger of biological and nonbiological thinking media but, more importantly, one of method and organization of thinking.

Nanobot technology will be able to expand our minds in virtually any imaginable way. Our brains today are relatively fixed in design. Although we do add patterns of interneuronal connections and neurotransmitter concentrations as a normal part of the learning process, the current overall capacity of the human brain is highly constrained, restricted to a mere trillion connections. Brain implants based on massively distributed intelligent nanobots will ultimately expand our memories a trillionfold, and otherwise vastly improve all of our sensory, pattern recognition, and cognitive abilities. Since the nanobots will be communicating with each other over a wireless local area network, they will be able to create any set of new neural connections, break existing connections (by suppressing neural firing), and create new hybrid biological-nonbiological networks, as well as add vast new nonbiological networks.

Using nanobots as brain extenders will be a significant improvement over surgically installed neural implants, which are beginning to be used today (e.g., ventral posterior nucleus, subthalamic nucleus, and ventral lateral thalamus neural implants to counteract Parkinson's disease and tremors from other neurological disorders; cochlear implants; and the like). Nanobots will be introduced without surgery, essentially just by injecting or even swallowing them. They can all be directed to leave, so the process is easily reversible. They are programmable, in that they can provide virtual reality one minute and a variety of brain extensions the next. They can change their configuration and alter their software. Perhaps most importantly, they are massively distributed and therefore can take up billions or trillions of positions throughout the brain, whereas a surgically introduced neural implant can be placed in only one or at most a few locations.

...AND DANGERS

Needless to say, we have already experienced technology's downside. One hundred million people were killed in two world wars during the last century — a scale of mortality made possible by technology. The crude technologies of the first industrial revolution have crowded out many of the species on our planet that existed a century ago. Our centralized technologies (e.g., buildings, cities, airplanes, and power plants) are demonstrably insecure.

The NBC (nuclear, biological, and chemical) technologies of warfare were all used, or threatened to be used, in our recent past.[22] The far more powerful GNR technologies pose what philosopher of science Nick Bostrom calls "existential risks," referring to potential threats to the viability of human civilization itself.[23]

If we manage to get past the concerns about genetically altered designer pathogens, followed by self-replicating entities created through nanotechnology, we will next encounter robots whose intelligence will rival and ultimately exceed our own. Such robots may make great assistants, but who's to say that we can count on them to remain reliably friendly to mere humans?

In my view, "strong AI" (artificial intelligence at human levels and beyond) promises to continue the exponential gains of human civilization. But the dangers are also more profound precisely because of this amplification of intelligence. Intelligence is inherently impossible to control, so the various strategies that have been devised to control nanotechnology won't work for strong AI. There have been discussions and proposals to guide AI development toward "friendly AI."[24] These are useful for discussion, but it is impossible to devise strategies today that will absolutely ensure that future AI embodies human ethics and values.

RELINQUISHMENT

In his *Wired* essay and subsequent presentations, Bill Joy eloquently describes the plagues of centuries past and how new self-replicating technologies, such as mutant bioengineered pathogens and "nanobots" run amok, may bring back long-forgotten pestilence. Of course, as Joy graciously acknowledges, it has also been technological advances, such as antibiotics and improved sanitation, that have freed us from the prevalence of such plagues. Suffering in the world continues and demands our steadfast attention. Should we tell the millions of people afflicted with cancer and other devastating conditions that we are canceling the development of all bioengineered treatments because there is a risk that these same technologies may someday be used for malevolent purposes?

Relinquishment is Bill's most controversial recommendation and personal commitment. I do feel that relinquishment at the right levels is part of a responsible and constructive response to these genuine perils. The issue, however, is exactly this: at what level are we to relinquish technology?

Ted Kaczynski would have us renounce all of it.[25] This, in my view, is neither desirable nor feasible, and the futility of such a position is only underscored by the senselessness of Kaczynski's deplorable tactics. There are other voices, less reckless than Kaczynski, who are nonetheless arguing for broad-based relinquishment of technology. The environmentalist Bill McKibben takes the position that "environmentalists must now grapple squarely with the idea of a world that has enough wealth and enough technological capability, and should not pursue more."[26] In my view, this position ignores the extensive suffering that remains in the human world, which we will be in a position to alleviate through continued technological progress.

Another level of relinquishment, one recommended by Joy, would be to forgo certain fields — nanotechnology, for example — that might be regarded as too dangerous. But such sweeping strokes of relinquishment are equally untenable. As I pointed out above, nanotechnology is simply the inevitable end result of the persistent trend toward miniaturization that pervades all of technology. It is far from a single centralized effort, but is being pursued by a myriad of projects with many diverse goals.

One observer wrote:

> A further reason why industrial society cannot be reformed … is that modern technology is a unified system in which all parts are dependent on one another. You can't get rid of the "bad" parts of technology and retain only the "good" parts. Take modern medicine, for example. Progress in medical science depends on progress in chemistry, physics,

> biology, computer science and other fields. Advanced medical treatments require expensive, high-tech equipment that can be made available only by a technologically progressive, economically rich society. Clearly you can't have much progress in medicine without the whole technological system and everything that goes with it.

The observer, again, is Ted Kaczynski.[27] Although one will properly resist Kaczynski as an authority, I believe he is correct on the deeply entangled nature of the benefits and risks. He and I clearly part company on our overall assessment of the relative balance between the two, however. Bill Joy and I have debated this issue both publicly and privately, and we both believe that technology will and should progress and that we need to be actively concerned with the dark side. If he and I disagree, it's on the granularity of relinquishment that is both feasible and desirable.

Abandonment of broad areas of technology would only push them underground, where development would continue unimpeded by ethics and regulation. In such a situation, it would be the less stable, less responsible practitioners (e.g., terrorists) who would have all the expertise. I do think that relinquishment at the right level needs to be part of our ethical response to the dangers of twenty-first-century technologies. One constructive example is the ethical guideline proposed by the Foresight Institute, founded by nanotechnology pioneer Eric Drexler, along with Christine Peterson, that nanotechnologists agree to relinquish the development of physical entities that can self-replicate in a natural environment.[28] Another is a ban on self-replicating physical entities that contain their own codes for self-replication. In what nanotechnologist Ralph Merkle calls the "broadcast architecture," such entities would have to obtain these codes from a centralized secure server, which would guard against undesirable replication.[29] I discuss these guidelines further below.

The broadcast architecture is impossible in the biological world, so it represents at least one way in which nanotechnology can be made safer than biotechnology. In other ways, nanotechnology is potentially more dangerous because nanobots can be physically stronger, and more intelligent, than protein-based entities. It will eventually be possible to combine the two by having nanotechnology provide the codes within biological entities (replacing DNA), in which case biological entities will be able to use the much safer broadcast architecture.

As responsible technologists, our ethics should include such "fine-grained" relinquishment among our professional ethical guidelines. Protections must also include oversight by regulatory bodies, the development of technology-specific "immune" responses, and computer-assisted surveillance by law enforcement organizations. Many people are not aware that our intelligence agencies already use advanced technologies such as automated word spotting to monitor a substantial flow of telephone conversations. As we go forward, balancing our cherished rights of privacy with our need to be protected from the malicious use of powerful twenty-first-century technologies will be one of many profound challenges. This is one reason such issues as an encryption "trap door" (in which law enforcement authorities would have access to otherwise secure information) and the FBI's Carnivore email-snooping system have been controversial.[30]

We can take a small measure of comfort from how our society has dealt with one recent technological challenge. There exists today a new form of fully non-biological self-replicating entity that did not exist just a few decades ago: the computer virus. When this form of destructive intruder first appeared, strong concerns were voiced that as these software pathogens became more sophisticated, they would have the potential to destroy the computer network medium they live in. Yet the "immune system" that has evolved in response to this challenge has been largely effective. Although destructive self-replicating software entities do cause damage from time to time, the injury is but a small fraction of the benefit we receive from the computers and communication links that harbor them. No one would suggest we do away with computers, local area networks, and the Internet because of software viruses.

One might counter that computer viruses do not have the lethal potential of biological viruses or of destructive nanotechnology. This is not always the case; we rely on software to monitor patients in critical care units, to fly and land airplanes, to guide intelligent weapons in wartime, and to perform other "mission-critical" tasks. To the extent that this assertion is true, however, it only strengthens my argument. The fact that computer viruses are not usually deadly to humans only means that more people are willing to create and release them. It also means that our response to the danger is that much less intense. Conversely, when it comes to self-replicating entities that are potentially lethal on a large scale, our response on all levels will be vastly more serious, as we have seen since September 11.

THE DEVELOPMENT OF DEFENSIVE TECHNOLOGIES AND THE IMPACT OF REGULATION

Bill Joy's *Wired* treatise is effective because he paints a picture of future dangers as if they were released on today's unprepared world. The reality is that the sophistication and power of our defensive technologies and knowledge will grow along with the dangers. When we have "gray goo" (unrestrained nanobot replication), we will also have "blue goo" ("police" nanobots that combat the "bad" nanobots). The story of the twenty-first century has not yet been written, so we cannot say with assurance that we will successfully avoid all misuse. But the surest way to prevent the development of defensive technologies would be to relinquish the pursuit of knowledge in broad areas. We have been able to largely control harmful software virus replication because the requisite knowledge is widely available to responsible practitioners. Attempts to restrict this knowledge would have created a far less stable situation. Responses to new challenges would have been far slower, and the balance would have likely shifted toward the more destructive applications (e.g., software viruses).

As we compare the success we have had in controlling engineered software viruses to the coming challenge of controlling engineered biological viruses, we are struck with one salient difference: the software industry is almost completely unregulated. The same is obviously not true for biotechnology. We require scientists developing defensive technologies to follow the existing regulations, which slow down the innovation process at every step. A bioterrorist, however, does not need to put his "innovations" through the FDA. Moreover, under existing regulations and ethical standards, it is impossible to test defenses against bioterrorist agents. Extensive discussion is already under way regarding modifying these regulations to allow for animal models and simulations, since human trials are infeasible.

For reasons I have articulated above, stopping these technologies is not feasible, and pursuit of such broad forms of relinquishment will only distract us from the vital task in front of us. It is quite clearly a race. There is simply no alternative. We cannot relinquish our way out of this challenge. In the software field, defensive technologies have remained a step ahead of offensive ones. With the extensive regulation in the medical field slowing down innovation at each stage, we cannot have the same confidence with regard to the abuse of biotechnology.

In the current environment, when one person dies in gene therapy trials, research can be severely restricted.[31] There is a legitimate need to make biomedical research as safe as possible, but our balancing of risks is completely off. The millions of people who desperately need the advances that will result from gene therapy and other breakthrough biotechnologies appear to carry little political weight against a handful of well-publicized casualties from the inevitable risks of progress.

This equation will become even more stark when we consider the emerging dangers of bioengineered pathogens. What is needed is a change in public attitude in terms of tolerance for needed risk.

Hastening defensive technologies is absolutely vital to our security. We need to streamline regulatory procedures to achieve this, and we also need to greatly increase our investment in defensive technologies explicitly. In the biotechnology field, this means the rapid development of antiviral medications. We will not have time to develop specific countermeasures for each new challenge that comes along. We are close to developing more generalized antiviral technologies, and these need to be accelerated.

I have addressed here the issue of biotechnology because that is the threshold and challenge that we now face. As the threshold for nanotechnology comes closer, we will then need to invest specifically in the development of defensive technologies in that area, including the creation of a nanotechnology-based immune system. Bill Joy and other observers have pointed out that such an immune system would itself be a danger because of the potential of "autoimmune" reactions (i.e., the immune system using its powers to attack the world it is supposed to be defending).[32]

This observation is not a compelling reason to avoid the creation of an immune system, however. No one would argue that humans would be better off without an immune system because of the possibility of autoimmune diseases. Although the immune system can itself be a danger, humans would not last more than a few weeks (barring extraordinary efforts at isolation) without one. The development of a technological immune system for nanotechnology will happen even without explicit efforts to create one. We have effectively done this with regard to software viruses. We created a software virus immune system not through a formal grand design, but rather through our incremental responses to each new challenge and through the development of heuristic algorithms for early detection. We can expect the same thing will happen as challenges from nanotechnology-based dangers emerge. The point for public policy will be to specifically invest in these defensive technologies.

It is premature to develop specific defensive nanotechnologies today, since we can have only a general idea of what we are trying to defend against. It would be similar to the engineering world creating defenses against software viruses before the first virus had been created. There is already fruitful dialogue and discussion on anticipating these issues, however, and significantly expanded investment in these efforts is to be encouraged.

As one example, as I mentioned above, the Foresight Institute has devised a set of ethical standards and strategies for ensuring the development of safe nanotechnology.[33] These guidelines include:

- Artificial replicators must not be capable of replication in a natural, uncontrolled environment.
- Evolution within the context of a self-replicating manufacturing system is discouraged.
- MNT (molecular nanotechnology) designs should specifically limit proliferation and provide traceability of any replicating systems.
- Distribution of molecular manufacturing development capability should be restricted, whenever possible, to responsible actors that have agreed to the guidelines. No such restriction need apply to end products of the development process.

Other strategies proposed by the Foresight Institute include:

- Replication should require materials not found in the natural environment.
- Manufacturing (replication) should be separated from the functionality of end products. Manufacturing devices can create end products but should not be able to replicate themselves, and end products should have no replication capabilities.
- Replication should require codes that are encrypted and time-limited. (The broadcast architecture mentioned earlier is an example.)

These guidelines and strategies are likely to be effective for preventing accidental release of dangerous self-replicating nanotechnology entities. But the intentional design and release of such entities is a more complex and challenging problem. A sufficiently determined and destructive opponent could possibly defeat each of these layers of protections. Take, for example, the broadcast architecture. When properly designed, each entity is unable to replicate without first obtaining replication codes. These codes are not passed on from one replication generation to the next. A modification to such a design could bypass the destruction of the codes, however, and thereby pass them on to the next generation. To counteract that possibility, it has been recommended that the memory for the replication codes be limited to only a subset of the full code so that insufficient memory exists to pass the full set of codes along. But this guideline could be defeated by expanding the size of the replication code memory to incorporate the entire code. Another suggestion is to encrypt the codes and build protections such as time limits into the decryption systems. We can see how easy it has been to defeat protections against unauthorized replications of intellectual property such as music files, however. Once replication codes and protective layers are stripped away, the information can be replicated without these restrictions. My point is not that protection is impossible. Rather, we need to realize that any level of protection will only work to a certain level of sophistication. The meta-lesson here is that we will need to place society's highest priority during the twenty-first century on continuing to advance defensive technologies and on keeping them one or more steps ahead of destructive technologies. We have seen analogies to this in many areas, including technologies for national defense as well as our largely successful efforts to combat software viruses.

The broadcast architecture won't protect us against abuses of strong AI. The barriers of the broadcast architecture rely on the nanoengineered entities' lacking the intelligence to overcome the built-in restrictions. By definition, intelligent entities have the cleverness to easily overcome such barriers. Inherently, there will be no absolute protection other than dominance by friendly AI. Although the argument is subtle, I believe that maintaining an open system for incremental scientific and technological progress, in which each step is subject to market *acceptance*, will provide the most constructive environment for technology to embody widespread human values. Attempts to control these technologies in dark government programs, along with inevitable underground development, would create an unstable environment in which the dangerous applications would likely become dominant.

One profound trend already well under way that will provide greater stability is the movement from centralized technologies to distributed ones, and from the real world to the virtual world discussed above. Centralized technologies involve an aggregation of resources such as people (e.g., cities and buildings), energy (e.g., nuclear power plants, liquid natural gas and oil tankers, and energy pipelines), transportation (e.g., airplanes and trains), and other resources. Centralized technologies are subject to disruption and disaster. They also tend to be inefficient, wasteful, and harmful to the environment.

Distributed technologies, on the other hand, tend to be flexible, efficient, and relatively benign in their environmental effects. The quintessential distributed technology is the Internet. Despite concerns about viruses, these information-based pathogens are mere nuisances. The Internet is essentially indestructible. If any hub or channel goes down, the information simply routes around it. The Internet is remarkably resilient, a quality that continues to increase with its continued exponential growth.

In energy, we need to move rapidly toward the opposite end of the spectrum of contemporary energy sources, away from the extremely concentrated energy installations we now depend on. In one example of a trend in the right direction, Integrated Fuel Cell Technologies is pioneering microscopic fuel cells that use microelectromechanical systems (MEMS) technology. The fuel cells are manufactured like electronic chips but are actually batteries with an energy-to-size ratio vastly exceeding conventional technology. Ultimately, forms of energy along these lines could power everything from our cell phones to our cars and homes, and would not be subject to disaster or disruption.

As these technologies develop, our need to aggregate people in large buildings and cities will diminish and people will spread out, living where they want and gathering together in virtual reality.

But we don't need to look past today to see the intertwined promise and peril of technological advancement. If we imagine describing the dangers that exist today to people who lived a couple of hundred years ago, they would think it mad to take such risks. On the other hand, how many people in the year 2000 would really want to go back to the short, brutish, disease-filled, poverty-stricken, disaster-prone lives that 99 percent of the human race struggled through a couple of centuries ago?[34] We may romanticize the past, but up until fairly recently, most of humanity lived extremely fragile lives where one all-too-common misfortune could spell disaster. Two hundred years ago, life expectancy for females in the record-holding country (Sweden) was roughly thirty-five years, compared to the longest life expectancy today — almost eighty-five years — enjoyed by Japanese women. Life expectancy for males was roughly thirty-three years to the current seventy-nine years in the record-holding countries.[35] It took half the day to prepare the evening meal, and hard labor characterized most human activity. There were no social safety nets. Substantial portions of our species still live in this precarious way, which is at least one reason to continue technological progress and the economic enhancement that accompanies it.

People often go through three stages in considering future technologies: awe and wonderment at their potential to overcome age-old problems; a sense of dread at the new set of grave dangers that accompany the new technologies; and, finally (and hopefully), the realization that the only viable and responsible path is to set a careful course that can reap the benefits while managing the dangers.

Technology will remain a double-edged sword. It represents vast power to be used for all humankind's purposes. We have no choice but to work hard to apply these quickening technologies to advance our human values, despite what often appears to be a lack of consensus on what those values should be.

ENDNOTES

1. In the 1960s, the U.S. government conducted an experiment in which they asked three recently graduated physics students to build a nuclear weapon using only publicly available information. The result was successful; the three students built one in about three years (www.pimall.com/nais/nl/n.nukes.html). Plans for how to build an atomic bomb are available on the Internet and have been published in book form by a national laboratory. In 2002, the British Ministry of Defense released measurements, diagrams, and precise details on bomb-building to the Public Record Office, since removed (www.news.bbc.co.uk/1/hi/uk/1932702.stm).

2. *Free Speech in America*, ABC News, March 23, 2000, www.abcnews.go.com/onair/abcnewsspecials/transcripts/Stossel_000323_freespeech_trans.html.
3. The web contains extensive information, including military manuals, on how to build bombs, weapons, and explosives. Some of this information is erroneous, but accurate information on these topics continues to be accessible despite efforts to remove it. In June 1997 Congress passed an amendment (the Feinstein Amendment, SP 419) to a Defense Department appropriations bill banning the dissemination of instructions on building bombs. See Helmenstine, Anne Marie, "How to Build a Bomb," www.chemistry.about.com/library/weekly/aa021003a.htm, posted February 10, 2003. Information on toxic industrial chemicals is widely available on the web and in libraries, as is information and tools for cultivating bacteria and viruses. Techniques for creating computer viruses and hacking into computers and networks are also widely available on the Internet. Note that it is my policy not to provide specific examples of such information, since it might be helpful to destructive individuals and groups. I realize that even stating the availability of such information has this potential, but I feel that the benefit of open dialogue about this issue outweighs this concern. Moreover, the widespread availability of this type of information has been widely discussed in the media and other venues.
4. Ray Kurzweil, *The Age of Intelligent Machines* (Cambridge, MA: MIT Press, 1990).
5. Ken Alibek, *Biohazard* (New York: Dell Publishing, 1999).
6. Ray Kurzweil, *The Age of Spiritual Machines* (New York: Viking, 1999).
7. John Searle, "I Married a Computer," and Ray Kurzweil, "Locked in His Chinese Room," in Jay Richards, et al., *Are We Spiritual Machines?: Ray Kurzweil vs. the Critics of Strong AI* (Seattle, WA: Discovery Institute, 2002), www.kurzweilai.net/meme/frame.html?main=/articles/art0502.html.
8. This panel resulted in a book: Jay Richards, et al., *Are We Spiritual Machines?: Ray Kurzweil vs. the Critics of Strong AI* (Seattle, WA: Discovery Institute, 2002), www.kurzweilai.net/meme/frame.html?m=19. The book includes point-counterpoint essays by John Searle, Thomas Ray, Michael Denton, William A. Dembski, and George Gilder, and responses by Ray Kurzweil.
9. Bill Joy, "Why the Future Doesn't Need Us." *Wired* (April 2000), www.wired.com/wired/archive/8.04/joy.html.
10. Handbooks on gene splicing such as A. J. Harwood, ed., *Basic DNA and RNA Protocols* (Totowa, NJ.: Humana Press, 1996), along with reagents and kits required to do gene splicing, are generally available. Even if access to these materials were limited in the West, a large number of Russian companies could provide equivalent materials.
11. The Internet, as measured by the number of nodes, was doubling every year during the 1980s, but was only tens of thousands of nodes in 1985. It grew to tens of millions of nodes by 1995.
12. Intel Corporation, "Expanding Moore's Law," www.intel.com/labs/eml, accessed 2003.
13. Electromechanical card-reading equipment was used in the 1890 U.S. census. Alan Turing and his colleagues built the first special-purpose computer out of electromagnetic relays in 1942. Vacuum tubes were used in the 1952 computer that predicted the election of Eisenhower, the first time the networks had used a computer to predict a presidential election. Computers based on discrete transistors were used in the space flights during the 1960s. Integrated circuits began to be used in the late 1960s.

14. Silicon nanowires and carbon nanotubes are now the candidate nanoscale technologies that could begin to replace standard transistors in the decade after 2010. Because nanotubes are small, a large number can be packed in a given space and will enable very high-density processors and memory-storage devices. IBM (www.ibm.com/news/2001/04/27.phtml) and other institutions are currently researching this technology. For a survey of recent developments in nanotubes and other recent breakthroughs, see "Top KurzweilAI.net News of 2002," www.kurzweilai.net/meme/frame.html?main=articles/arto550.html, posted February 5, 2003.
15. These trends are discussed in more detail in Ray Kurzweil, "The Law of Accelerating Returns," www.kurzweilai.net/law, posted March 7, 2001.
16. Based on U.S. Department of Commerce Bureau of Economic Analysis data, www.bea.doc.gov/.
17. Despite these weaknesses in the productivity statistical methods, the gains in productivity are now reaching the steep part of the exponential curve. Labor productivity grew at 1.6 percent per year until 1994, then rose at 2.4 percent per year, and is now growing even more rapidly. In the quarter ending July 30, 2000, labor productivity grew at 5.3 percent. Manufacturing productivity grew at 4.4 percent annually from 1995 to 1999, durables manufacturing at 6.5 percent per year.
18. Steven E. Landsburg, "I've Got to Admit It's Getting Better." *Slate* (February 10, 2000), slate.msn.com/id/74544.
19. Robert A. Freitas, Jr., *Nanomedicine series.* (Georgetown, TX: Landes Bioscience, 1999), www.nanomedicine.com.
20. Thomas Hoffman, "Smart Dust." *Computerworld* (March 24. 2003).
21. Phillip F. Schewe, and Ben Stein, "Physics News Update." *The American Institute of Physics Bulletin of Physics News* (August 7, 1995), www.aip.org/enewsphysnews/1995/physnews.236.htm.
22. Steve Bowman, and Helit Barel, *Weapons of Mass Destruction — The Terrorist Threat.* Washington, D.C.: Congressional Research Service, 1999, www.cnie.org/nle/crsreports/international/inter-75.pdf.
23. Nick D. Bostrom, "Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards," www.kurzweilai.net/meme/frame.html?main=/articles/arto194.html, posted May 29, 2001.
24. Eliezer S. Yudkowsky, "What Is Friendly AI?" www.kurzweilai.net/meme/frame.html?main=/articles/art0172.html, posted May 3, 2001.
25. Ted Kaczynski, "The Unabomber's Manifesto," www.kurzweilai.net/meme/frame.html?main=/articles/art0182.html, posted May 14, 2001.
26. Bill McKibben, "Are We Becoming an Endangered Species? Technology and Ethics in the 21st Century," www.kurzweilai.net/meme/frame.html?main=/articles/art0351.html, posted November 8, 2001; Bill McKibben, *Enough: Staying Human in an Engineered Age* (New York: Times/Holt, 2003).
27. Kaczynski, "The Unabomber's Manifesto."
28. Foresight Institute and the Institute for Molecular Manufacturing (IMM), *Foresight Guidelines on Molecular Nanotechnology*, www.foresight.org/guidelines/current.html, accessed February 21, 1999; Christine Peterson, "Molecular Manufacturing: Social Implications of Advanced Nanotechnology," www.kurzweilai.net/meme/frame.html?main=articles/art0557.html, posted April 9, 2003; Chris Phoenix, and Mike

Treder, "Safe Utilization of Advanced Nanotechnology," www.kurzweilai.net/meme/frame.html?main=/articles/art0547.html, posted January 28, 2003.

29. Ralph C. Merkle, "Self-Replicating Systems and Low-Cost Manufacturing." In M. E. Welland, and J. K. Gimzewski, eds., *The Ultimate Limits of Fabrication and Measurement* (Dordrecht, The Netherlands: Kluwer, 1994), pp. 25–32, www.zyvex.com/nanotech/selfRepNATO.html.
30. "Carnivore Diagnostic Tool," www.fbi.gov/hq/lab/carnivore/carnivore.htm.
31. Larry Thompson, "Human Gene Therapy: Harsh Lessons, High Hopes." *FDA Consumer* (September–October 2000), www.fda.gov/fdac/features/2000/500_gene.html.
32. Joy, "Why the Future Doesn't Need Us."
33. Foresight and IMM, *Foresight Guidelines on Molecular Nanotechnology*.
34. George DeWan, "Diary of a Colonial Housewife." *Newsday*. [EDI] www.newsday.com/extras/lihistory/3/hs331a.htm.
35. Jim Oeppen, and James W. Vaupel, "Broken Limits to Life Expectancy." *Science* 296 (5570):1029–1031 (May 10, 2002). Female life expectancy in the record-holding country has risen linearly for 160 years at almost three months per year. Extrapolating backward from 1840 to 1800, female life expectancy would be 35. "The gap between female and male levels has grown from 2 to 6 years."

PART IV

Contemporary Technological Dilemmas: Climate Change

This and the following two parts of *Technology and the Future* are devoted to current issues, problems, and opportunities involving science and technology as causes, potential solutions or both. Some of these issues are being addressed — or should be addressed — by governments and public policies. Others are matters that are best addressed by individuals or within families. Some are international in scope; others are more specific to the United States. All are having or soon could have wide-ranging impacts on society. Part IV deals with climate change. Parts V and VI, are devoted to life science issues and information and communications technology (ICT), respectively.

A great deal has been written and said about climate change in recent years. There are those who still regard the question of whether the earth is warming (or if it is warming, whether the warming is caused by human activities) as unsettled. Indeed, some even claim the whole business is a fraud perpetrated on the public by some vast conspiracy. Nevertheless, the overwhelming consensus among scientists and many policymakers throughout the United States and the international community is that the problem is real and that it must be dealt with effectively and soon lest the world face increasingly dire consequences. The selection included here as chapter 12 is a review of the scientific evidence that explains why climatologists are confident that human activities are causing the earth to warm and that the warming is a very serious problem. The authors are five scientists who participated in one of the working groups that wrote the 2007 report of the Nobel Prize-winning Intergovernmental Panel on Climate Change. This chapter is followed by a concise critique by one of the most respected dissenters from the scientific consensus. The author, a professor of meteorology at

M.I.T., believes that the mathematical models used by most climate experts are flawed, that they have been "tweaked" to maximize the appearance of warming and that apparent changes we are seeing in our climate are part of a natural cycle.

The authors of chapters 14 and 15 do not question the notion that the earth's climate is changing. Rather, they approach the phenomenon as a technological problem to be solved. The approach discussed (and questioned) in both of these articles is "geoengineering" — deploying technological systems on a planetary scale with the aim of intervening in the processes that are causing global temperatures to increase, sea levels to rise, and otherwise changing our climate. Geoengineering is a "technological fix," as proposed nearly half a century ago by Alvin Weinberg and discussed in chapter 4 of this book. If one of the geoengineering proposals were to be implemented successfully, it could address the climate problem without the need to make the kind of massive changes in human behavior that most alternative approaches would require. If the possibilities of geoengineering are substantial, however, so are the risks. The authors of chapter 14 focus primarily on the international political dimensions of the issue, systematically examining the pros and cons and possible ways in which geoengineering might play out. The author of chapter 15 generally feels that it is a bad idea and cites a long list of reasons. The cure, he believes, could well be worse than the disease.

12

The Science Behind Climate Change

WILLIAM COLLINS, ROBERT COLMAN, JAMES HAYWOOD, MARTIN R. MANNING, AND PHILIP MOTE

In December 2007, the Nobel Peace Prize was awarded jointly to former U.S. Vice President Al Gore and the Intergovernmental Panel on Climate Change (IPCC). Gore, whose film An Inconvenient Truth *helped raise public awareness of climate change, received the lion's share of media attention, but it was the IPCC that sorted through the vast amount of data that has come out of climate research to determine what is happening to the Earth's climate, what the causes are, and where things are headed.*

It was for this purpose — to conduct an authoritative assessment of the available scientific and technical knowledge about climate change — that two United Nations bodies (the World Meteorological Organization and the U.N. Environment Program) established the IPCC in 1988. Hundreds of scientists from more than 130 countries have been involved in the IPCC's work. Over the past two decades, as evidence of climate change has accumulated, the IPCC's reports have reflected increasing certainty about the warming trend and its causes. The Fourth Assessment Report, *published in 2007, states that warming of the Earth's climate is unequivocal and that it is almost certainly caused by human activity. Working Group I, which prepared the "Summary for Policymakers" on which this chapter is based, involved over 600 authors from 40 countries and* *hundreds of reviewers. Although the reputations of the IPCC and some of its participants were somewhat tarnished by the publicity surrounding the November 2009 release of e-mails stolen from a server at the University of East Anglia (UK) known as "climategate," the major conclusions of the* Fourth Assessment Report *were not brought into question. The* Fifth Assessment Report *is now underway with an expected completion date of 2014.*

The authors were among the participants in Working Group I. William Collins is a professor in residence in the Department of Earth and Planetary Science at the University of California, Berkeley, and a senior scientist and head of the Climate Science Department at Lawrence Berkeley National Laboratory. Robert Colman is a senior research scientist in the Climate Dynamics Group at the Australian Bureau of Meteorology Research Centre in Melbourne. James Haywood is a research fellow in the Observation Based Research Group and the Chemistry, Climate, and Ecosystem Group at the Met Office in Exeter, England. Martin R. Manning was director of the IPCC Working Group I Support Unit at the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory in Boulder, Colorado. He is currently professor and research fellow in climate change at Victoria University of Wellington, New Zealand. Philip Mote is a research scientist at the University of Washington, in the Climate Impacts Group (CIG), and an affiliate professor in the Department of Atmospheric Sciences. He was formerly the Washington State Climatologist. The full report of the IPCC, including the "Summary for Policymakers," is available online on the IPCC web site: http://www.ipcc.ch/. *Information about the* Fifth Assessment Report *can also be found there.*

For a scientist studying climate change, "eureka" moments are unusually rare. Instead progress is generally made by a painstaking piecing together of evidence from every new temperature measurement, satellite sounding or climate-model experiment. Data get checked and rechecked, ideas tested over and over again. Do the observations fit the predicted changes? Could there be some alternative explanation? Good climate scientists, like all good scientists, want to ensure that the highest standards of proof apply to everything they discover.

And the evidence of change *has* mounted as climate records have grown longer, as our understanding of the climate system has improved and as climate models have become ever more reliable. Over the past 20 years, evidence that humans are affecting the climate has accumulated inexorably, and with it has come ever greater certainty across the scientific community in the reality of recent climate change and the potential for much greater change in the future. This increased certainty is starkly reflected in the latest report of the Intergovernmental Panel on Climate Change (IPCC), the fourth in a series of assessments of the state of knowledge on the topic, written and reviewed by hundreds of scientists worldwide.

The panel released a condensed version of the first part of the report, on the physical science basis of climate change, in February 2007. Called the "Summary for Policymakers," it delivered to policymakers and ordinary people alike an unambiguous message: scientists are more confident than ever that humans have interfered with the climate and that further human-induced climate change is on the way. Although the report finds that some of these further changes are now inevitable, its analysis also confirms that the future, particularly in the longer term, remains largely in our hands — the magnitude of expected change depends on what humans choose to do about greenhouse gas emissions.

The physical science assessment focuses on four topics: drivers of climate change, changes observed in the climate system, understanding cause-and-effect relationships, and projection of future changes. Important advances in research into all these areas have occurred since the IPCC assessment in 2001. In the pages that follow, we lay out the key findings that document the extent of change and that point to the unavoidable conclusion that human activity is driving it.

DRIVERS OF CLIMATE CHANGE

Atmospheric concentrations of many gases — primarily carbon dioxide, methane, nitrous oxide and halocarbons (gases once used widely as refrigerants and spray propellants) — have increased because of human activities. Such gases trap thermal energy (heat) within the atmosphere by means of the well-known greenhouse effect, leading to global warming. The atmospheric concentrations of carbon dioxide, methane and nitrous oxide remained roughly stable for nearly 10,000 years, before the abrupt and rapidly accelerating increases of the past 200 years [*see right illustrations in box on page 155*]. Growth rates for concentrations of carbon dioxide have been faster in the past 10 years than over any 10-year period since continuous atmospheric monitoring began in the 1950s, with concentrations now roughly 35 percent above preindustrial levels (which can be determined from air bubbles trapped in ice cores). Methane levels are roughly two and a half times preindustrial levels, and nitrous oxide levels are around 20 percent higher.

How can we be sure that humans are responsible for these increases? Some greenhouse gases (most of the halocarbons, for example) have no natural source. For other gases, two important observations demonstrate human influence. First, the geographic differences in concentrations reveal that sources occur predominantly over land in the more heavily populated Northern Hemisphere. Second, analysis of isotopes, which can distinguish among sources of emissions, demonstrates that the majority of the increase in carbon dioxide comes from combustion of fossil fuels (coal, oil and natural gas). Methane and nitrous oxide increases derive from agricultural practices and the burning of fossil fuels.

Climate scientists use a concept called radiative forcing to quantify the effect of these increased concentrations on climate. Radiative forcing is the change that is caused in the global energy balance of the earth relative to preindustrial times. (Forcing is usually expressed as watts per square meter.) A positive forcing induces warming; a negative forcing induces cooling. We can determine the radiative forcing associated with the long-lived greenhouse gases fairly precisely, because we know their atmospheric concentrations, their spatial distribution and the physics of their interaction with radiation.

Climate change is not driven just by increased greenhouse gas concentrations; other mechanisms — both natural and human-induced — also play a part. Natural

Box 12.1 Key Concepts

- Scientists are confident that humans have interfered with the climate and that further human-induced climate change is on the way.
- The principal driver of recent climate change is greenhouse gas emissions from human activities, primarily the burning of fossil fuels.
- The report of the Intergovernmental Panel on Climate Change places the probability that global warming has been caused by human activities at greater than 90 percent. The previous report, published in 2001, put the probability at higher than 66 percent.
- Although further changes in the world's climate are now inevitable, the future, particularly in the longer term, remains largely in our hands — the magnitude of expected change depends on what humans choose to do about greenhouse gas emissions.

drivers include changes in solar activity and large volcanic eruptions. The report identifies several additional significant human-induced forcing mechanisms — microscopic particles called aerosols, stratospheric and tropospheric ozone, surface albedo (reflectivity) and aircraft contrails — although the influences of these mechanisms are much less certain than those of greenhouse gases [*see left illustration in box on page 155*].

Investigators are least certain of the climatic influence of something called the aerosol cloud albedo effect, in which aerosols from human origins interact with clouds in complex ways and make the clouds brighter, reflecting sunlight back to space. Another source of uncertainty comes from the direct effect of aerosols from human origins: How much do they reflect and absorb sunlight directly as particles? Overall these aerosol effects promote cooling that could offset the warming effect of long-lived greenhouse gases to some extent. But by how much? Could it overwhelm the warming? Among the advances achieved since the 2001 IPCC report is that scientists have quantified the uncertainties associated with each individual forcing mechanism through a combination of many modeling and observational studies. Consequently, we can now confidently estimate the total human-induced component. Our best estimate is some 10 times larger than the best estimate of the natural radiative forcing caused by changes in solar activity.

This increased certainty of a net positive radiative forcing fits well with the observational evidence of warming discussed next. These forcings can be visualized as a tug-of-war, with positive forcings pulling the earth to a warmer climate and negative ones pulling it to a cooler state. The result is a no contest; we know the strength of the competitors better than ever before. The earth is being pulled to a warmer climate and will be pulled increasingly in this direction as the "anchorman" of greenhouse warming continues to grow stronger and stronger.

Box 12.2 Jargon Buster

RADIATIVE FORCING, as used in the box on page 155, is the change in the energy balance of the earth from preindustrial times to the present.

LONG-LIVED GREENHOUSE GASES include carbon dioxide, methane, nitrous oxide and halocarbons. The observed increases in these gases are the result of human activity.

OZONE is a gas that occurs both in the earth's upper atmosphere and at ground level. At ground level ozone is an air pollutant. In the upper atmosphere, an ozone layer protects life on the earth from the sun's harmful ultraviolet rays.

SURFACE ALBEDO is the reflectivity of the earth's surface: a lighter surface, such as snow cover, reflects more solar radiation than a darker surface does.

AEROSOLS are airborne particles that come from both natural (dust storms, forest fires, volcanic eruptions) and man-made sources, such as the burning of fossil fuels.

CONTRAILS, or vapor trails, are condensation trails and artificial clouds made by the exhaust of aircraft engines.

TROPOSPHERE is the layer of the atmosphere close to the earth. It rises from sea level up to about 12 kilometers (7.5 miles).

STRATOSPHERE lies just above the troposphere and extends upward about 50 kilometers.

OBSERVED CLIMATE CHANGES

The many new or improved observational data sets that became available in time for the 2007 IPCC report allowed a more comprehensive assessment of changes than was possible in earlier reports. Observational records indicate that 11 of the past 12 years are the warmest since reliable records began around 1850. The odds of such warm years happening in sequence purely by chance are exceedingly small. Changes in three important quantities — global temperature, sea level and snow cover in the Northern Hemisphere [*see box on page 157*] — all show evidence of warming, although the details vary. The previous IPCC assessment reported a warming trend of 0.6 ± 0.2 degree Celsius over the period 1901 to 2000. Because of the strong recent warming, the updated trend over 1906 to 2005 is now 0.74 ± 0.18 degree C. Note that the 1956 to 2005 trend alone is 0.65 ± 0.15 degree C, emphasizing that the majority of twentieth century warming occurred in the past 50 years. The climate, of course, continues to vary around the increased averages, and extremes have changed consistently with these averages — frost days and cold days and nights have become less common, while heat waves and warm days and nights have become more common.

The properties of the climate system include not just familiar concepts of averages of temperature, precipitation, and so on but also the state of the ocean and the cryosphere (sea ice, the great ice sheets in Greenland and Antarctica, glaciers, snow, frozen ground, and ice on lakes and rivers). Complex interactions among different parts of the climate system are a fundamental part of climate change — for example, reduction in sea ice increases the absorption of heat by the ocean and the heat flow between the ocean and the atmosphere, which can also affect cloudiness and precipitation.

A large number of additional observations are broadly consistent with the observed warming and reflect a flow of heat from the atmosphere into other components of the climate system. Spring snow cover, which decreases in concert with rising spring temperatures in northern midlatitudes, dropped abruptly around 1988 and has remained low since. This drop is of concern because snow cover is important to soil moisture and water resources in many regions.

In the ocean, we clearly see warming trends, which decrease with depth, as expected. These changes indicate that the ocean has absorbed more than 80 percent of the heat added to the climate system: this heating is a major contributor to sea-level rise. Sea level rises because water expands as it is warmed and because water from melting glaciers and ice sheets is added to the oceans. Since 1993, satellite observations have permitted more precise calculations of global sea-level rise, now estimated to be 3.1 ± 0.7 millimeters per year over the period 1993 to 2003. Some previous decades displayed similarly fast rates, and longer satellite records will be needed to determine unambiguously whether sea-level rise is accelerating. Substantial reductions in the extent of Arctic sea ice since 1978 (2.7 ± 0.6 percent per decade in the annual average, 7.4 ± 2.4 percent per decade for summer), increases in permafrost temperatures, and reductions in glacial extent globally and in Greenland and Antarctic ice sheets have also been observed in recent decades. Unfortunately, many of these quantities were not well monitored until recent decades, so the starting points of their records vary.

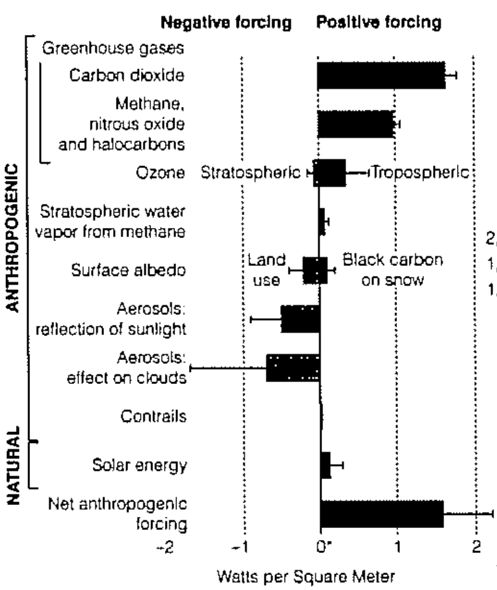
Hydrological changes are broadly consistent with warming as well. Water vapor is the strongest greenhouse gas; unlike other greenhouse gases, it is controlled principally by temperature. It has generally increased since at least the 1980s. Precipitation is very variable locally but has increased in several large regions of the world, including eastern North and South America, northern Europe, and northern and central Asia. Drying has been observed in the Sahel, the Mediterranean, southern Africa and parts of southern Asia. Ocean salinity can act as a massive rain gauge. Near-surface waters of the oceans have generally freshened in middle and high latitudes, while they have become saltier in lower latitudes, consistent with changes in large-scale patterns of precipitation.

Reconstructions of past climate — paleoclimate — from tree rings and other proxies provide important additional insights into the workings of the climate system with and without human influence. They indicate that the warmth of the past half a century is unusual in at least the previous 1,300 years. The warmest period between A.D. 700 and 1950 was probably A.D. 950 to 1100, which was several tenths of a degree C cooler than the average temperature since 1980.

Box 12.3 Influences on Climate

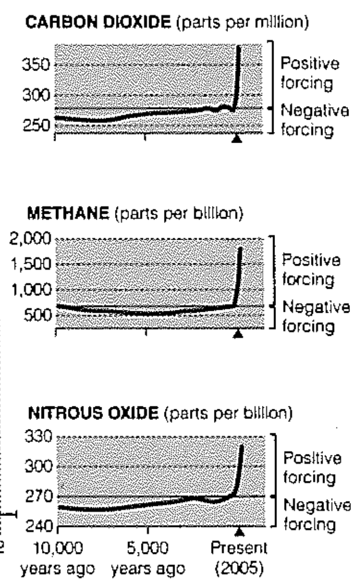
A tug-of-war between positive forcings (influences that cause the climate to grow warmer) and negative forcings (those that cause it to grow cooler) is a hands down "victory" for the predominantly human-induced forces that lead to warming (left graph). The dominant human-induced forcings are from the long-lived greenhouse gases in the atmosphere, whose concentrations have soared in the past 200 years or so (right graphs).

Radiative Forcing: The Overview



*Zero represents the preindustrial energy balance.

Greenhouse Gases: The Major Forcings



Estimates for global averages of radioactive forcing in 2005 are shown for the major mechanisms. The black error bars indicate the level of certainty associated with each forcing: it is 90 percent likely that values lie within the error bars. The radioactive forcing of the greenhouse gases, for example, is quite certain, as opposed to the uncertainty associated with the aerosol effects. (Volcanic aerosols are not included in the graph because of their episodic nature.)

Carbon dioxide, methane and nitrous oxide concentrations of the past were derived from ice cores: those for recent times come from samples of the atmosphere. Large recent increases can be attributed to human activities.

ATTRIBUTION OF OBSERVED CHANGES

Although confidence is high both that human activities have caused a positive radiative forcing and that the climate has actually changed, can we confidently link the two? This is the question of attribution: Are human activities primarily responsible for observed climate changes, or is it possible they result from some other cause, such as some natural forcing or simply spontaneous variability within the climate system? The 2001 IPCC report concluded it was *likely* (more than 66 percent probable) that most of the warming since the mid-twentieth century was attributable to humans. The 2007 report goes significantly further, upping this to *very likely* (more than 90 percent probable).

The source of the extra confidence comes from a multitude of separate advances. For a start, observational records are now roughly five years longer, and the global temperature increase over this period has been largely consistent with IPCC projections of greenhouse gas–driven warming made in previous reports dating back to 1990. In addition, changes in more aspects of the climate have been considered, such as those in atmospheric circulation or in temperatures within the ocean. Such changes paint a consistent and now broadened picture of human intervention. Climate models, which are central to attribution studies, have also improved and are able to represent the current climate and that of the recent past with considerable fidelity. Finally, some important apparent inconsistencies noted in the observational record have been largely resolved since the last report.

The most important of these was an apparent mismatch between the instrumental surface temperature record (which showed significant warming over recent decades, consistent with a human impact) and the balloon and satellite atmospheric records (which showed little of the expected warming). Several new studies of the satellite and balloon data have now largely resolved this discrepancy — with consistent warming found at the surface and in the atmosphere.

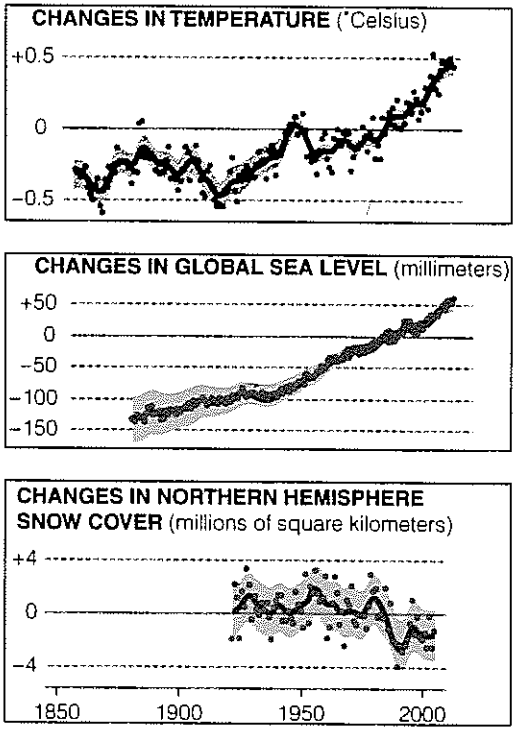
An experiment with the real world that duplicated the climate of the twentieth century with constant (rather than increasing) greenhouse gases would be the ideal way to test for the cause of climate change, but such an experiment is of course impossible. So scientists do the next best thing: they simulate the past with climate models.

Two important advances since the last IPCC assessment have increased confidence in the use of models for both attribution and projection of climate changes. The first is the development of a comprehensive, closely coordinated ensemble of simulations from 18 modeling groups around the world for the historical and future evolution of the earth's climate. Using many models helps to quantify the effects of uncertainties in various climate processes on the range of model simulations. Although some processes are well understood and well represented by physical equations (the flow of the atmosphere and ocean or the propagation of sunlight and heat, for example), some of the most critical components of the climate system are less well understood, such as clouds, ocean eddies and transpiration by vegetation. Modelers approximate these components using simplified representations called parameterizations. The principal reason to develop a multimodel ensemble for the IPCC assessments is to understand how this lack of certainty affects attribution and prediction of climate change. The ensemble for the latest assessment is unprecedented in the number of models and experiments performed.

The second advance is the incorporation of more realistic representations of climate processes in the models. These processes include the behavior of atmospheric aerosols, the dynamics (movement) of sea ice, and the exchange of water

Box 12.4 Observed Evidence

Observations of global average surface temperature, sea level and snow cover for the Northern Hemisphere in March and April document increased warming. Heavy black lines represent values averaged over a decade, with the gray shading indicating the range of uncertainty; dots show yearly values. All measures are relative to 1961–1990.



and energy between the land and the atmosphere. More models now include the major types of aerosols and the interactions between aerosols and clouds.

When scientists use climate models for attribution studies, they first run simulations with estimates of only "natural" climate influences over the past 100 years, such as changes in solar output and major volcanic eruptions. They then run models that include human-induced increases in greenhouse gases and aerosols. The results of such experiments are striking. Models using only natural forcings are unable to explain the observed global warming since the mid-twentieth century, whereas they can do so when they include anthropogenic factors in addition to natural ones. Large-scale *patterns* of temperature change are also most consistent between models and observations when all forcings are included.

Two patterns provide a fingerprint of human influence. The first is greater warming over land than ocean and greater warming at the surface of the sea than in the deeper layers. This pattern is consistent with greenhouse gas–induced warming by the overlying atmosphere: the ocean warms more slowly because of its large thermal inertia. The warming also indicates that a large amount of heat is being taken up by the ocean, demonstrating that the planet's energy budget has been pushed out of balance. A second pattern of change is that while the troposphere (the lower region of the atmosphere) has warmed, the stratosphere, just above it, has cooled. If solar changes provided the dominant forcing, warming would be expected in both atmospheric layers. The observed contrast, however, is just that expected from the combination of greenhouse gas increases and stratospheric ozone decreases. This collective evidence, when subjected to careful statistical analyses, provides much of the basis for the increased confidence that human influences are behind the observed global warming. Suggestions that cosmic rays could affect clouds, and thereby climate, have been based on correlations using limited records; they have generally not stood up when tested with additional data, and their physical mechanisms remain speculative.

What about at smaller scales? As spatial and temporal scales decrease, attribution of climate change becomes more difficult. This problem arises because natural small-scale temperature variations are less "averaged out" and thus more readily mask the change signal. Nevertheless, continued warming means the signal is emerging on smaller scales. The report has found that human activity is likely to have influenced temperature significantly down to the continental scale for all continents except Antarctica.

Human influence is discernible also in some extreme events such as unusually hot and cold nights and the incidence of heat waves. This does not mean, of course, that individual extreme events (such as the 2003 European heat wave) can be said to be simply "caused" by human-induced climate change — usually such events are complex, with many causes. But it does mean that human activities have, more likely than not, affected the *chances* of such events occurring.

PROJECTIONS OF FUTURE CHANGES

How will climate change over the twenty-first century? This critical question is addressed using simulations from climate models based on projections of future emissions of greenhouse gases and aerosols. The simulations suggest that, for greenhouse gas emissions at or above current rates, changes in climate will very likely be larger than the changes already observed during the twentieth century. Even if emissions were immediately reduced enough to stabilize greenhouse gas concentrations at current levels, climate change would continue for centuries. This inertia in the climate results from a combination of factors. They include the heat capacity of the world's oceans and the millennial timescales needed for the circulation to mix heat and carbon dioxide throughout the deep ocean and thereby come into equilibrium with the new conditions.

To be more specific, the models project that over the next 20 years, for a range of plausible emissions, the global temperature will increase at an average rate of about 0.2 degree C per decade, close to the observed rate over the past 30 years. About half of this near-term warming represents a "commitment" to future climate change arising from the inertia of the climate system response to current atmospheric concentrations of greenhouse gases.

The long-term warming over the twenty-first century, however, is strongly influenced by the future rate of emissions, and the projections cover a wide variety of scenarios, ranging from very rapid to more modest economic growth and from more to less dependence on fossil fuels. The best estimates of the increase in global temperatures range from 1.8 to 4.0 degrees C for the various emission scenarios, with higher emissions leading to higher temperatures. As for regional impacts, projections indicate with more confidence than ever before that these will mirror the *patterns* of change observed over the past 50 years (greater warming over land than ocean, for example) but that the *size* of the changes will be larger than they have been so far.

The simulations also suggest that the removal of excess carbon dioxide from the atmosphere by natural processes on land and in the ocean will become less efficient as the planet warms. This change leads to a higher percentage of emitted carbon dioxide remaining in the atmosphere, which then further accelerates global warming. This is an important positive feedback on the carbon cycle (the exchange of carbon compounds throughout the climate system). Although models agree that carbon-cycle changes represent a positive feedback, the range of their responses remains very large, depending, among other things, on poorly understood changes in vegetation or soil uptake of carbon as the climate warms. Such processes are an important topic of ongoing research.

The models also predict that climate change will affect the physical and chemical characteristics of the ocean. The estimates of the rise in sea level during the twenty-first century range from about 30 to 40 centimeters, again depending on emissions. More than 60 percent of this rise is caused by the thermal expansion of the ocean. Yet these model-based estimates do not include the possible acceleration of recently observed increases in ice loss from the Greenland and Antarctic ice sheets. Although scientific understanding of such effects is very limited, they could add an additional 10 to 20 centimeters to sea-level rises, and the possibility of significantly larger rises cannot be excluded. The chemistry of the ocean is also affected, as the increased concentrations of atmospheric carbon dioxide will cause the ocean to become more acidic.

Some of the largest changes are predicted for polar regions. These include significant increases in high-latitude land temperatures and in the depth of thawing in permafrost regions and sharp reductions in the extent of summer sea ice in the Arctic basin. Lower latitudes will likely experience more heat waves, heavier precipitation, and stronger (but perhaps less frequent) hurricanes and typhoons. The extent to which hurricanes and typhoons may strengthen is uncertain and is a subject of much new research.

Some important uncertainties remain, of course. For example, the precise way in which clouds will respond as temperatures increase is a critical factor

Box 12.5 The Mechanics of the IPCC

The IPCC was established by governments in 1988 to provide assessments of available scientific and technical information on climate change. The process used to produce these assessments is designed to ensure their high credibility in both science and policy communities.

Comprehensive assessments were published in 1990, 1995, 2001, and 2007.

Three separate "working groups" examine the physical science of climate change, the effects on nature and society, and methods for mitigation.

Lead authors, who are active participants in relevant research, are nominated by governments. Care is taken to balance points of view as well as geography, gender and age.

A review process tests the authors' assessment against views in the broader expert community. More than 600 expert reviewers provided over 30,000 comments on the report of Working Group I, on which this article is based.

Each of the three working groups also issues a "Summary for Policymakers," which is done in cooperation with government delegates to ensure that the language used is clear to policymakers.

governing the overall size of the projected warming. The complexity of clouds, however, means that their response has been frustratingly difficult to pin down, and, again, much research remains to be done in this area.

We are now living in an era in which both humans and nature affect the future evolution of the earth and its inhabitants. Unfortunately, the crystal ball provided by our climate models becomes cloudier for predictions out beyond a century or so. Our limited knowledge of the response of both natural systems and human society to the growing impacts of climate change compounds our uncertainty. One result of global warming is certain, however. Plants, animals and humans will be living with the consequences of climate change for at least the next thousand years.

13

The Climate Science Isn't Settled

RICHARD S. LINDZEN

The IPCC report and the large majority of scientists who hold views consistent with most of its conclusions have many critics. Some of the critics have vested interests in industries and technologies that would be affected by policies that have been proposed to adapt to or counter global climate change. Others adhere to what has become virtually a political ideology of climate change denial. Still others are scientists who are critical of the way their colleagues have selected and interpreted climate data. There are relatively few such scientists and fewer still who are distinguished scholars of climate science. Richard Lindzen is one of the latter. He is Alfred P. Sloan Professor of Meteorology in the Department of Earth, Atmospheric and Planetary Sciences at MIT. He holds a Ph.D. from Harvard and is a member of the National Academy of Sciences. He has received awards from the American Meteorological Society and the American Geophysical Union. An atmospheric physicist, Lindzen has published more than 200 scientific papers and books and was the lead author of the chapter on "Physical Climate Processes and Feedbacks" in the third IPCC assessment report. Lindzen is also a skeptic of the scientific consensus on climate change. In this article, originally published as an op-ed in The Wall Street Journal *a few weeks after "climategate" hit the news media, he argues that the conclusions of the IPCC fourth assessment report are flawed and that the models on which they are based are too weak to support the notion that we are facing a climate "catastrophe." Although the vast majority of climate scientists disagree with Lindzen and specific rebuttals can be found online, his arguments deserve to be taken seriously.*

Is there a reason to be alarmed by the prospect of global warming? Consider that the measurement used, the globally averaged temperature anomaly (GATA), is always changing. Sometimes it goes up, sometimes down, and occasionally — such as for the last dozen years or so — it does little that can be discerned.

Richard S. Lindzen, "The Climate Science Isn't Settled," *The Wall Street Journal* (November 30, 2009). Reprinted by permission of the author.

Claims that climate change is accelerating are bizarre. There is general support for the assertion that GATA has increased about 1.5 degrees Fahrenheit since the middle of the 19th century. The quality of the data is poor, though, and because the changes are small, it is easy to nudge such data a few tenths of a degree in any direction. Several of the emails from the University of East Anglia's Climate Research Unit (CRU) that have caused such a public ruckus dealt with how to do this so as to maximize apparent changes.

The general support for warming is based not so much on the quality of the data, but rather on the fact that there was a little ice age from about the 15th to the 19th century. Thus it is not surprising that temperatures should increase as we emerged from this episode. At the same time that we were emerging from the little ice age, the industrial era began, and this was accompanied by increasing emissions of greenhouse gases such as CO_2, methane and nitrous oxide. CO_2 is the most prominent of these, and it is again generally accepted that it has increased by about 30%.

The defining characteristic of a greenhouse gas is that it is relatively transparent to visible light from the sun but can absorb portions of thermal radiation. In general, the earth balances the incoming solar radiation by emitting thermal radiation, and the presence of greenhouse substances inhibits cooling by thermal radiation and leads to some warming.

That said, the main greenhouse substances in the earth's atmosphere are water vapor and high clouds. Let's refer to these as major greenhouse substances to distinguish them from the anthropogenic minor substances. Even a doubling of CO_2 would only upset the original balance between incoming and outgoing radiation by about 2%. This is essentially what is called "climate forcing."

There is general agreement on the above findings. At this point there is no basis for alarm regardless of whether any relation between the observed warming and the observed increase in minor greenhouse gases can be established. Nevertheless, the most publicized claims of the U.N.'s Intergovernmental Panel on Climate Change (IPCC) deal exactly with whether any relation can be discerned. The failure of the attempts to link the two over the past 20 years bespeaks the weakness of any case for concern.

The IPCC's Scientific Assessments generally consist of about 1,000 pages of text. The Summary for Policymakers is 20 pages. It is, of course, impossible to accurately summarize the 1,000-page assessment in just 20 pages; at the very least, nuances and caveats have to be omitted. However, it has been my experience that even the summary is hardly ever looked at. Rather, the whole report tends to be characterized by a single iconic claim.

The main statement publicized after the last IPCC Scientific Assessment two years ago was that it was likely that most of the warming since 1957 (a point of anomalous cold) was due to man. This claim was based on the weak argument that the current models used by the IPCC couldn't reproduce the warming from about 1978 to 1998 without some forcing, and that the only forcing that they could think of was man. Even this argument assumes that these models adequately deal with natural internal variability — that is, such naturally occurring cycles as El Nino, the Pacific Decadal Oscillation, the Atlantic Multidecadal Oscillation, etc.

Yet articles from major modeling centers acknowledged that the failure of these models to anticipate the absence of warming for the past dozen years was due to the failure of these models to account for this natural internal variability. Thus even the basis for the weak IPCC argument for anthropogenic climate change was shown to be false.

Of course, none of the articles stressed this. Rather they emphasized that according to models modified to account for the natural internal variability, warming would resume — in 2009, 2013 and 2030, respectively.

But even if the IPCC's iconic statement were correct, it still would not be cause for alarm. After all we are still talking about tenths of a degree for over 75% of the climate forcing associated with a doubling of CO_2. The potential (and only the potential) for alarm enters with the issue of climate sensitivity — which refers to the change that a doubling of CO_2 will produce in GATA. It is generally accepted that a doubling of CO_2 will only produce a change of about two degrees Fahrenheit if all else is held constant. This is unlikely to be much to worry about.

Yet current climate models predict much higher sensitivities. They do so because in these models, the main greenhouse substances (water vapor and clouds) act to amplify anything that CO_2 does. This is referred to as positive feedback. But as the IPCC notes, clouds continue to be a source of major uncertainty in current models. Since clouds and water vapor are intimately related, the IPCC claim that they are more confident about water vapor is quite implausible.

There is some evidence of a positive feedback effect for water vapor in cloud-free regions, but a major part of any water-vapor feedback would have to acknowledge that cloud-free areas are always changing, and this remains an unknown. At this point, few scientists would argue that the science is settled. In particular, the question remains as to whether water vapor and clouds have positive or negative feedbacks.

The notion that the earth's climate is dominated by positive feedbacks is intuitively implausible, and the history of the earth's climate offers some guidance on this matter. About 2.5 billion years ago, the sun was 20%–30% less bright than now (compare this with the 2% perturbation that a doubling of CO_2 would produce), and yet the evidence is that the oceans were unfrozen at the time, and that temperatures might not have been very different from today's. Carl Sagan in the 1970s referred to this as the "Early Faint Sun Paradox."

For more than 30 years there have been attempts to resolve the paradox with greenhouse gases. Some have suggested CO_2 — but the amount needed was thousands of times greater than present levels and incompatible with geological evidence. Methane also proved unlikely. It turns out that increased thin cirrus cloud coverage in the tropics readily resolves the paradox — but only if the clouds constitute a negative feedback. In present terms this means that they would diminish rather than enhance the impact of CO_2.

There are quite a few papers in the literature that also point to the absence of positive feedbacks. The implied low sensitivity is entirely compatible with the small warming that has been observed. So how do models with high sensitivity manage to simulate the currently small response to a forcing that is almost as large as a doubling of CO_2? Jeff Kiehl notes in a 2007 article from the National Center for Atmospheric Research, the models use another quantity that the IPCC lists as poorly known (namely aerosols) to arbitrarily cancel as much greenhouse warming as needed to match the data, with each model choosing a different degree of cancellation according to the sensitivity of that model.

What does all this have to do with climate catastrophe? The answer brings us to a scandal that is, in my opinion, considerably greater than that implied in the hacked emails from the Climate Research Unit (though perhaps not as bad as their destruction of raw data): namely the suggestion that the very existence of warming or of the greenhouse effect is tantamount to catastrophe. This is the grossest of "bait and switch" scams. It is only such a scam that lends importance to the machinations in the emails designed to nudge temperatures a few tenths of a degree.

The notion that complex climate "catastrophes" are simply a matter of the response of a single number, GATA, to a single forcing, CO_2 (or solar forcing for that matter), represents a gigantic step backward in the science of climate. Many disasters associated with warming are simply normal occurrences whose existence is falsely claimed to be evidence of warming. And all these examples involve phenomena that are dependent on the confluence of many factors.

Our perceptions of nature are similarly dragged back centuries so that the normal occasional occurrences of open water in summer over the North Pole, droughts, floods, hurricanes, sea-level variations, etc. are all taken as omens, portending doom due to our sinful ways (as epitomized by our carbon footprint). All of these phenomena depend on the confluence of multiple factors as well.

Consider the following example. Suppose that I leave a box on the floor, and my wife trips on it, falling against my son, who is carrying a carton of eggs, which then fall and break. Our present approach to emissions would be analogous to deciding that the best way to prevent the breakage of eggs would be to outlaw leaving boxes on the floor. The chief difference is that in the case of atmospheric CO_2 and climate catastrophe, the chain of inference is longer and less plausible than in my example.

14

The Geoengineering Option: A Last Resort Against Global Warming?

DAVID G. VICTOR, M. GRANGER MORGAN, JAY APT, JOHN STEINBRUNER, AND KATHARINE RICKE

Global climate change is one of the major technological dilemmas — probably the technological dilemma of our time. Although there is broad agreement among scientists and policymakers in many countries that it is real and that it is anthropogenic (that is, caused by human activities), there are those who disagree, and even among scientists and policymakers who share the consensus, there is no clear agreement on what should be done about it. Views range along a spectrum of potential actions. At one end are those who consider current trends to be part of a natural climate cycle and advocate doing nothing in the expectation that the problem will eventually take care of itself. Others regard climate change as inevitable and believe that the best course is to adapt to it (by building seawalls to control flooding or changing agricultural crops and practices, for example). Still others support stronger actions — the mostly widely discussed of which is to limit emissions of greenhouse gases. Finally there some who believe that we should explore the possibility of deploying systems on a global scale aimed at intervening in climate processes. The last of these approaches, known as "geoengineering" — or, as one science writer calls it, "hacking the planet,"[1] — is

From David G. Victor, M. Granger Morgan, Jay Apt, John Steinbrunner, and Katharine Ricke. "The Geoengineering Option: A Last Resort Against Global Warming?" *Foreign Affairs* (Vol. 88, No. 2), March/April 2009, pp. 64–76. Reprinted with permission.

probably the most risky, but it could also be the most effective means of meeting the challenge.

The risks, as well as the scale and unpredictability of potential impacts, have made geoengineering a highly controversial notion. Until the past few years, discussions of geoengineering projects took place mostly within military circles (where their potential as weapons were examined) or on the fringes of scientific legitimacy. More recently, however, they have moved into the mainstream and the discussions, involving some of the more important figures in climate science and foreign policy, have become more focused and more serious.

This article, in the highly respected journal, Foreign Affairs, *by a group of distinguished engineers, technologists, political scientists, and public policy experts, is a prime example. Part of the journal's "syllabus on climate change," it takes a wide-ranging look at the ramifications of unprecedented technological projects that would seek to alter the earth's climate.*

The first author, David G. Victor, is professor at the School of International Relations and Pacific Studies of the University of California at San Diego, and director of the School's new Laboratory on International Law and Regulation. At the time this article was written, he was professor at Stanford Law School. M. Granger Morgan is Lord Chair Professor in Engineering, and head of the Department of Engineering and Public Policy at Carnegie Mellon University. Jay Apt is professor of technology in the Tepper School of Business and the Department of Engineering and Public Policy at Carnegie Mellon University. He is a former astronaut and flew four missions on NASA's space shuttle. John Steinbruner is professor of public policy at the School of Public Policy and director of the Center for International and Security Studies at the University of Maryland. Katharine Ricke is a Ph.D. candidate in the Department of Engineering and Public Policy at Carnegie Mellon.

Each year, the effects of climate change are coming into sharper focus. Barely a month goes by without some fresh bad news: ice sheets and glaciers are melting faster than expected, sea levels are rising more rapidly than ever in recorded history, plants are blooming earlier in the spring, water supplies and habitats are in danger, birds are being forced to find new migratory patterns.

The odds that the global climate will reach a dangerous tipping point are increasing. Over the course of the twenty-first century, key ocean currents, such as the Gulf Stream, could shift radically, and thawing permafrost could release huge amounts of additional greenhouse gases into the atmosphere. Such scenarios, although still remote, would dramatically accelerate and compound the consequences of global warming. Scientists are taking these doomsday scenarios seriously because the steady accumulation of warming gases in the atmosphere is forcing change in the climate system at rates so rapid that the outcomes are extremely difficult to predict.

Eliminating all the risks of climate change is impossible because carbon dioxide emissions, the chief human contribution to global warming, are unlike conventional air pollutants, which stay in the atmosphere for only hours or days. Once carbon dioxide enters the atmosphere, much of it remains for over a hundred years. Emissions from anywhere on the planet contribute to the global problem, and once headed in the wrong direction, the climate system is slow to respond to attempts at reversal. As with a bathtub that has a large faucet and a small drain, the only practical way to lower the level is by dramatically cutting the inflow. Holding global warming steady at its current rate would require a worldwide 60–80 percent cut in emissions, and it would still take decades for the atmospheric concentration of carbon dioxide to stabilize.

Most human emissions of carbon dioxide come from burning fossil fuels, and most governments have been reluctant to force the radical changes necessary to reduce those emissions. Economic growth tends to trump vague and elusive global aspirations. The United States has yet to impose even a cap on its emissions, let alone a reduction. The European Union has adopted an emissions-trading scheme that, although promising in theory, has not yet had much real effect because carbon prices are still too low to cause any significant change in behavior. Even Norway, which in 1991 became one of the first nations to impose a stiff tax on emissions, has seen a net increase in its carbon dioxide emissions. Japan, too, has professed its commitment to taming global warming. Nevertheless, Tokyo is struggling to square the need for economic growth with continued dependence on an energy system powered mainly by conventional fossil fuels. And China's emissions recently surpassed those of the United States, thanks to coal-fueled industrialization and a staggering pace of economic growth. The global economic crisis is stanching emissions a bit, but it will not come close to shutting off the faucet.

The world's slow progress in cutting carbon dioxide emissions and the looming danger that the climate could take a sudden turn for the worse require policymakers to take a closer look at emergency strategies for curbing the effects of global warming. These strategies, often called "geoengineering," envision deploying systems on a planetary scale, such as launching reflective particles into the atmosphere or positioning sunshades to cool the earth. These strategies could cool the planet, but they would not stop the buildup of carbon dioxide or lessen all its harmful impacts. For this reason, geoengineering has been widely shunned by those committed to reducing emissions.

Serious research on geoengineering is still in its infancy, and it has not received the attention it deserves from politicians. The time has come to take it seriously. Geoengineering could provide a useful defense for the planet — an emergency shield that could be deployed if surprisingly nasty climatic shifts put vital ecosystems and billions of people at risk. Actually raising the shield, however, would be a political choice. One nation's emergency can be another's opportunity, and it is unlikely that all countries will have similar assessments of how to balance the ills of unchecked climate change with the risk that geoengineering could do more harm than good. Governments should immediately begin to undertake serious research on geoengineering and help create international norms governing its use.

THE RAINMAKERS

Geoengineering is not a new idea. In 1965, when President Lyndon Johnson received the first-ever U.S. presidential briefing on the dangers of climate change, the only remedy prescribed to counter the effects of global warming was geoengineering. That advice reflected the scientific culture of the time, which imagined that engineering could fix almost any problem.

By the late 1940s, both the United States and the Soviet Union had begun exploring strategies for modifying the weather to gain battlefield advantage. Many schemes focused on "seeding" clouds with substances that would coax them to drop more rain. Despite offering no clear advantage to the military, "weather makers" were routinely employed (rarely with much effect) to squeeze more rain from clouds for thirsty crops. Starting in 1962, U.S. government researchers for Project Stormfury tried to make tropical hurricanes less intense through cloud seeding, but with no clear success. Military experts also dreamed of using nuclear explosions and other interventions to create a more advantageous climate. These applications were frightening enough that in 1976 the United Nations adopted the Convention on the Prohibition of Military or Any Other Hostile Use of Environmental Modification Techniques to bar such projects. By the 1970s, after a string of failures, the idea of weather modification for war and farming had largely faded away.

Today's proposals for geoengineering are more likely to have an impact because the interventions needed for global-scale geoengineering are much less subtle than those that sought to influence local weather patterns. The earth's climate is largely driven by the fine balance between the light energy with which the sun bathes the earth and the heat that the earth radiates back to space. On average, about 70 percent of the earth's incoming sunlight is absorbed by the atmosphere and the planet's surface; the remainder is reflected back into space. Increasing the reflectivity of the planet (known as the albedo) by about one percentage point could have an effect on the climate system large enough to offset the gross increase in warming that is likely over the next century as a result of a doubling of the amount of carbon dioxide in the atmosphere. Making such tweaks is much more straightforward than causing rain or fog at a particular location in the ways that the weather makers of the late 1940s and 1950s dreamed of doing.

In fact, every few decades, volcanoes validate the theory that it is possible to engineer the climate. When Mount Pinatubo, in the Philippines, erupted in 1991, it ejected plumes of sulfate and other fine particles into the atmosphere, which reflected a bit more sunlight and cooled the planet by about 0.5 degrees Celsius over the course of a year. Larger eruptions, such as the 1883 eruption of Krakatau, in Indonesia, have caused even greater cooling that lasted longer. Unlike efforts to control emissions of greenhouse gases, which will take many years to yield a noticeable effect, volcano-like strategies for cooling the planet would work relatively promptly.

Another lesson from volcanoes is that a geoengineering system would require frequent maintenance, since most particles lofted into the stratosphere would disappear after a year or two. Once a geoengineering project were under way, there would be strong incentives to continue it, since failure to keep the shield in place could allow particularly harmful changes in the earth's climate, such as warming so speedy that ecosystems would collapse because they had no time to adjust. By carefully measuring the climatic effects of the next major volcanic eruption with satellites and aircraft, geoengineers could design a number of climate-cooling technologies.

ALBEDO ENHANCERS

Today, the term "geoengineering" refers to a variety of strategies designed to cool the climate. Some, for example, would slowly remove carbon dioxide from the atmosphere, either by manipulating the biosphere (such as by fertilizing the ocean with nutrients that would allow plankton to grow faster and thus absorb more carbon) or by directly scrubbing the air with devices that resemble big cooling towers. However, from what is known today, increasing the earth's albedo offers the most promising method for rapidly cooling the planet.

Most schemes that would alter the earth's albedo envision putting reflective particles into the upper atmosphere, much as volcanoes do already. Such schemes offer quick impacts with relatively little effort. For example, just one kilogram of sulfur well placed in the stratosphere would roughly offset the warming effect of several hundred thousand kilograms of carbon dioxide. Other schemes include seeding bright reflective clouds by blowing seawater or other substances into the lower atmosphere. Substantial reductions of global warming are also possible to achieve by converting dark places that absorb lots of sunlight to lighter shades — for example, by replacing dark forests with more reflective grasslands. (Engineered plants might be designed for the task.) More ambitious projects could include launching a huge cloud of thin refracting discs into a special space orbit that parks the discs between the sun and the earth in order to bend just a bit of sunlight away before it hits the planet.

So far, launching reflective materials into the upper stratosphere seems to be the easiest and most cost-effective option. This could be accomplished by using high-flying aircraft, naval guns, or giant balloons. The appropriate materials could include sulfate aerosols (which would be created by releasing sulfur dioxide gas), aluminum oxide dust, or even self-levitating and self-orienting designer particles engineered to migrate to the Polar Regions and remain in place for long periods. If it can be done, concentrating sunshades over the poles would be a particularly interesting option, since those latitudes appear to be the most sensitive to global warming. Most cost estimates for such geoengineering strategies are preliminary and unreliable. However, there is general agreement that the strategies are cheap; the total expense of the most cost-effective options would amount to perhaps as little as a few billion dollars, just one percent (or less) of the cost of dramatically cutting emissions.

Cooling the planet through geoengineering will not, however, fix all of the problems related to climate change. Offsetting warming by reflecting more sunlight back into space will not stop the rising concentration of carbon dioxide in the atmosphere. Sooner or later, much of that carbon dioxide ends up in the oceans, where it forms carbonic acid. Ocean acidification is a catastrophe for marine ecosystems, for the 100 million people who depend on coral reefs for their livelihoods, and for the many more who depend on them for coastal protection from storms and for biological support of the greater ocean food web. Over the last century, the oceans have become markedly more acidic, and current projections suggest that without a serious effort to control emissions, the concentration of carbon dioxide will be so high by the end of the century that many organisms that make shells will disappear and most coral reef ecosystems will collapse, devastating the marine fishing industry. Recent studies have also suggested that ocean acidification will increase the size and depth of "dead zones," areas of the sea that are so oxygen depleted that larger marine life, such as squid, are unable to breathe properly.

Altering the albedo of the earth would also affect atmospheric circulation, rainfall, and other aspects of the hydrologic cycle. In the six to 18 months following the eruption of Mount Pinatubo, rainfall and river flows dropped, particularly in the tropics. Understanding these dangers better would help convince government leaders in rainfall-sensitive regions, such as parts of China and India (along with North Africa, the Middle East, and the desert regions of the southwestern United States), not to prematurely deploy poorly designed geoengineering schemes that could wreak havoc on agricultural productivity. Indeed, some climate models already suggest that negative outcomes — decreased precipitation over land (especially in the tropics) and increased precipitation over the oceans — would accompany a geoengineering scheme that sought to lower average temperatures by raising the planet's albedo. Such changes could increase the risk of major droughts in some regions and have a major impact on agriculture and the supply of fresh water. Complementary policies — such as investing in better water-management schemes — may be needed.

The highly uncertain but possibly disastrous side effects of geoengineering interventions are difficult to compare to the dangers of unchecked global climate change. Chances are that if countries begin deploying geoengineering systems, it will be because calamitous climate change is near at hand. Yet the assignment of blame after a geoengineering disaster would be very different from the current debates over who is responsible for climate change, which is the result of centuries of accumulated emissions from activities across the world. By contrast, the side effects of geoengineering projects could be readily pinned on the geoengineers themselves. That is one reason why nations must begin building useful international norms to govern geoengineering in order to assess its dangers and decide when to act in the event of an impending climatic disaster.

LONE RANGERS

An effective foreign policy strategy for managing geoengineering is difficult to formulate because the technology involved turns the normal debate over climate change on its head. The best way to reduce the danger of global warming is, of course, to cut emissions of carbon dioxide and other greenhouse gases. But success in that venture will require all the major emitting countries, with their divergent interests, to cooperate for several decades in a sustained effort to develop and deploy completely new energy systems with much lower emissions. Incentives to defect and avoid the high cost of emissions controls will be strong.

By contrast, geoengineering is an option at the disposal of any reasonably advanced nation. A single country could deploy geoengineering systems from its own territory without consulting the rest of the planet. Geoengineers keen to alter their own country's climate might not assess or even care about the dangers their actions could create for climates, ecosystems, and economies elsewhere. A unilateral geoengineering project could impose costs on other countries, such as changes in precipitation patterns and river flows or adverse impacts on agriculture, marine fishing, and tourism. And merely knowing that geoengineering exists as an option may take the pressure off governments to implement the policies needed to cut emissions.

At some point in the near future, it is conceivable that a nation that has not done enough to confront climate change will conclude that global warming has become so harmful to its interests that it should unilaterally engage in geoengineering. Although it is hardly wise to mess with a poorly understood global climate system using instruments whose effects are also unknown, politicians must take geoengineering seriously because it is cheap, easy, and takes only one government with sufficient hubris or desperation to set it in motion. Except in the most dire climatic emergency, universal agreement on the best approach is highly unlikely. Unilateral action would create a crisis of legitimacy that could make it especially difficult to manage geoengineering schemes once they are under way.

Although governments are the most likely actors, some geoengineering options are cheap enough to be deployed by wealthy and capable individuals or corporations. Although it may sound like the stuff of a future James Bond movie, private-sector geoengineers might very well attempt to deploy affordable geoengineering schemes on their own. And even if governments manage to keep freelance geoengineers in check, the private sector could emerge as a potent force by becoming an interest group that pushes for deployment or drives the direction of geoengineering research and assessment. Already, private companies are running experiments on ocean fertilization in the hope of sequestering carbon dioxide and earning credits that they could trade in carbon markets. Private developers of technology for albedo modification could obstruct an open and transparent research environment as they jockey for position in the potentially lucrative market for testing and deploying geoengineering systems. To prevent such scenarios and to establish the rules that should govern the use of geoengineering technology for the good of the entire planet, a cooperative, international research agenda is vital.

FROM SCIENCE FICTION TO FACTS

Despite years of speculation and vague talk, peer-reviewed research on geoengineering is remarkably scarce. Nearly the entire community of geoengineering scientists could fit comfortably in a single university seminar room, and the entire scientific literature on the subject could be read during the course of a transatlantic flight. Geoengineering continues to be considered a fringe topic.

Many scientists have been reluctant to raise the issue for fear that it might create a moral hazard: encouraging governments to deploy geoengineering rather than invest in cutting emissions. Indeed, geoengineering ventures will be viewed with particular suspicion if the nations funding geoengineering research are not also investing in dramatically reducing their emissions of carbon dioxide and other greenhouse gases. Many scientists also rightly fear that grants for geoengineering research would be subtracted from the existing funds for urgently needed climate-science research and carbon-abatement technologies. But there is a pressing need for a better understanding of geoengineering, rooted in theoretical studies and empirical field measurements. The subject also requires the talents of engineers, few of whom have joined the small group of scientists studying these techniques.

The scientific academies in the leading industrialized and emerging countries — which often control the purse strings for major research grants — must orchestrate a serious and transparent international research effort funded by their governments. Although some work is already under way, a more comprehensive understanding of geoengineering options and of risk-assessment procedures would make countries less trigger-happy and more inclined to consider deploying geoengineering systems in concert rather than on their own. (The International Council for Science, which has a long and successful history of coordinating scientific assessments of technical topics, could also lend a helping hand.) Eventually, a dedicated international entity overseen by the leading academies, provided with a large budget, and suffused with the norms of transparency and peer review will be necessary.

In time, international institutions such as the Intergovernmental Panel on Climate Change could be expected to synthesize the findings from the published research. The IPCC, which shared the Nobel Peace Prize in 2007 for its pivotal role in building a consensus around climate science, has not considered geoengineering so far because the topic is politically radioactive and there is a dearth of peer-reviewed research on it. The IPCC's fifth assessment report on climate change, which is being planned right now, should promise to take a closer look at geoengineering. Attention from the IPCC and the world's major scientific academies would help encourage new research.

A broad and solid foundation of research would help on three fronts. First, it would transform the discussion about geoengineering from an abstract debate into one focused on real risk assessment. Second, a research program that was backed by the world's top scientific academies could secure funding and political cover for essential but controversial experiments. (Field trials of engineered aerosols, for example, could spark protests comparable to those that accompanied trials of genetically modified crops.) Such experiments will be seen as more acceptable if they are designed and overseen by the world's leading scientists and evaluated in a fully transparent fashion. Third, and what is crucial, a better understanding of the dangers of geoengineering would help nations craft the norms that should govern the testing and possible deployment of newly developed technologies. Scientists could be influential in creating these norms, just as nuclear scientists framed the options on nuclear testing and influenced pivotal governments during the Cold War.

If countries were actually to contemplate the deployment of geoengineering technologies, there would inevitably be questions raised about what triggers would compel the use of these systems. Today, nobody knows which climatic triggers are most important for geoengineering because research on the harmful effects of climate change has not been coupled tightly enough with research on whether and how geoengineering might offset those effects.

Although the international scientific community should take the lead in developing a research agenda, social scientists, international lawyers, and foreign policy experts will also have to play a role. Eventually, there will have to be international laws to ensure that globally credible and legitimate rules govern the deployment of geoengineering systems. But effective legal norms cannot be imperiously declared. They must be carefully developed by informed consensus in order to avoid encouraging the rogue forms of geoengineering they are intended to prevent.

Those who worry that such research will cause governments to abandon their efforts to control emissions, including much of the environmental community, are prone to seek a categorical prohibition against geoengineering. But a taboo would interfere with much-needed scientific research on an option that might be better for humanity and the world's ecosystems than allowing unchecked climate change or reckless unilateral geoengineering. Formal prohibition is unlikely to stop determined rogues, but a smart and scientifically sanctioned research program could gather data essential to understanding the risks of geoengineering strategies and to establishing responsible criteria for their testing and deployment.

BRAVE NEW WORLD

Fiddling with the climate to fix the climate strikes most people as a shockingly bad idea. Many worry that research on geoengineering will make governments less willing to regulate emissions. It is more likely, however, that serious study will reveal the many dangerous side effects of geoengineering, exposing it as a true option of last resort. But because the option exists, and might be used, it would be dangerous for scientists and policymakers to ignore it. Assessing and managing the risks of geoengineering may not require radically different approaches from those used for other seemingly risky endeavors, such as genetic engineering (research on which was paused in the 1970s as scientists worked out useful regulatory systems), the construction and use of high-energy particle accelerators (which a few physicists suggest could create black holes that might swallow the earth), and the development of nanotechnology (which some worry could unleash self-replicating nanomachines that could reduce the world to "gray goo"). The option of eliminating risk altogether does not exist. Countries have kept smallpox samples on hand, along with samples of many other diseases, such as the Ebola and Marburg viruses, despite the danger of their inadvertent release. All of these are potentially dangerous endeavors that governments, with scientific support, have been able to manage for the greater good.

Humans have already engaged in a dangerous geophysical experiment by pumping massive amounts of carbon dioxide and other greenhouse gases into the atmosphere. The best and safest strategy for reversing climate change is to halt this buildup of greenhouse gases, but this solution will take time, and it involves myriad practical and political difficulties. Meanwhile, the dangers are mounting. In a few decades, the option of geoengineering could look less ugly for some countries than unchecked changes in the climate. Nor is it impossible that later in the century the planet will experience a climatic disaster that puts ecosystems and human prosperity at risk. It is time to take geoengineering out of the closet — to better control the risk of unilateral action and also to know the costs and consequences of its use so that the nations of the world can collectively decide whether to raise the shield if they think the planet needs it.

ENDNOTE

1. Eli Kintisch, *Hack the Planet: Science's Best Hope — or Worst Nightmare — for Averting Climate Catastrophe* (Hoboken: NJ: John Wiley & Sons, 2010).

15

20 Reasons Why Geoengineering May Be a Bad Idea

ALAN ROBOCK

The authors of the previous article advocate a pragmatic approach to geoengineering. They have serious reservations about the technology and the risks it poses, but they are at least as worried about the potential impacts of climate change.

They are concerned about the politics of climate change and the possibility that some nations may choose to take unilateral action. In the end they say it is time "to take geoengineering out of the closet" and discuss it openly at the international level.

Alan Robock has a more negative view. His concerns "address unknowns in climate system response; effects on human quality of life; and the political, ethical, and moral issues" raised by geoengineering. Robock is a distinguished professor in the Department of Environmental Sciences at Rutgers University, where he is also associate director of the Center for Environmental Prediction, director of the Meteorology Undergraduate Program, and a member of the Graduate Program in Atmospheric Science. He has many honors and distinctions, including serving as an American Association for the Advancement of Science (AAAS) Congressional Fellow in 1986 and as snow forecaster for Montgomery County (Maryland) Public Schools in the early 1980s (when two of my children were students there and eagerly awaited his forecasts of snow as well as — they hoped — subsequent school closings). He was a participant in the Intergovernmental Panel on Climate

From Alan Robock, "20 Reasons Why Geoengineering May Be a Bad Idea," *Bulletin of the Atomic Scientists* (Vol. 64, No. 2), May/June 2008, pp. 14–18, 59. Reprinted with permission.

Change and is a Fellow of the American Meteorological Society. Robock holds a B.A. in meteorology from the University of Wisconsin at Madison and an S.M. and a Ph.D. in meteorology from M.I.T.

Robock worries that while implementation of geoengineering approaches to limiting solar radiation would perhaps limit temperature increases, they would not address other important impacts of carbon emissions — acid deposition, ozone depletion, and regional climate effects. His article, originally published in the Bulletin of the Atomic Scientists, *includes two interesting sidebars: one on carbon sequestration by Kirsten Jerch, assistant editor at the* Bulletin *and another on ethical aspects of geoengineering, by Martin Bunzl, a philosophy professor at Rutgers.*

The stated objective of the 1992 U.N. Framework Convention on Climate Change is to stabilize greenhouse gas concentrations in the atmosphere "at a level that would prevent dangerous anthropogenic interference with the climate system." Though the framework convention did not define "dangerous," that level is now generally considered to be about 450 parts per million (ppm) of carbon dioxide in the atmosphere; the current concentration is about 385 ppm. up from 280 ppm before the Industrial Revolution.

In light of society's failure to act concertedly to deal with global warming in spite of the framework convention agreement, two prominent atmospheric scientists recently suggested that humans consider geoengineering — in this case, deliberate modification of the climate to achieve specific effects such as cooling — to address global warming. Nobel laureate Paul Crutzen, who is well regarded for his work on ozone damage and nuclear winter, spearheaded a special August 2006 issue of *Climatic Change* with a controversial editorial about injecting sulfate aerosols into the stratosphere as a means to block sunlight and cool Earth. Another respected climate scientist, Tom Wigley, followed up with a feasibility study in *Science* that advocated the same approach in combination with emissions reduction.[1]

The idea of geoengineering traces its genesis to military strategy during the early years of the Cold War, when scientists in the United States and the Soviet Union devoted considerable funds and research efforts to controlling the weather. Some early geoengineering theories involved damming the Strait of Gibraltar and the Bering Strait as a way to *warm* the Arctic, making Siberia more habitable.[2] Since scientists became aware of rising concentrations of atmospheric carbon dioxide, however, some have proposed artificially altering climate and weather patterns to reverse or mask the effects of global warming.

Some geoengineering schemes aim to remove carbon dioxide from the atmosphere, through natural or mechanical means. Ocean fertilization, where iron dust is dumped into the open ocean to trigger algal blooms; genetic modification of crops to increase biotic carbon uptake; carbon capture and storage techniques such as those proposed to outfit coal plants; and planting forests are such examples. Other schemes involve blocking or reflecting incoming solar radiation, for example by spraying seawater hundreds of meters into the air to seed the formation of stratocumulus clouds over the subtropical ocean.[3]

Two strategies to reduce incoming solar radiation — stratospheric aerosol injection as proposed by Crutzen and space-based sun shields (i.e., mirrors or shades placed in orbit between the sun and Earth) — are among the most widely discussed geoengineering schemes in scientific circles. While these schemes (if they could be built) would cool Earth, they might also have adverse consequences. Several papers in the August 2006 *Climatic Change* discussed some of these issues, but here I present a fairly comprehensive list of reasons why geoengineering might be a bad idea, first written down during a two-day NASA-sponsored conference on Managing Solar Radiation (a rather audacious title) in November 2006.[4] These concerns address unknowns in climate system response; effects on human quality of life; and the political, ethical, and moral issues raised.

1. Effects on regional climate. Geoengineering proponents often suggest that volcanic eruptions are an innocuous natural analog for stratospheric injection of sulfate aerosols. The 1991 eruption of Mount Pinatubo on the Philippine island of Luzon, which injected 20 megatons of sulfur dioxide gas into the stratosphere, produced a sulfate aerosol cloud that is said to have caused global cooling for a couple of years without adverse effects. However, researchers at the National Center for Atmospheric Research showed in 2007 that the Pinatubo eruption caused large hydrological responses, including reduced precipitation, soil moisture, and river flow in many regions.[5] Simulations of the climate response to volcanic eruptions have also shown large impacts on regional climate, but whether these are good analogs for the geoengineering response requires further investigation.

Scientists have also seen volcanic eruptions in the tropics produce changes in atmospheric circulation, causing winter warming over continents in the Northern Hemisphere, as well as eruptions at high latitudes weaken the Asian and African monsoons, causing reduced precipitation.[6] In fact, the eight-month-long eruption of the Laki fissure in Iceland in 1783–1784 contributed to famine in Africa, India, and Japan.

If scientists and engineers were able to inject smaller amounts of stratospheric aerosols than result from volcanic eruptions, how would they affect summer wind and precipitation patterns? Could attempts to geoengineer isolated regions (say, the Arctic) be confined there? Scientists need to investigate these scenarios. At the fall 2007 American Geophysical Union meeting, researchers presented preliminary findings from several different climate models that simulated geoengineering schemes and found that they reduced precipitation over wide regions, condemning hundreds of millions of people to drought.

2. Continued ocean acidification. If humans adopted geoengineering as a solution to global warming, with no restriction on continued carbon emissions, the ocean would continue to become more acidic, because about half of all excess carbon dioxide in the atmosphere is removed by ocean uptake. The ocean is already 30 percent more acidic than it was before the Industrial Revolution, and continued acidification threatens the entire oceanic biological chain, from coral reefs right up to humans.[7]

3. Ozone depletion. Aerosol particles in the stratosphere serve as surfaces for chemical reactions that destroy ozone in the same way that water and nitric acid aerosols in polar stratospheric clouds produce the seasonal Antarctic ozone hole.[8] For the next four decades or so, when the concentration of anthropogenic ozone-depleting substances will still be large enough in the stratosphere to

Box 15.1 Capitalizing on Carbon

Without market incentives, geoengineering schemes to reflect solar heat are still largely confined to creative thought and artists' renderings. But a few ambitious entrepreneurs have begun to experiment with privatizing climate mitigation through carbon sequestration. Here are a few companies in the market to offset your carbon footprint:

California-based technology startups Planktos and Climos are perhaps the most prominent groups offering to sell carbon offsets in exchange for performing ocean iron fertilization, which induces blooms of carbon-eating phytoplankton. Funding for Planktos dried up in early 2008 as scientists grew increasingly skeptical about the technique, but Climos has managed to press on, securing $3.5 million in funding from Braemar Energy Ventures as of February.

Also in the research and development phase is Sydney, Australia-based Ocean Nourishment Corporation, which similarly aims to induce oceanic photosynthesis, only it fertilizes with nitrogen-rich urea instead of iron, Atmocean, based in Santa Fe, New Mexico, takes a slightly different tack: It's developed a 200-meter deep, wave-powered pump that brings colder, more biota-rich water up to the surface where lifeforms such as tiny, tube-like salps sequester carbon as they feed on algae.

Related in mission if not in name, stationary carbon-capture technologies, which generally aren't considered geoengineering, are nonetheless equally inventive: Skyonic, a Texas-based startup, captures carbon dioxide at power plants (a relatively well-proven technology) and mixes it with sodium hydroxide to render high-grade baking soda. A pilot version of the system is operating at the Brown Stream Electric Station in Fairfield, Texas. To the west in Tucson, Arizona, Global Research Technologies, the only company in the world dedicated to carbon capture from ambient air, recently demonstrated a working "air extraction" prototype — a kind of carbon dioxide vacuum that stands upright and is about the size of a phone booth. Meanwhile, GreenFuel Technologies Corporation, in collaboration with Arizona Public Service Company, is recycling carbon dioxide emissions from power plants by using it to grow biofuel stock in the form of — what else? — algae.

Kirsten Jerch

produce this effect, additional aerosols from geoengineering would destroy even more ozone and increase damaging ultraviolet flux to Earth's surface.

4. Effects on plants. Sunlight scatters as it passes through stratospheric aerosols, reducing direct solar radiation and increasing diffuse radiation, with important biological consequences. Some studies, including one that measured this effect in trees following the Mount Pinatubo eruption, suggest that diffuse radiation allows plant canopies to photosynthesize more efficiently, thus increasing their capacity as a carbon sink.[9] At the same time, inserting aerosols or reflective disks into the atmosphere would reduce the total sunlight to reach Earth's surface. Scientists need to assess the impacts on crops and natural vegetation of reductions in total, diffuse, and direct solar radiation.

5. More acid deposition. If sulfate is injected regularly into the stratosphere, no matter where on Earth, acid deposition will increase as the material passes through the troposphere — the atmospheric layer closest to Earth's surface. In 1977, Russian climatologist Mikhail Budyko calculated that the additional acidity caused by sulfate injections would be negligibly greater than levels that resulted from air pollution.[10] But the relevant quantity is the *total* amount of acid that reaches the ground, including both wet (acid rain, snow, and fog) and dry deposition (acidic gases and particles). Any additional acid deposition would harm the ecosystem, and it will be important to understand the consequences of exceeding different biological thresholds. Furthermore, more acidic particles in the troposphere would affect public health. The effect may not be large compared to the impact of pollution in urban areas, but in pristine areas it could be significant.

6. Effects of cirrus clouds. As aerosol particles injected into the stratosphere fall to Earth, they may seed cirrus cloud formations in the troposphere.[11] Cirrus clouds affect Earth's radiative balance of incoming and outgoing heat, although the amplitude and even direction of the effects are not well understood. While evidence exists that some volcanic aerosols form cirrus clouds, the global effect has not been quantified.[12]

7. Whitening of the sky (but nice sunsets). Atmospheric aerosols close to the size of the wavelength of light produce a white, cloudy appearance to the sky. They also contribute to colorful sunsets, similar to those that occur after volcanic eruptions. The red and yellow sky in *The Scream* by Edvard Munch was inspired by the brilliant sunsets he witnessed over Oslo in 1883, following the eruption of Krakatau in Indonesia.[13] Both the disappearance of blue skies and the appearance of red sunsets could have strong psychological impacts on humanity.

8. Less sun for solar power. Scientists estimate that as little as a 1.8 percent reduction in incoming solar radiation would compensate for a doubling of atmospheric carbon dioxide. Even this small reduction would significantly affect the radiation available for solar power systems — one of the prime alternate methods of generating clean energy — as the response of different solar power systems to total available sunlight is not linear. This is especially true for some of the most efficiently designed systems that reflect or focus direct solar radiation on one location for direct heating.[14] Following the Mount Pinatubo eruption and the 1982 eruption of El Chichón in Mexico, scientists observed a direct solar radiation decrease of 25–35 percent.[15]

9. Environmental impacts of implementation. Any system that could inject aerosols into the stratosphere, i.e., commercial jetliners with sulfur mixed into their fuel, 16-inch naval rifles firing 1-ton shells of dust vertically into the air, or hoses suspended from stratospheric balloons, would cause enormous environmental damage. The same could be said for systems that would deploy sun shields. University of Arizona astronomer Roger P. Angel has proposed putting a fleet of 2-foot-wide reflective disks in a stable orbit between Earth and the sun that would bend sunlight away from Earth.[16] But to get the needed *trillions* of disks into space, engineers would need 20 electromagnetic launchers to fire missiles with stacks of 800,000 disks every five minutes for twenty years. What would be the atmospheric effects of the resulting sound and gravity waves? Who would want to live nearby?

10. Rapid warming if deployment stops. A technological, societal, or political crisis could halt a project of stratospheric aerosol injection in mid-deployment.

Such an abrupt shift would result in rapid climate warming, which would produce much more stress on society and ecosystems than gradual global warming.[17]

11. There's no going back. We don't know how quickly scientists and engineers could shut down a geoengineering system — or stem its effects — in the event of excessive climate cooling from large volcanic eruptions or other causes. Once we put aerosols into the atmosphere, we cannot remove them.

12. Human error. Complex mechanical systems never work perfectly. Humans can make mistakes in the design, manufacturing, and operation of such systems. (Think of Chernobyl, the *Exxon Valdez*, airplane crashes, and friendly fire on the battlefield.) Should we stake the future of Earth on a much more complicated arrangement than these, built by the lowest bidder?

13. Undermining emissions mitigation. If humans perceive an easy technological fix to global warming that allows for "business as usual," gathering the national (particularly in the United States and China) and international will to change consumption patterns and energy infrastructure will be even more difficult.[18] This is the oldest and most persistent argument against geoengineering.

14. Cost. Advocates casually claim that it would not be too expensive to implement geoengineering solutions, but there have been no definitive cost studies, and estimates of large-scale government projects are almost always too low. (Boston's "Big Dig" to reroute an interstate highway under the coastal city, one of humankind's greatest engineering feats, is only one example that was years overdue and billions over budget.) Angel estimates that his scheme to launch reflective disks into orbit would cost "a few trillion dollars." British economist Nicholas Stern's calculation of the cost of climate change as a percentage of global GDP (roughly $9 trillion) is in the same ballpark; Angel's estimate is also orders of magnitude greater than current global investment in renewable energy technology. Wouldn't it be a safer and wiser investment for society to instead put that money in solar power, wind power, energy efficiency, and carbon sequestration?

15. Commercial control of technology. Who would end up controlling geoengineering systems? Governments? Private companies holding patents on proprietary technology? And whose benefit would they have at heart? These systems could pose issues analogous to those raised by pharmaceutical companies and energy conglomerates whose products ostensibly serve the public, but who often value shareholder profits over the public good.

16. Military use of the technology. The United States has a long history of trying to modify weather for military purposes, including inducing rain during the Vietnam War to swamp North Vietnamese supply lines and disrupt antiwar protests by Buddhist monks.[19] Eighty-five countries, including the United States, have signed the U.N. Convention on the Prohibition of Military or Any Other Hostile Use of Environmental Modification Techniques (ENMOD), but could techniques developed to control global climate forever be limited to peaceful uses?

17. Conflicts with current treaties. The terms of ENMOD explicitly prohibit "military or any other hostile use of environmental modification techniques having widespread, long-lasting or severe effects as the means of destruction, damage, or injury to any other State Party." Any geoengineering scheme that adversely affects regional climate, for example, producing warming or drought, would therefore violate ENMOD.

18. Control of the thermostat. Even if scientists could predict the behavior and environmental effects of a given geoengineering project, and political leaders could muster the public support and funding to implement it, how would the world agree on the optimal climate? What if Russia wants it a couple of degrees warmer, and India a couple of degrees cooler? Should global climate be reset to preindustrial temperature or kept constant at today's reading? Would it be possible to tailor the climate of each region of the planet independently without affecting the others? If we proceed with geoengineering, will we provoke future climate wars?

19. Questions of moral authority. Ongoing global warming is the result of inadvertent climate modification. Humans emit carbon dioxide and other greenhouse gases to heat and cool their homes; to grow, transport, and cook their food; to run their factories; and to travel — not intentionally, but as a byproduct of fossil fuel combustion. But now that humans are aware of their effect on climate, do they have a moral right to continue emitting greenhouse gases? Similarly, since scientists know that stratospheric aerosol injection, for example, might impact the ecosphere, do humans have a right to plow ahead regardless? There's no global agency to require an environmental impact statement for geoengineering. So, how should humans judge how much climate control they may try?

20. Unexpected consequences. Scientists cannot possibly account for all of the complex climate interactions or predict all of the impacts of geoengineering. Climate models are improving, but scientists are discovering that climate is changing more rapidly than they predicted, for example, the surprising and unprecedented extent to which Arctic sea ice melted during the summer of 2007. Scientists may never have enough confidence that their theories will predict how well geoengineering systems can work. With so much at stake, there is reason to worry about what we don't know.

The reasons why geoengineering may be a bad idea are manifold, though a moderate investment in *theoretical* geoengineering research might help scientists to determine whether or not it *is* a bad idea. Still, it's a slippery slope: I wouldn't advocate actual small-scale stratospheric experiments unless comprehensive climate modeling results could first show that we could avoid at least all of the potential consequences we *know* about. Due to the inherent natural variability of the climate system, this task is not trivial. After that there are still the unknowns, such as the long-term effects of short-term experiments — stratospheric aerosols have an atmospheric lifetime of a couple years.

Solving global warming is not a difficult technical problem. As Stephen Pacala and Robert Socolow detail with their popular wedge model, a combination of several specific actions can stabilize the world's greenhouse gas emissions — although I disagree with their proposal to use nuclear power as one of their "wedges."[20] Instead, the crux of addressing global warming is political. The U.S. government gives multibillion-dollar subsidies to the coal, oil, gas, and nuclear industries, and gives little support to alternative energy sources like solar and

Box 15.2 An Ethical Assessment of Geoengineering

While there are many questions about the feasibility, cost, and effectiveness of geoengineering plans, my colleague Alan Robock has been the most systematic and persistent of a number of scientists in raising ethical quandaries about the enterprise. But just how serious are these ethical quandaries?

Most science poses risks of unintended consequences, and lots of science raises issues of commercial and military control. At issue here is whether there is any reason to believe *ex ante* that these are special or unusually large risks. Merely asserting them does not ground an objection *per se*.

Not all of Robock's concerns involve ethics, but of those that do, some involve issues of procedural justice (such as who decides) while others involve matters of distributive justice (such as uneven benefit and harm). To simplify things, let's assume that injecting aerosols into the stratosphere successfully cooled Earth without any untoward effects and with evenly distributed benefits. One might still object that there are issues of procedural justice involved — who decides and who controls. But such concerns don't get much traction when everyone benefits.

Let's pull back from this idealization to imagine an outcome that involves untoward consequences and an uneven distribution of benefits. We deal with consequences by balancing them against the benefits of our interventions. The issue is whether or not we can obtain reliable estimates of both risks and benefits without full-scale implementation of the planned intervention. We already know from modeling that the impact of any such intervention will be uneven, but again, without knowing what the distribution of benefit and harm would be, it's hard to estimate how much this matters. Let's differentiate two circumstances under which going ahead with the intervention might be judged: One is where everyone benefits, while the other is a circumstance in which something less is the case. A conservative conclusion would be to say that beyond modeling and controlled, low-level tests (if the modeling justifies it), we shouldn't sanction any large-scale interventions unless they are in everyone's interest. A slightly eased condition, proposed by the philosopher Dale Jamieson, would be that at least nobody is worse off. That may not be as farfetched a condition as one might think, since, in the end, we are considering this intervention as a means to balance a risk we all face — global warming.

But suppose there are isolated livelihoods that only suffer negative effects of geoengineering. Then numbers begin to matter. In the case that a geoengineering scheme were to harm the few, we should have the foresight to be able to compensate, even if doing so requires something as drastic as relocating populations. I don't mean to oversimplify a complicated issue, but objection to any negative consequences whatsoever isn't a strong enough argument to end discussion.

More trenchant is the worry that the mere possibility of geoengineering would undermine other efforts to decrease our carbon output. Such moral hazard is a familiar worry, and we don't let it stop us in other areas: Antilock braking systems and airbags may cause some to drive more recklessly, but few would let that argument outweigh the overwhelming benefits of such safety features.

As Robock correctly asserts, the crux of addressing global warming may be a political — not a scientific — problem, but it doesn't follow that we may not need geoengineering to solve it. If it is a political problem, it is a *global* political problem, and getting global agreement to curb greenhouse gases is easier said than done.

Box 15.2 An Ethical Assessment of Geoengineering *(continued)*

With geoengineering, in principle, one nation or agent could act, but a challenge arises if the intervention is certain to have uneven impacts among nations. At this early stage, there is no cost associated with improving our ability to quantify and describe what those inequalities would look like. Once we have those answers in hand, then we can engage in serious ethical consideration over whether or not to act.

Martin Bunzl

wind power that could contribute to a solution. Similarly, the federal government is squashing attempts by states to mandate emissions reductions. If global warming is a political problem more than it is a technical problem, it follows that we don't need geoengineering to solve it.

The U.N. Framework Convention on Climate Change defines "dangerous anthropogenic interference" as *inadvertent* climate effects. However, states must also carefully consider geoengineering in their pledge to prevent dangerous anthropogenic interference with the climate system.

ENDNOTES

1. Paul Crutzen, "Albedo Enhancement by Stratospheric Sulfur Injections: A Contribution to Solve a Policy Dilemma?" *Climatic Change*, vol. 77, pp. 211–19 (2006); Tom M. L. Wigley, "A Combined Mitigation/Geoengineering Approach to Climate Stabilization," *Science*, vol. 314, pp. 452–54 (2006).
2. See the chapter on climate modification schemes in Spencer R. Weart, *The Discovery of Global Warming* (2007), available at http://www.aip.org/history/climate/RainMake.htm; a long history of geoengineering proposals in James R. Fleming, "Fixing the Weather and Climate: Military and Civilian Schemes for Cloud Seeding and Climate Engineering," in Lisa Rosner, ed., *The Technological Fix* (New York: Routledge, 2004), pp. 175–200; and James R. Fleming, "The Pathological History of Weather and Climate Modification," *Historical Studies in the Physical Sciences*, vol. 37, pp. 3–25 (2006). See also N. Rusin and L. Flit, *Man Versus Climate* (Moscow: Peace Publishers, 1960); Mikhail I. Budyko, *Climatic Changes* (Washington, D.C.: American Geophysical Union, 1977); Ralph J. Cicerone et al., "Global Environmental Engineering," *Nature*, vol. 356, p. 472 (1992); Edward Teller et al., *Global Warming and Ice Ages: I. Prospects for Physics-Based Modulation of Global Change* (Lawrence Livermore National Laboratory Publication UCRL-JC-128715, 1997); David W. Keith, "Geoengineering the Climate: History and Prospect," *Annual Review of Energy and the Environment*, vol. 25, pp. 245–84 (2000).
3. John Latham first raised this idea in two articles that appeared in *Nature*, vol. 347, no. 6291: "Control of Global Warming," pp. 330–40, and "Effect on Global Warming of Wind-Dependent Aerosol Generation at the Ocean Surface," pp. 372–73 (1990). Keith Bower offers a numerical evaluation in "Computational Assessment of a Proposed Technique for Global Warming Mitigation Via Albedo-Enhancement of Marine Stratocumulous Clouds," *Atmospheric Research*, vol. 82, pp. 328–36 (2006).

4. See Lee Lane, Ken Caldeira, Robert Chatfield, and Stephanie Langhoff, eds., "Workshop Report on Managing Solar Radiation," NASA/CP-2007–214558 (2007).
5. Kevin E. Trenberth and Aiguo Dai, "Effects of Mount Pinatubo Volcanic Eruption on the Hydrological Cycle as an Analog of Geoengineering," *Geophysical Research Letters*, vol. 34, no. 16 (2007).
6. For more on warming over continents of the Northern Hemisphere, see Alan Robock, "Volcanic Eruptions and Climate," *Reviews of Geophysics*, vol. 38, pp. 191–219 (2000): Georgiy Stenchikov et al., "Arctic Oscillation Response to Volcanic Eruptions in the IPCC AR4 Climate Models," *Journal of Geophysical Research*, vol. 111 (2006). For more on the effects of Asian and African monsoons, see Luke Oman et al., "Climatic Response to High-Latitude Volcanic Eruptions," *Journal of Geophysical Research*, vol. 110 (2005); Luke Oman et al., "High-Latitude Eruptions Cast Shadow Over the African Monsoon and the Flow of the Nile," *Geophysical Research Letters*, vol. 33 (2006).
7. Royal Society, *Ocean Acidification Due to Increasing Atmospheric Carbon Dioxide*, June 30, 2005, available at royalsociety.org/displaypagedoc.asp?id=13539
8. Susan Solomon et al., "The Role of Aerosol Variations in Anthropogenic Ozone Depletion at Northern Midlatitudes," *Journal of Geophysical Research*, vol. 101 (1996); Susan Solomon. "Stratospheric Ozone Depletion: A Review of Concepts and History," *Reviews of Geophysics*, vol. 37 (1999).
9. L. Gu et al., "Responses of Net Ecosystem Exchanges of Carbon Dioxide to Changes in Cloudiness: Results from Two North American Deciduous Forests," *Journal of Geophysical Research*, vol. 104, no. 31, pp. 421–31, 434 (1999); L. Gu et al., "Advantages of Diffuse Radiation for Terrestrial Ecosystem Productivity," *Journal of Geophysical Research*, vol. 107 (2002); L. Gu et al., "Response of a Deciduous Forest to the Mount Pinatubo Eruption: Enhanced Photosynthesis," *Science*, vol. 299, pp. 2,035–38 (2003).
10. Budyko, *Climatic Changes*.
11. Richard P. Turco et al., "A Study of Mesospheric Rocket Contrails and Clouds Produced by Liquid-Fueled Rockets," *Space Solar Power Review*, vol. 3, pp. 223–34 (1982); V. A. Mohnen, "Stratospheric Ion and Aerosol Chemistry and Possible Links With Cirrus Cloud Microphysics — A Critical Assessment," *Journal of Atmospheric Science*, vol. 47, pp. 1,933–48 (1990).
12. K. Sassen et al., "The 5–6 December 1991 FIRE IFO II Jet Stream Cirrus Case Study: Possible Influences of Volcanic Aerosols," *Journal of Atmospheric Science*, vol. 52, pp. 97–123 (1993).
13. D. W. Olsen et al., "When the Sky Ran Red: The Story Behind *The Scream*," *Sky & Telescope*, February 2004, pp. 29–35.
14. For the estimate for reducing incoming solar radiation, see Balan Govindasamy and Ken Caldeira, "Geoengineering Earth's Radiation Balance to Mitigate CO_2-Induced Climate Change," *Geophysical Research Letters*, vol. 27, pp. 2,141–44 (2000). For the response of solar power systems, see Michael C. MacCracken, "Geoengineering: Worthy of Cautious Evaluation?" *Climatic Change*, vol. 77, pp. 235–43 (2006).
15. Robock, "Volcanic Eruptions and Climate," pp. 191–219.

16. Roger P. Angel, "Feasibility of Cooling the Earth with a Cloud of Small Spacecraft Near the Inner Lagrange Point (L_1)," *Proceedings of the National Academy of Sciences*, vol. 103, pp. 17,184–89 (2006).
17. See Figure 1 in Wigley, "A Combined Mitigation/Geoengineering Approach to Climate Stabilization," pp. 452–54, and Figure 3 in H. Damon Matthews and Ken Caldeira, "Transient Climate-Carbon Simulations of Planetary Geoengineering," *Proceedings of the National Academy of Sciences*, vol. 104, pp. 9,949–54 (2007).
18. See for example Stephen H. Schneider, "Earth Systems: Engineering and Management," *Nature*, vol. 409, pp. 417–19, 421 (2001), and Ralph J. Cicerone, "Geoengineering: Encouraging Research and Overseeing Implementation," *Climatic Change*, vol. 77, pp. 221–26 (2006).
19. James R. Fleming writes eloquently about the militaristic history of climate modification schemes in "The Climate Engineers," *Wilson Quarterly*, Spring 2007, pp. 46–60. See also Fleming, "Fixing the Weather and Climate," and Fleming, "The Pathological History of Weather and Climate Modification."
20. Stephen W. Pacala and Robert Socolow, "Stabilization Wedges: Solving the Climate Problem for the Next 50 Years with Current Technologies," *Science*, vol. 305, pp. 968–72 (2004); Alan Robock, "Nuclear Power's Costs and Perils" (Letter to the Editor), *Physics Today*, vol. 60, no. 1, p. 14 (2007).

PART V

Contemporary Technological Dilemmas: The New Biology

The twentieth century is often regarded as the century of physics. Discoveries in physics including quantum theory, general and special relativity, the basic constituents of the atom, the nature of subatomic forces, the photon and the nature of light, and the like led to the incredible technological progress that took place during that century: computers, wireless communication, lasers, television and so many other technologies that are now part of our daily lives.

The twenty-first century, scientists, technologists, and futurists agree, will most likely be known as the century of biology.

The explosion of knowledge began with the discovery of the structure of DNA in the mid-twentieth century. The understanding of the double helix led to the growing ability to manipulate our genes and those of other organisms. Research in biology now depends on computers and information technology, modeling, imaging technologies, materials science — technologies that are transforming the life sciences and in turn laying the foundation for enormous progress in the technologies of medicine, in understanding of the human brain, as well as in such fields as energy, agriculture, and environmental science.

Part V is devoted to the nature and implications of this new biology. It opens with a discussion of perhaps the most morally and politically divisive area that has entered the science, technology and policy arena in recent times — human embryonic stem cell research. Christopher Thomas Scott provides a balanced introduction to the possibilities offered by research in this contentious area as well as the arguments of its advocates and its opponents. While stem cell research has been debated for well over a decade, synthetic biology — the subject of the next two chapters — is relatively new, at least outside the realm of fiction.

The ability to create a self-replicating genome from laboratory chemicals, demonstrated for the first time in 2010, has stimulated widespread debate about the potential and limits of synthetic biology. Articles from the Presidential Commission for the Study of Bioethical Issues and from the journal *Nature* are included here to help frame the emerging issues in this field. Next, Michael Sandel, a political philosopher at Harvard who teaches one of the most popular courses offered at that institution, argues "The Case against Perfection," exploring the moral issues raised by the technologies of human genetic enhancement in such realms as physical strength, memory, height, and sex selection of babies. Finally, Stanford law professor Henry Greely looks at the profound and difficult issues raised by current and potential future developments in neuroscience, one of the fastest growing areas of the new biology. His discussion intersects with Sandel's in some areas, such as the enhancement of the brain's capabilities, but it also goes into such areas as lie detection in criminal justice and even detection of racial or ethnic bias. In all, these chapters suggest the exciting possibilities — and the complex dilemmas — that the century of biology has in store.

16

Stem Cell Research: The Great Moral Divide

CHRISTOPHER THOMAS SCOTT

The debate over use of human embryonic stem cells in medical research and, ultimately, in therapy has been raging in the United States for over a decade, since James Thomson at the University of Wisconsin, Madison, became the first person to isolate and culture this unique type of cell. Under proper conditions, these cells are capable of developing into any of the more than 200 cell types in the human body. In other words, if a number of difficult hurdles are overcome through research, such cells open up possibilities for what is being called "regenerative medicine" or tissue replacement, potentially leading to cures for Parkinson's disease, diabetes, spinal cord injuries, and many more causes of human suffering. The catch is that the process of creating a stem cell line (a live culture maintained artificially after being separated from its original source) destroys a fertilized human egg — a blastocyst or pre-embryo — that, if implanted in a woman's uterus, could grow into a person. Balancing these two considerations is at the root of the stem cell controversy. Some countries have policies encouraging stem cell research; others have banned it or declared moratoriums.

In the United States, stem cell research is one of the most politicized areas of science. While stem cell research conducted without government funding is legal almost everywhere in the country, federal funding remains in dispute. President George W. Bush issued an executive order in August 2001 limiting federal funding to research using stem cell lines already in existence on the date and time of the speech announcing the policy. Early in his presidency, in March 2009, President Barack Obama fulfilled a campaign promise and pleased scientists and patient advocates by overturning the Bush policy. Under the Obama

Administration, the National Institutes of Health initiated a modest program of research grants. About a year and a half later, in August 2010, a federal judge, in response to a suit brought by two scientists and a group called the Alliance Defense Fund, issued a temporary injunction stopping the stem cell research program in its tracks. The judge ruled that the Obama policy violated a law banning "research in which a human embryo or embryos are destroyed, discarded or knowingly subjected to risk of injury or death." The ruling caught the biomedical research community, which had thought the policy was settled by President Obama's executive order, by surprise. Although an appeals court suspended implementation of the injunction two weeks later, allowing federally funded embryonic stem cell research to continue, as of this writing the matter remains in dispute and may well end up in the Supreme Court. In the following article, Christopher Thomas Scott looks at the ethical issues underlying the stem cell debate, quoting sources from all sides of this complex issue.

Scott is a senior research scholar in the Center for Biomedical Ethics at the Stanford University School of Medicine. He is director of the Program on Stem Cells in Society and has served as a member of the Stem Cell Research Advisory Panel at the medical school. He holds degrees in biological sciences as well as policy and has a wide range of publications, including Stem Cell Now: From the Experiment That Shook the World to the New Politics of Life, *from which this article was taken.*

Nobody, I imagine, will credit me with a desire to limit the
empire of physical science, but I really feel bound to
confess that a great many very familiar and, at the
same time, extremely important phenomena lie
quite beyond its legitimate limits.

THOMAS HUXLEY[1]

The newly fertilized egg, a corpuscle one two-hundredth of an inch in diameter, is not a human being. It is a set of instructions set adrift into the cavity of the womb, wrote E. O. Wilson, a Pulitzer prize–winning zoologist.[2] But where E. O. Wilson seems unperturbed, Charles Krauthammer, M.D., a syndicated columnist and member of the President's Council on Bioethics (PCBE), a 17-member group charged [in the administration of George W. Bush] with advising the president on the ethics of biological innovation, is deeply concerned. "We will, slowly and by increments, have gone from stem cells to embryo farms to factories with fetuses hanging (metaphorically) on meat hooks waiting to be cut open and used by the already born."[3]

The two quotes reflect opposite poles of an emotionally charged debate about the use of human embryonic stem cells (hESCs) for research and medicine. The rhetoric often reveals from which side of the debate the person speaks. Supporters of hESC research sometimes describe the human embryo as nothing more than a ball of cells; opponents mention the embryo as if it were a newborn baby. Others portray the blastocyst as something in between: neither a clump of cells nor a person. All who consider the question of using embryos for research share one thing: passion. Embryonic stem cells touch us deeply, not just because they might cure disease. It is because forms of human life are at stake — living embryos and living persons.

Many ethical questions touch hESC research. What are our moral obligations to the sick among us who could benefit from embryonic stem cell research? What roles do family, religion, and society play, and how do they inform our opinions and decisions? How do we properly explain the benefits and risks of using embryos and cells to the individuals and parents who donate them? Who should benefit from the first therapies, and how will we pay for them? Tackling these questions could fill an entire book. But the ethical question on most people's minds (including people in the federal government) is whether a four- to six-day-old human embryo — the blastocyst — should be used to make lines of hESCs. Adult stem cells taken from a consenting person's body don't raise the same ethical concerns. Adult stem cells made from embryonic stem cells do.

A much more incendiary issue is whether somatic cell nuclear transfer (SCNT) should be used to make cloned humans for medical or other purposes. Here, the ethical positions are unanimous: both sides of the debate overwhelmingly condemn human reproductive cloning. Making a cloned human would require surmounting all the difficulties inherent in nuclear transfer, many of which may never be fully understood. Attempting to overcome these problems in order to achieve a normal birth or to use embryos to make fetuses or human beings for spare parts would amount to the worst forms of illegal and unethical human experimentation. Our moral codes are designed to protect people in our pursuit of knowledge, and humans or fetuses cloned for research purposes fall squarely and unambiguously into this category. Despite the universal agreement on prohibiting human cloning, some, like Charles Krauthammer, worry that if we use embryos for research, unscrupulous persons will use them to create more developed forms of life, including humans. Because of this, they argue, it is best to err on the side of safety and not use embryos at all.

PHILOSOPHY MEETS THE NEW BIOLOGY

Philosophers, theologians, and entire societies are wrestling with these questions, so much so that a new field, bioethics, has emerged to study the advances of medicine and technology and their impact on humankind. Bioethics is a child of the field of medical ethics, the qualities embodied by practitioners of medicine, who are enjoined to heal and not to harm; to (if asked) refuse to extinguish life; and to strictly keep the confidence of their patients. But, 40 years ago, the "new biology" entered the arena and, with it, an avalanche of discovery. Scientists could insert foreign genes into living things and transplant human organs. No longer do we think only about the pros and cons of traditional medical care; we now face the ability of science to change our human future — and change it radically.

Bioethics is still a young field, and the rapid pace of stem cell biology is just one area challenging ethics to stay in step. In academia, dozens of programs in bioethics, most within medical schools, grapple with a wide range of rapidly evolving issues. One of the early thinkers in the field, Albert Jonsen, a specialist in medical ethics and former presidential advisor, writes, "Only half of bioethics counts as an ordinary academic discipline: the half that has original and borrowed theory, principles, and methods. But only part of bioethics lies within the academy, where scholars worry about whether they have a discipline to teach and promote. The other half of bioethics is the public discourse: people of all sorts and professions talking and arguing about bioethical questions."[4]

Jonsen is right about public discourse. Our moral universe is changing and challenging us, impinged upon and steered by ethics debates. Over time classic ethics endeavored to describe our world and the best prescriptions for living in it. The results often translate into laws and public policy — the voice of governments and institutions. When it comes to stem cells, the positions taken by scholars and philosophers are strewn across the map. It is worth exploring the views of a handful of prominent bioethicists — including members of President George W. Bush's ethics council — to see how the discourse on stem cells is shaping up.

THE MORAL STATUS OF THE EMBRYO

Of all the moral dimensions of stem cell research, the human embryo looms largest. Recall that viable embryos may be created two ways: naturally, through the process of conception, or artificially, through either nuclear transfer or *in vitro* fertilization (IVF). In the latter case, removing the inner cell mass (ICM) in order to produce stem cells destroys the embryo. The cells may survive indefinitely, but the embryo is gone. Fertilized eggs kept in a freezer or blastocysts grown in a lab will never develop into a fetus or human because the outer layer of cells, the trophoblast, is missing. Normal development and a live birth require a viable embryo's successful implantation into the uterus, where it will grow to term.

What should we do with leftover embryos from IVF procedures? Are we justified in using them for the purposes of research? Should we create embryos using nuclear transfer so that we may use them for therapies? Do we treat embryos the same as a newborn or a child or even as an adult? The moral status of an embryo — whether we consider it a mere object, a human being, or somewhere in between — is the subject of thousands of pages of opinion, essays, and research. A rough sorting divides the issue into two camps. One group believes that embryos deserve protection and should not be used for research. The other group believes that embryos can be used and embryonic stem cell research should proceed. The divining rod prompting each stance is the moral weight given to the embryo.

[An early stage embryo (known as a blastocyst) is between 0.1 and 0.2 mm in diameter.] If we place a culture dish under a microscope and peer at the embryo at this stage, should we protect it from harm? If not, should it be used to save others from disability, disease, and death? The debate pivots around these two questions.

PROTECT THE EMBRYO

The answers depend, in part, on whether we believe the embryo has a soul or is a person; in essence, whether it is a human being. If so, then we should protect it. [Former] presidential council members Robert P. George and Alfonso Gómez-Lobo — the first a political scientist and the second a philosopher — are clear about where they stand. They write, "A human embryo is a whole living member of the species *Homo sapiens* in the earliest stage of his or her natural development.... The human embryo, then, is a whole (though immature) and distinct human organism — a human being."[5] Their position is based on the fact that human development proceeds along a continuum. Each of us was once a one-celled zygote, embryo, fetus, and an adult. To George and Gómez-Lobo, each stage represents a different version of the same person. If the zygote has the "natural capacity" to become a person, they argue, then it's an either/or matter: "a 'thing' either is or is not a whole human being."

Conservative Christian thinkers agree. Gilbert Meilaender, a Lutheran theologian and member of the president's council, considers the embryo "the weakest and least advantaged of our fellow human beings."[6] Meilaender believes that the embryo's mere presence is enough to make it a person, though we may be barely aware of it during the early stages of a pregnancy or when it is stored for use at a reproductive clinic. He warns that some of his fellow philosophers define the concept of "person" too narrowly, "trying to find some capacity — perhaps self-awareness, reasoning power, or sense of oneself as having a history — that marks the point at which human beings become persons (or cease to be persons)."[7] The Valparaiso theologian might disagree with the English philosopher John Locke (1632–1704), who described a person as "a thinking, intelligent being, that has reason and reflection, that can consider itself as itself, the same thinking thing, in different times and places."[8] The traits Locke describes can be broadly interpreted: humans who have lost or never acquired these abilities may not be persons; animals with these traits could be construed as persons.

"I think it is accurate to state that a person is simply a 'someone who' — a someone who has a history," Meilaender says. "That history begins before we have some of the highest human capacities and, for many of us, it may continue after we have lost those capacities. But we remain the same person throughout the trajectory of our life." Meilaender's concept of personal history — and *personhood* — thus begins before we are conscious of it and continues after we have lost consciousness. Acknowledging this history gives the embryo the same amount of respect that we would give a fully conscious adult human being.

Some philosophers use a different rationale to argue for the protection of embryos. They say that our uncertainties about embryos are enough to give us pause. Although council advisors George and Gómez-Lobo characterize human development as a seamless series of the "same person," those who aren't sure about that claim say that it is impossible to pinpoint when the metaphysical characteristics of soul and personhood begin. Perhaps the blastocyst has a scintilla of personhood at four days — how can we really know? If an embryo is *possibly* a person, they contend, then it follows that destroying an embryo could possibly mean killing a person. Given this doubt, we should leave it alone. The Christian ethicist Robert Song believes that using embryos for research and medicine requires a "burden of proof " that they are *not* persons. "The merely possible personhood of the embryo may *seem* abstract or theoretical in comparison with the concrete hopes for clinical treatments." But it is proof enough, he says. "[F]or all one knows, they are persons, and should be treated as persons."[9]

A similar argument rests on the fact that a one-celled embryo *will* become a person, circumstances permitting. It uses a concept rich with meaning in stem cell biology: potential. A fertilized egg has the potential to become a fully realized person and, when an embryo is destroyed, its potential to become human is destroyed, too. When something dies naturally, as in the case of a fertilized egg that never implants in the uterus, or intentionally, when a less-than-perfect zygote is discarded after an IVF procedure, the occurrence causes us to reflect and wonder what that embryo could have been. President George W. Bush's former chief advisor on bioethics, Leon Kass, puts it this way. "[I] must acknowledge that the human blastocyst is (1) human in origin and (2) *potentially* a mature human being, if all goes well."[10]

The official positions of conservative religions consider the embryo as a human being. The Southern Baptist Convention, noted for their long-standing position on abortion, strenuously opposes human embryo research and calls upon research centers to "cease and desist from research which destroys human embryos, the most vulnerable members of the human community."[11] Both the Orthodox Christian and Methodist churches assert that embryo research is a fundamental violation of human life. The Anglican view, prepared as a brief for the House of Lords in 2002, elevates the moral status of the embryo as sacred, containing "the very beginning of each human being."[12]

The Catholic Church declares that God bestows personhood and a soul at the moment of conception. The official position of the Vatican, however, has shifted over the years. Saint Thomas Aquinas (1225–1274) believed that a human soul was not present at conception, but appeared between 40 and 90 days later.[13] In 1974, the Vatican wrote in its *Declaration on Procured Abortion*, "Respect for human life is called for from the time that the process of generation begins. From the time that the ovum is fertilized, a life is begun which is neither that of the father nor of the mother, it is rather the life of a new human being with his own growth."[14] The church went on to define an embryo as a person in 1987. The instruction, called the *Donum Vitae* (The Gift of Life), explains, "From the time that the ovum is fertilized … [it] demands the unconditional respect that is morally due to the human being in his bodily and spiritual totality … [and] his rights as a person must be recognized, among which in the first place is the inviolable right of every innocent human being to life."[15]

For the Vatican, the unique combination of genes resulting from meiosis and fertilization is enough to trigger personhood. Using the *Donum Vitae* as its guide, the church condemns the use of embryos for research. In 2000, it wrote that removing the inner cell mass from an embryo "is a gravely immoral act, and consequently is gravely illicit."[16]

USE THE EMBRYO

On the other side of the moral divide are those who believe that embryos are not people and maintain that we can — indeed must — use them to help humankind. Not surprisingly, many scientists, patients, and doctors hold this view. Moderate and left-leaning philosophers and theologians tend to be in this camp and disagree — at times vehemently — with their conservative colleagues on the moral status of the embryo.

If polls are any indication of American religious sentiment, they show a disagreement between those who profess faith and those who institutionalize it. The fracture is particularly notable among mainstream American Protestants and Catholics. In a Harris Poll conducted in August 2004, 73 percent of Catholics voted in favor of embryonic stem cell research; only 11 percent were against it. The margin of Protestants in favor of research was even larger — 8 to 1. Even among the "evangelical" or "born-again" Christians so important to the conservative political right, only one in five voted against using embryos.[17]

Other exceptions to the "bright line" religious views include the American Presbyterian Church. Its resolution places the respect of sick persons above the respect due the embryo, affirming the use of stem cells "for research that may result in the restoring the health of those suffering from serious illness."[18] A growing number of centrist Catholic theologians don't consider the embryo before two weeks of age an individual human entity. Other moderate Catholics use an ethical middle ground called "proportionate reason," which tolerates a "lesser evil" to bring about a "greater good." Michael Mendiola, a Catholic and professor of Christian ethics in the Graduate Theological Union at Berkeley, California, suggests one version of this compromise: a requirement that researchers consistently seek to move beyond the use of human embryos while allowing their use in the near term.[19]

Lutheran Theologians Ted Peters and Gaymon Bennett use a bioethical principle of beneficence — a moral obligation to act for the benefit of others — to make a case for supporting embryonic stem cell research.[20] They recall the parable of the Good Samaritan, who pursues one goal: to heal a suffering stranger. The duo contrasts positive acts to help others against an instinct to protect the embryo, a position they summarize as "better safe than sorry." Bennett, an affable young religious scholar who studies at the Center for Theology and the Natural Sciences in Berkeley, California, doesn't fit the image of a gray-haired philosopher with a furrowed brow.* He says he wants a just world — where people suffer less.

---

* Bennett is now (2011) the Communications Coordinator for the Science and Religion Course Program at the Center for Theology and the Natural Sciences, at Berkeley.— *ed.*

"I think about how science can help me bring about that kind of world. But I also consider how science can get in the way of achieving that aim."[21] He notes that for most people, the stem cell debate comes down to three questions: What are our responsibilities for (1) people who are sick and injured, (2) the embryo's protection, and (3) the unseen consequences of the research? On question (3), Bennett says, "We shouldn't believe 'anything goes.' Ethics argues against doing science purely for science's sake. We need a philosophy that guides science and we must realize we won't always have a perfect match of scientific and ethical truth."

Even with these cautions in mind, Peters and Bennett say the potential benefits of medicine should come first. Opponents of hESC research maintain that we should uphold human dignity and, as a result, protect the embryo. The two theologians point out that we aren't certain about when an embryo or fetus becomes human. In the meantime, we are certain there is plenty of human suffering, and these people could benefit from discoveries made with embryonic stem cells. Because of this certainty, they argue, we should pursue hESC research.[22]

The Jewish faith's assessment of the early embryo's moral position stands in stark contrast to conservative Christian dogma. Laurie Zoloth is a Jewish philosopher and professor of medical ethics at Northwestern University. Jewish religious law — called *halachah* — uses a defense of justice in all of its deliberations about medicine and health-care. Zoloth describes the mandate *pikauch nefesh*, the duty to save a life and to heal. If healing is mandatory then embryonic stem cells *must* be used. She explains that with the exceptions of the ill-begotten use of murder, adultery, or idolatry, "If one can save a life, one must save a life; if one can heal, one must heal."[23]

Is breaking apart an embryo in the lab murder? *Halachah* says no, it is not considered murder, nor is it considered killing. Zoloth uses an example of Judaic case law to illustrate. "Suppose two men are fighting and as a consequence a pregnant woman is injured as an innocent bystander. If she has a miscarriage, we consider the harm has come to the woman — not to the potential child." The rabbinical text treats pregnancy as a continuous process, with the moral status of the embryo, fetus, and newborn gradually increasing along the way. Pregnancy — along with a significant moral imprint — isn't officially recognized until after 40 days. "The old texts describe the developmental period before 40 days like water," Zoloth says. "It makes practical sense if you think about it. If a pregnancy ended before 40 days, you wouldn't be able to detect a recognizable human form with the naked eye." The same reasoning applies to embryos in a laboratory dish. They don't have true moral status unless they are implanted in a woman and survive until 40 days. Zoloth points out, however, that we should not be careless with embryos and that they should be treated with special consideration and respect.

Jewish religious leaders emphasize the importance of the duty to heal. In a written statement prepared for a government commission on bioethics in 2000, Rabbi Elliot Dorff of the University of Judaism contends, "[I]n light of our divine mandate to seek to maintain life and health, one might even argue that from a Jewish perspective, we have a duty to proceed with that research."[24] In the same session, Yeshiva University's Rabbi Moshe Dovid Tendler extends a warning, "[A] fence that prevents the cure of fatal diseases must not be erected, for then the loss is greater than the benefit. In the Judeo-biblical legislative tradition, a fence that causes pain and suffering is dismantled. Even biblical law is superseded by the duty to save lives."[25]

Interpretations of Islamic law also find support for the use of embryos for research. As in many cultures, Islam's debate about the status of extracorporeal tissue arises from its discussion about cells taken from aborted fetuses for therapies. Opinion and law about embryos and stem cells are sparse, but in a review of both Shi'ite and Sunni interpretations of the Koran, the Islamic scholar Abdulaziz Sachedina mentions that most opinions state that the soul is formed sometime after the early stages of embryogenesis. As a rule, Islam usually accepts abortion before the eightieth day. Like Judaism, Islam permits interventions in nature that will further a greater good for humanity. But Sachedina also notes that there is disagreement among Muslim jurists about whether embryos in IVF clinics have a predefined sanctity, ensuring their protection, or no sanctity, allowing their destruction.[26]

PHILOSOPHICAL ARGUMENTS

Philosophers and ethicists describe the concept of *sentience*, the capacity to feel psychic or physical pleasure or pain. In order to be sentient, an animal or person must be conscious of its existence. Permanently unconscious persons — including those classified neurologically as being in a "persistent vegetative state" — lack sentience. Though we think of them as human, we also acknowledge that they lack an essential quality of "being in the world," that is, relating, speaking, and thinking like other members of society. Such experiences make certain creatures not just a thing, but also a being. In other words, consciousness, whether it is explained behaviorally or neurochemically, is the experience of *being a being*.[27] Sentient beings deserve our respect and demand that we're attentive to the possibility of causing them pain, fear, or discomfort.

Biology isn't adept at defining when a soul becomes present or when a person emerges during development. However, biology can be helpful when we think about sentience. The threshold of sentience can be estimated in developmental and biochemical terms, when the nervous system has matured. When are the sensations of pleasure and pain and an inkling of consciousness possible? At this juncture, it's best to ask a neuroscientist.

Michael Gazzaniga has spent a career at the edge of brain biology and is another member of the PCBE. The formation of the brain is a *tour de force*, a long process that begins early in development. The neuroscientist disagrees with some of his fellow council members about the status of an embryo, and bases his position on what he understands about brain biology. He points out that the human brain isn't viable until gestation has reached six months — about the same time a fetus could survive a premature birth in a hospital's neonatal unit.[28] Working backward through time, the brain's first electrical activity is detected around week six and the major divisions of the brain don't form until week four. The primitive streak — an indentation on the embryo that signals the scant beginnings of a nervous system — appears at 15 days.

In contrast to those who consider an embryo the same moral person at different stages, Gazzaniga believes that moral rights of the developing human begin at six months, coincident with its survival outside the womb. Just as Zoloth describes the early stages of embryogenesis "like water," Gazzaniga uses a perceptual (and emotional) cue to imagine brain function at eight weeks, when the head of the fetus begins to look human. He writes, "I am reacting to a sentiment that wells up in me, a perceptual moment that is stark, defining, and real. The brain at [eight weeks] is hardly a brain that could sustain any serious mental life."[29]

There are other, nuanced arguments made by philosophers and commentators who support hESC research. James Wilson, an emeritus professor of public policy at the University of California, Los Angeles, resigned from the PCBE in May 2005. He wrote in the council's 2002 report, "My view is that people endow a thing with humanity when it appears, or even begins to appear, human — that is, when it resembles a human creature. The more an embryo resembles a person, the more claims it exerts on our moral feelings."[30] Others note that had a modern version of "technological ethics" been present a century earlier, we could well be without air travel, space flight, or the benefits derived from other risky areas of science.[31] A number of scholars respond to religious "sanctity of life" views by pointing out that up to half of fertilized eggs never implant in the uterus or, if implanted, fail to develop. If the Vatican contends that life begins at conception, then, the argument goes, how could God create a system for making babies that destroys half of them?[32] Zoloth says, "Of all the things I don't understand about God's plan, I do understand this: part of the way the world is structured is that every time you become pregnant, you do not produce a living being. I know that one can be a woman of faith and experience a miscarriage and understand that you have not failed."

The assertion that an embryo should not be destroyed because it has the potential to become a person or possibly is a person is countered by a world full of real people, each with a fully realized potential. Professor of ethics and theology James Petersen writes, "How can we let patients who are unmistakably people die to protect embryos that, even if implanted, may or may not turn out to someday become persons? We should not kill people to benefit others, but we should also not let people die to protect human tissue such as sperm or ova, even though such gametes have great potential."[33]

THE PURSUIT OF PERFECTION

Conservative Christian thinkers and commentators, including a number of the president's advisors, brook no argument when it comes to morality and human forms of life. The message from members of this group is clear: "Leave the embryo alone." They believe that carelessly using an embryo for scientific research is the same as carelessly using a person for scientific research. Destroying an embryo is equal to killing a person.

It is not simply a case of Christian morality at work among those who object to hESC research. The connection to personhood recalls longstanding moral instructions from philosophers like Immanuel Kant (1724–1804), who said that persons should never be used as a means for someone else's ends. The Kantian ethic of how we should treat each other is an essential ingredient of our professional codes of medicine — doctors must treat patients with dignity and respect and, above all, "do no harm." If an embryo is a person, then the Kantian imperatives apply.

Whichever side of the debate one chooses, dead philosophers help us only up to a point. It is no exaggeration to say that looking at early embryonic life in the glare of modern technology provokes a visceral response in some, and the feeling translates into a values-laden moral opinion. For example, Michael Gazzaniga uses his "gut feeling" and what he knows about science to determine that the embryo is not a person. The neuroscientist's political colleague was Leon Kass, former chairman of the PCBE. To Kass, the "gut feeling" is a sinking one. His reaction doesn't have much to do with whether the embryo has a moral status — Kass maintains he is "agnostic" on that point. A University of Chicago philosopher with degrees in medicine and biochemistry, Kass says our *hubris* is the danger — Americans have developed a scientific swagger that threatens our humanity. Because of our overconfident attitude, we find ourselves teetering on the precipice of a "slippery slope." Kass writes, "In leading laboratories, academic and industrial, new creators are confidently amassing their powers and quietly honing their skills, while on the street their evangelists are zealously prophesying a posthuman future."[34]

What lies at the bottom of Kass's stance that we have embarked on a perilous decent? An existence akin to Aldous Huxley's *Brave New World*, a futuristic place with an amoral society living in perfect health. *New World*'s inhabitants have technology to blame for the lack of their humanity; children are born in embryo factories and tranquilizing drugs distributed by technocrats control adult behavior. Kass finds a deep resonance with Huxley's dystopic vision and compares it to our devotion to biomedical technologies, including those that deal with the human embryo. He begins with *in vitro* fertilization. The reproductive clinician, Kass believes, subverts family values by becoming a third partner along with the parents of the newly implanted embryo.[35] He is fearful that if left unchallenged, scientists will genetically engineer babies, grow "newborns" in laboratory vats, and use them as helpless factories to produce medical products and human parts.[36] The embryos in his posthuman future are treated as "mere meat" and "human caviar."[37]

Kass admits his images are meant to provoke — Laurie Zoloth describes his views as a "philosophy that demands engagement." Though Kass avers that the embryo probably deserves less respect than a human being, he thinks we should treat it with dignity. After all, each of us begins life as an embryo and if *we* are worthy of dignity, then we are worthy of such respect at every stage of our existence. To experiment with an embryo is to meddle with nature, or worse, to play God. In this worldview, nature and God are intertwined and inseparable. Genetic engineering, genetic testing, *in vitro* fertilization, mood-enhancing drugs, and embryonic stem cells — indeed, any technology that pushes beyond "natural limits" — are listed as dangerous outcomes of our headlong pursuit of perfection.[38] In this very important respect, Kass is in agreement with conservative Christians about the embryo and hESC research. The trouble with biotechnology, writes fellow council member Michael Sandel, is that it represents "the one-sided triumph of willfulness over giftedness, of dominion over reverence, of molding over beholding." It leads, according to Sandel, a Harvard University professor of government, to a confusion "of our role with God's."[39] There is a distinction between using science for therapeutic reasons (antibiotics to overcome an infection) and using science for human enhancement (increasing our lifespans or our capacity for memory) concludes Meilaender. "Here is where the limitlessness of our desire enters in," he warns. "Enhancement deliberately tries to move beyond the normal condition of our species — and there is no end to the ways in which we might like to be 'better.' Either we get a grip on that desire or it will increase its already strong hold on us."*

A cadre of commentators who worry about scientific "progress" echoes the *Brave New World* philosophy. If we aren't careful, they say, Huxley's fictional world may come to pass. Among the group is Daniel Callahan, a co-founder of the Hastings Center for Bioethics, a think tank nestled in a tree-lined acreage [about 50 miles north] of New York City. In his 2003 book, *What Price Better Health: Hazards of the Research Imperative*, Callahan writes that biomedical research has become a "moral imperative" with its promise to win the battle against the "diseases of aging." The spoils of war are longer, healthier lives. Our attempts to find ways of living longer shouldn't be a mandate. Callahan contends, "We ought to act in a beneficent way toward our fellow citizens, but there are many ways of doing that, and medical research can claim no more of us than many other worthy ways of spending our time and resources."[40] It's a fair argument: scarce resources can be better spent on preventing illness and disability and teaching people how to take care of their own health. Regenerative medicine seems to fall squarely within Callahan's crosshairs. He writes, "Unless someone can come up with a plausible case that the nation needs everyone to live much longer, and longer than the present steady gain of normalization will bring, there is no reason whatever for government-supported research aimed at maximizing or optimizing lifespans."

Joining Callahan and Kass is Francis Fukuyama. The professor of international political economy at Johns Hopkins University writes in the introduction to his book, *Our Posthuman Future: Consequences of the Biotechnology Revolution*, "The aim of this book is to argue that Huxley was right, that the most significant threat posed by contemporary biotechnology is the possibility that it will alter human nature and thereby move us into a 'posthuman' stage of history."[41]

*For more on enhancement see Chapter 19.