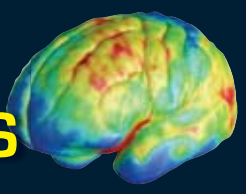
Why Prediction Markets Beat Political Polls (*page 38*)

SCIENTIFIC AMERICAN

Brain Power
BEYOND NEURONS
page 54



March 2008 www.SciAm.com

The End of Cosmology

Evidence of the big bang is disappearing as the universe expands

Space Wars

How Weapons in Orbit Put the Earth at Risk

Quantum Computers

Problems That Even They Can't Solve

Bluefin Tuna

Are We Eating These Fish to Extinction?

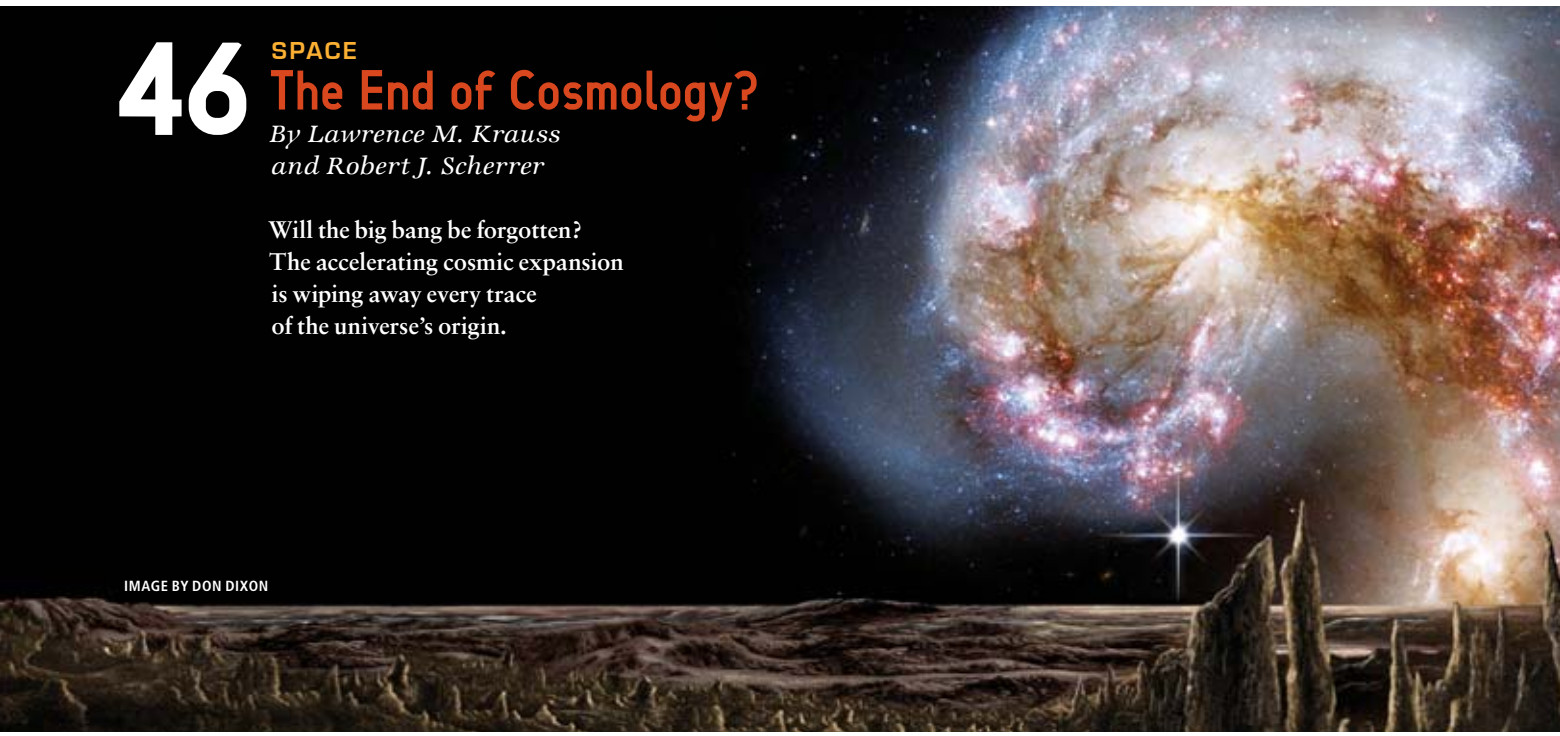
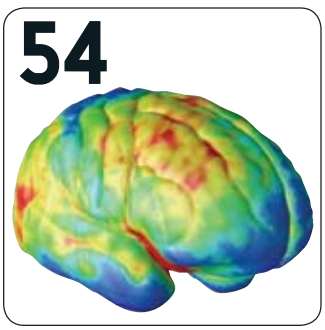
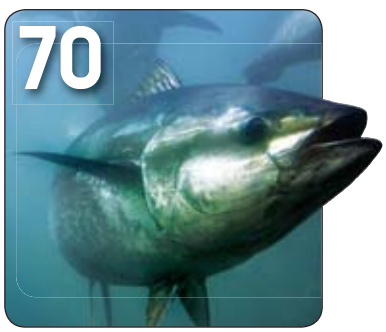


IMAGE BY DON DIXON





ON THE COVER
If the universe's expansion continues to accelerate, evidence of the big bang will become ever more elusive. Image by Phil Saunders, Space Channel Ltd.

MORE FEATURES

GO TO SCIAM.COM

BIODIESEL TAKES TO THE SKY ▶
An unmodified Czechoslovakian jet flew burning nothing but cooking oil.
More at **www.SciAm.com/ontheweb**

COURTESY OF RUDI WIEDEMANN *BioDiesel Solutions*

In Focus
Culture Speeds Up Human Evolution
Analysis of common patterns of genetic variation reveals that humans have been evolving faster in recent history.

News
A New Way to Help Networks Handle Ever Heavier Data Loads
Researchers discover an acoustic way to briefly store data that alleviates traffic bottlenecks.

Podcast
Docs Use Fake Pills as Real Meds
A new study finds that a significant number of physicians will on occasion knowingly prescribe a placebo.

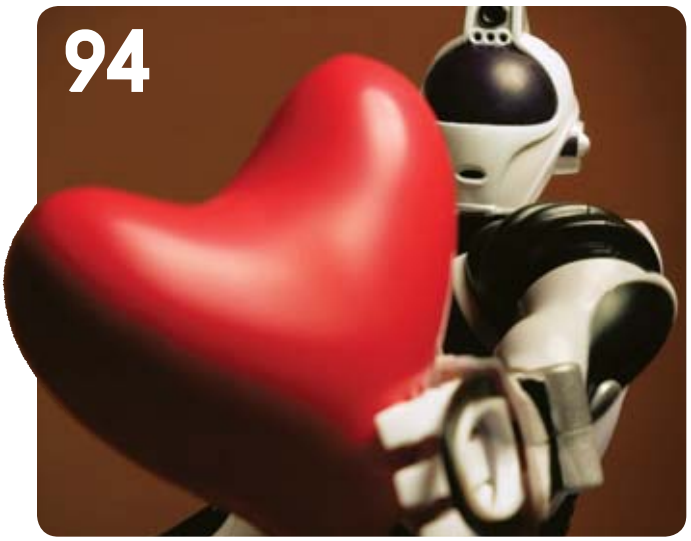
Fact or Fiction
Leaving Computers On Helps Them Last
What's the best way to extend a PC's life?

Video
Interview with Bill Gates
We catch up with the outgoing chairman of Microsoft.

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111.

Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 40012504. Canadian BN No. 127387652RT; QST No. Q1015332537. Publication Mail Agreement #40012504. Return undeliverable mail to Scientific American, P.O. Box 819, Stn Main, Markham, ON L3P 8A2. Subscription rates: one year $34.97, Canada $49 USD, International $55 USD. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. **Reprints available:** write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; (212) 451-8877; fax: (212) 355-0408. **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 248-7684. Send e-mail to sacust@sciam.com** Printed in U.S.A.

DEPARTMENTS



SION TOUHIG

SCI AM **A Q&A on Robot Love**
with David Levy is available at **www.SciAm.com/ontheweb**

Looking Both Ways

The past and the future can be equally hard to interpret

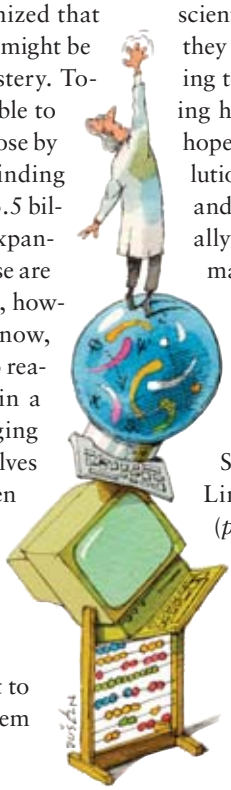
Predicting the future is usually a sucker's bet. Conversely, making bets might be an unusually sound way of predicting the future. That premise lies behind the plethora of futures markets that have sprung up in recent years, inspired by the success of commodities markets at determining the best time to go long on soybeans. Some political observers have begun to trust that if futures markets can anticipate the price of pork bellies, they should foresee the outcome of congressional races, too—possibly better than traditional polls can. Yet these markets are still far from infallible, so theoreticians continue trying to determine how to design smarter, more reliable markets. Senior writer Gary Stix writes about the efforts to turn price tickers into crystal balls, beginning on page 38.

When physicist Lawrence M. Krauss and cosmologist Robert J. Scherrer peered into the future—the fantastically distant future, to be sure—they recognized that any sentient beings of that time might be in the grip of an insoluble mystery. Today's astronomers have been able to reconstruct how the universe arose by looking deep into space and finding telltale evidence of a big bang 13.5 billion years ago (*page 46*). The expansion and evolution of the universe are gradually erasing that evidence, however: a billion centuries from now, naive astronomers would see no reason to doubt that they lived in a small, empty, almost unchanging cosmos. Should we count ourselves privileged to live in an era when the facts are still accessible to us? Or should we ponder whether unsuspected truths about the origins of space and time are already lost to us, too?

Even when the facts relevant to a problem still exist, finding them amid the flotsam of confounding information and assembling them into a cogent argument can be forbiddingly tough. That challenge is at the heart of the story written by investigative reporter Carole Bass, "Solving a Massive Worker Health Puzzle" (*page 86*).

Since 2001 various authorities have been looking into a statistically unusual cluster of brain cancer cases among employees of a manufacturing plant in Connecticut. Determining whether those cancers had a shared industrial cause is a type of problem with which epidemiologists, industrial chemists and other specialists are increasingly tasked, because such answers are the basis for safety regulations, not to mention legal findings of culpability. As Bass's account demonstrates, however, the scale of the problem—which involves a quarter of a million people experiencing unmeasured exposures to unidentified chemical and radiological agents over decades—can be overwhelming.

Sometimes better instruments can help scientists past obstacles; sometimes they can't. In recent years the emerging technology of quantum computing has become a lodestone for the hopes of many people looking for solutions to problems in cryptanalysis and other fields that would be literally or effectively impossible for normal computers. As a result, quantum computing has started to acquire a popular reputation as a nearly magical method for solving any conceivable problem. Nevertheless, it is not, as Scott Aaronson describes in "The Limits of Quantum Computing" (*page 62*). Some answers will stay beyond easy reach for a long time to come; that remains a safe bet. ■



FLYNN LARSEN (*Rennie*); DUŠAN PETRIČIĆ (*drawing*)

JOHN RENNIE
editor in chief

Among Our Contributors

CAROLE BASS
is a Connecticut-based investigative journalist specializing in public health and the environment. She was previously a reporter and editor for the *New Haven Advocate.*

RICHARD ELLIS
is a conservationist and author of several books, including a forthcoming one on bluefin tuna. He is also recognized as the world's foremost painter of marine natural history.

R. DOUGLAS FIELDS
is a section chief for the National Institute of Child Health and Human Development. In the August 2007 issue of this magazine he wrote about the electric sense of sharks.

THERESA HITCHENS
directs the Center for Defense Information in Washington, D.C., and is author of *Future Security in Space: Charting a Cooperative Course.* She is a past editor of the *Defense News* weekly.

LAWRENCE M. KRAUSS
is a cosmologist at Case Western Reserve University and a best-selling author with seven books to his credit. His article in this issue is his fifth for *Scientific American.*

ROBERT J. SCHERRER
chairs the department of physics and astronomy at Vanderbilt University, where he studies cosmology. He is also the published author of seven science-fiction stories.

SCIENTIFIC AMERICAN®
Established 1845

Environment and ALS ▪ The Genetic Code ▪ Vitamin D

NOVEMBER 2007

"Epidemiological studies of amyotrophic lateral sclerosis have so far found that the only environmental factors to have any correlation are increased age, being male and participating in the first Gulf War."
—Dougal MacDonald EDMONTON, ALBERTA

Lifestyle Link?

"Playing Defense against Lou Gehrig's Disease," by Patrick Aebischer and Ann C. Kato, was an excellent and hopeful summary of current research on amyotrophic lateral sclerosis (ALS). But the last paragraph, which suggests that lifestyle may play a role in the development of ALS (and which mentions that regular exercise offers some protection against neurodegenerative diseases), seems to have little to do with the research described in the article.

Many of those affected by ALS tend to be fit, lean, active people who lead a healthy life. This lean-active correlation is more likely an effect than a cause: something about various genetic predispositions may also tend to militate against becoming sedentary. And the epidemiological studies of ALS that have been done—albeit with small samples—have so far found that the only environmental factors to have any correlation are increased age, being male and participating in the first Gulf War.

Although lifestyle may be a factor to investigate, the fact that about 50 percent of motor neurons can die before symptoms become noticeable makes it more likely that the most effective approach to treatment and a cure will be to find a set of reliable biomarkers that can diagnose the problem and its specific starting event or events years earlier and to develop a cocktail of appropriate drugs to direct at the initial and root causes.

Dougal MacDonald
Edmonton, Alberta

Code Crack Credit

In asserting Marshall W. Nirenberg as the discoverer of the genetic code, "The Forgotten Code Cracker," by Ed Regis [Insights], might lead an unwary reader to conclude that Francis Crick considered himself to have discovered it. Such an inference should be checked against the available records in the Crick Papers at the Wellcome Library in London. There one can find correspondence in which Crick, referring to the "fuss" made by the British press over his and his colleagues' 1961 *Nature* article establishing major features of the genetic code, reassures Nirenberg "that it is your discovery which was the real breakthrough."

The issue of priority was still a concern for Crick when *Scientific American* was to publish review articles by both him and Nirenberg, in that order. Crick stressed to the editor that he should not be published first, as "it should not appear to anyone that we wish to claim more than is our due. However ingenious and elegant our experiments are, it must be realized that it is the biochemical work on the cell-free system which will be crucial." Crick also explained that Nirenberg and Johann Matthaei's basic discovery came before Crick and his colleagues received the triple mutants that clinched their own conclusions. These reservations surely support the view that he would not have wanted to be called *the* discoverer of the code.

Robert Olby
University of Pittsburgh

Sunbed Supplement?

"Cell Defenses and the Sunshine Vitamin," by Luz E. Tavera-Mendoza and John H. White, describes evidence of a possible link between vitamin D deficiency and certain autoimmune diseases such as multiple sclerosis (MS) and refers to studies that have found therapeutic possibilities for using vitamin D to control those diseases. The authors also describe how vitamin D is produced in human skin when it is exposed to ultraviolet B (UVB) light from the sun.

I have been affected by MS since 1975. I do not have adequate opportunities to expose myself to sunshine because I work in a closed environment and my condition makes it difficult for me to walk around. Would exposure to a tanning lamp produce vitamin D equivalent to that created by true sunshine?

Massimo Cannata
Varese, Italy

SKIN JOB: Vitamin D is manufactured in human skin with the aid of sunlight.

WHITE REPLIES: *Tanning lamps are not equivalent to UV radiation. Such lamps are generally a much stronger source of UVA light and a weaker source of UVB light than the sun. (It is only the latter that induces cutaneous vitamin D synthesis.) Certain manufacturers claim to offer portable UV sources designed to generate UVB, but I do not know much about them.*

In any case, vitamin supplements present a safer option. (In Montreal, I personally take 4,000 international units of Vitamin D_3 daily from October until mid-May, when it becomes warm enough to be outside in shirtsleeves.) You should also have your 25-hydroxyvitamin D levels checked: ideally they should be close to 50 nanograms per milliliter (and at least 40 ng/ml).

Wasteful Warheads

"Build Diplomacy, Not Bombs" [Perspectives] was right on about the unnecessary nature of a program to replace nuclear warheads. We are financing far too many government programs just to keep a few scientists and other high-paid employees available and trained in case there is ever a need for their outdated services. In my experience as a government engineer, I found that the differences of opinion among scientists advocating mission callback capabilities, hardened missile sites and highly mobile launch vehicles used up a huge amount of our tax-supported resources.

Furthermore, a show of force can never create anything other than an opposing show of force. The road to diplomacy must be carefully paved with a desire to implement free trade for needed resources, a common agreement on monetary policy and well-monitored negotiations among companies that agree to that policy.

Ray D. Close
Scottville, Ill.

Population Pollution

One can only wonder whether the politicians who will be participating in the new international agreement on greenhouse gas reduction, referred to by Jeffrey D. Sachs in "Climate Change and the Law" [Sustainable Developments], do any arithmetic. World population is still growing, and energy demand is growing even more rapidly. Every year more fossil fuel is burned. Large new coal mines are being opened to exploit lower-quality coal deposits, and the capacity of new fossil-fuel plants swamps that of new wind and solar installations. Furthermore, millions of cars per year are being built in or shipped to countries with huge populations that until now have had few private vehicles. Without drastic population reduction, all our other efforts are fatuous.

Don Hirschberg
Horseshoe Bend, Ark.

ERRATUM The Calabi-Yau space image on the November 2007 cover appears courtesy of Wikipedia and is available at http://tinyurl.com/2gzzb2

Letters to the Editor

Scientific American
415 Madison Ave.
New York, NY 10017-1111
or editors@SciAm.com

Letters may be edited for length and clarity. We regret that we cannot answer all correspondence.

SCIENTIFIC AMERICAN®
Established 1845

CHAIRMAN: Brian Napack
PRESIDENT: Steven Yee
VICE PRESIDENT AND MANAGING DIRECTOR, INTERNATIONAL: Dean Sanderson
VICE PRESIDENT: Frances Newburg
CHAIRMAN EMERITUS: John J. Hanley

ASSOCIATE PUBLISHER, CIRCULATION: Simon Aronin
CIRCULATION DIRECTOR: Christian Dorbandt
RENEWALS MANAGER: Karen Singer
FULFILLMENT AND DISTRIBUTION MANAGER: Rosa Davis

VICE PRESIDENT AND PUBLISHER: Bruce Brandfon
SALES DEVELOPMENT MANAGER: David Tirpack
SALES REPRESENTATIVES: Jeffrey Crennan, Stephen Dudley, Stan Schmidt

ASSOCIATE PUBLISHER, STRATEGIC PLANNING: Laura Salant
PROMOTION MANAGER: Diane Schube
RESEARCH MANAGER: Aida Dadurian
PROMOTION DESIGN MANAGER: Nancy Mongelli

VICE PRESIDENT, FINANCE, AND GENERAL MANAGER: Michael Florek
BUSINESS MANAGER: Marie Maher
MANAGER, ADVERTISING ACCOUNTING AND COORDINATION: Constance Holmes

DIRECTOR, SPECIAL PROJECTS: Barth David Schwartz

MANAGING DIRECTOR AND VICE PRESIDENT, ONLINE: Mina C. Lux
ASSOCIATE PUBLISHER, ONLINE: Paul Rothkopf
DIRECTOR, WEB TECHNOLOGIES, ONLINE: Vincent Ma
SALES REPRESENTATIVE, ONLINE: Gary Bronson

DIRECTOR, ANCILLARY PRODUCTS: Diane McGarvey
PERMISSIONS MANAGER: Linda Hertz

How to Contact Us

SUBSCRIPTIONS
For new subscriptions, renewals, gifts, payments, and changes of address: U.S. and Canada, 800-333-1199; outside North America, 515-248-7684 or www.SciAm.com

REPRINTS
To order reprints of articles: Reprint Department, Scientific American, 415 Madison Ave., New York, NY 10017-1111; 212-451-8877, fax: 212-355-0408; reprints@SciAm.com

PERMISSIONS
For permission to copy or reuse material: Permissions Department, Scientific American, 415 Madison Ave., New York, NY 10017-1111; www.SciAm.com/permissions or 212-451-8546 for procedures. Please allow three to six weeks for processing.

ADVERTISING
www.SciAm.com has electronic contact information for sales representatives of Scientific American in all regions of the U.S. and in other countries.

JAMES PORTO

Atomic Economics ▪ Hammondsport Flight Test ▪ Exactitude and Fashion

Compiled by Daniel C. Schlenoff

MARCH 1958

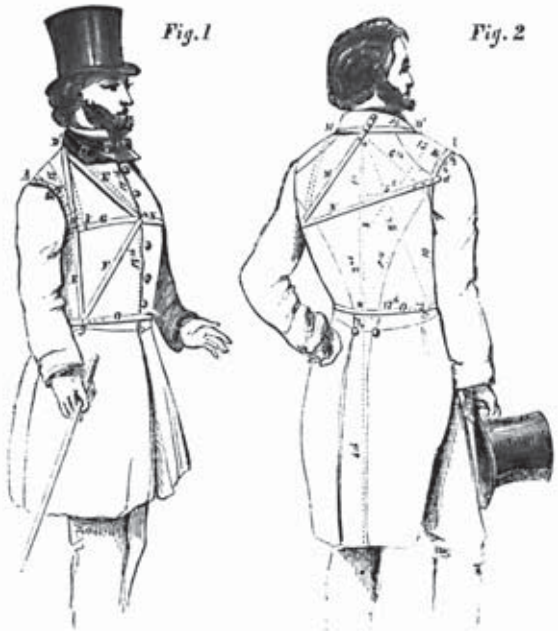
ENERGY, ASH, MONEY—"If atomic power is to be developed on an important scale, methods will have to be found for safely disposing of the vast quantities of radioactive 'ashes' that will be produced by nuclear reactors. Last month a committee of the National Academy of Sciences reported on waste disposal, pointing out that the costs of storing radioactive fission products temporarily to 'cool' them, of extracting long-lived isotopes and of shipping waste to distant points for ultimate disposal will have a major influence on the economics of nuclear power."

KRYPTON AND MARRIAGE— "The precious bar of platinum and iridium at Sèvres, France, against which all the world's distance scales are theoretically checked, may soon be melted down for wedding rings. An international advisory committee for the definition of the meter has now recommended the adoption of an atomic standard of length—an orange spectral line of krypton 86. The meter is to be defined as 1,650,763.73 times this wavelength. The krypton line is the sharpest of those currently available for length measurement."

MARCH 1908

RED WING—"The Aerial Experiment Association, which was formed last summer by Dr. Alexander Graham Bell, has been actively engaged during the past three months in constructing and testing an aeroplane. Mr. F. W. Baldwin operated the aeroplane [the "Red Wing"] in its initial test. The motor employed was an 8-cylinder Curtiss engine of 40 horse-power. The idea of mounting the aeroplane upon runners and testing it upon the ice seems to be an excellent one. Owing to the warm weather and the melting of the ice on Lake Keuka, near Hammondsport, N.Y., where the test was held, it was feared that it would be impossible to try the machine. Fortunately, however, a cold snap gave the experimenters a chance to make the trial; and on the 12th of March, upon its first test, the aeroplane flew a distance of 318 feet 11 inches."

The entire article from 1908 is available at
www.SciAm.com/ontheweb



BESPOKE—and scientifically precise—tailoring, 1858

RENEWAL—"In the old days of '49 the ships which called at Yerba Buena—as San Francisco was then called—found themselves stranded, for the crews would desert *en masse* for the gold diggings. Among these ships was the 'Niantic,' built in a Maine shipyard. After being deserted, the ship was pulled ashore at Clay and Sansome streets and converted into a lodging house. The shallow water at her stern was gradually filled up with sand. The Niantic Block of apartments was erected over her timbers, but perished in the conflagration of April, 1906. Recently, in digging the foundations for a new Niantic Building, the keel and ribs of the old ship were found, stuck fast in the mud and sand."

FOURTH DIMENSION—"Mathematics is the most exact and the most thoroughly grounded of the sciences. And yet, in the very field explored by this rigorous and tedious method, have arisen fantastic and fairy-like structures of the imagination, which transcend all our experience. They have arrived at the conception of the fourth and higher dimensions. It would be impossible to confine a person having the secret of this dimension by the six surfaces of a prison cell. His slightest movement in the direction of a fourth dimension would put him at once out of three-dimensional space. It would be well for him to take care just what he did when in four-dimensional space, as upon coming back into space of three dimensions he might be much changed."

MARCH 1858

MAN IS THE MEASURED—"When it is recollected that the human figure is the very acme of symmetry and grace, it is somewhat astonishing to see so many awkward and ungraceful-looking persons walking our streets, and we are forced into the conclusion that their tailor did not do them justice when he cut their clothes. A slovenly disregard is just as culpable; and as we must wear coats and the like, it is correct and proper that we should have them fit. Here we show an instrument patented by Simeon Corley, of Lexington, S.C., for the purpose of taking accurate measurements of the body, and of afterwards drafting the garment on cloth."

SCIENTIFIC AMERICAN, VOL. XIII, NO. 29, MARCH 28, 1858

Hoyle's Missing Equation ▪ Flu Sites ▪ Artificial Capillaries ▪ Stubborn Scores

Edited by Philip Yam

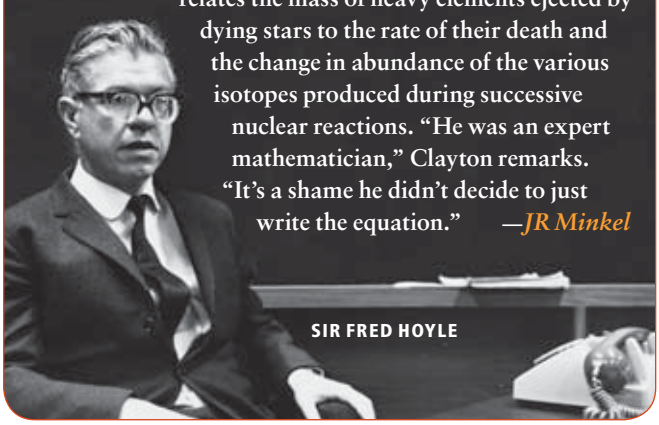
Not Showing His Work

The failure to include a key equation may have kept Sir Fred Hoyle from getting the recognition he deserved for a paper on the formation of elements in stars. Hoyle, who died in 2001 at the age of 86, was something of a tragic scientific figure. He never accepted the big bang theory, preferring instead the idea of a steady-state cosmos; later, he embraced the view that life on Earth originated in outer space. These attitudes probably cost him a Nobel Prize [see the profile of Hoyle, "The Return of the Maverick"; SciAm, March 1995].

But before his scientific infamy, Hoyle made what should have been a lasting contribution with a 1954 *Astrophysical Journal* paper laying out a process by which stars heavier than 10 suns would burn the hydrogen and helium at their cores into heavier elements through a progressively hotter series of nuclear fusion reactions. When such a star finally exploded in a supernova, it would scatter these elements into space, where they would seed still-forming star systems. Prior to Hoyle's work, many experts believed that the elements must have been born during the big bang.

Instead of citing the study, researchers discussing the formation of elements, or nucleosynthesis, typically reference a 1957 paper, co-authored by Hoyle, which focused on other facets of the problem, says Donald Clayton of Clemson University, who was a colleague of Hoyle's. Clayton found that of 30 major nucleosynthesis papers published between 1960 and 1973, 18 cited the 1957 work and only one gave the nod to Hoyle's 1954 paper.

Writing in the December 21, 2007, issue of *Science*, Clayton attributes the misplaced citation to the lack of a relatively straightforward equation that was implicit in the 1954 study. "Hoyle's equation," as Clayton calls it, relates the mass of heavy elements ejected by dying stars to the rate of their death and the change in abundance of the various isotopes produced during successive nuclear reactions. "He was an expert mathematician," Clayton remarks. "It's a shame he didn't decide to just write the equation." —*JR Minkel*



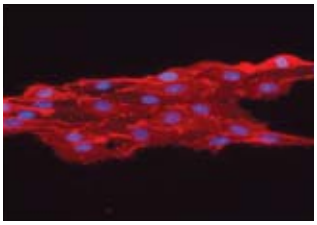
SIR FRED HOYLE

Shaping up for the Flu

The next influenza pandemic will most likely occur when a change in an avian or swine flu virus enables it to infect human cells. In the January *Nature Biotechnology*, scientists report that the critical change will be the virus developing an ability to latch onto a particular variety of the so-called alpha 2-6 glycan receptors on lung cells. These receptors come in umbrella and cone shapes; in humans, the umbrella version, which is more plentiful, opens the door to serious illness. This finding answers such questions as why some flu strains dock at cells but do not infect humans very well (because the viruses have latched onto cone-shaped receptors), and it could provide a way to track flu adaptations as well as to uncover new treatments.

Growing Blood Vessels

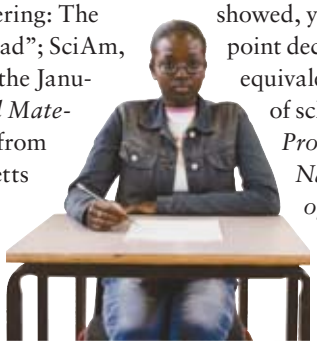
As part of the goal of growing replacement organs, researchers have relied on scaffolds that encourage cells to form in particular ways. But they still need to perfect the scaffold materials and geometry, among other factors [see "Tissue Engineering: The Challenges Ahead"; SciAm, April 1999]. In the January 1 *Advanced Materials*, scientists from the Massachusetts Institute of Technology describe how 600-nanometer-wide ridges and grooves on a silicone-based substrate provided pathways for flourishing endothelial progenitor cells. Grown three-dimensionally, these cells formed tubes that could serve as capillaries. The challenge remains in getting such constructs into the body.



GROWING CELLS have followed a patterned substrate.

Lasting Disparity

A study of some 2,000 children observed over seven years supports the nurture side of the IQ debate [see "Unsettled Scores"; SciAm, February 2007]. Researchers tracked residents of six severely disadvantaged neighborhoods in Chicago—almost all African-American—and found significant and stubborn declines in their verbal ability compared with peers in better neighborhoods. Children who moved out of the disadvantaged areas still showed, years later, four-point declines in IQ, equivalent to a year of schooling. The *Proceedings of the National Academy of Sciences USA* published the findings online December 19, 2007.



HULTON-DEUTSCH COLLECTION *Corbis (Hoyle)*; KARL GRUPE *Getty Images (student)*; CHRISTOPHER BETTINGER *Massachusetts Institute of Technology (cell scaffold)*

SECURITY

Perilous Pursuit

With missile defense, India turns the thumbscrews on unsettled Pakistan BY MADHUSREE MUKERJEE

The assassination of opposition leader Benazir Bhutto, the defiance of militants and public unease with President Pervez Musharraf's government have raised questions about the stability of Pakistan and the security of its nuclear armament. Exacerbating these concerns is a nervous neighbor. In January, weeks after an Indian missile successfully crashed into another missile over the Bay of Bengal, an official announced that India could deploy a defense shield against ballistic missiles by 2011.

By seeking to fend off its tenacious rival, India may have inadvertently increased the risk of a regional nuclear exchange. Furthermore, "missile defense will make it likely that greater damage will be inflicted on India" if such a war breaks out, argues Theodore Postol, a defense analyst at the Massachusetts Institute of Technology.

Experts have long considered the Indian subcontinent to be the likeliest flash point for the world's first nuclear exchange [see "India, Pakistan and the Bomb," by M. V. Ramana and A. H. Nayyar; SCIENTIFIC AMERICAN, December 2001]. Ever since India began the nuclear arms race in 1974, Pakistan has responded tit-for-tat to every development. The effort seems to have paid off for Pakistan: in 1999, when it sent paramilitary forces across the border, India repelled the attackers but did not pursue them home, reportedly because of threats of atomic retaliation.

India's planners have now decided that a missile defense shield is the answer to their self-inflicted predicament. They have an ally in the Bush administration, which last year forged a deal that will allow India to purchase uranium from international sources for its civilian reactors. If the agreement goes through, India will be able to process all its domestic uranium for the military, adding 60 to 100 kilograms of weapons-grade plutonium annually to its current stock (estimated at 600 kilograms). Earlier, in 2005, the U.S. had offered to share military technology, including that of missile defense, with India. Officials and military contractors from around the world have since been thronging to New Delhi in the hope of selling components of a defense system.

INCOMING: India's president Pratibha Patil (*in ivory sari*) visited a defense research base in January. In a test, an Akash missile—named after the Sanskrit word for "sky"—intercepted another missile in midair last December, aggravating the arms race with Pakistan.

MAHESH KUMAR A/AP Photo

Many scientists have pointed out the inherent shortcomings of ballistic-missile defense [see "Holes in the Missile Shield," by Richard L. Garwin; SCIENTIFIC AMERICAN, November 2004]. Defense systems of this kind cannot, for instance, distinguish decoys from real threats. In South Asia, short distances magnify the problems. "It's pretty unlikely that you can expect to reliably intercept anything," opines physicist Zia Mian of Princeton University, who studies nuclear proliferation and global security. India's Defense Research and Development Organization claims that its planned defense shield will destroy an enemy missile three minutes after the missile's detection by radar. Early-warning radar, such as one that India has im-

DEFENSE MINISTRY *AP Photo*

HARD TO HIT: Pakistan's Shaheen ballistic missile has fins that may give the payload section maneuverability after it separates from the main stage.

ported from Israel, could detect the missile and determine its course within 110 seconds after its launch. But a ballistic missile launched on a low trajectory from, say, a Pakistani air base could reach New Delhi in as little as five minutes, according to Mian and physicists M. V. Ramana of the Center for Interdisciplinary Studies in Environment and Development in Bangalore and R. Rajaraman of Jawaharlal Nehru University in Delhi.

That could leave technicians with too little time to figure out if the warning is real. U.S. scientists once spent eight minutes determining that a warning of a Soviet launch was false. Indeed, false alarms are frequent when it comes to early-warning systems. A flight of geese or an incidence of atmospheric turbulence can fool radar, and anomalous reflections of the sun can trick satellite-based infrared detectors. Between 1977 and 1988—the only period for which data have been released—the U.S. recorded an annual average of 2,600 false alarms of ballistic-missile launches from the Soviet Union. Even if India responds to a false alarm just by launching an interceptor missile, that action could be interpreted by Pakistan as an attack.

Moreover, soon after India acquired early-warning radar Pakistan tested a cruise missile, which Postol believes was reverse-engineered from an American Tomahawk missile that fell in Pakistan during a 1999 attack on terrorist training camps in Afghanistan. Powered throughout its flight, such a missile can hug the ground to evade radar. And in an apparent response to the India-U.S. nuclear deal, Pakistan has begun building a reactor for producing plutonium, which may yield a warhead small enough to fit onto a cruise missile. Pakistan also possesses ballistic missiles fitted with small fins on their forward, payload sections. These structures can add maneuverability, making the warheads exceedingly difficult to catch.

Still, "the attacker is always concerned that missile defense might work better than he thinks," Postol points out, and will launch more projectiles than necessary to ensure that at least a few get through. He suggests an alternative way of avoiding a nuclear holocaust. Both nations should disperse and hide their missiles and authorize a designated general in a remote outpost to launch retaliatory strikes should the political leadership be taken out in a first strike—and let the other side know. That way the rulers of both countries can be sure that a nuclear misadventure will lead to their homeland becoming history. Unlike missile defense, mutually assured destruction is at least a time-tested way to keep nuclear weapons in their holsters.

Madhusree Mukerjee, a former editor at Scientific American, *is writing about India's World War II experience in a forthcoming book.*

ARCHAEOGENETICS

Lovers, Not Fighters?

New genetic signs that modern humans mated with *Homo erectus* BY JOHN WHITFIELD

The question of whether modern humans made love or war with our ancestors has swung back and forth over decades of often acrimonious debate. At present, most researchers trying to read prehistory in our genomes believe that we contain no trace of species past and that we are all descended from a group that left Africa within the past 100,000 years and replaced all other humans, such as Neandertals, without interbreeding.

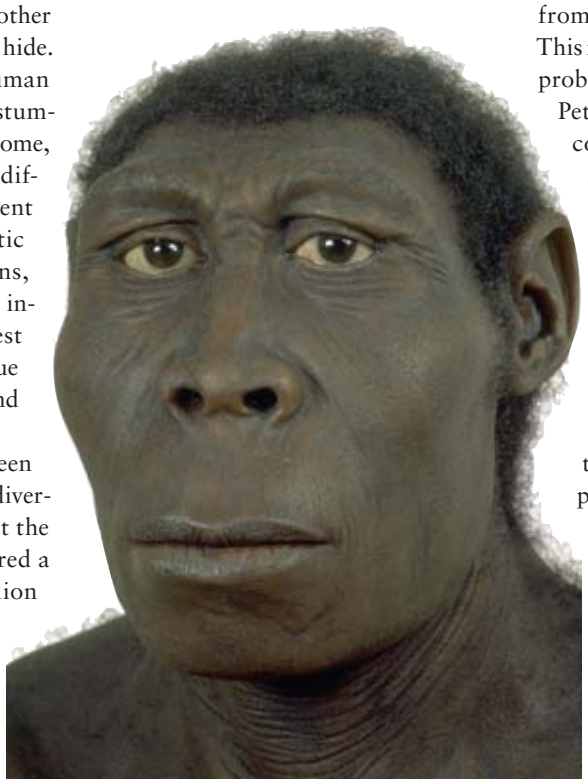
Those who favor the alternative view feel that the issue is moving their way, however. "Things are coming to the surface that don't fit that model" of a single recent migration out of Africa, says Murray Cox of the University of Arizona. Cox believes that he and his colleagues have found the clearest sign so far that modern humans mated with *Homo erectus*, a species that originated about two million years ago that many believe has a place on the lineage leading to ourselves.

The case for the all-conquering Africans is based mainly on studies of the Y chromosome and the mitochondrion, an energy-generating structure within the

cell that has its own small genome and passes down the female line. All modern variation in these sequences traces back to Africa and shows relatively recent common ancestors. (The absence of obviously hybrid fossils also supports this idea.)

But, Cox says, there are plenty of other places for genes from our relatives to hide. Studying the genetics of modern human populations, he and his colleagues stumbled across a region of the X chromosome, called *RRM2P4*, that shows large differences between people in different places—a sign of an ancient genetic split. To pin down the gene's origins, the researchers sequenced some 250 individuals, half from Africa, the rest from China, Central Asia, the Basque Country of southwestern Europe, and the Pacific Islands.

Translating the differences between gene sequences into a date for their divergence, the researchers conclude that the various forms of *RRM2P4* last shared a common ancestor about two million years ago—around when *H. erectus* migrated from Africa into Asia. And the oldest variant seems to have originated in Asia, where it is now almost exclusively found.



RÉGIS BOSSU *Sygma/Corbis*

HOT OR NOT? Modern humans may have mated with *Homo erectus*, according to a new genetic study.

The combination of great antiquity and Asian origins is compelling, Cox believes. Fossils suggest that *H. erectus* may have survived in Asia up until about 30,000 years ago, overlapping with modern humans by about 15,000 years. "This gene is most common exactly where you find *Homo erectus* fossils," says Cox, who described the finding in the January edition of *Genetics*. The team also used new statistical techniques to show that the gene is significantly more likely to have arisen in Asia than in Africa. The possibility that *H. erectus* and modern humans interbred is all the more surprising, he adds, because most researchers think there is no evidence for our having swapped genes with the more closely related Neandertals.

Other researchers note that many factors can create deep genetic divisions among human groups. No one disputes that the large majority of our genome has a recent African origin. But that dominance makes it difficult to tell what might be the legacy of a small amount of breeding with other lineages and what might be chance. Across the whole genome, some groups will show large differences even if they all left Africa at about the same time, explains geneticist Peter Underhill of Stanford University. "When you see these outliers, are they really emblematic of something dramatic, or are they just the extremes of the normal distributions? To me, it's still an open question," he says.

Alternatively, the Asian form of *RRM2P4* might have been present in a group that left Africa but later died out in that continent. Or the researchers might simply have not looked at enough Africans to find it; in fact, the DNA sequence that Cox and his colleagues think came from *H. erectus* existed in one African. This finding suggests that Asian *RRM2P4* probably did originate in Africa, argues Peter Forster, who studies what has become known as archaeogenetics at Anglia Ruskin University in Cambridge, England. "It spoils the story considerably to find this gene in an African," he says, adding that "I don't see it as convincing evidence" of ancient mixing between humans.

Cox counters that the lone African's *RRM2P4* sequence is identical to the Asian group, suggesting that this person probably descended from recent Asian immigrants. The team also has preliminary data on a second DNA region that shows an equally ancient split, where one group seems to be exclusively Asian.

No one gene can settle the matter, says geneticist Rosalind Harding of the University of Oxford. Nevertheless, several studies suggest that the human genome records some strikingly deep splits between populations. Her group, for example, has found that part of the gene for hemoglobin seems very old. If discoveries of ancient sequences continue to stack up, archaeogeneticists might eventually be convinced that we all have a little *erectus* in our blood.

John Whitfield is based in London.

Finding Ancestors in Our Genes

Researchers have long debated over two general theories of the origin of modern humans—whether our ancestors left Africa in one wave less than 100,000 years ago or whether we result from several waves out of Africa that interbred. One thing everyone agrees on is that reading our history in our genes is fraught with pitfalls. Getting lots of data has become easy, but working out which sequences to use, whom to sample, and how to identify the genetic legacies of natural selection, migration and population bottlenecks is extremely tough. Even relatively simple analyses involve assumptions and educated guesswork and can take thousands of hours of computer time.

SPACEFLIGHT

Stirling in Deep Space

To cut back on radioisotope fuel, NASA goes back 200 years BY MARK WOLVERTON

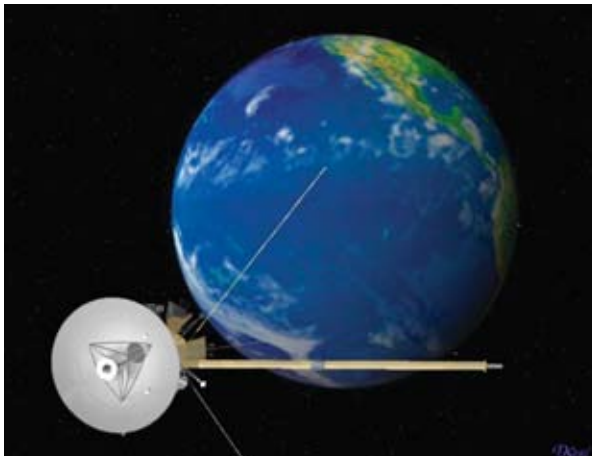
For more than 30 years now, NASA's deep-space probes have relied on radioisotope thermoelectric generators (RTGs), devices that use decaying plutonium 238 to warm thermocouples and generate electricity. Now the space agency is poised to replace those heavy, expensive and inefficient RTGs with a system that provides more power with much less radioactive fuel—technology based on a 19th-century invention.

Patented in 1816 by an intellectually restless Scottish minister named Robert Stirling, the Stirling engine is simplicity itself: two chambers or cylinders, one cold and one hot, containing a "working fluid" (commonly air, helium or hydrogen) with a regenerator or heat exchanger between the two. Differences in temperature and pressure between the two cylinders cause the working fluid to expand and contract, passing back and forth through the exchanger and moving a piston. The process thereby converts thermal energy (in NASA's case, supplied by radioactive decay) into mechanical energy.

"Stirling is actually something we've been investing in for almost the past three decades at some level, and it's now reached the point where we're ready to make the next step forward," says Dave Lavery, one of the directors of the Solar System Exploration program at NASA headquarters in Washington, D.C.

Lockheed Martin is putting the finishing touches on an engineering test unit that should be ready this spring: the advanced Stirling radioisotope generator. Two Stirling converters inside the generator drive pistons within a linear alternator, generating about 100 watts of electrical power. The unit will be less than a yard long and a foot wide, small enough to fit in the backseat of a subcompact car and weighing just over 40 pounds—less than half the weight of a typical RTG. It will boast a conversion efficiency of 20 to 30 percent, compared with the measly 6 to 7 percent of RTGs—while requiring only one-fourth the amount of radioisotope fuel.

Those characteristics translate into important advantages for spaceflight. Because the Stirling unit will be less massive and thus cheaper to launch, it will allow a spacecraft to carry a larger payload. The fourfold reduction in radioactive fuel—from 20 pounds in an RTG to five in a Stirling—also saves money while considerably reducing safety concerns involved in a worst-case scenario of a launch vehicle exploding in midair. NASA is keenly aware of public concerns about radiological safety, and as Lavery puts it, "for any nuclear-based system, we go through the entire National Environmental Policy Act process," which requires NASA to collect public comments before any final launch decision.



NEW TRAJECTORY: NASA plans to replace radioisotope thermoelectric generators, used on deep-space probes such as Cassini, with Stirling engines, which need one-fourth as much fuel. The artist's rendition shows Cassini's flyby of Earth.

NASA/JPL CALTECH/DAVID SEAL

Richard Shaltens, chief of the Thermal Energy Conversion Branch at the NASA Glenn Research Center, explains that once Lockheed completes initial testing, NASA Glenn will put the device through extended evaluations to begin its transition to flight status. "We're planning to go forward with the potential use of this technology on future missions in probably the 2012–2013 time frame," Shaltens says. He also points out that in more than 100,000 hours of lab testing in various environments, the Stirling converters "have demonstrated that they perform as predicted and have the potential for long life" comparable to RTGs.

NASA is so confident in the Stirling radioisotope generator that the agency has already invited the space science community to submit Stirling-based planetary mission concepts. Lavery emphasizes that the generator's inaugural mission will not be decided until at least 2009, but possible jobs include flights to the outer planets and manned missions to the moon or Mars. "Their overall design right now is to be compatible with either a deep-space interplanetary environment or planetary-surface environment with either atmosphere or vacuum," Lavery says.

Eventually Stirling technology may phase out RTGs completely. Lavery expects that it "would be the beginning of a new family of [radioisotope power systems] that are significantly more efficient and significantly less costly than the solutions we had available so far." Reverend Stirling could hardly have imagined that his ingenious invention might well become the prime mover that powers the next great era of solar system exploration.

Mark Wolverton is based in Bryn Mawr, Pa.

HIGH-ENERGY PHYSICS

Fiasco at Fermilab

Last-minute budget cuts stun U.S. physicists BY MARK ALPERT

In recent years the U.S. national laboratories have laid out an ambitious research agenda for particle physics. About 170 scientists and engineers at Fermi National Accelerator Laboratory in Batavia, Ill., have been developing designs and technologies for the International Linear Collider (ILC), a proposed machine that would explore the frontiers of high-energy physics by smashing electrons into their antimatter counterparts [see "Building the Next-Generation Collider," by Barry Barish, Nicholas Walker and Hitoshi Yamamoto; SCIENTIFIC AMERICAN, February 2008]. Another 80 researchers at Fermilab have been finalizing the plans for NOvA, a giant detector in northern Minnesota that could answer fundamental questions about the neutrino, a particle that is ubiquitous but maddeningly elusive. But on December 17, 2007—a date that scientists quickly dubbed "Black Monday"—Congress unexpectedly slashed funding for the ILC and NOvA, throwing the entire future of American physics into doubt.

What made the cutbacks so devastating was that President George W. Bush and Congress had promised substantial budget increases for the physical sciences earlier in 2007. In the rush to trim the 2008 spending bill enough to avert a presidential veto, however, legislative leaders excised $88 million from the U.S. Department of Energy's funding for high-energy physics. Fermilab's 2008 budget abruptly shrank from $372 million to $320 million. In addition to targeting the ILC and NOvA, Congress eliminated the $160-million U.S. contribution to ITER, the international project to build an experimental fusion reactor.

DESY

THREATENED BY THE AX: Prototype beam-line cooling modules designed for the International Linear Collider may never make it to the U.S., because recent budget cuts have reduced the chances that the particle collider will be built at Fermi National Accelerator Laboratory in Batavia, Ill. (The modules shown are at the DESY research center in Hamburg, Germany.)

Fermilab director Pier Oddone announced that the lab would need to lay off 200 employees, or about 10 percent of its workforce, and that the remaining researchers would have to take off two unpaid days per month. These measures would allow the lab to keep operating the Tevatron, its phenomenally successful proton-antiproton collider, which is now racing to find evidence of new particles and extra dimensions before the more powerful European accelerator, the Large Hadron Collider (LHC), begins operations later this year. But the ILC and NOvA were expected to become the major focuses of research at Fermilab after the shutdown of the Tevatron, due to occur by 2010, and now the investigators in those projects must be assigned to different efforts or dismissed. "The greatest impact is on the future of the lab," Oddone says. "We have no ability now to develop our future."

A big part of that envisioned future is the proposed ILC, a 31-kilometer-long facility that would be able to detail the properties of any new particles discovered by the Tevatron or the LHC. American physicists had taken a leading role in the international effort to develop the collider, but the sudden cutoff in funding reduces the chances that the machine will be built on U.S. soil. "The ILC will go forward, but the U.S. will fall behind," says Barry Barish, director of the global design effort for the collider. The project is expected to yield technological advances that could benefit medical accelerators and materials science, and Barish says the U.S. may become less competitive in these fields if American support for the ILC is not restored.

The NOvA project is further along than the ILC; in fact, before the funding cuts were announced, the program managers had planned to upgrade the roads to their Minnesota site this spring so that they could begin delivering construction mate-

rials for their enormous neutrino detector, which will weigh 15,000 tons when completed. Neutrinos come in three flavors—electron, muon and tau—and the particles constantly oscillate from one flavor to another; the NOvA detector is intended to measure how many of the muon neutrinos generated at Fermilab transform to electron neutrinos by the time they reach northern Minnesota. The results could reveal the answer to a long-standing mystery: why matter rather than antimatter dominates our universe.

Although NOvA has not been canceled, the suspension of funding may lead some of its scientists to abandon the effort. "The question is whether you can put a project on mothballs for a year and bring it back again," says Mark Messier of Indiana University, a spokesperson for NOvA. "The signal this sends is, 'Go do your research somewhere else.' "

Sharing the Pain

Fermilab isn't the only physics facility devastated by the recent budget cuts. Congress also axed the 2008 funding for the Stanford Linear Accelerator Center (SLAC), which was collaborating with Fermilab in the planning of the International Linear Collider. The cuts will force SLAC to lay off 125 employees and to prematurely end its BaBar experiment (also known as the B-factory), which is looking for violations of charge and parity symmetry in the decay of short-lived particles called B mesons.

Q&A WITH H. JEFF KIMBLE

Here and There

Why quantum teleportation is nothing like getting beamed up by Scotty BY JR MINKEL

The sci-fi dream (or utter fantasy) of getting from one place to another instantaneously continued this February 14, with the opening of Doug Liman's film *Jumper,* based on the novel by Steven Gould. We asked quantum physicist H. Jeff Kimble of the California Institute of Technology to explain how physicists understand quantum teleportation, which turns out to be more relevant to computing than to commuting.

What's the biggest misconception about teleportation in physics?

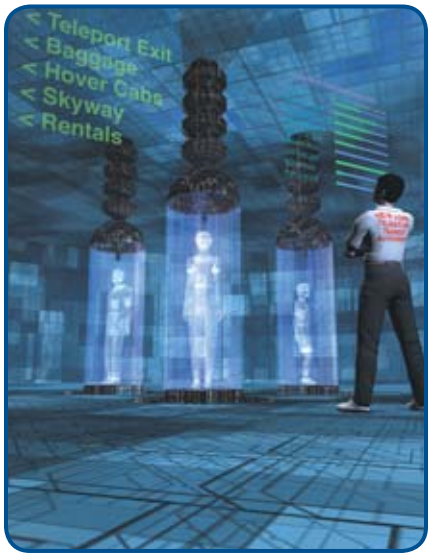
That the object itself is being sent. We're not sending around material stuff. If I wanted to send you a Boeing 757, I could send you all the parts, or I could send you a blueprint showing all the parts, and it's much easier to send a blueprint. Teleportation is a protocol about how to send a quantum state—a wave function—from one place to another.

Is that hard to do?

The most straightforward way to do it would be to imagine it was an electron: just shoot the electron from point A to point B, and it takes its quantum state with it. But that's not always so good, because the state gets messed up in the process.

How does teleportation get around that problem?

The special resource that enables teleportation is entanglement. You're Alice (location A), and I hand you an electron in an unknown quantum state. Your job is to send the quantum state (not the electron) to location B, which is Bob. If you try to measure it directly, you necessarily disturb it.



TELEPORTATION may inspire images of instant travel, but physicists today see it more as a component of quantum computers.

PHIL SAUNDERS *Space Channel Ltd.*

You and Bob also share a pair of electrons—you have one, Bob has the other—and they're in an entangled state such that if yours is spinning up, his is spinning down, and conversely. You make a joint measurement of two electrons—the one I handed you and the one you're sharing with Bob. And that gives you two bits of information. You call up Bob on the cell phone and give him those two bits, and he uses them to manipulate his electron. And bingo, in the ideal case he can perfectly re-create the state of the electron that I handed you.

Why would you want to transmit a quantum state? What are the applications?

Imagine you want to build a quantum computer. The quantum memory's got to talk to the quantum processor. Teleportation is just a fancy quantum wire.

So how has quantum teleportation advanced since *SciAm*'s 1997 story on it?

In 1998 my team demonstrated teleportation of a beam of light. I would say that

was the first bona fide demonstration. A few years ago [in 2004] a group led by David J. Wineland at the National Institute of Standards and Technology in Boulder, Colo.—and simultaneously with that, a group led by Rainer Blatt in Innsbruck, Austria—teleported the internal spin of a trapped ion. It's the first time teleportation had been done with the state of a massive particle. More recently [in 2006], the group of Eugene S. Polzik at the University of Copenhagen teleported the quantum state of light directly into a material system.

Do these demonstrations have any practical value?

It has practical implications, because a quantum computer is going to be a hybrid system. Light is good for propagating from one place to the other with very low loss, but it's really hard to store light.

Switching gears—this new movie *Jumper* is about a kid, and some other people, who teleport from place to place.

I didn't know that.

If you saw *X-Men 2,* with Nightcrawler…

I haven't seen *X-Men* either.

Do you watch *Heroes* on NBC?

No. I watch some of the football play-offs.

But you know Captain Kirk…

I have some advice. Just don't talk about teleporting people in your story. There's a really incredibly exciting frontier in science that didn't exist 15 or 20 years ago, and it's this quantum information science, which brings together traditional computer science and quantum mechanics. There's stuff going on that is just titillating.

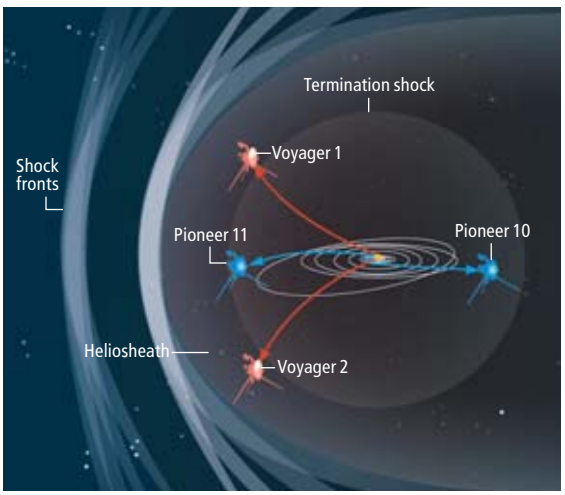
ASTRONOMY

Voyagers to the End

The solar system may be dented at the bottom BY CHRISTINA REED

Last year, after 30 years in space, Voyager 2 plowed through waves of charged particles as it crossed the termination shock, the first boundary marking the end of the solar system. As it joins its twin, Voyager 1, which is on a more northerly route to the stars, researchers have announced that, based on the two spacecraft readings, the solar system is "squashed" on one side—specifically, that the solar wind does not go as far to the south as it does to the north before being blunted and changing direction.

Astronomers define the termination shock zone as the place where the solar wind can no longer hold its own against the vast ocean of interstellar space particles. The solar wind consists of a supersonic river of charged particles traveling radially out away from the sun at 400 kilometers per second, faster than any other wave along the sun's magnetic field. (Space can transmit sonic waves, which travel in the solar system at speeds of about 50 to 70 kilometers per second; even so, in space no one can hear you scream because it is such a rarefied medium that the amplitude of any sound waves will be extremely weak.) Only as the solar wind approaches the termination shock does it begin to slow down, to 300 kilometers per second—the result of cosmic-ray particles swimming upstream into the solar wind from the heliosheath, which is defined as the immediate area on the other side of the shock.



ROUGH WINDS: As the sun plows through the galaxy, the solar wind crashes into interstellar particles, creating shock fronts. The Voyagers have crossed the termination shock (as the Pioneer craft probably did—contact has been lost) and entered the heliosheath. The artwork is not to scale.

LUCY READING-IKKANDA

At the termination shock, the solar wind cuts its speed almost in half as it falls to 150 kilometers per second and mixes with wisps of plasma coming from the wind of other stars. The result, as Voyager 2 discovered, is a surf of energetic ions. The undaunted spacecraft crested five waves of high-speed charged particles between August 30 and September 1 as it crossed the termination shock into the lull of the heliosheath. There the slower and diluted solar wind eddies back and trails in the wake of the sun's own orbit through the Milky Way galaxy. Before Voyager 2, astronomers had categorized the speed of the solar wind on the other side of the termination shock as subsonic. "One of the surprises was that the solar wind doesn't slow down as much as we expected," says Voyager mission scientist Ed Stone of the California Institute of Technology.

Like dolphins on either side of the prow of a ship, the Voyagers flank the ecliptic, cruising the front wave of the solar system. Voyager 2 hit the termination shock at a distance of 84 astronomical units (AU) from the sun—one billion miles

closer to the sun than Voyager 1's encounter with the termination shock at 94 AU in 2004 (1 AU is the average distance from Earth to the sun). The asymmetry of the termination shock indicates that for some reason the solar system is heeling to the north, exposing more of its south-facing hull to the interstellar wind. "We need to know why," Stone says.

As any good sailor knows, the winds encountered play a significant role in how a ship handles at sea. Voyager magnetometer specialist Leonard Burlaga of the NASA Goddard Space Flight Center interprets the short distance as an indication that the interstellar magnetic field is pushing harder against the southern hemisphere of the solar system. At the same time, the sun is also trimming the sails a bit through variations in the solar cycle.

Voyager 2, which unlike its twin still has a functioning plasma detector, also reported another surprise. The reduced velocity of solar wind at the termination shock should have been converted to heat. "We expected to find ions in the heliosheath with temperatures upward of one million kelvins," Stone says, "and instead the numbers were around 100,000 to 200,000—a factor of five to 10 cooler than we expected." Astronomers suspect cosmic rays may have pirated the energy for their own accelerating purposes. As Burlaga puts it, "ions bounce off the magnetic perturbations from the wind, and the energy of the solar wind ends up in these ions." How far away from the termination shock and into the heliosheath such accelerations take place remains unknown.

The answers may come in time as the Voyagers continue their race through the heliosheath. They will get help with this summer's launch of an Earth-orbiting craft designed to collect particles that come through the termination shock. In the meantime, astronomers are reconfiguring their model of the solar system. "Our current magnetohydrodynamic models do not fully describe what is happening," Stone says.

Christina Reed is based in Seattle.

Collecting Atoms from Beyond

Mapping the thickness of the heliosheath will be the goal of the new spacecraft IBEX, planned for a mid-July launch. While the Voyagers provide on-site evaluation of the heliosheath, IBEX will sweep Earth's orbit for so-called energetic neutral atoms formed in the inner heliosheath. The atoms start out with a positive charge but become neutral after stealing an electron from another particle. This neutrality enables the particles to travel straight, like Jet Skis across the solar magnetic field. They should provide a global view of how the solar system interacts with the rest of the Milky Way galaxy. Both Pioneer 10 and 11, the other two spacecraft leaving the solar system, stopped communicating short of reaching the termination shock. Ground stations last heard from Pioneer 10 in 2003 at about 82 AU away and last located Pioneer 11 in 2000 at 54 AU away.

BIOLOGY

Eye-Opening Sex

Sight returns to cavefish blind for one million years BY CHARLES Q. CHOI

BLINDERS ON: Skin grows over what remains of the eyes of a blind cavefish, shown here with two sighted relatives.

COURTESY OF NEW YORK UNIVERSITY

Anatomical features that took millennia to evolve can revert in a single generation. Specifically, sex between blind cavefish, if done right, can lead to progeny that can see.

The blind, albino, cave-dwelling form of the Mexican tetra (*Astyanax mexicanus*) evolved from ancestors living near the water's surface whose eyesight withered after they descended into complete darkness roughly one million years ago. These cavefish, which thrive in the freshwater caves of northeastern Mexico, can reach about 12 centimeters in length, and skin grows over their useless eyes.

Studies revealed that eye loss evolved independently at least three times among the cavefish; in each case, blindness resulted from mutations at several gene sites. At least some of the mutations responsible for this loss of vision differed between the 29 populations of cavefish known to exist. This mutational variety suggested that breeding different cavefish lineages together might result in offspring that could see, because the genetic deficiencies of one group might be compensated for by working genes in another.

In the January 8 *Current Biology*, evolutionary biologist Richard Borowsky of New York University found that the hybrid progeny of different cave populations all indeed had some fish with smaller than normal yet functional eyes capable of following a series of moving stripes. The farther apart the home caves of the blind parents were, the more likely their progeny had sight, Borowsky adds.

This pattern supports the notion that lineages separated by greater distances are probably more distantly related and so have less overlap in terms of the genes behind their blindness. Identifying the specific mutations underlying vision loss in the fish could help illuminate human eye development and blindness.

Charles Q. Choi is a frequent contributor.

In Brief

NOT PICTURING IT

Memory fades with age, and now imagination seems to disappear with it, too. Harvard University researchers asked volunteers in their 20s and those around 70 to construct within three minutes a future event using as much detail as possible. The younger adults created significantly richer scenarios. The results, presented in the January *Psychological Science*, support the notion that picturing what is to come requires the ability to recall past experiences and piece them together to form a coherent scenario. Imagine that—if you can. —*Philip Yam*

TESTOSTERONE TRAVAIL

As men grow older, their testosterone levels gradually but progressively wane, a decline linked with an increase in fat and drops in strength, cognition and bone mass. Unfortunately, testosterone supplements seem to do little to thwart these changes. Researchers found that testosterone supplements did decrease body fat and increase lean body mass in older men with low testosterone levels. But the subjects were no stronger and showed no improvement in mobility, cognition or bone mineral density. The January 2 *Journal of the American Medical Association* describes the outcomes. —*Charles Q. Choi*

MUTANTS FROM THE AIR

Mice kept downwind from two steel mills and a major highway developed 60 percent more mutations in their sperm than their brethren inhaling HEPA-filtered air did. The sperm stem cells became damaged after just three weeks of exposure, perhaps because of oxidative stress triggered by the particulates. Because the sperm were still functional, the mutations could be passed on to offspring. The findings show up in the January 15 *Proceedings of the National Academy of Sciences USA*. —*Philip Yam*

NANOTECH

Magnetic Control of Cells

To sense their environment, cells rely on the receptor proteins that stud their surface. These receptors latch onto specific molecules, triggering a cascade of biochemical events that lead to cell behaviors, such as the secretion of hormones or the destruction of pathogens. But before receptors can switch on, they often have to bump into one another. Donald Ingber of Harvard Medical School and his colleagues demonstrated that they could control this activation using particles of iron oxide attached to dinitrophenyl (DNP) molecules, which attach to the receptors on histamine-producing mast cells. When magnetized, the 30-nanometer-wide beads would attract one another, forcing the receptors to huddle and activate. The researchers detected a spike in the calcium levels inside the cells, which is the first step in histamine secretion. The technique could lead to lighter-weight, lower-power biosensors for detecting pathogens or to novel ways of delivering drugs in the body. The work appears in the January *Nature Nanotechnology*. —*JR Minkel*

INFECTIOUS DISEASE

New World, New Disease

New genetic evidence supports the view that Columbus introduced syphilis to Europe. The first recorded syphilis epidemic happened in 1495, fueling centuries of debate as to whether the germ came from the Americas or existed previously in the Old World but had not been distinguished from other skin-lesion diseases until 1500. To uncover syphilis's origins, scientists at Emory University and their colleagues genetically compared strains of the microbe from around the world with related bacteria. They found that syphilis's closest kin were South American variants of yaws, a disease spread by skin contact and limited to hot and humid areas. One theory suggests that syphilis became sexually transmitted only after it reached Europe, where more clothing and cooler climes limited the ways it could otherwise spread. Ultimately the progenitors of syphilis may be as old as humanity, hitching a ride with migrants to the Americas millennia before Columbus, speculate researchers in the January 15 *PLoS Neglected Tropical Diseases*. —*Charles Q. Choi*

ENERGY

Better Ethanol through Grass

For making ethanol, switchgrass appears to be a feasible choice—and a better one than corn. Working with the U.S. Department of Agriculture, farmers grew and monitored the native North American perennial, which often naturally grows on the borders of croplands. Specifically, they tracked the seed used to establish the plant, the fertilizer used to boost its growth, the fuel consumed to farm it and the overall rainfall that the areas received. The five-year study showed that switchgrass grown on plots three to nine hectares in size would yield from 5.2 to 11.1 metric tons of grass bales per hectare, depending on rainfall. If processed by appropriate biorefineries (now being built), the yields would have delivered 540 percent more energy than was used to produce them, compared with the at most 25 percent more energy returned by corn-based ethanol. The January 15 *Proceedings of the National Academy of Sciences USA* has the findings. —*David Biello*

SWITCHGRASS beats corn when it comes to making ethanol efficiently.

IMAGE SOURCE/CORBIS (*man*); CHARLIE NEIBERGALL *AP Photo* (*grass*)

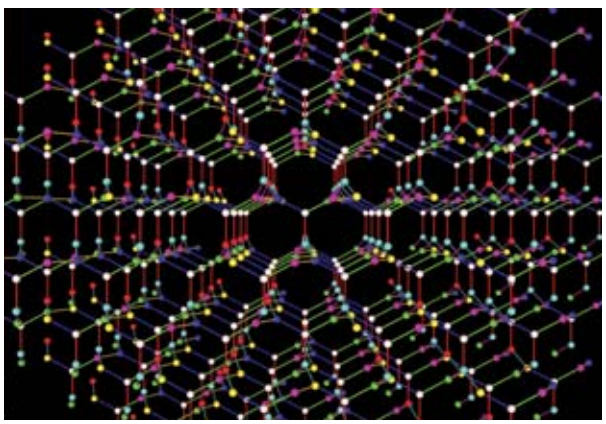
GEOMETRY

Crystal Math

Diamonds are rarities not just on earth but also mathematically. The crystal structure of diamond has two key distinguishing properties, notes mathematician Toshikazu Sunada of Meiji University in Japan. It has maximal symmetry, which means that its components cannot be rearranged to make it any more symmetrical than it is, and a strong isotropic property, which means that it looks the same when viewed from the direction of any edge. In the February *Notices of the American Mathematical Society,* Sunada finds that out of an infinite universe of crystals that can exist mathematically, just one other shares these properties with diamond. Whereas diamond is a web of hexagonal rings, its cousin is made of 10-sided rings.

Sunada had originally thought that no one had described this object before (which he had dubbed K_4). But it turns out that "I rediscovered the crystal structure mathematically in rather an accidental way" while working on another problem, Sunada says. After his paper was published, chemists and crystallographers informed him that they had long known about the crystal, which was called (10,3)-a by A. F. Wells in 1977. Diamond's mathematical twin can exist in a slightly distorted form as an arrangement of silicon atoms in strontium silicide.

—Charles Q. Choi



TWO OUT OF INFINITY: Diamond and the K4, or (10,3)-a, crystal shown cannot be any more symmetrical, and they look the same when viewed from any edge.

BEHAVIOR

Choosiness for Cooperation

To explain why cooperation with nonrelatives arises and persists in populations, British researchers developed a computational model in which players have varying degrees of cooperativeness (a willingness to allow partners to accrue benefits at their own expense) and choosiness (a willingness to leave partners based on their teamwork). After each round of play, an individual receives a payoff that reflects the effort both teammates exert. Individuals do best, however, if they manage to get their partners to do most of the work. After the payoff, players can continue together to the next round or decide to divorce the other, in which case each would be randomly paired with another partner. Because players that are repeatedly divorced get slapped with greater costs than those that stick together, cooperation and choosiness rise in tandem when many rounds are played—the equivalent of long life spans. Choose the January 10 *Nature* for the complete payoff.

—Philip Yam

Data Points

Cleared for Takeoff

Last year NASA failed to release useful data from its study on aviation safety, which the agency thought would undermine public confidence. Another study at least shows that mishaps caused by pilot error among U.S. carriers declined between 1983 and 2002. The analysis, by Susan P. Baker of Johns Hopkins University and her colleagues, traced the gains to improved training and crew communication as well as to advanced technology that provides more accurate data on such variables as aircraft position and rough weather. The overall mishap rate has remained steady, however, perhaps because of increased mistakes by ground personnel and air-traffic controllers, who must handle many more flights than in the past.

TOTAL MISHAPS ANALYZED IN STUDY: **558**

MISHAPS CAUSED BY PILOT ERROR: **180**

MISHAPS FROM OTHER CAUSES: **378**

PILOT ERROR MISHAPS PER 10 MILLION FLIGHTS IN:

1983–1987: 14.2

1998–2002: 8.5

DECLINE IN PILOT ERROR MISHAPS PER 10 MILLION FLIGHTS RELATED TO:

FLAWED DECISIONS: FROM 6.2 TO 1.8

POOR CREW INTERACTION: FROM 2.8 TO 0.9

MISHANDLED WIND OR RUNWAY CONDITIONS: FROM 2.5 TO 0.54

SOURCE: Aviation, Space, and Environmental Medicine, *January 2008*

Read More . . .

News Scan stories with this icon have extended coverage on **www.SciAm.com/ontheweb**

HISASHI NAITO (*crystal*); ILLUSTRATION BY MATT COLLINS

SciAm Perspectives

Fishing Blues

Without limits on industrial-scale catches, marine populations will continue to collapse

BY THE EDITORS

If there is any benefit to be salvaged from the disastrous overfishing of the bluefin tuna (see "The Bluefin in Peril," by Richard Ellis, on page 70), it's the spotlight that it shines on the plundering of the world's marine life. It has been 16 years since the demand for cod led to the collapse of the once superabundant cod fisheries in the North Atlantic off Newfoundland. Disappearing with them were some 40,000 jobs. Seafood Watch, an online information clearinghouse run by the Monterey Bay Aquarium in California, has placed all Atlantic populations of flatfish, including flounder, halibut, plaice and sole, on a list of fishes that it urges consumers to avoid. The list goes on.

You can't entirely blame the fishers. Yes, a lot of pirates are out there, taking fish illegally, underreporting their catches, fishing under the flags of countries not party to international fishing agreements. But for many cultures, fishing is a way of life—and sadly, because of overfishing, a hard way to carry on. The lure of dollars—or euros or yen—becomes all but irresistible when the alternatives become ever more limited. As Ellis reports, a single bluefin tuna fetched $173,600 in Tokyo, and prices of a sushi dinner for two in New York City can reach $1,000.

With that kind of money at stake, it is hardly surprising that industrial-scale technology has caught on, big time. Hooks are paid out on "long-lines" more than 50 miles in length. Factory ships that can hold 1,000 tons of fish store and process the catches. Fishing on such a massive scale can quickly exhaust a fishing ground, but when that happens, the factory ships just move on. As a result, fisheries themselves are becoming ever more remote.

The bottom of what is known as the continental slope, between 600 and 6,000 feet deep, is home to several species that swim in schools and grow as long as two to three feet. Their presence opened up the continental slope to industrial deep-sea fishing that pays off handsomely. The usual method, known as bottom trawling, is to drag a large cone-shaped net, weighted with 15 tons of gear, across the seabed. The net catches everything in its path, and the gear crushes any 1,000-year-old coral that stands in its way.

What are the environmental costs? No one really knows—and that is part of the problem. According to Richard L. Haedrich, an ichthyologist writing in a recent issue of *Natural History*, catch quotas for deep-sea fishes were set "essentially by guesswork, relying on … knowledge of shallow-water species. They took no account of the far slower turnover rates in a typical population of deep-sea fishes." The predictable result is that two deep-sea species have already been depleted: the orange roughy, formerly known as the slimehead, and the Chilean sea bass, aka Patagonian or Antarctic toothfish. When they're gone, Big Fishing will pack up and move on once again.

What is to be done? Biologists must have the chance to study fish populations before sustainability levels are set and fish are taken. Laws, treaties, police work and stiff penalties are essential to curb the pirates and keep honest fishers in business. But market forces are ultimately to blame, and market forces will determine the outcome. Consumers who vote with their pocketbooks can turn the tide of demand.

The first step is finding out what is safe to buy. Numerous Internet sites such as Seafood Watch, mentioned earlier, give basic information about the sustainability of various marine populations. The second step is determining the provenance of a fish on the market: Where does it come from, and how can you know the information is reliable? The provenance system is already in place for wine and in some countries for beef. A similar system of tracking fish from catch to consumer could drive down demand, and hence price, for endangered, uncertified products.

With provenance determined, fish lovers could harness the power of the Internet. Does a restaurant or supermarket persist in selling "red card" fish? A comment in an online review might get results. Publicity and shame are powerful tools, not to be used lightly or without warning. But those who knowingly trade in fish that are demonstrably at risk lose their right to be ignored. ■

SCI AM

Want to Take Action?

Learn more about how fish populations can be sustained. Go to **www.SciAm.com/ontheweb**

MATT COLLINS

Sustainable Developments

Climate Change after Bali

Do the math: affordable new technologies can prevent global warming while fostering growth

BY JEFFREY D. SACHS

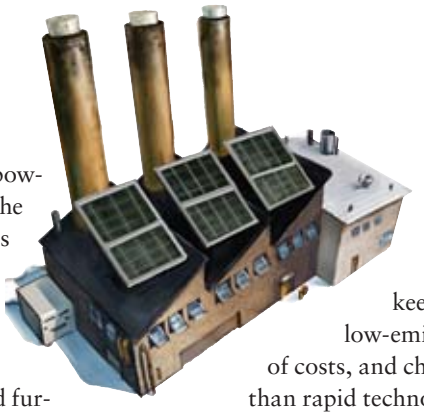
BRUCE GILBERT/EARTH INSTITUTE

Last December's agreement in Bali to launch a two-year negotiation on climate change was good news, a rare example of international cooperation in a world seemingly stuck in a spiral of conflict. Cynics might note that the only accomplishment was an agreement to talk some more, and their cynicism may yet be confirmed. Nevertheless, the growing understanding that serious climate-control measures are feasible at modest cost is welcome.

The arithmetic is becoming clearer. If the rich nations continue to grow in income and the poor ones systematically narrow the income gap with successful development, by 2050 the global economy might increase sixfold and global energy use roughly fourfold. Today's anthropogenic carbon dioxide (CO_2) emissions are around 36 billion tons annually, of which 29 billion are the result of fossil-fuel combustion and industrial processes, and another seven billion or so are the result of tropical deforestation. Roughly speaking, every 30 billion tons of emissions raises CO_2 levels by around two parts per million (ppm). The current atmospheric concentration of CO_2 is around 380 ppm, up from 280 ppm at the start of the industrial era in 1800. Thus, to arrive at 440 ppm by midcentury—a plausibly achievable "safe" level in terms of its likely climate change consequences but only 60 ppm more than the current one—cumulative emissions should be kept to roughly 900 billion tons, or roughly 21 billion tons a year on average until 2050. This goal can be achieved by ending deforestation (on a net basis) and by cutting our current fossil-fuel-based emissions by one third.

So here is the challenge. Can the world economy use four times more primary energy while lowering emissions by one third?

A promising core strategy seems to be the following: Electricity needs to be made virtually emission-free, through the mass mobilization of solar and nuclear power and the capture and sequestra-



tion of carbon dioxide from coal-burning power plants. With a clean power grid, most of the other emissions can also be controlled. In less than a decade, plug-in hybrid automobiles recharged on the grid will probably get 100 miles per gallon. Clean electricity could produce hydrogen for fuel-cell-powered vehicles and replace on-site boilers and furnaces for residential heating. The major industrial emitters could be required (or induced through taxation for tradable permits) to capture their CO_2 emissions or to convert part of their processes to run on power cells and clean electricity.

Carbon capture and sequestration at coal-fired power plants might raise costs for electricity as little as one to three cents per kilowatt-hour, according to a special report of the Intergovernmental Panel on Climate Change. The mass conversion of the U.S. to solar power might involve an incremental cost of roughly four cents per kilowatt-hour, with overall electricity costs on the order of eight to nine cents per kilowatt-hour. These incremental costs imply far less than 1 percent of the world's annual income to convert to a clean power grid. The costs in the other sectors will also be small. The fuel savings of low-emissions cars could easily pay for batteries or fuel cells. Residential heating by electricity (or cogenerated heat) rather than by home boilers will generally yield a net savings, especially when combined with improved insulation.

The Bali negotiations will succeed if the world keeps its eye on supporting the speedy adoption of low-emissions technologies. Issues of blame, allocation of costs, and choice of control mechanisms are less important than rapid technological development and deployment, backed by a control mechanism chosen by each country.

If the less polluting technologies pan out at low cost, as seems possible, the rich countries will be able to afford to clean up their own energy systems while also bearing part of the costs to enable the poor to make the needed conversions. Climate control is not a morality play. It is mainly a practical and solvable technological challenge, which, if met correctly, can be combined with the needs and aspirations for a growing global economy. ■

Jeffrey D. Sachs is director of the Earth Institute at Columbia University (www.earth.columbia.edu).

SCI AM An extended version of this essay is available at **www.SciAm.com/ontheweb**

MATT COLLINS

Forum

NASA's Flimsy Argument for Nuclear Weapons

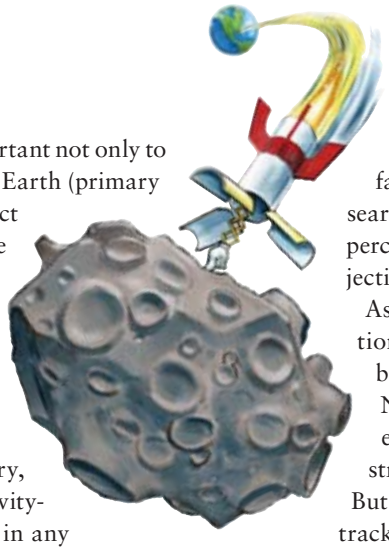
Nukes will not be needed to guard against dangers from space

BY THOMAS GRAHAM, JR., AND RUSSELL L. SCHWEICKART

On January 4, 2007, the *Wall Street Journal* published an op-ed entitled "A World Free of Nuclear Weapons," written by an impressive array of statesmen: former secretary of state George Shultz, former secretary of defense William Perry, former secretary of state Henry Kissinger and former senator Sam Nunn of Georgia. In the article the authors worried that the likelihood of international terrorists acquiring nuclear weapons is increasing. They asserted that "unless urgent new actions are taken, the U.S. soon will be compelled to enter a new nuclear era that will be more precarious, psychologically disorienting and economically even more costly than was Cold War deterrence." Invoking President Ronald Reagan's call in the 1980s for the abolition of all nuclear weapons, they endorsed "setting the goal of a world free of nuclear weapons and working energetically on the actions required to reach that goal."

Recently, however, a counterargument has been advanced—by NASA. In 2005 Congress ordered the space agency to analyze the alternatives that it could employ to divert a near-Earth object (NEO)—an asteroid or comet—if one was found to be on a collision course with our planet. Last March, NASA submitted a report entitled "Near-Earth Object Survey and Deflection Analysis of Alternatives," having first coordinated its response with the White House, the Department of Defense and the Department of Energy. In its report NASA chose to analyze only the highly improbable threat posed by large NEOs, which very rarely strike Earth, in lieu of the more realistic danger of a collision with one of the cohort of smaller NEOs, which are far more numerous. What is more, the report emphasized the effectiveness of nuclear explosions in providing the force to deflect an NEO from a collision course, but it completely neglected the need for precision in such a procedure.

COURTESY OF THOMAS GRAHAM, JR. (*top*); COURTESY OF RUSSELL L. SCHWEICKART (*bottom*)



MATT COLLINS

This analysis is seriously flawed. It is important not only to deflect an NEO from a collision course with Earth (primary deflection) but also to avoid knocking the object into a potential return orbit that would cause it to come back a few years later (secondary deflection). Nuclear explosions are not controllable in this way. But a nonnuclear kinetic impact—that is, simply smashing a spacecraft into an NEO—can provide the primary deflection for the vast majority of objects, and a precise secondary deflection, if necessary, could be performed by an accompanying gravity-tractor spacecraft, which would be needed in any event to observe the NEO deflection and its aftermath [see "Gravitational Tractor for Towing Asteroids," by Edward T. Lu and Stanley G. Love, in *Nature;* November 10, 2005].

Nuclear explosives would be needed only for deflecting the largest NEOs, which are the least common and most easily detectable objects. Scientists are not concerned about a collision with an extremely large NEO—say, 10 kilometers in diameter—because all these objects have been discovered and none currently threatens Earth. Big things are easy for astronomers to find; the smaller objects are what we have to worry about. Of the estimated 4,000 NEOs with diameters of 400 meters or more—which includes all objects that might conceivably require nuclear explosives to divert them—researchers have so far identified about 1,500. And if NASA meets the search goals mandated by Congress, it will locate 98 percent of these objects and calculate 100-year projections of their orbits by 2020.

As NASA continues to find big NEOs, the calculations of risk change accordingly. A decade ago, before astronomers began to systematically locate NEOs larger than 400 meters in diameter, they estimated that we faced a statistical risk of being struck by such an object once every 100,000 years. But now that researchers have identified and are tracking about 37 percent of these NEOs, the frequency of being hit by one of the remaining large objects has dropped to once in 160,000 years. Unless NASA finds a large NEO on an immediate collision course by 2020 (a *very* unlikely event), the frequency of a collision with one of the 80 still undiscovered objects (2 percent of 4,000) will drop to once every five million years.

Thus, the probability that nuclear explosives might be needed to deflect an NEO is extremely small. And even this minuscule probability will diminish to the vanishing point as researchers improve nonnuclear interception technologies. After 2020 the need to keep nuclear devices on standby to defend against an NEO virtually disappears. As a result, the decision to move toward the

worldwide elimination of nuclear weapons can be made strictly on the basis of human threats to global security. Extraterrestrial dangers need not be considered. ■

Thomas Graham, Jr., served as special representative of the president for arms control in the 1990s and now chairs Thorium Power Ltd., which develops proliferation-resistant reactor fuel. Russell L. Schweickart, a former astronaut who flew on Apollo 9, *heads the B612 Foundation, which champions the testing of spacecraft designs that can deflect NEOs.*

Skeptic

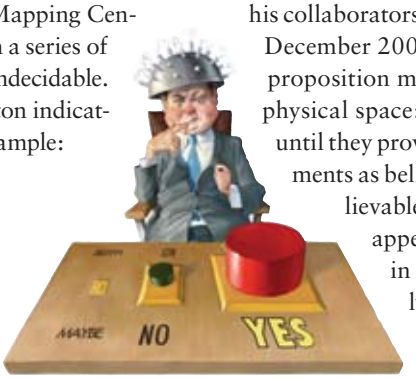
Adam's Maxim and Spinoza's Conjecture

Belief, disbelief and uncertainty generate different neural pathways in the brain

BY MICHAEL SHERMER

During an early episode of the über-pyrotechnic television series *MythBusters,* Adam Savage was busted by the camera crew for misremembering his predictions of the probability of an axle being ripped out of a car, à la *American Graffiti.* When confronted with the unmistakable video evidence of his error, Adam sardonically rejoined: "I reject your reality and substitute my own."

Skepticism is the fine art and technical science of understanding why rejecting everyone else's reality and substituting your own almost always results in a failed belief system. Where in the brain do such belief processes unfold? To find out, neuroscientists Sam Harris, Sameer A. Sheth and Mark S. Cohen employed functional magnetic resonance imaging to scan the brains of 14 adults at the University of California, Los Angeles, Brain Mapping Center. The researchers presented the subjects with a series of statements designed to be plainly true, false or undecidable. In response, the volunteers were to press a button indicating their belief, disbelief or uncertainty. For example:



Mathematical:
(2 + 6) + 8 = 16.
62 can be evenly divided by 9.
$1.2^{57} = 32608.5153$.

Factual:
Most people have 10 fingers and 10 toes.
Eagles are common pets.
The Dow Jones Industrial Average rose 1.2% last Tuesday.

Ethical:
It is bad to take pleasure at another's suffering.
Children should have no rights until they can vote.
It is better to lie to a child than to an adult.

The findings were revealing. First, there were significant reaction time differences in evaluating statements; responses to belief statements were significantly shorter than responses to both disbelief and uncertainty statements (but no difference was detected between disbelief and uncertainty statements). Second, contrasting belief and disbelief in the brain scans yielded a spike in neural activity in the ventromedial prefrontal cortex, associated with decision making and learning in the context of rewards. Third, contrasting disbelief and belief showed increased brain response in the left inferior frontal gyrus, the anterior insula and the dorsal anterior cingulate, all associated with responses to negative stimuli, pain perception and disgust. Finally, contrasting uncertainty with both belief and disbelief revealed elevated neural action in the anterior cingulate cortex, a region associated with conflict resolution.

What do these results tell us? "Several psychological studies appear to support [17th-century Dutch philosopher Benedict] Spinoza's conjecture that the mere comprehension of a statement entails the tacit acceptance of its being true, whereas disbelief requires a subsequent process of rejection," report Harris and his collaborators on the study in their paper, published in the December 2007 *Annals of Neurology.* "Understanding a proposition may be analogous to perceiving an object in physical space: We seem to accept appearances as reality until they prove otherwise." So subjects assessed true statements as believable faster than they judged them as unbelievable or undecidable. Further, because the brain appears to process false or uncertain statements in regions linked to pain and disgust, especially in judging tastes and odors, this study gives new meaning to a claim passing the "taste test" or the "smell test."

As for the neural correlates of belief and skepticism, the ventromedial prefrontal cortex is instrumental in linking higher-order cognitive factual evaluations with lower-order emotional response associations, and it does so in evaluating all types of claims. Thus, the assessment of the ethical statements showed a similar pattern of neural activation, as did the evaluation of the mathematical and factual statements. People with damage in this area have a difficult time feeling an emotional difference between good and bad decisions, and they are susceptible to confabulation—mixing true and false memories and conflating reality with fantasy.

This research supports Spinoza's conjecture that most people

PHOTOGRAPH BY BRAD SWONETZ; ILLUSTRATION BY MATT COLLINS

have a low tolerance for ambiguity and that belief comes quickly and naturally, whereas skepticism is slow and unnatural. The scientific principle of the null hypothesis—that a claim is untrue unless proved otherwise—runs counter to our natural tendency to accept as true what we can comprehend quickly. Given the chance, most of us would like to invoke Adam's Maxim because it is faster and feels better. Thus, it is that we should reward skepticism and disbelief and champion those willing to change their mind in the teeth of new evidence. ■

Michael Shermer is publisher of Skeptic *(www.skeptic.com). His new book is* The Mind of the Market.

Anti Gravity

Nothing to Sneeze At

And other interesting results from researchers with some operating room

BY STEVE MIRSKY

It's a problem faced by Yogi Berra, welders and surgeons: How do you sneeze with a mask covering your face? Catchers and welders, however, only have to deal with the unpleasant bounce-back effect. Surgeons need to worry about ejecting multitudinous microbes directly into the gaping hole they've carved in a patient. Not good. And with "uh-oh" being among the worst words a surgeon can say at work ("Where's my watch?" is also bad, as is the simple and direct "oops"), how best then to avoid an uh-oh following an achoo?

The answer to this and other pressing questions in science and medicine can be found in the year-end issue of the *British Medical Journal,* well known for its unusual array of offbeat articles. (Had the Puritans never left Britain for New England, they might later have fled the *British Medical Journal* to found the *New England Journal of Medicine*.)

First, the case of the surgical sneeze. The accepted wisdom was that the surgeon should in fact sneeze facing the area being operated on—because the mask will redirect the ejecta and send it backward out of the sides of the mask, away from the open wound. But two plastic surgeons from a British hospital checked the literature and found no actual evidence that the masked sneeze did in fact fling the phlegm sideways. They thus phlegmatically set out to test the hypothesis, using high-speed photography and some finely ground pepper to encourage sneezing by masked volunteers.

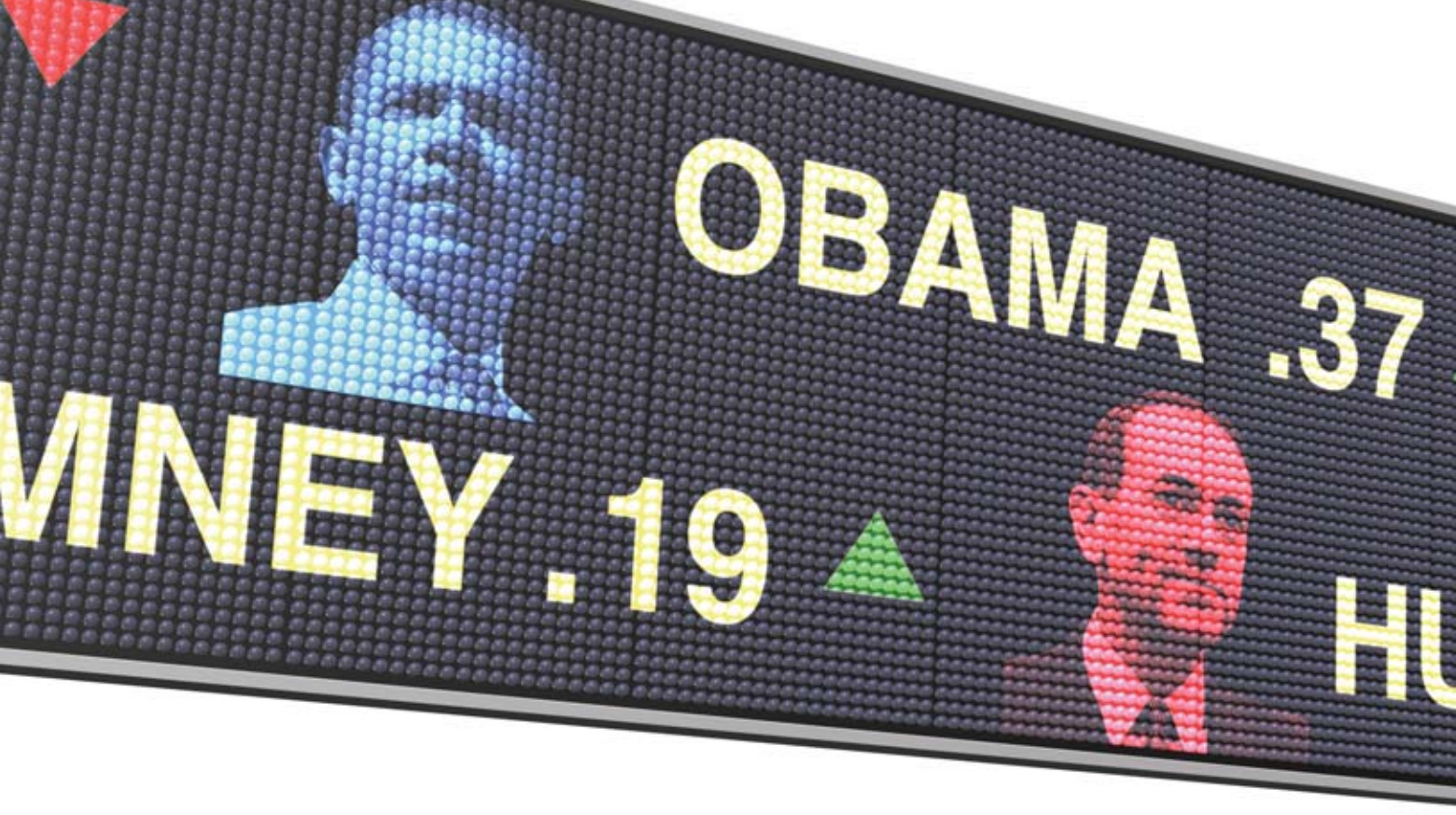
The result: very little of the blast escapes out the sides, and a bit sneaks out of the bottom, onto the surgeon's upper chest. Most of the debris appears to stay safely within the doctor's domain, leaving the patient pristine. Unable to offer any clear direction to surgeons, the authors offer these clear directions: "Surgeons should follow their instincts when sneezing during operations." One might call such instruction the gesundheit of reason.

BMJ also featured a review of commonly held medical myths that showed that the drowsiness commonly induced by the Thanksgiving meal is probably not a function of the tryptophan in the turkey. That amino acid has been getting a bad rap for years for Uncle Dave's open-belted couch coma during the second game of the turkey-day NFL doubleheader. But a given weight of turkey actually has the same amount of tryptophan as does chicken and beef, whereas pork and cheese have even more. In reality, any big meal diverts blood, and therefore oxygen, from the brain, inducing sleepiness. And, as the authors point out, "wine may also play a role."

In a short item entitled "A Day in the Life of a Doctor: The PowerPoint Presentation," two British physicians reveal that "the main purpose of a PowerPoint presentation is entertainment. Intellectual content is an unwarranted distraction." They go on to advise that "the more lines of writing that can be coerced onto a slide and the smaller the font, the lower the risk of anyone criticising any data which has accidentally been included" and that "the number of slides you can show in your allotted time is inversely proportional to the number of awkward questions which can be asked at the end."

Then there was a study that questioned the efficacy and purpose of the intensive screening of travelers at airports. The researchers, from Harvard University, the Massachusetts Institute of Technology and the Washington University School of Medicine in St. Louis, note that no scientific evaluation has ever been done of the "screening tools currently in place." They ask the arch yet brief question, "Can you hide anything in your shoes that you cannot hide in your underwear?" And they point out that spending on "airport security ($9 per passenger) is 1,000 times higher than for railway security ($0.01 per passenger), even though the number of attacks on trains is similar to that in planes." Which, they explain, is "analogous to committing mammography resources to screening only the left breast." Indeed, whenever I fly and see signs at the airport claiming that the risk of a terrorist attack that day is "high," I think, "Compared to what?" I don't *say* it out loud, of course, because I want to be allowed to board my flight. ■

PHOTOGRAPH BY FLYNN LARSEN; ILLUSTRATION BY MATT COLLINS



When Markets Beat

KEY CONCEPTS

- In 1988 the University of Iowa launched an experiment to test whether a market using securities for presidential candidates could predict the outcome of the election.
- In presidential elections from 1988 to 2004, the Iowa Electronic Markets have predicted final results better than the polls three times out of four.
- Despite the track record of the Iowa market, a fundamental understanding of how prediction markets work remains elusive, and economists are still trying to develop a body of theory to provide definitive answers.

—The Editors

In late March 1988 three economists from the University of Iowa were nursing beers at a local hangout in Iowa City, when conversation turned to the news of the day. Jesse Jackson had captured 55 percent of the votes in the Michigan Democratic caucuses, an outcome that the polls had failed to intimate. The ensuing grumbling about the unreliability of polls sparked the germ of an idea. At the time, experimental economics—in which economic theory is tested by observing the behavior of groups, usually in a classroom setting—had just come into vogue, which prompted the three drinking partners to deliberate about whether a market might do better than the polls.

A market in political candidates would serve as a novel way to test an economic theory asserting that all information about a security is reflected in its price. For a stock or other financial security, the price summarizes, among other things, what traders know about the factors influencing whether a company will achieve its profit goals in the coming quarter or whether sales may plummet. Instead of recruiting students to imitate "buyers" or "sellers" of goods and services, as in other economics experiments, participants in this election market would trade contracts that would provide payoffs depending on what percentage of the vote George H. W. Bush, Michael Dukakis or other candidates received.

If the efficient-market hypothesis, as the theory relating to securities is known, applied to contracts on political candidates as well as shares of General Electric, it might serve as a tool for discerning who was leading or trailing during a political campaign. Maybe an election market could have foretold Jackson's win. Those beer-fueled musings appear to have produced one of the most notable successes in experimental economics—and have blossomed into a subdiscipline devoted to studying prediction markets that allow investing or betting (pick the term you like best) not just on elections but on the future of climate change, movie box-office receipts and the next U.S. military incursion.

Make Your Best Bet

When the three academics—George R. Neumann, Robert Forsythe and Forrest Nelson—sought support from the university, the dean of its business college, a free-market advocate,

PHOTOILLUSTRATION BY GEORGE RETSECK; RICHARD PHIBBS (*Clinton*)

Internet-based financial markets appear to forecast elections better than polls do. They also probe how well the next George Clooney drama will do at the box office and how bad the next flu season will be

the Polls

By Gary Stix

could not contain his enthusiasm. On the other hand, the dean of the college of arts and sciences, a political scientist, characterized the proposal as "the stupidest thing he had ever heard of," Neumann recalls. "At best, it would be a shadow of the polls," he was told.

With the business school dean onboard, the three pressed forward. They wanted to use real money as an incentive for participants to take the exercise seriously. But they needed permission to allow students and faculty to gamble legally on campus. The university's general counsel resisted, but Iowa's state attorney general let the real-money market go ahead under a state law that permits office-betting pools.

The World Wide Web was still a glint in the eye of Tim Berners-Lee when the Iowa Political Stock Market opened on June 1, 1988. Nearly 200 students and faculty members began buying contracts on George H. W. Bush, Dukakis and others using the relatively primitive tools of the pre-Web Internet. A Bush or Dukakis contract was bought or sold in a futures market, the same type in which Iowa hog farmers trade pork bellies. Instead of pigs, however, the investors in the Iowa Political Stock Market were trading contracts on the share of the vote that a candidate would receive on Election Day.

Up until the morning of the election, traders carried out their transactions, although a rule stipulated that no one could invest more than $500. Taking a simplified example, a Bush contract in the vote-share market paid $0.53, corresponding to Bush's 53 percent of the vote, and a Dukakis contract paid $0.45, tied to the Democrat's popular vote percentage. If you had bought a Bush security at $0.50 before the market closed the morning of the election, you would have made a gain of $0.03 [*see box on next two pages*].

To the three economists, finding out who won or lost money—or the election—was less important than whether this exercise answered the question posed in the barroom: Would the expected share of the votes represented by the market's closing prices on Election Day match the actual share the candidates obtained more closely than the polls would? The experiment worked. The final market price corresponded to Bush's and Dukakis's market shares better than Gallup, Harris, CBS/*New York Times* and three other major polls.

In 1992 the Iowa Political Stock Market was

[A PRIMER]

HOW TO TRADE FOR PRESIDENT

The Iowa Electronic Markets allow anyone with an Internet connection and $5, even a trader in Dhaka or Novosibirsk, to buy and sell securities in elections. The example below depicts, in simplified form, how a market run during the 2004 presidential election operated. A trader initially purchases

MARKET OPENS
Trading has just begun on the Iowa Electronic Markets with futures contracts that will provide payoffs based on a percentage of the vote a candidate receives.

BUYING A PORTFOLIO
IEM sells an individual portfolio for $1 each, which consists of one contract for both candidates. Joe Citizen decides to buy a portfolio.

JOE'S REASONING
Joe looks at the prices on the IEM and decides that they are out of line with what they should be. Kerry's chances of winning mean that his contract should be priced higher than Bush's.

redubbed the Iowa Electronic Markets (IEM), and trading was opened to anyone from Dubuque to Beijing who could come up with the requisite minimum of $5. The Commodity Futures Trading Commission (CFTC) had granted the University of Iowa an exemption from regulation because the IEM is mainly run for research purposes (only minor sums are transacted).

The election exchange has continued to beat the polls consistently for presidential elections and at times has prevailed in congressional and international races. A paper being prepared for publication by several Iowa professors compares the performance of the IEM as a predictor of presidential elections from 1988 to 2004 with 964 polls over that same period and shows that the market was closer to the outcome of an election 74 percent of the time. The market, moreover, does better than the polls at predicting the outcome not just around Election Day but as long as 100 days before [*see box on page 42*].

The IEM will never be the New York Stock Exchange. But even with the CFTC trading restrictions, it has flourished. The number of contracts traded expanded from 15,286 in 1988 (a dollar volume of $8,123) to 339,222 ($46,237) in the 2004 elections. And another IEM market that furnishes a payoff only to those who picked the winner of an election had even more activity in 2004 (1,106,722 contracts totaling $327,385). Television commentators have recognized this new barometer of voter sentiment by sometimes mentioning market prices in the months running up to an election. The IEM's status has risen among those who contribute to the incessant blog-based chatter that has become a cornerstone of contemporary political discourse. And after a spike in trading during the 2004 election, the IEM office received e-mails charging that über-financier George Soros was trying to manipulate the market to create a bandwagon effect for Democratic presidential candidate John Kerry, an assertion for which there was never any proof.

HOW YOU CAN PLAY

IEM | Iowa Electronic Markets

Anyone with $5 to invest can trade on the Iowa Electronic Markets. Don't call your broker, though. Trades can only be placed by going to the IEM Web site (www.biz.uiowa.edu/iem), where detailed instructions can be found about how to buy a futures contract for Hillary Clinton or John McCain.

The How and Why

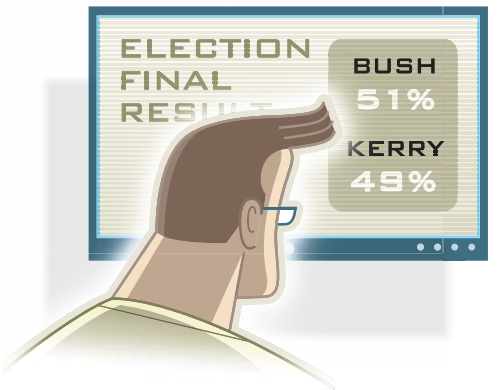
The IEM continues to serve not only as a forecasting tool but as an energizing environment for students to learn about markets and, perhaps most important, as a testing ground for experimental economists to probe theories of how and why markets appear to make accurate predictions. Its track record provides arguably the best empirical evidence to date to justify the case for prediction markets. But when researchers have tried to backtrack, looking for theories of why markets serve as effective means of forecasting, straightforward answers have not been forthcoming. Some of these analyses have even called into question the basic assumption that a market does a good job of foretelling what lies ahead.

At first, the idea that a market can prophesize the outcome of an election does not seem partic-

SAMUEL VELASCO *5W Infographics*, WITH ILLUSTRATION ASSISTANCE BY JOHN O. LEDYARD *California Institute of Technology*

a portfolio for $1, thus obtaining one contract for Bush and one for Kerry. After the election, the payout is based on the share of the vote each candidate receives. Despite the small sums transacted, market prices before an election have proved to be surprisingly good at predicting a race's outcome.

CONSOLIDATING HIS POSITION
Joe chooses immediately to sell the Bush contract that he thinks is priced too high and buy an additional Kerry, which he considers undervalued.



THE BIG MOMENT
If Kerry did get the expected 60 percent, Joe's holding would be worth $1.20 ($0.60 × 2), netting him $0.20 above the $1 he spent. But tallies (rounded off) show Bush wins.

THE PAYOUT
IEM pays $0.98 for Joe's two Kerry contracts ($0.49 × 2). Joe loses money. After subtracting the net cost of $1 from the payout, Joe has a net loss of $0.02.

ularly startling. After all, the chairman of the Federal Reserve or the chief economist at Goldman Sachs will routinely look at the price of stocks or commodities as a guide to making forecasts about the economy, and the futures market for orange juice concentrate predicts Florida weather better than the National Weather Service does.

Developers of the IEM and other prediction markets contrast a poll with a market by saying that the latter takes a reading not of whom people are going to vote for but of whom they think will win—and cash wagered indicates the strength of those beliefs. You might have voted for Kerry in the 2004 election because you opposed the Iraq War, but after watching news shows and talking to neighbors, you may have decided that George W. Bush was going to win. When putting money down, you might have picked Bush.

The question, though, of how one individual's belief—that IBM's stock will rise or that a Bush will be elected—gets combined with those of every other trader and then translated into a price that is an accurate predictor continues to provoke heated debates in the research community. Economic theoreticians have yet to understand precisely why this novel means of forecasting elections should work better than well-tested social science methods.

On close inspection, the characteristics of IEM traders would drive a statistician batty. Early on it became clear that the traders are by no means a representative sampling of the population at large, the prerequisite for any poll. And a survey of them in the 2004 presidential election market underscored the point: most were found to be well educated, affluent, white, male Republicans who tended to have a high opinion of their own political insight into the face-off between Bush and Kerry, a grouping that does not fit the definition of a well-designed sample. In about

DOUG BENTON *Fisheye*

FOUNDING FATHERS of the Iowa Electronic Markets—George R. Neumann, Forrest Nelson and Robert Forsythe—came up with the idea for trading on elections in a bar in Iowa City after Democratic candidate Jesse Jackson's unexpected victory in the 1988 Michigan caucuses, which the polls failed to predict.

[A LONG TRACK RECORD]

MARKETS VS. POLLS

The Iowa Electronic Markets have usually been more accurate than the polls in predicting candidates' share of the vote in presidential elections. The table shows whether the poll or market was most accurate for each of the polls taken for U.S. presidential races beginning in 1988. In 2004, for instance, the polls were more accurate overall in 110 instances and the market trumped the polls 258 times.

All days from the beginning of the market

MOST ACCURATE	1988	1992	1996	2000	2004	All
Poll	25	43	21	56	110	255
Market	34	108	136	173	258	709
Market percentage	**58%**	**72%**	**87%**	**76%**	**70%**	**74%**

More than 100 days before election

Poll	1	20	3	2	66	92
Market	13	49	30	47	129	268
Market percentage	**93%**	**71%**	**91%**	**96%**	**66%**	**74%**

Last 5 days before election

Poll	0	1	4	8	12	25
Market	6	5	7	17	18	53
Market percentage	**100%**	**83%**	**64%**	**68%**	**60%**	**68%**

DIFFERENCES BETWEEN A POLL AND A MARKET

POLL

- Takes a representative sample
- Indicates a margin of error
- Expresses voter preferences for a particular candidate on the day of the poll

MARKET

- Accepts anyone who wants to trade
- Relies only on the fluctuation of prices
- Uses prices to represent the probability of a candidate winning or receiving a given percentage of the vote on Election Day
- Provides, unlike a poll, a monetary incentive to make the best choice

one in five transactions, traders had no personal opinions or beliefs at all about the Swift Boat smear campaign or prisoners being held in Guantánamo. Rather those buying or selling were "robots"—automated trading programs that buy and sell when the software perceives that a security is too high or low. Automated programs routinely execute trades on Wall Street. And IEM election market researchers are still plumbing what a machine's trading patterns add to the market's ability to deduce the outcome of an election.

As early as the aftermath of the 1988 presidential race, the Iowa team began to probe deeply into why the IEM seems to predict election outcomes with such precision. Discounting pure luck and the possibility that traders somehow constitute a representative sample of the population, the team analyzed trading patterns and found a select group of "marginal traders" who would buy and sell actively when the share price was not valued properly. This group might have bought, say, Bush securities if the price was way under what the members thought was the likely percentage of votes the Republican would attract.

These traders were the Warren Buffetts of the 1988 race, investing an average of $56, twice the level of less active participants who might have simply bought and held contracts for the candidate they liked best, without making a careful judgment about that candidate's prospects. The wallflowers would typically make nothing from their trades, whereas marginal traders took home 9.6 percent returns (a whopping $5.38; the reason such small sums act as an incentive to traders—or the use of play money in other markets—is also closely studied).

The identification of marginal traders, described in a 1992 paper in the *American Economic Review*, has sometimes elicited phone calls from Wall Street types interested in new insight into the perennial question of the traits of a person who can beat the market. Other than noting that most of those investing are male, the Iowa researchers have not succeeded in identifying more specific qualities of this special class of trader.

One possibility is that they do not exist. James Surowiecki, a *New Yorker* columnist who wrote *The Wisdom of Crowds*, a book first published in 2004 that brought attention to prediction markets and other novel means of group decision making, thinks that the marginal trader is a myth. No individual or subgroup in a market has the financial wherewithal to sway prices in the way the marginal-trader hypothesis suggests—an opinion that is echoed by some economists.

Just a Word Argument

Perhaps the most incisive critique of prediction markets has come from Charles F. Manski, an econometrician at Northwestern University whose academic research focuses on how people assign probabilities to future events, such as the possibility that they might lose their job. Manski started wondering a few years ago about the theoretical basis for statements made repeatedly in the popular press that markets can predict an election better than polls and experts can.

Advocates of prediction markets often invoke Austrian-born economist Friedrich Hayek, who argued in 1945 that prices aggregate information held by a group—"dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess." That knowledge is combined into a price that expresses the relative desirability of a commodity or public sentiment at a given moment, whether it be a pork belly or a candidate for the U.S. presidency. Manski went back to Hayek's original work to examine the quantitative underpinnings of his ideas. No hard numbers supported the notion of the collective wisdom of crowds. "It's a very loose argument," he says. "There's no theory in the modern sense of the word. It's just a word argument."

So Manski set out to explore whether he could build a mathematical model that would confirm Hayek's notion of the market as an information

SAMUEL VELASCO *SW/Infographics*

aggregation mechanism and, secondarily, bolster the empirical findings taken from the IEM. Manski created a model of a diverse group of traders using the IEM's winner-take-all market in which a trader buys a contract for a candidate that pays $1 for a victory and nothing for a loss. If the market worked in accordance with the way that proponents of prediction markets have interpreted Hayek, the price would represent the average, or mean, value of traders' belief that a particular candidate would win. A Kerry contract selling for $0.49 would mean that there would be a 49 percent probability that Kerry would win.

But Manski's model did not confirm this conjecture. In many instances, the mean did not necessarily coincide with the price and could even diverge sharply, a finding suggesting that the market would not serve as a particularly accurate prediction tool. If, for instance, the price was $0.50, the mean of traders' beliefs could be anything from a 25 to 75 percent chance that Kerry would win. Manski remarks that even if the price and the mean were the same, it would not be certain that the mean would correspond to a reasonable probability of a candidate's chances.

Manski is a respected economist, and his finding caused a minor furor because it appeared to contradict an emerging consensus about the value of these markets for making predictions about anything from elections to public policy. But two subsequent papers offered a way to reconcile the dispute. They also compared prices with the mean but factored in a variable called risk aversion—which measures how traders react to uncertainty in the market. In the revised model, said by the authors to offer a more realistic scenario, the price and the mean were about the same, which seemed to confirm that a price is, indeed, a good measure of a probability.

But the debate has never been resolved, and exactly how the markets achieve success remains unclear. Manski, for his part, suspects that his critics' models do not account for all the actual ways prediction markets operate in the real world. "There isn't going to be a simple interpretation of that market price that always works as a prediction," he observes. "It really depends on the beliefs and the attitudes toward risk of those trading." Manski also remains unsatisfied with the IEM's proponents' reliance on its record of consistently besting the polls. "Comparison to the polls is not the best comparison," he says. "Everyone knows there are all kinds of problems with the polls, and they're just one piece of information." In fact, Manski notes, IEM traders may be taking the polls into account as one of many factors in making decisions about when to buy or sell.

Oft-cited statistics about election markets beating the polls have come under scrutiny from other quarters. A 2005 analysis by political scientists Robert S. Erikson of Columbia University and Christopher Wlezien of Temple University insisted that polls and election markets do not serve the same functions and so do not merit direct comparison. The authors contended that the polls identify vote preferences on the day each poll is taken, whereas the IEM market prices forecast what is to happen on the day of the election. In their analysis, they made a series of mathematical adjustments to the polls, which they then found to be more accurate in projecting Election Day outcomes than both the IEM's vote-share and winner-take-all markets.

Analyses by some economists have contradicted the emerging consensus about the value of markets for making predictions for anything from elections to public policy.

HULTON-DEUTSCH COLLECTION/CORBIS

BETTING on elections was ubiquitous in the early part of the 20th century on informal exchanges and among individuals. Instead of paying off a losing bet with cash, the losers—consider these unfortunates who bet on John W. Davis instead of Calvin Coolidge in 1924—had to sometimes perform stunts. In this instance, Davis supporters had to pull Muriel Gordon, the winner of the bet, in a hansom cab down Fifth Avenue.

Controversy again ensued. One dissenter, Justin Wolfers, an economist at the Wharton School of the University of Pennsylvania who has done extensive analyses of prediction markets, criticized Erikson and Wlezien's results, saying that their study only compared a few elections and polls. Wolfers also objects because the 2005 analysis "adjusts polls but doesn't make a corresponding adjustment of prediction markets."

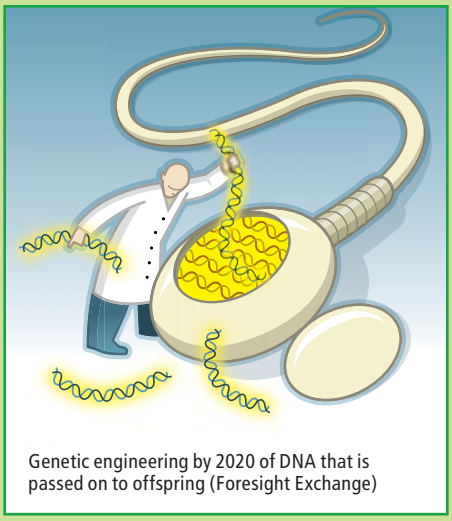
The Triumph of the Market

It will take years to put these debates to rest. In spite of persistent wrangling, the IEM has inspired formation of other prediction markets, many of them outside an academic setting. On

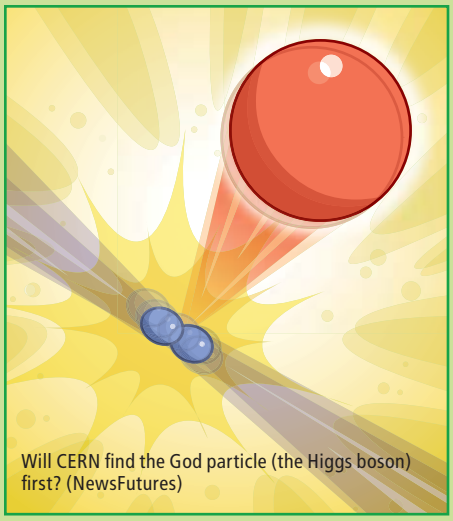
PICK YOUR MARKET

The Iowa Electronic Markets have inspired the establishment of other prediction markets that do not

Chinese moon landing by 2020 (Foresight Exchange)



Genetic engineering by 2020 of DNA that is passed on to offspring (Foresight Exchange)



Will CERN find the God particle (the Higgs boson) first? (NewsFutures)

the Hollywood Stock Exchange, traders speculate on box-office sales for new movies. NewsFutures trades in current events. Some markets allow traders to buy and sell securities on the prospects for new ideas or technologies. Without the CFTC exemption accorded to the IEM, other U.S. markets use virtual play money on the Internet. In Ireland, which lacks similar restrictions, TradeSports and Intrade, both part of the same company, accept real cash for trading on sports, elections or other events. Intrade, for instance, provides a contract that will furnish a payoff if the U.S. or Israel executes an air strike against Iran by March 31. Another contract will provide recompense if the U.S. economy goes into recession during 2008.

The place accorded markets in U.S. society, along with the revolution in new forms of information sharing afforded by the World Wide Web, has meant that prediction markets are now being increasingly adopted as innovative decision-making tools in both government and private institutions. The ardor for market-based answers can at times border on the hyperbolic. Robin Hanson, a professor of economics at George Mason University, has advocated that if trading patterns on prediction markets suggest that implementation of a particular policy will cause the economy to grow and unemployment to shrink, then policy officials should, by fiat, adopt that policy—an interest rate cut or a public works project, perhaps. Hanson reasons that the collective information held by traders is superior to the analyses that can be marshaled by a panel of economists or other experts. Hanson has even proposed a form of government called futarchy, based on policy-making markets.

The idea of a "terrorist futures market" proved repulsive, but new markets that assist in corporate or government decision making continue to emerge.

Such utopian leanings have sometimes led advocates to push too far too fast. Several years ago the Defense Advanced Research Projects Agency (DARPA) began planning for a project called the Policy Analysis Market, which would have allowed investors to trade on geopolitical events, not unlike the Intrade Iran contract, including assassinations, wars and the next al-Qaeda attack. If the market—for which Hanson was an adviser—bid up a contract that would pay off if a terrorist attack occurred, the Department of Homeland Security might then decide to raise the threat condition status from yellow to red.

Or so went the rationale. The idea of a "terrorist futures market" repulsed many in Washington, and the market died quickly, even forcing the resignation of DARPA head John Poindexter (but not before TradeSports launched a market to speculate on the prospects of his ouster). Senator Barbara Boxer of California fumed when she learned about the Policy Analysis Market: "There is something very sick about it."

But not everyone experienced the same distaste. Some argued that a prediction market able to serve as an efficient intelligence-gathering mechanism just might avert a pending crisis. Writing in the *Washington Post*, Wolfers and his colleague Eric Zitzewitz speculated that a contract on whether Niger had made a sale of uranium to Saddam Hussein would have been trading at low levels in early 2003, reflecting the actual intelligence consensus that the transaction never occurred and thereby undercutting one of the Bush administration's rationales for going to war in Iraq.

The attacks on the Policy Analysis Market ultimately doomed the project, although the hoop-

SAMUEL VELASCO *SW/Infographics*

trade on elections but rather on virtually any conceivable event, from the bombing of Iran to whether a network anchor is about to be fired.

Katie Couric departs from CBS News (Intrade)


Central Park's seasonal snowfall to be more than 60 inches (Intrade)


Human-to-human transmission of avian flu (Avian Influenza Prediction Market)

la managed to boost public awareness of prediction markets. DARPA's project became an informal tutorial that broadened public awareness of prediction markets. "It actually took the DARPA thing to get people's attention," comments Joyce Berg, a professor of accounting and IEM's interim director.

New types of markets intended to assist in formulating government or internal corporate decision making have continued to emerge. Here again the University of Iowa has been a leader. Its markets for predicting influenza outbreaks serve as an example. In one, which ran for seven months, beginning in mid-September 2004, an IEM spinoff sold influenza futures contracts to a set of 62 health care professionals in Iowa to predict influenza activity for each week of the flu season. If a contract for the third week of January accurately forecast flu prevalence—gauged by a Centers for Disease Control and Prevention scale (ranked as no activity, sporadic, local, regional or widespread)—it would pay $1. The market accurately predicted the beginning, the peak and the end of the influenza season two to four weeks ahead of the CDC reports on influenza activity.

"Prediction markets will never replace traditional surveillance systems, but they may provide an efficient and relatively inexpensive source of information to supplement existing disease surveillance systems," says Philip M. Polgreen, a physician and professor at the University of Iowa's Carver College of Medicine, who helped to run the market. The university has more recently begun a market, in collaboration with Pro-MED mail, an electronic disease-reporting system, that is intended to predict events related to the H5N1 "bird flu" virus.

Attracted by the markets' apparent soothsaying powers, companies such as Hewlett-Packard (HP), Google and Microsoft have established internal markets that allow employees to trade on the prospect of meeting a quarterly sales goal or a deadline for release of a new software product. As in other types of prediction markets, traders frequently seem to do better than the internal forecasts do.

HP has refined the running of prediction markets to make them effective for groups that might be too small to make accurate predictions. Before a market is launched, HP gauges the expertise level of participants and their attitude toward risk—factors that are then used to mathematically adjust the predictions made when participants place their bets on some future outcome. "Our mechanism basically distills the wisdom of the crowd from a very small group," says Bernardo Huberman, director of the social computing laboratory at HP. This filtering process achieves better results than does a market alone or the predictions of the most knowledgeable members of the group.

The burgeoning interest in prediction markets evokes the prepoll era of the early 20th century, when betting on election results was ubiquitous. Newspapers would routinely run stories on the odds for a particular candidate, reports that often proved to be surprisingly prescient. In that sense, prediction markets may truly hark back to the future. "My long-run prediction is that newspapers in 2020 will look like newspapers in 1920," Wharton School's Wolfers says. If that happens, the wisdom of crowds will have arrived at a juncture that truly rivals the musings of the most seasoned pundits. ■

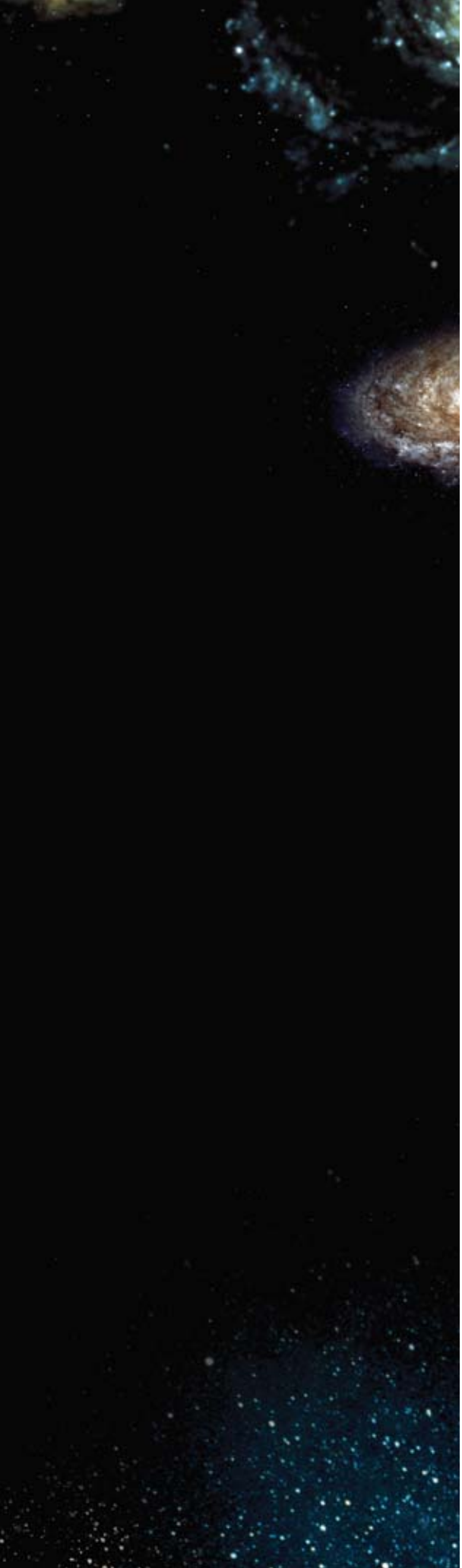
MORE TO EXPLORE

The Wisdom of Crowds. James Surowiecki. Anchor Books, 2005.

Information Markets: A New Way of Making Decisions. Edited by Robert W. Hahn and Paul C. Tetlock. AEI-Brookings Joint Center for Regulatory Studies, Washington, D.C., 2006. Available at **www.aei-brookings.org/admin/authorpdfs/page.php?id=1261**

Infotopia: How Many Minds Produce Knowledge. Cass R. Sunstein. Oxford University Press, 2006.

the End of Cosmology?



LONELY PLANET: As space empties out because of the quickening cosmic expansion, the galaxy that Earth inhabits will come to be surrounded by a total void.

SLIM FILMS

An accelerating universe wipes out traces of its own origins

By Lawrence M. Krauss and Robert J. Scherrer

One hundred years ago a *Scientific American* article about the history and large-scale structure of the universe would have been almost completely wrong. In 1908 scientists thought our galaxy constituted the entire universe. They considered it an "island universe," an isolated cluster of stars surrounded by an infinite void. We now know that our galaxy is one of more than 400 billion galaxies in the observable universe. In 1908 the scientific consensus was that the universe was static and eternal. The beginning of the universe in a fiery big bang was not even remotely suspected. The synthesis of elements in the first few moments of the big bang and inside the cores of stars was not understood. The expansion of space and its possible curvature in response to matter was not dreamed of. Recognition of the fact that all of space is bathed in radiation, providing a ghostly image of the cool afterglow of creation, would have to await the development of modern technologies designed not to explore eternity but to allow humans to phone home.

It is hard to think of an area of intellectual inquiry that has changed more in the past century than cosmology, and the shift has transformed how we view the world. But must science in the future always reflect more empirical knowledge than existed in the past? Our recent work suggests that on cosmic timescales, the answer is no. We may be living in the only epoch in the history of the universe when scientists can achieve an accurate understanding of the true nature of the universe.

A dramatic discovery almost a decade ago

KEY CONCEPTS

- A decade ago astronomers made the revolutionary discovery that the expansion of the universe is speeding up. They are still working out its implications.
- The quickening expansion will eventually pull galaxies apart faster than light, causing them to drop out of view. This process eliminates reference points for measuring expansion and dilutes the distinctive products of the big bang to nothingness. In short, it erases all the signs that a big bang ever occurred.
- To our distant descendants, the universe will look like a small puddle of stars in an endless, changeless void.
- What knowledge has the universe already erased?

—The Editors

motivated our study. Two different groups of astronomers traced the expansion of the universe over the past five billion years and found that it appears to be speeding up. The source of this cosmic antigravity is thought to be some new form of "dark energy" associated with empty space. Some theorists, including one of us (Krauss), had actually anticipated this new result based on indirect measurements, but in physics it is direct observations that count. The acceleration of the universe implies that empty space contains almost three times as much energy as all the cosmic structures we observe today: galaxies, clusters and superclusters of galaxies. Ironically, Albert Einstein first postulated such a form of energy to keep the universe static. He called it the cosmological constant [see "Cosmological Antigravity," by Lawrence M. Krauss; Scientific American, January 1999].

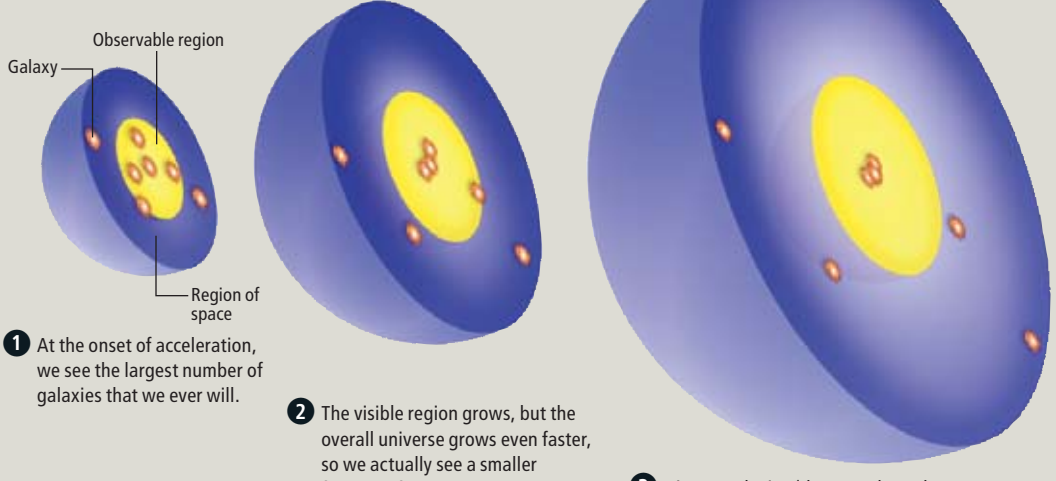
Dark energy will have an enormous impact on the future of the universe. With cosmologist Glenn Starkman of Case Western Reserve University, Krauss explored the implications for the fate of life in a universe with a cosmological constant. The prognosis: not good. Such a universe becomes a very inhospitable place. The cosmological constant produces a fixed "event horizon," an imaginary surface beyond which no matter or radiation can reach us. The universe comes to resemble an inside-out black hole, with matter and radiation trapped outside the horizon rather than inside it. This finding means that the observable universe contains only a finite amount of information, so information processing (and life) cannot endure forever [see "The Fate of Life in the Universe," by Lawrence M. Krauss and Glenn D. Starkman; Scientific American, November 1999].

Long before this information limit becomes a problem, all the expanding matter in the universe will be driven outside the event horizon. This process has been studied by Abraham Loeb and Kentaro Nagamine, both then at Harvard University, who found that our so-called Local Group of galaxies (the Milky Way, Andromeda and a host of orbiting dwarf galaxies) will collapse into a single enormous supercluster of stars. All the other galaxies will disappear into the oblivion beyond the event horizon. This process takes about 100 billion years, which may seem long but is fairly short compared to the wilderness of eternity.

STEVE GREEN *Vanderbilt University (Krauss and Scherrer)*; SLIM FILMS *(sphere illustrations)*

EXPANDING UNIVERSE, SHRINKING VIEW

The universe may be infinite, but consider what happens to the patch of space around us (*purple sphere*), of which we see only a part (*yellow inner sphere*). As space expands, galaxies (*orange spots*) spread out. As light has time to propagate, we observers on Earth (or our predecessors or descendants) can see a steadily increasing volume of space. About six billion years ago, the expansion began to accelerate, carrying distant galaxies away from us faster than light.



1 At the onset of acceleration, we see the largest number of galaxies that we ever will.

2 The visible region grows, but the overall universe grows even faster, so we actually see a smaller fraction of what is out there.

3 Distant galaxies (those not bound to us by gravity) move out of our range of view. Meanwhile, gravity pulls nearby galaxies together.

NOTE:
Because space is expanding uniformly, alien beings in other galaxies see this same pattern.

[THE AUTHORS]

Lawrence M. Krauss (*right*) **and Robert J. Scherrer** (*left*) **began working together two years ago, when Krauss spent a sabbatical year at Vanderbilt University and came to know every honky-tonk in Nashville. Krauss is a cosmologist at Case Western Reserve University and director of its Center for Education and Research in Cosmology and Astrophysics. He is the author of seven books and an activist for the public understanding of science. Scherrer is a cosmologist, chair of the Department of Physics and Astronomy at Vanderbilt and a published science-fiction author. They both enjoy doing cosmology while there is still time left.**

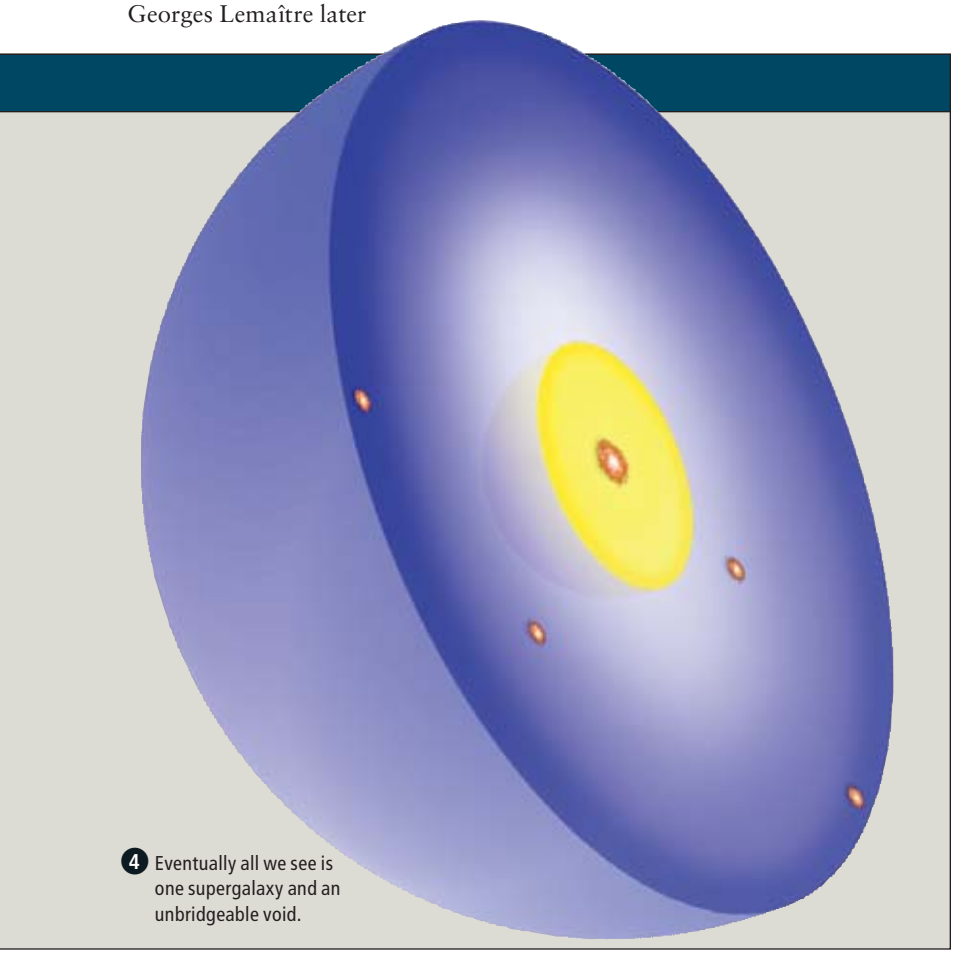
Collapsing Pillars

What will astronomers of the far future, living in this supercluster, conclude about the history of the universe? To think about this question, recall the pillars on which our current understanding of the big bang is based.

The first is Einstein's general theory of relativity. For nearly 300 years Newton's theory of universal gravitation served as the basis for almost all of astronomy. Newton's theory does an excellent job of predicting the motions of objects on scales from the terrestrial to the galactic, but it is completely incapable of dealing with infinitely large collections of matter. General relativity overcomes this limitation. Shortly after Einstein published the theory in 1916, Dutch physicist Willem de Sitter solved the equations of general relativity for a simplified universe incorporating Einstein's cosmological constant. De Sitter's work appeared to reproduce the prevailing view of the universe at the time: an island galaxy embedded in a largely empty, static void.

Cosmologists soon realized that the stasis was a misinterpretation. In fact, the de Sitter universe is eternally expanding. As Belgian physicist Georges Lemaître later made clear, Einstein's equations predict that an infinite, homogeneous, static universe is impossible. The universe has to expand or contract. From this realization, the big bang theory, as it would later be called, was born.

In 100 billion years, Hubble's crucial discovery of the expanding universe will become irreproducible.



4 Eventually all we see is one supergalaxy and an unbridgeable void.

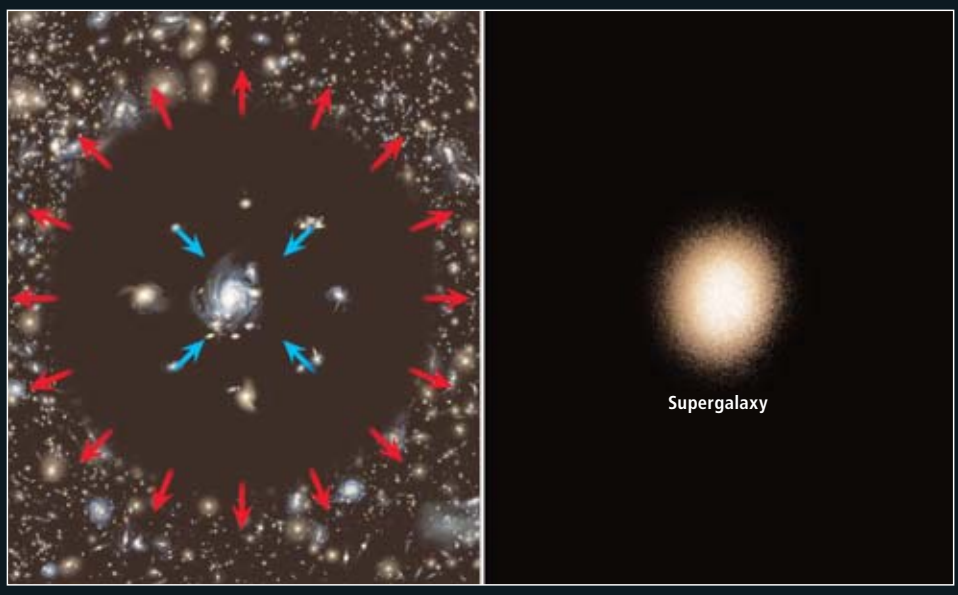
The next pillar came in the 1920s, when astronomers detected the expansion of the universe. The first person to provide observational evidence for expansion was American astronomer Vesto Slipher, who used the spectra of stars to measure the velocities of nearby galaxies. Waves of light from a star moving toward Earth are compressed, shortening the wavelength and making the light bluer. Light waves from an object moving away from us are stretched, making the wavelength longer and the light redder. By measuring the lengthening or compression of the light waves from distant galaxies, Slipher was able to determine whether they were moving toward us or away from us and at what speed. (At the time, astronomers were not even sure whether the fuzzy patches of light that we call "galaxies" were actually independent bodies of stars or simply gas clouds inside our own galaxy.) Slipher found that almost all these galaxies were moving away from us. We seemed to be sitting at the center of a runaway expansion.

The person who is generally credited for discovering the expansion of the universe is not Slipher but American astronomer Edwin Hubble. (When was the last time you read about the Slipher Space Telescope?) Hubble determined not just the velocities of nearby galaxies but also their distances. His measurements led to two conclusions that justify his fame. First, Hubble showed that galaxies were so far away that they really were independent collections of stars, just like our own galaxy. Second, he discovered a simple relation between the distance to galaxies and their velocities. The velocity was directly proportional to its distance from us: a galaxy twice as far away as another was moving twice as fast. This relation between distance and velocity is exactly what happens when the universe is expanding. Hubble's measurements have since been refined, most recently by the observations of distant supernovae, which led to the discovery of dark energy.

The third pillar is the faint glow of the cosmic microwave background, discovered serendipitously in 1965 by Bell Labs physicists Arno Penzias and Robert Wilson as they tracked down sources of radio interference. This radiation was quickly recognized to be a relic left over from the

THE APOCALYPSE OF KNOWLEDGE

The accelerating cosmic expansion is beginning to undermine the three observational pillars of the big bang theory: the motion of galaxies away from one another, the cosmic microwave background radiation, and the relative quantities of light chemical elements such as hydrogen and helium.



TODAY all three pillars are prominent. We see distant galaxies recede from us (*red arrows*) as nearby ones pull tighter (*blue*); background radiation suffuses space; and cosmic gas largely retains the chemical mix produced early in the big bang.

BILLIONS OF YEARS LATER nearby galaxies have merged and distant ones have receded from view. The background radiation is undetectably dilute. Multiple generations of stars have contaminated the original chemical mix.

COSMIC AMNESIA

The current accelerating cosmic expansion is not the only way that the universe destroys records of its past.

COSMIC INFLATION

Expansion probably accelerated early in cosmic history as well, erasing almost all traces of the preexisting universe, including whatever transpired at the big bang itself.

BLACK HOLES

These cosmic sinkholes swallow not only things but also the information those things embody. This information may be lost forever.

QUANTUM MEASUREMENTS

Whenever we measure a quantum system, we typically force it into a particular state, erasing evidence of the many possible configurations the object may have been occupying.

early stages of the expansion of the universe. It indicates that the universe began hot and dense and has since cooled and thinned out.

The final observational pillar of the big bang is that the hot, dense early universe was a perfect location for nuclear fusion. When the temperature of the universe was one billion to 10 billion kelvins, lighter nuclei could fuse into heavier nuclei, a process known as big bang nucleosynthesis. This process can occur for only a few minutes as the universe expands and cools, so fusion was limited to the lightest elements. Most of the helium in the universe was produced then, as was deuterium, or heavy hydrogen. The measured abundances of helium and deuterium match the predictions of big bang nucleosynthesis, providing further evidence for the theory as well as an accurate estimate of the abundance of protons and neutrons in the universe.

Dark Skies

What will the scientists of the future see as they peer into the skies 100 billion years from now? Without telescopes, they will see pretty much what we see today: the stars of our galaxy. The largest and brightest stars will have burned up their nuclear fuel, but plenty of smaller stars will still light up the night sky. The big difference will occur when these future scientists build telescopes capable of detecting galaxies outside our own. They won't see any! The nearby galaxies will have merged with the Milky Way to form one large galaxy, and essentially all the other galaxies will be long gone, having escaped beyond the event horizon.

The disappearance of distant galaxies is not immediate but gradual. The redshift of these galaxies becomes infinitely large as they approach the horizon. Krauss and Starkman calculated that this redshift will exceed 5,000 for all galaxies by 100 billion years, rising to an unfathomable 10^{53} by 10 trillion years—at which time even the highest-energy cosmic rays will have redshifted so much that their wavelength will be larger than the horizon size. These objects will then be truly and completely invisible to us.

As a result, Hubble's crucial discovery of the expanding universe will become irreproducible.

DON DIXON

All the expanding matter in the universe will have visually disappeared beyond the horizon, and everything remaining will be part of a gravitationally bound cluster of stars. For these future astronomers, the observable universe will closely resemble the "island universe" of 1908: a single enormous collection of stars, static and eternal, surrounded by empty space.

Our own experience demonstrates that even when we have data, the correct cosmological model is not so obvious. For example, from the 1940s to the mid-1960s, with the edifice of observational cosmology resting only on Hubble's discovery of the expanding universe, some astrophysicists resurrected the idea of an eternal universe: the steady-state universe, in which matter is created as the universe expands, so that the universe as a whole does not really change with time. This idea proved to be an intellectual dead end, but it does demonstrate the kind of mistaken notion that can develop in the absence of adequate observational data.

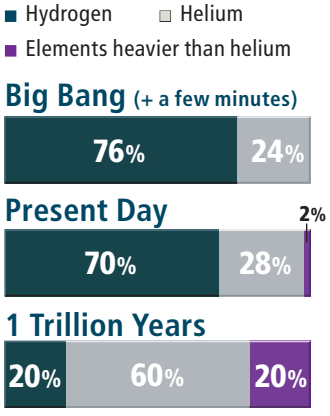
Where else might astronomers of the future search for evidence of the big bang? Would the cosmic microwave background allow them to probe the dynamics of the universe? Alas, no. As the universe expands, the wavelengths of the background radiation stretch and the radiation becomes more diffuse. When the universe is 100 billion years old, the peak wavelengths of the microwave radiation will be on the scale of meters, corresponding to radio waves instead of microwaves. The intensity of the radiation will be diluted by a factor of one trillion and might never be seen.

Even further into the future, the cosmic background will become truly unobservable. The space between stars in our galaxy is filled with an ionized gas of electrons. Low-frequency radio waves cannot penetrate such a gas; they are absorbed or reflected. A similar effect is the reason that AM radio stations can be picked up far from their cities of origin at night; the radio waves reflect off the ionosphere and back down to the ground. The interstellar medium can be thought of as one big ionosphere filling the galaxy. Any radio waves with frequencies below about one kilohertz (a wavelength of greater than 300 kilometers) cannot penetrate into our galaxy. Radio astronomy below one kilohertz is forever impossible inside our galaxy. When the universe is about 25 times its present age, the microwave background will be stretched beyond this wavelength and become undetectable to the residents of the galaxy. Even before then, the subtle patterns in this background radiation, which have provided so much useful information to today's cosmologists, will become too muted to study.

If this article survives in an archive, it might be the only way future civilizations will know about the big bang. Whether they will believe it is another question.

LOSING THE CHEMICAL CLUES

The universe consists almost entirely of hydrogen and helium produced in the first three minutes of the big bang. Stars have converted some of the hydrogen to helium, but not much—yet. The relative quantities of these elements have been an essential observational clue to the big bang. In the future, as stars continue to convert material, that will no longer be the case.



LISA APFELBACHER

Burning Up

Would observations of the abundances of chemical elements lead cosmologists of the distant future to a knowledge of the big bang? Once again, the answer is likely to be no. The problem is that our ability to probe big bang nucleosynthesis hinges on the fact that the abundances of deuterium and helium have not evolved very much since they were produced 14 billion years ago. Helium produced in the early universe, for example, makes up about 24 percent of the total matter. Although stars produce helium in the course of their fusion reactions, they have increased this abundance by no more than a few percent. Astronomers Fred Adams and Gregory Laughlin of the University of Michigan at Ann Arbor have suggested that this fraction could increase to as much as 60 percent after many generations of stars. An observer in the distant future would find the primordial helium swamped by the helium produced in later generations of stars.

Currently the cleanest probe of big bang nucleosynthesis is the abundance of deuterium. Our best measurements of the primordial deuterium abundance come from observations of hydrogen clouds backlit by quasars, extremely distant and bright beacons thought to be powered by black holes. In the far future of the universe, however, both these hydrogen clouds and quasars will have passed beyond the event horizon and will be forever lost to view. Only galactic deuterium might be observable. But stars destroy deuterium, and little will survive. Even if astronomers of the future observe deuterium, they might not ascribe it to the big bang. Nuclear reactions involving high-energy cosmic rays, which have been studied today as a possible source of at least some of the observed deuterium, might seem more plausible.

Although the observational abundance of light elements will not provide any direct evidence for a fiery big bang, it will nonetheless make one aspect of future cosmology different from the illusory cosmology of a century ago. Astronomers and physicists who develop an understanding of nuclear physics will correctly conclude that stars burn nuclear fuel. If they then conclude (incorrectly) that all the helium they observe was produced in earlier genera-

FADE TO BLACK

The night sky on Earth (assuming it survives) will change dramatically as our Milky Way galaxy merges with its neighbors and distant galaxies recede beyond view.

DIFFUSE BAND stretching across the sky is the disk of the Milky Way. A few nearby galaxies, such as Andromeda and the Magellanic Clouds, are visible to the naked eye. Telescopes reveal billions more.

ANDROMEDA has been moving toward us and now nearly fills the sky. The sun swells to red giant size and subsequently burns out, consigning Earth to a bleak existence.

COSMIC MILESTONES

10^{-30} second
Cosmic inflation occurs

100 seconds
Deuterium and helium are created

400,000 years
Microwave background is released

8 billion years
Expansion begins to accelerate

13.7 billion years
Today

20 billion years
Milky Way and Andromeda collide

100 billion years
All other galaxies are invisible

1 trillion years
Primordial isotopes are lost or diluted

100 trillion years
Last star burns out

tions of stars, they will be able to place an upper limit on the age of the universe. These scientists will thus correctly infer that their galactic universe is not eternal but has a finite age. Yet the origin of the matter they observe will remain shrouded in mystery.

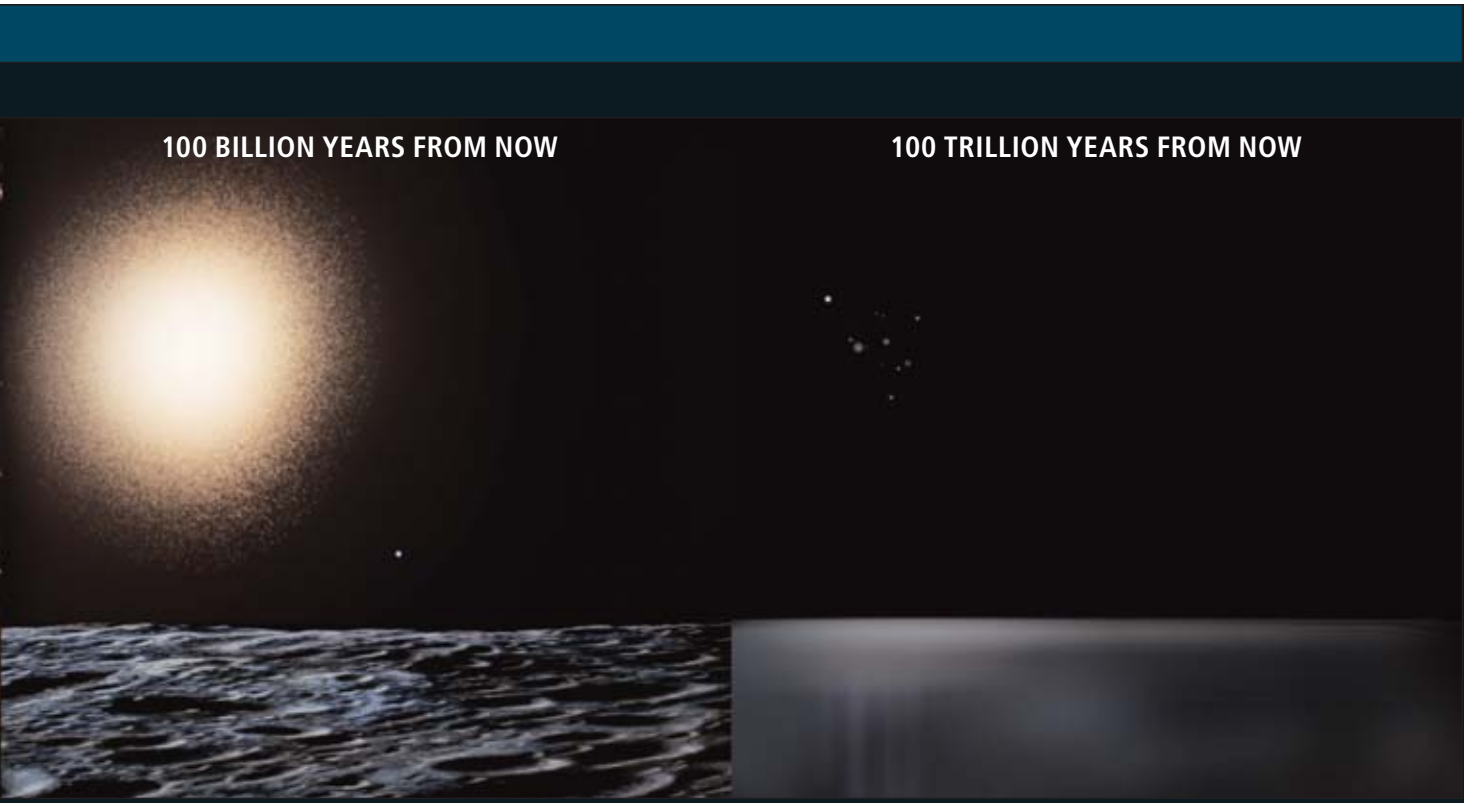
What about the idea with which we began this article, namely that Einstein's theory of relativity predicts an expanding universe and therefore a big bang? The denizens of the far future of the universe should be able to discover the theory of general relativity from precision measurements of gravity in their own solar system. Using this theory to infer a big bang, however, rests on observations about the large-scale structure of the universe. Einstein's theory predicts an expanding universe only if the universe is homogeneous. The universe that our descendants survey will be anything but homogeneous. It will consist of an island of stars embedded in a vast emptiness. It will, in fact, resemble de Sitter's island universe. The ultimate future of the observable universe is to collapse into a black hole, precisely what will in fact occur to our galaxy in the distant future.

Alone in the Void

Is there no way at all for our descendants to perceive an expanding universe? One telltale effect of acceleration would indeed remain within our observational horizon, at least according to our current understanding of general relativity. Just as the event horizon of a black hole emits radiation, so, too, does our cosmological event horizon. Yet the temperature associated with this radiation is unmeasurably small, about 10^{-30} kelvin. Even if astronomers were able to detect it, they would probably attribute it to some other, far larger local source of noise.

Ambitious future observers might also send out probes that escape the supergalaxy and could serve as reference points for detecting a possible cosmic expansion. Whether it would occur to them to do so seems unlikely, but in any event it would take billions of years at the very least for the probe to reach the point where the expansion noticeably affected its velocity, and the probe would need the energy output comparable to that of a star to communicate back to its builders from such a great distance. That the science-funding agencies of the future would

DON DIXON



SUCCESSOR to the Milky Way is a ball-like supergalaxy, and Earth may float forlornly through its distant outskirts. Other galaxies have disappeared from view.

LIGHTS OUT: The last stars burn out. Apart from dimly glowing black holes and any artificial lighting that civilizations have rigged up, the universe goes black. The galaxy later collapses into a black hole.

support such a shot-in-the-dark is unlikely, at least if our own experience is any guide.

Thus, observers of the future are likely to predict that the universe ultimately ends with a localized big crunch, rather than the eternal expansion that the cosmological constant produces. Instead of a whimper, their limited universe will end with a bang.

We are led inexorably to a very strange conclusion. The window during which intelligent observers can deduce the true nature of our expanding universe might be very short indeed. Some civilizations might hold on to deep historical archives, and this very article might appear in one—if it can survive billions of years of wars, supernovae, black holes and countless other perils. Whether they will believe it is another question. Civilizations that lack such archives might be doomed to remain forever ignorant of the big bang.

Why is the present universe so special? Many researchers have tried to argue that the existence of life provides a selection effect that might explain the coincidences associated with the present time [see "The Anthropic Principle," by George Gale; Scientific American, December 1981]. We take different lessons from our work.

First, this would quite likely not be the first time that information about the universe would be lost because of an accelerating expansion. If a period of inflation occurred in the very early universe, then the rapid expansion during this era drove away almost all details of the preexisting matter and energy out of what is now our observable universe. Indeed, one of the original motivations for inflationary models was to rid the universe of pesky cosmological objects such as magnetic monopoles that may once have existed in profusion.

More important, although we are certainly fortunate to live at a time when the observational pillars of the big bang are all detectable, we can easily envisage that other fundamental aspects of the universe are unobservable today. What have we already lost? Rather than being self-satisfied, we should feel humble. Perhaps someday we will find that our current careful and apparently complete understanding of the universe is seriously wanting. ■

MORE TO EXPLORE

Life, the Universe and Nothing: Life and Death in an Ever-Expanding Universe. Lawrence Krauss and Glenn Starkman in *Astrophysical Journal,* Vol. 531, No. 22, pages 22–30; March 2000. Available at **www.arxiv.org/abs/astro-ph/9902189**

The Five Ages of the Universe: Inside the Physics of Eternity. Fred C. Adams and Greg Laughlin. Free Press, 2000.

Atom: A Single Oxygen Atom's Journey from the Big Bang to Life on Earth … and Beyond. Lawrence M. Krauss. Back Bay Books, 2002.

The Return of a Static Universe and the End of Cosmology. Lawrence M. Krauss and Robert J. Scherrer in *Journal of General Relativity and Gravitation,* Vol. 39, No. 10, pages 1545–1550; October 2007. **www.arxiv.org/abs/0704.0221**

White Matter

Although scientists have long regarded the brain's white matter as passive infrastructure, new work shows that it actively affects learning and mental illness

By R. Douglas Fields

KEY CONCEPTS

- White matter, long thought to be passive tissue, actively affects how the brain learns and dysfunctions.
- Although gray matter (composed of neurons) does the brain's thinking and calculating, white matter (composed of myelin-coated axons) controls the signals that neurons share, coordinating how well brain regions work together.
- A new type of magnetic resonance technology, called diffusion tensor imaging (DTI), has for the first time shown white matter in action, revealing its underappreciated role.
- Myelin is only partially formed at birth and gradually develops in different regions throughout our 20s. The timing of growth and degree of completion can affect learning, self-control (and why teenagers may lack it), and mental illnesses such as schizophrenia, autism and even pathological lying.

—The Editors

Imagine if we could peek through the skull to see what makes one brain smarter than another. Or to discover whether hidden traits might be driving a person's schizophrenia or dyslexia. A new kind of imaging technique is helping scientists observe such evidence, and it is revealing a surprise: intelligence, and a variety of mental syndromes, may be influenced by tracts within the brain made exclusively of white matter.

Gray matter, the stuff between your ears your teachers chided you about, is where mental computation takes place and memories are stored. This cortex is the "topsoil" of the brain; it is composed of densely packed neuronal cell bodies—the decision-making parts of nerve cells, or neurons. Underneath it, however, is a bedrock of "white matter" that fills nearly half of the human brain—a far larger percentage than found in the brains of other animals. White matter is composed of millions of communications cables, each one containing a long, individual wire, or axon, coated with a white, fatty substance called myelin. Like the trunk lines that connect telephones in different parts of a country, this white cabling connects neurons in one region of the brain with those in other regions.

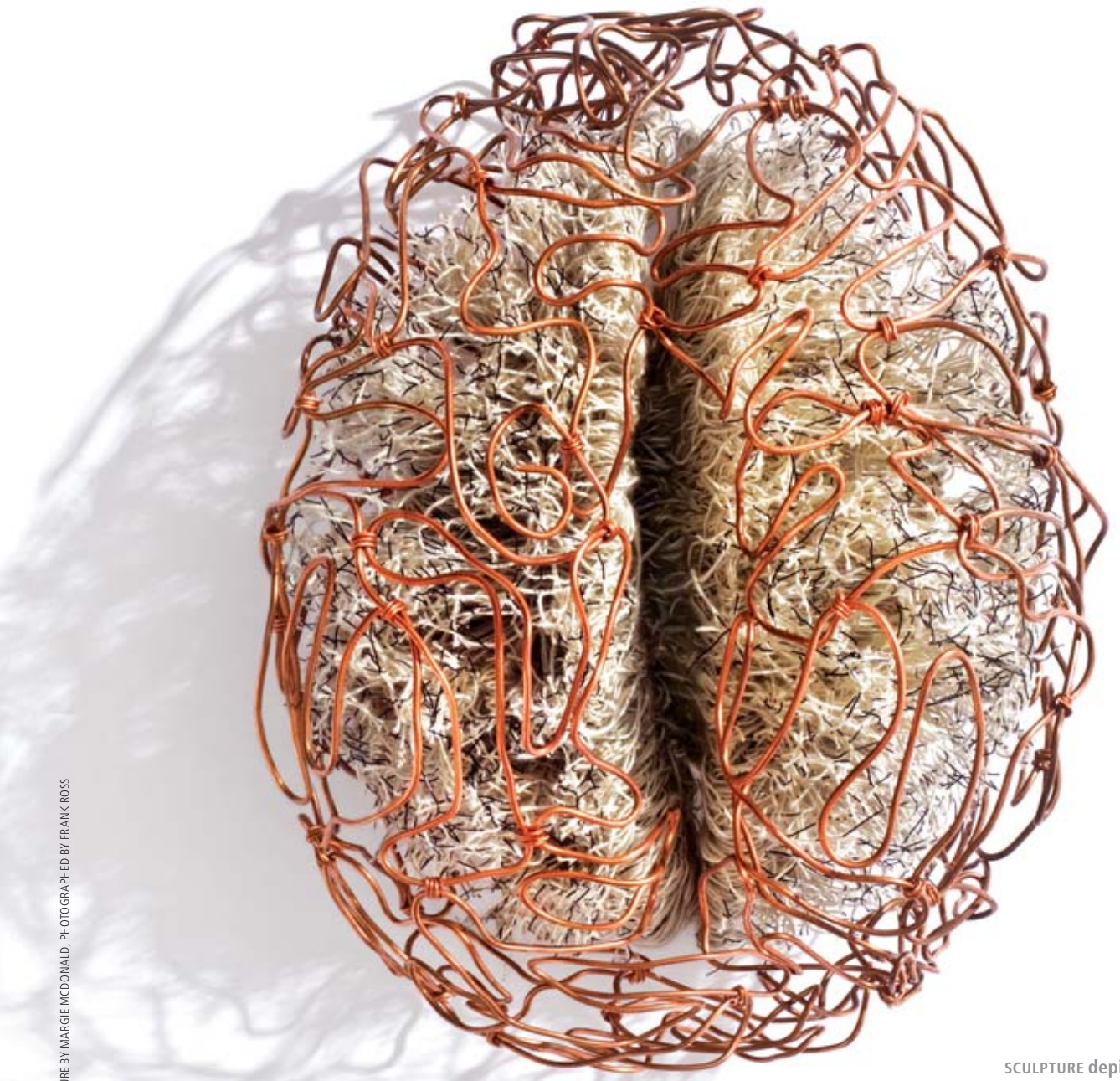
For decades neuroscientists exhibited little interest in white matter. They considered the myelin to be mere insulation and the cables inside it little more than passive passageways. Theories about learning, memory and psychiatric disorders centered on molecular action inside the neurons and at the famous synapses—the tiny contact points between them. But scientists are now realizing that we have underestimated the importance of white matter in the proper transfer of information among brain regions. New studies show that the extent of white matter varies in people who have different mental experiences or who have certain dysfunctions. It also changes within one person's brain as he or she learns or practices a skill such as playing the piano. Even though the neurons in gray matter execute mental and physical activities, the functioning of white matter may be just as critical to how people master mental and social skills, as well as to why it is hard for old dogs to learn new tricks.

More with Mastery

The myelin that gives white matter its color has always posed mysteries. For more than a century scientists looked at neurons through their microscopes and saw long fibers, the axons, extending from a neuronal cell body to a neighboring one, like an outstretched, elongated finger. Each axon was found to be coated with a thick crystalline gel. Anatomists surmised that the fatty covering must insulate axons like rubber sheathing along a copper wire. Strangely, however, many axons, especially the smaller filaments, were not coated at all. And even along insulated fibers, gaps in the insulation appeared every millimeter or so. The bare spots came to be known as nodes of Ranvier, after French anatomist Louis-Antoine Ranvier, who first described them.

Modern investigation has revealed that nerve impulses race down axons on the order of 100 times faster when they are coated with myelin—

Matters



SCULPTURE BY MARGIE MCDONALD, PHOTOGRAPHED BY FRANK ROSS

SCULPTURE depicts overhead view of brain's cortex (*copper*) and white matter core.

and that myelin is laid on axons somewhat like electrical tape, wrapped up to 150 times between every node. The substance is manufactured in sheets by two types of glial cells. These cells are not neurons, but they are prevalent in the brain and nervous system [see "The Other Half of the Brain," by R. Douglas Fields; SCIENTIFIC AMERICAN, April 2004]. An octopus-shaped glial cell called an oligodendrocyte does the wrapping. Electrical signals, unable to leak out through the sheath, jump swiftly down the axon from node to node. In nerves outside the brain and spinal cord, a sausage-shaped glial cell called a Schwann cell forms myelin.

Without myelin, the signal leaks and dissipates. For maximum conduction velocity, the insulation thickness must be strictly proportional to the diameter of the fiber inside. The optimal ratio of bare axon diameter divided by the total fiber diameter (including the myelin) is 0.6. We have no idea how oligodendrocytes "know" whether 10 or 100 layers of insulation are required to create the proper thickness on axons of different diameters. But recently biologist Klaus-Armin Nave of the Max Planck Institute for Experimental Medicine in Göttingen, Germany, discovered that Schwann cells detect a protein called neuregulin that coats axons, and if the amount of this protein is augmented or inhibited, the Schwann cell will wrap more or fewer sheets of myelin around the axon. Interestingly, many people who suffer bipolar disorder or schizophrenia have a defect in the gene that regulates production of this protein.

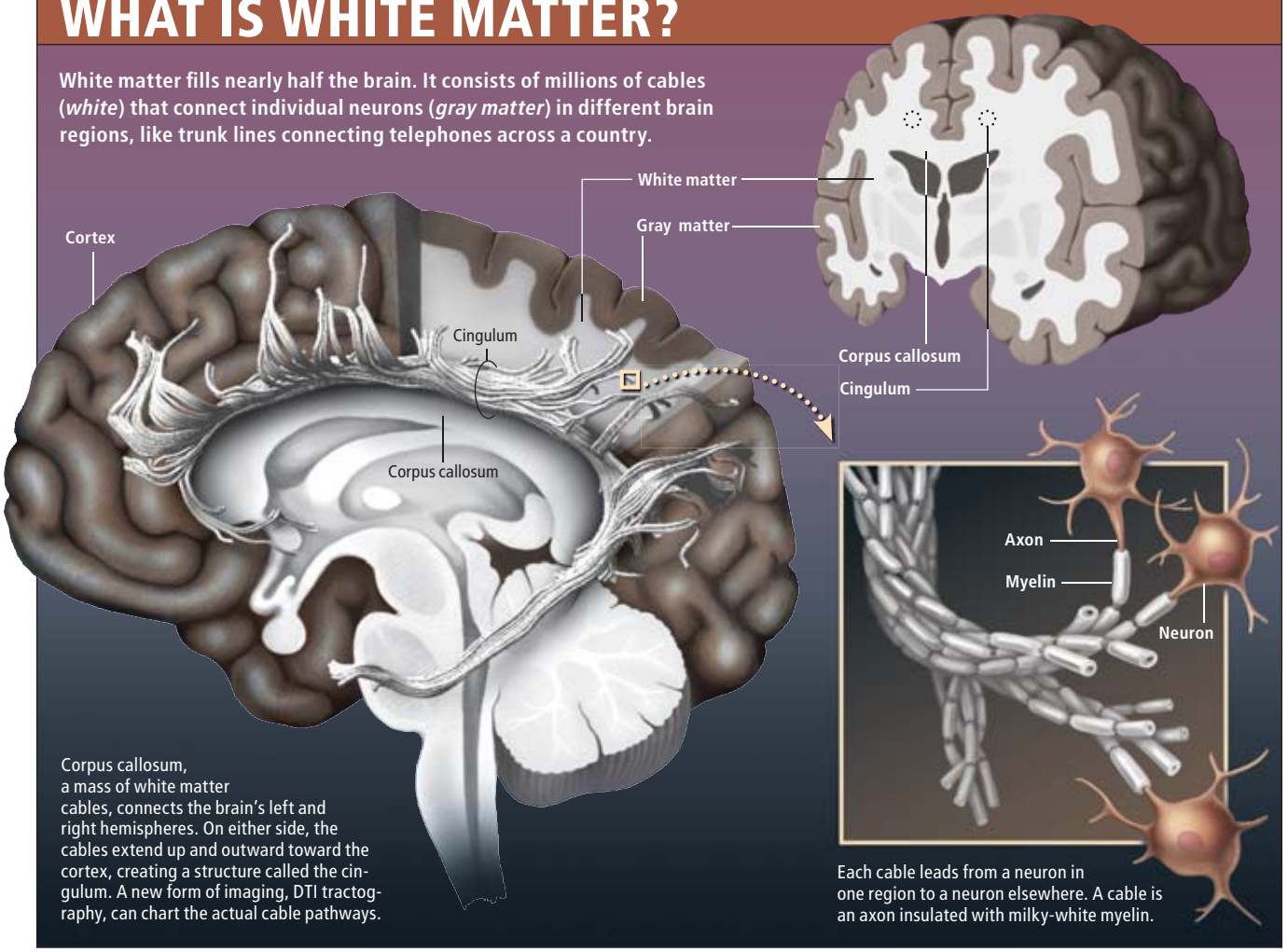
Myelin is laid down until age 25 or so, one reason teenagers do not have adult decision-making abilities.

The wrapping occurs at different ages. Myelin is prevalent only in a few brain regions at birth, expands in spurts and is not fully laid until age 25 or 30 in certain places. Myelination generally proceeds in a wave from the back of the cerebral cortex (shirt collar) to its front (forehead) as we grow into adulthood. The frontal lobes are the last places where myelina-

[BASICS]

WHAT IS WHITE MATTER?

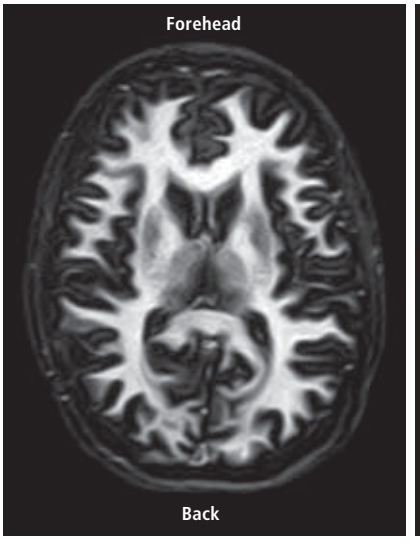
White matter fills nearly half the brain. It consists of millions of cables (*white*) that connect individual neurons (*gray matter*) in different brain regions, like trunk lines connecting telephones across a country.

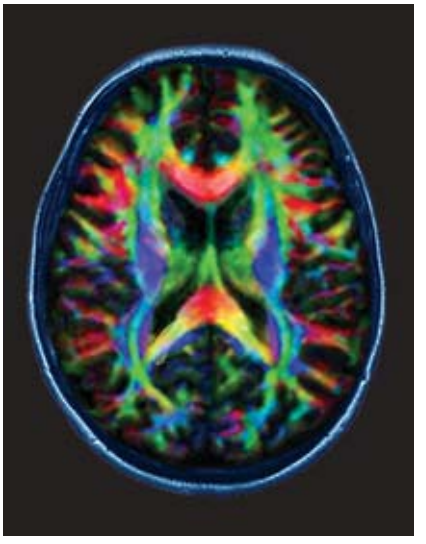


Corpus callosum, a mass of white matter cables, connects the brain's left and right hemispheres. On either side, the cables extend up and outward toward the cortex, creating a structure called the cingulum. A new form of imaging, DTI tractography, can chart the actual cable pathways.

Each cable leads from a neuron in one region to a neuron elsewhere. A cable is an axon insulated with milky-white myelin.

JEN CHRISTIANSEN (*illustration*); DEREK JONES *Cardiff University (cingulum tractography)*





CONVENTIONAL MRI machine (*top*) can roughly depict white matter (*bottom left, white areas*). But a new MRI process called DTI shows structure in greater detail (*bottom right*); red and yellow indicate more highly organized white matter.

tion occurs. These regions are responsible for higher-level reasoning, planning and judgment—skills that only come with experience. Researchers have speculated that skimpy forebrain myelin is one reason that teenagers do not have adult decision-making abilities. Such observations suggest that myelin is important to intelligence.

Presumably the brain does not finish wrapping human axons until early adulthood because, throughout that time, axons continue to grow, gain new branches and trim others in response to experience. Once axons are myelinated, the changes they can undergo become more limited. Still, for a long time a question remained: Is myelin formation totally programmed, or do our life experiences alter the degree of wrapping and thus how well we learn? Does myelin actually build cognitive ability, or is cognition simply limited in regions where it has not yet formed?

Piano virtuoso Fredrik Ullén decided to find out. Ullén also happens to be an associate professor at the Stockholm Brain Institute in Sweden. In 2005 he and his colleagues used a new brain-scanning technology called diffusion tensor imaging (DTI) to investigate the brains of professional pianists. DTI is done with the same kind of magnetic resonance imaging machines found in hospitals but involves a different type of magnetic field and different algorithms to create the many brain-image slices that are assembled into a three-dimensional picture. The slices display the vectors (mathematically defined as tensors) of water that diffuses in tissue. In gray matter the DTI signals are low because water diffuses symmetrically. But water diffuses asymmetrically along bundles of axons; this irregular pattern illuminates white matter, exposing the major highways of information that flow among brain regions. The more tightly packed and heavily coated with myelin the fibers are, the stronger the DTI signal.

Ullén found that in professional pianists, certain white matter regions are more highly developed than in nonmusicians. These regions connect parts of the cerebral cortex that are crucial to coordinated movement of the fingers with areas involving other cognitive processes that operate when making music.

He also found that the more hours a day a musician had practiced over time, the stronger the DTI signals were in these white matter tracts; the axons were more heavily myelinated or tightly packed. Of course, the axons could simply have expanded, requiring more myelin to maintain the optimal 0.6 ratio. Without performing an autopsy, the question remains open. The discovery is important, however, because it shows that when learning a complex skill, noticeable changes occur in white matter—a brain structure that contains no neuronal cell bodies or synapses, only axons and glia. Studies on animals, in which brains can be physically examined, show myelin can change in response to mental experience and a creature's developmental environment. Recently neurobiologist William T. Greenough of the University of Illinois at Urbana-Champaign confirmed that rats raised in "enriched" environments (with access to abundant toys and social interaction) had more myelinated fibers in the corpus callosum—the hefty bundle of axons that connects the brain's two hemispheres.

[THE AUTHOR]

R. Douglas Fields is chief of the Nervous System Development and Plasticity Section at the National Institute of Child Health and Human Development. He has written several articles for *Scientific American*, most recently in the August 2007 issue on how sharks zero in on prey by sensing extremely weak electrical fields emitted by fish.

COURTESY OF R. DOUGLAS FIELDS (*Fields*); PHILIPS HEALTHCARE (*machine*); ZEPHYR/ PHOTO RESEARCHERS, INC. (*color scan*); JOE GALLIVAN *Ulster Medical and Surgical Specialists (black-and-white scan)*

KNOWN DISEASES

Faulty or missing myelin in the central nervous system can cause several debilitating diseases, including:

MULTIPLE SCLEROSIS
(degeneration of central nervous system)

CEREBRAL PALSY
(severely impaired muscle control)

ALEXANDER DISEASE
(destruction of central nervous system)

These results seem to jibe with DTI studies performed by neuroscientist Vincent J. Schmithorst of Cincinnati Children's Hospital, which compared white matter in children ages five to 18. A higher development of white matter structure, Schmithorst found, correlates directly with higher IQ. Other reports reveal that children who suffer severe neglect have up to 17 percent less white matter in the corpus callosum.

Stimulating Change

Such findings strongly suggest that experience influences myelin formation and that the resulting myelin supports learning and improvement of skills. But to be fully convinced of that conclusion, investigators need a plausible explanation of how abundant myelin can enhance cognition, as well as some direct evidence that defects can impair mental abilities.

My lab has uncovered several ways in which an individual's experiences can influence myelin formation. In the brain, neurons fire electrical impulses down axons; by growing neurons from fetal mice in culture dishes equipped with platinum electrodes, we can impose patterns of impulses on them. We found that these impulses can regulate specific genes in neurons. One of the genes causes production of a sticky protein called L1-CAM that is crucial for pasting the first layer of membrane around an axon as myelin begins to form.

We also found that glia can "listen in" on impulses shooting through axons and that the traffic heard alters the degree of myelination; a type of glial cell called an astrocyte releases a chemical factor when it senses increased impulse traffic. This chemical code stimulates oligodendrocytes to form more myelin. Children who succumb to Alexander disease, a fatal childhood disorder causing mental retardation and abnormal myelin, have a mutation of an astrocyte gene.

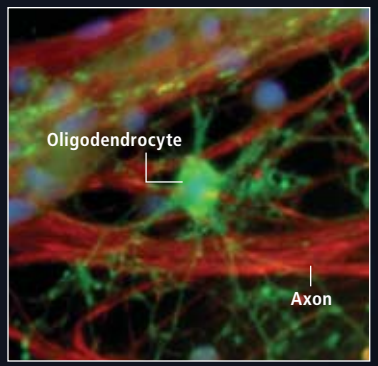
Logic, too, helps to explain how white matter can influence cognitive ability. It might seem that, by analogy to the Internet, all information in the brain should be transmitted as quickly as possible. That would mean all axons should be equally myelinated. But for neurons, faster is

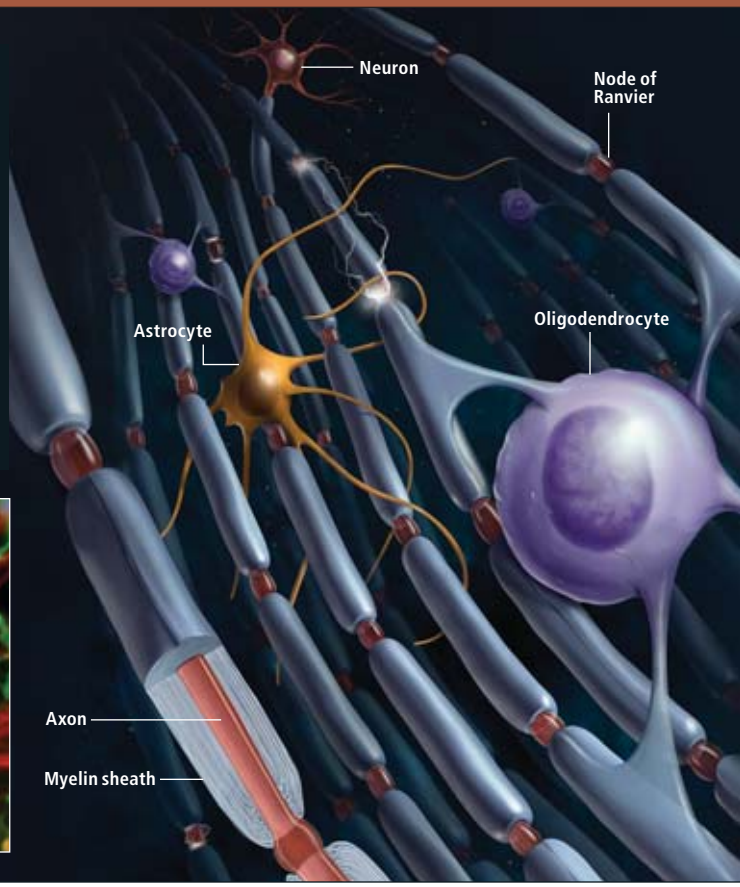
VARSHA SHUKLA (*micrograph*); ALAN HOOFRING *NIH Medical Arts (illustration)*

[HOW IT WORKS]

MYELIN FORMATION

Long axons insulated with myelin carry signals between neurons faster than unmyelinated axons. Oligodendrocyte cells manufacture the fatty membrane and wrap the axon with 10 to 150 layers. Different factors can stimulate the myelination process; often astrocyte cells "listen in" on the signals traveling along axons and relay chemical messages to the oligodendrocytes. Below, a microscope shows axons in red being wrapped.





not always better. Information must travel enormous distances between brain centers. Each center carries out its particular function and sends the output to another region for the next step of analysis. For complex learning, such as learning the piano, information must be shuttled back and forth among many regions; information flowing over different distances must arrive simultaneously at one place at a certain time. For such precision to occur, delays are necessary. If all axons transmitted information at the maximum rate, signals from distant neurons would always arrive later than signals from neighboring neurons. An impulse typically takes 30 milliseconds to travel from one cerebral hemisphere to the other through myelinated axons in the corpus callosum, compared with 150 to 300 milliseconds through unmyelinated axons. None of the corpus callosum's axons are myelinated at birth, and by adulthood 30 percent remain that way. The variation helps to coordinate transmission speeds.

Perhaps just as crucial are the nodes of Ranvier. In the past few years scientists have concluded that far from being mistakes, the nodes act as intricate, bioelectric repeaters—relay stations that generate, regulate and rapidly propagate electrical signals along an axon. By studying owls' excellent hearing, neurobiologists have shown that during myelination the oligodendrocytes insert more nodes than are optimal for fast signaling along certain axons to slow signals traveling along them.

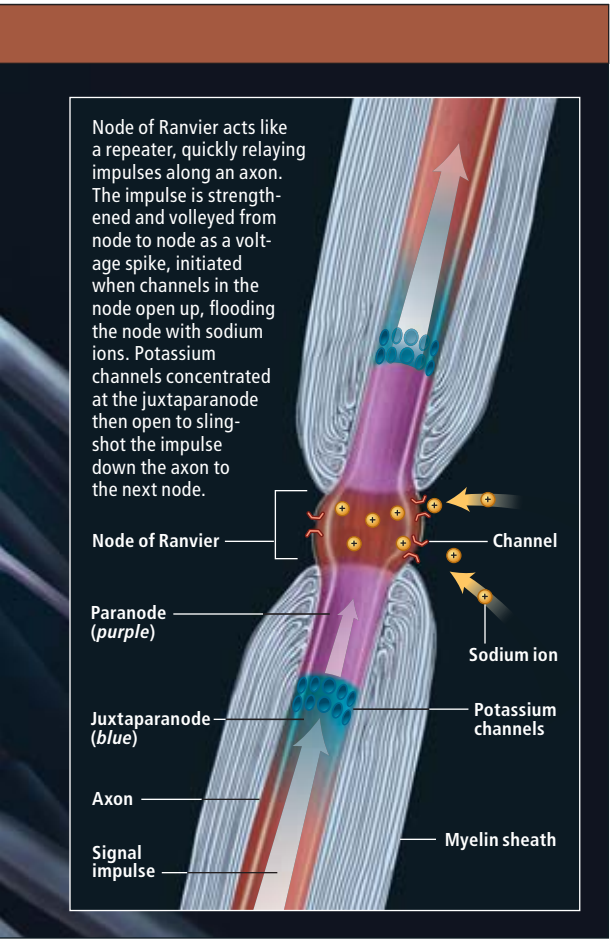
Clearly, the speed of impulse transmission is a vital aspect of brain function. We know that memory and learning occur when certain neuronal circuits connect more strongly. It seems likely that myelin affects this strength, by adjusting conduction velocity so that volleys of electrical impulses arrive at the same neuron simultaneously from multiple axons. When this convergence occurs, the individual voltage blips pile up, increasing the strength of the signal, thus making a stronger connection among the neurons involved. Much more research must be done to explore this theory, but there is no doubt that myelin responds to the environment and participates in learning skills.

Learning and Mental Illness

With this new perspective, it is not hard to imagine how faulty transmission could lead to mental challenges. After decades of searching gray matter for the causes of mental disabilities, neuroscientists now have circumstantial evidence suggesting that white matter plays a role. Dyslexia, for example, results from disrupted timing of information transmission in circuits required for reading; brain imaging has revealed reduced white matter in these tracts, which could cause such disruption. The white matter abnormalities are thought to reflect both defects in myelination and developmental abnormalities in neurons affecting these white matter connections.

Tone deafness results from defects in higher-level processing in the cerebral cortex where sounds are analyzed; psychologist Kristi L. Hyde of McGill University has found that white matter is reduced in a specific fiber bundle in the right forebrain of tone-deaf individuals. Furthermore, recent research by Leslie K. Jacobsen of Yale University indicates that exposure to tobacco smoke during late fetal development or adolescence, when this bundle is undergoing myelination, disrupts the white matter. The structure, as seen by DTI, correlates directly with performance on auditory tests. Nicotine is known to affect receptors on oligodendrocytes

ALAN HOOFRING *NIH Medical Arts* AND JEN CHRISTIANSEN (*inset*)



SUSPECTED ILLNESSES

Abnormal myelin formation is suspected of contributing to several mental illnesses, including:

SCHIZOPHRENIA
(delusions, hallucinations)

AUTISM
(impaired communication and emotional detachment)

BIPOLAR DISORDER
(periods of mania alternating with periods of depression)

DYSLEXIA
(spelling, reading or more general language disorder)

Myelin responds to the environment and participates in learning, in part by strengthening neuronal connections.

that regulate the cells' development. Exposure to environmental factors during crucial periods of myelination can have lifelong consequences.

Schizophrenia is now understood to be a developmental disorder that involves abnormal connectivity. The evidence is multifold. Doctors have always wondered why schizophrenia typically develops during adolescence—but recall that this is the primary age when the forebrain is being myelinated. The neurons there have largely been established, but the myelin is changing, making it suspect. In addition, nearly 20 studies in recent years have concluded that white matter is abnormal (possessing fewer oligodendrocytes than it should) in several regions of the schizophrenic brain. And when gene chips—tiny diagnostic devices that can survey thousands of genes at a time—recently became available, researchers were startled to discover that many of the mutated genes linked to schizophrenia were involved in myelin formation. White matter abnormalities have also been found in people affected by ADHD, bipolar disorder, language disorders, autism, cognitive decline in aging and Alzheimer's disease and even in individuals afflicted with pathological lying.

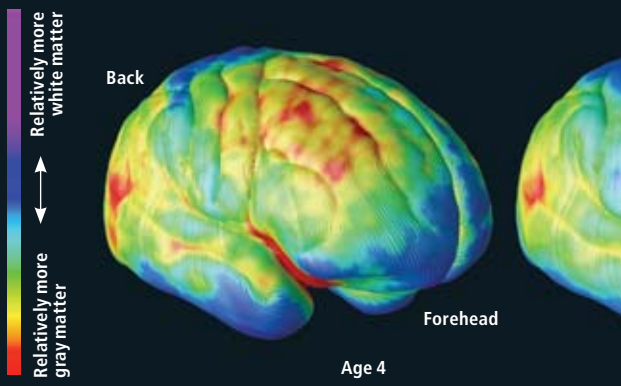
Of course, underdeveloped or withered myelin could be a result of poor signaling among neurons, not necessarily a cause. After all, cognitive function does depend on neuronal communication across synapses in the cortex's gray matter, where most psychoactive drugs act. Yet optimal communication among brain regions, which is also fundamental to proper cognition, depends on the white matter bedrock connecting the regions. In 2007 Gabriel Corfas, a neurologist at Children's Hospital Boston, showed that experimental disruption of genes in oligodendrocytes—not in neurons—of mice causes striking behavioral changes that mimic schizophrenia. And the behavioral effects involve one of the same genes, *neuregulin,* found to be abnormal in biopsies of schizophrenic brains.

The chicken-and-egg question of whether changes in myelin alter neurons or whether changing neuronal patterns alter myelin will be settled the same way such dilemmas always are: with the acknowledgment that there is a close interdependence between the two mechanisms. Myelinating glia can respond to changes in axon diameter, but they also regulate that diameter. And they can determine whether or not a given axon survives. In multiple sclerosis, for example, axons and neurons can die after myelin is lost as a result of the disease.

[DEVELOPMENT]

THE BRAIN MATURES

Few axons are covered with myelin at birth. More are insulated over time, from the back of the cerebral cortex to the front. The sequence here, by Paul Thompson of the University of California, Los Angeles, depicts the pruning of neurons and the relative increase in myelin. Basic functional areas such as vision (*back*) are completed before age 4, followed by language and, last, self-control (*forehead*).



To reach world-class status in certain intellectual or athletic skills, an individual must start young.

PROFESSIONAL PIANISTS have more highly developed white matter in certain regions than nonmusicians do, suggesting it affects learning. Furthermore, it is more extensive in pianists who begin regular practice before age 11 than in those who start during their teens or later, indicating that critical periods exist for superior skill acquisition.

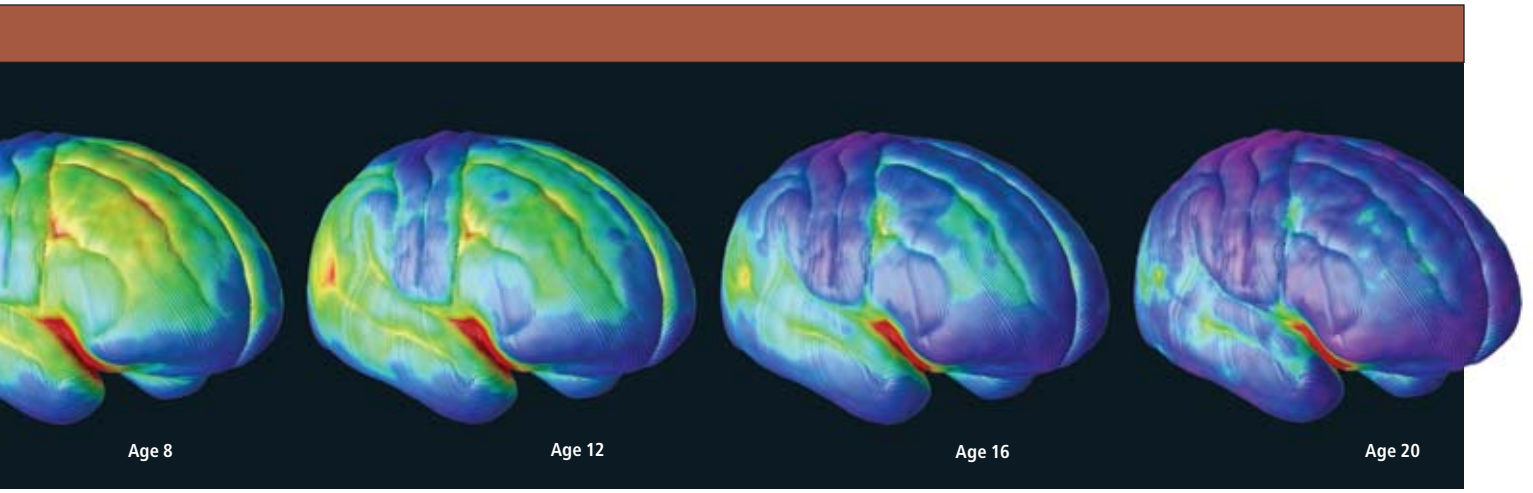
Remodeling Old Age

Whatever the mechanism, as our brain matures from childhood to adulthood the precision of connections among regions improves. How well the connections are made may dictate how well we can learn certain skills at certain ages.

Indeed, Ullén's studies of accomplished pianists revealed an additional finding: white matter was more highly developed throughout the brains of individuals who had taken up the instrument at an earlier age. In people who learned after adolescence, white matter development was increased only in the forebrain—the region that was still undergoing myelination.

This finding suggests that the insulating of

INTI ST. CLAIR Getty Images



PAUL THOMPSON U.C.L.A. (*gray matter scans*)

nerve fibers in part determines age limits for learning new skills—windows of opportunity, or critical periods, when certain learning can occur or at least can occur readily. Learn a foreign language after puberty, and you are destined to speak it with an accent; learn the language as a child, and you will speak it like a native. The difference occurs because the brain circuits that detect speech rewire according to the sounds we hear only as a child. We literally lose the connections that would allow us to hear sounds unique to foreign languages. In evolutionary terms, the brain has no reason to retain connections to detect sounds that it has never heard after years of childhood. Critical periods are also one of the main reasons adults do not recover as well from brain injuries as children do.

Specialists have identified specific protein molecules in myelin that stop axons from sprouting and forming new connections. Martin E. Schwab, a brain researcher at the University of Zurich, revealed the first of several myelin proteins that cause young sprouts from axons to wither instantly on contact. When this protein, which he named Nogo (now referred to as Nogo-A), is neutralized, animals with a spinal cord injury can repair their damaged connections and recover sensation and movement. Recently Stephen M. Strittmatter of Yale found that the critical period for wiring the brains of animals through experience could be reopened by blocking signals from Nogo. When the protein is disrupted in old mice, the critters can rewire connections for vision.

If myelination is largely finished in a person's 20s, however, does that contradict recent claims that the brain remains plastic throughout middle and old age? For example, studies show that mental exercise into a person's 60s, 70s and 80s helps to delay the onset of Alzheimer's. And how does a person's wisdom increase over the decades? Answers are still forthcoming. Researchers have not yet looked for myelin changes in older animals. Other experiments suggest myelination continues into our mid-50s but on a much subtler level.

Certainly white matter is key to types of learning that require prolonged practice and repetition, as well as extensive integration among greatly separated regions of the cerebral cortex. Children whose brains are still myelinating widely find it much easier to acquire new skills than their grandparents do. For a range of intellectual and athletic abilities, if an individual wants to reach world-class level he or she must start young. You built the brain you have today by interacting with the environment while you were growing up and your neural connections were still myelinating. You can adapt those abilities in many ways, but neither you nor I will become a world-class pianist, chess player or tennis pro unless we began our training when we were children.

Of course, old geezers can still learn, but they are engaged in a different kind of learning involving the synapses directly. And yet intensive training causes neurons to fire, so the potential exists for that firing to stimulate myelination. Perhaps someday, when we fully understand when and why white matter forms, we can devise treatments to change it, even as it grows old. To deliver on that speculation, we would need to find the signal that tells an oligodendrocyte to myelinate one axon and not another one nearby. That discovery, buried deep underneath the gray matter, awaits unearthing by future explorers. ■

MORE TO EXPLORE

Myelination: An Overlooked Mechanism of Synaptic Plasticity? R. Douglas Fields in *Neuroscientist,* Vol. 11, No. 5, pages 528–531; 2005.

Extensive Piano Practicing Has Regionally Specific Effects on White Matter Development. Sara L. Bengtsson et al. in *Nature Neuroscience,* Vol. 8, No. 9, pages 1148–1150; September 2005.

Astrocytes Promote Myelination in Response to Electrical Impulses. T. Ishibashi et al. in *Neuron,* Vol. 49, No. 6, pages 823–832; March 16, 2006.

How to Grow a Super Athlete. D. Coyle in *Play Magazine,* (*New York Times Sports*) March 4, 2007. Available at **www.nytimes.com/2007/03/04/sports/playmagazine/04play-talent.html**

THE LIMITS OF Quantum

By Scott Aaronson

Quantum computers would be exceptionally fast at a few specific tasks, but it appears that for most problems they would outclass today's computers only modestly. This realization may lead to a new fundamental physical principle

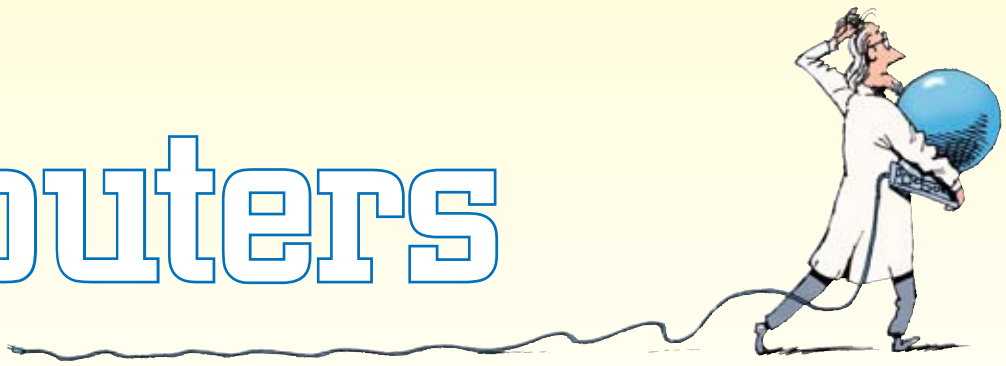
"Haggar Physicists Develop 'Quantum Slacks,' " read a headline in the satirical weekly the *Onion*. By exploiting a bizarre "Schrödinger's Pants" duality, the article explained, these non-Newtonian pants could paradoxically behave like formal wear and casual wear at the same time. *Onion* writers were apparently spoofing the breathless articles about quantum computing that have filled the popular science press for a decade.

A common mistake—see for instance the February 15, 2007, issue of the *Economist*—is to claim that, in principle, quantum computers could rapidly solve a particularly difficult set of mathematical challenges called NP-complete problems, which even the best existing computers cannot solve quickly (so far as anyone knows). Quantum computers would supposedly achieve this feat not by being formal and casual at the same time but by having hardware capable of processing every possible answer simultaneously.

If we really could build a magic computer capable of solving an NP-complete problem in a snap, the world would be a very different place: we could ask our magic computer to look for whatever patterns might exist in stock-market data or in recordings of the weather or brain activity. Unlike with today's computers, finding these patterns would be completely routine and require no detailed understanding of the subject of the problem. The magic computer could also automate mathematical creativ-

ILLUSTRATIONS BY DUŠAN PETRIČIĆ

Computers



ity. Given any holy grail of mathematics—such as Goldbach's conjecture or the Riemann hypothesis, both of which have resisted resolution for well over a century—we could simply ask our computer to search through all possible proofs and disproofs containing up to, say, a billion symbols. (If a proof were much longer than that, it is not clear that we would even want to read it.)

If quantum computers promised such godlike mathematical powers, maybe we should expect them on store shelves at about the same time as warp-drive generators and antigravity shields. But although we should not accept the usual hype, in my view it is equally misguided to dismiss quantum computing as science fiction. Instead we should find out what the limits of quantum computers are and what they could really do if we had them.

In the 26 years since physicist Richard Feynman first proposed the idea of quantum computing, computer scientists have made enormous progress in figuring out what problems quantum computers would be good for. According to our current understanding, they *would* provide dramatic speedups for a few specific problems—such as breaking the cryptographic codes that are widely used for monetary transactions on the Internet. For other problems, however—such as playing chess, scheduling airline flights and proving theorems—evidence now strongly suggests that quantum computers would suffer from many of the same algorithmic limitations as today's classical computers. These limitations are completely separate from the practical difficulties of building quantum computers, such as decoherence (unwanted interaction between a quantum computer and its environment, which introduces errors). In particular, the bounds on what it is mathematically possible to program a computer to do would apply even if physicists managed to build a quantum computer with no decoherence at all.

Hard, Harder, Hardest

How is it that a quantum computer could provide speedups for some problems, such as breaking codes, but not for others? Isn't a faster computer just a faster computer? The answer is no, and to explain why takes one straight to the intellectual core of computer science. For computer scientists, the crucial thing about a problem is how quickly the time needed to solve it grows as the problem size increases. The time is measured in the number of elementary steps required by the algorithm to reach a solution. For example, using the grade school method, we can multiply two n-digit numbers in an amount of time that grows like the number of digits squared, n^2 (an amount of time said to be "a polynomial in n"). But for factoring a number into primes, even the most advanced methods known take an amount of time that grows exponentially with the number of digits (in particular, like 2 to the cube root of n power). Thus, factoring seems intrinsically harder than multiplying—and when we get up to thousands of digits, this difference matters much more than the difference between a Commodore 64 and a supercomputer.

The kind of problems that computers can solve in a reasonable amount of time, even for large values of n, are those for which we have an algorithm that uses a number of steps that grows

KEY CONCEPTS

- Quantum computers would exploit the strange rules of quantum mechanics to process information in ways that are impossible on a standard computer.
- They would solve certain specific problems, such as factoring integers, dramatically faster than we know how to solve them with today's computers, but analysis suggests that for most problems quantum computers would surpass conventional ones only slightly.
- Exotic alterations to the known laws of physics *would* allow construction of computers that could solve large classes of hard problems efficiently. But those alterations seem implausible. In the real world, perhaps the impossibility of efficiently solving these problems should be taken as a basic physical principle.

—The Editors

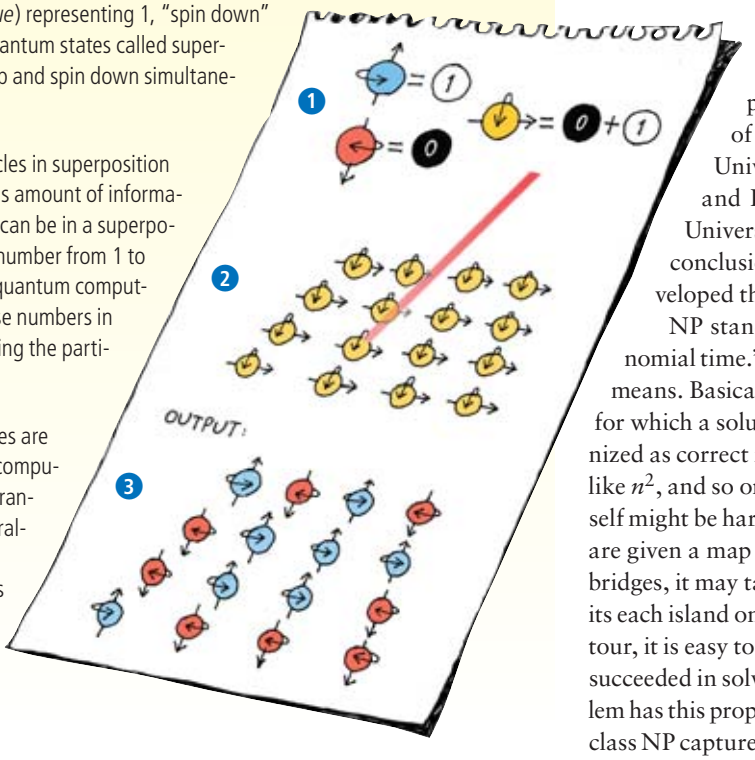
Quantum Computing 101

Physicists are hotly pursuing the construction of quantum computers, which would harness the quirks of quantum mechanics to perform certain computations more efficiently than a conventional computer.

1 The fundamental feature of a quantum computer is that it uses qubits instead of bits. A qubit may be a particle such as an electron, with "spin up" (*blue*) representing 1, "spin down" (*red*) representing 0, and quantum states called superpositions that involve spin up and spin down simultaneously (*yellow*).

2 A small number of particles in superposition states can carry an enormous amount of information: a mere 1,000 particles can be in a superposition that represents every number from 1 to $2^{1,000}$ (about 10^{300}), and a quantum computer would manipulate all those numbers in parallel, for instance, by hitting the particles with laser pulses.

3 When the particles' states are measured at the end of the computation, however, all but one random version of the 10^{300} parallel states vanish. Clever manipulation of the particles could nonetheless solve certain problems very rapidly, such as factoring a large number.



A good quantum computer algorithm ensures that **computational paths leading to a wrong answer cancel out** and that paths leading to a correct answer reinforce.

as n raised to a fixed power, such as n, n^2 or $n^{2.5}$. Computer scientists call such an algorithm efficient, and problems that can be solved by an efficient algorithm are said to be in the complexity class P, which stands for "polynomial time."

A simple example of a problem in P is: Given a road map, is every town reachable from every other town? P also contains some problems whose efficient solutions are not so obvious, such as: Given a whole number, is it prime (like 13) or composite (like 12)? Given a list of which men and women are willing to marry one another, is it possible to pair everyone off with a willing partner?

But now suppose you are given the dimensions of various boxes and you want a way to pack them in your trunk. Or suppose that you are given a map and you want to color each country red, blue or green so that no two neighboring countries are colored the same. Or that you are given a list of islands connected by bridges and you want a tour that visits each island exactly once. Although algorithms that are somewhat better than trying every possible solution are known for these problems, no algorithm is known that is fundamentally better. Every known algorithm will take an amount of time that increases exponentially with the problem size.

It turns out that the three problems I just listed have a very interesting property: they are all the "same" problem, in the sense that an efficient algorithm for any one of them would imply efficient algorithms for all the others. Stephen A. Cook of the University of Toronto, Richard Karp of the University of California, Berkeley, and Leonid Levin, now at Boston University, arrived at this remarkable conclusion in the 1970s, when they developed the theory of NP-completeness.

NP stands for "*n*ondeterministic *p*olynomial time." Do not worry about what that means. Basically, NP is the class of problems for which a solution, once found, can be recognized as correct in polynomial time (something like n^2, and so on)—even though the solution itself might be hard to find. As an example, if you are given a map with thousands of islands and bridges, it may take years to find a tour that visits each island once. Yet if someone shows you a tour, it is easy to check whether that person has succeeded in solving the problem. When a problem has this property, we say that it is in NP. The class NP captures a huge number of problems of practical interest. Note that all the P problems are also NP problems, or to put it another way, the class P is contained within the class NP. If you can solve a problem quickly you can also verify the solution quickly.

NP-complete problems are in essence the hardest of the NP problems. They are the ones with the property found by Cook, Karp and Levin: If an efficient algorithm for any one of them were found, it could be adapted to solve all the other NP problems as well.

An efficient algorithm for an NP-complete problem would mean that computer scientists' present picture of the classes P, NP and NP-complete was utterly wrong, because it would mean that every NP problem (including all the NP-complete ones) was actually a P problem. In other words, the class P would equal the class NP, which is written P = NP.

Does such an algorithm exist? Is P equal to NP? That is literally a million-dollar question—it carries a $1,000,000 reward from the Clay Math Institute in Cambridge, Mass.—and it has

played cameo roles on at least three TV shows (*The Simpsons*, *Futurama* and *NUMB3RS*).

In the half a century since the problem was recognized, no one has found an efficient algorithm for an NP-complete problem. Consequently, computer scientists today almost universally believe P does not equal NP, or $P \neq NP$, even if we are not yet smart enough to understand why this is or to prove it as a theorem.

What the Quantum Can Do

If we grant that $P \neq NP$, then only one hope remains for solving NP-complete problems in polynomial time: namely, to broaden what we mean by "computer." At first sight, quantum mechanics would appear to provide just the kind of resources needed. Quantum mechanics makes it possible to store and manipulate a vast amount of information in the states of a relatively small number of particles. To see how this comes about, imagine that we have 1,000 particles and that each particle, when measured, can be found to be either spinning up or spinning down. For our purposes, what it means for a particle to spin up or down is irrelevant; all that matters is that there is some property of the particle that has one of two values when measured.

To describe the quantum state of this collection of particles, one must specify a number for every possible result of measuring the particles. These numbers are called the amplitudes of the possible outcomes and relate to each outcome's probability, but unlike probabilities, quantum amplitudes can be positive or negative (in fact, they are complex numbers). For example, an amplitude is needed for the possibility that all 1,000 particles will be found spinning up, another amplitude for the possibility of finding that the first 500 particles are spinning up and that the remaining 500 are spinning down, and so on. There are $2^{1,000}$ possible outcomes, or about 10^{300}, so that is how many numbers are needed—more than there are atoms in the visible universe! The technical terminology for this situation is that the 1,000 particles are in a superposition of those 10^{300} states.

Put another way, we can store 10^{300} numbers on our 1,000 particles simultaneously. Then, by performing various operations on the particles and on some auxiliary ones—perhaps hitting them with a sequence of laser pulses or radio waves—we can carry out an algorithm that transforms all 10^{300} numbers (each one a potential solution) at the same time. If at the end of doing that we could read out the particles' final quantum state accurately, we really would have a magic computer: it would be able to check 10^{300} possible solutions to a problem, and at the end we could quickly discern the right one.

Unfortunately, there is a catch. When the particles are measured (as is necessary to read out their final state), the rules of quantum mechanics dictate that the measurement will pick out just one of the 10^{300} possibilities at random and that all the others will then disappear. (To go back to the quantum slacks developed at Haggar, if you tried to wear them you would find yourself in either formal or casual attire, not both.) We would seem to be no better off than if we used a classical computer and tried out one randomly chosen possible solution—in either case, we end up knowing about only one such possible solution.

Happily, we still have tricks we can play to wring some advantage out of the quantum particles. Amplitudes can cancel out when positive ones combine with negative ones, a phenomenon

The "**killer app**" **for quantum computers** will most likely be simulating quantum physics.

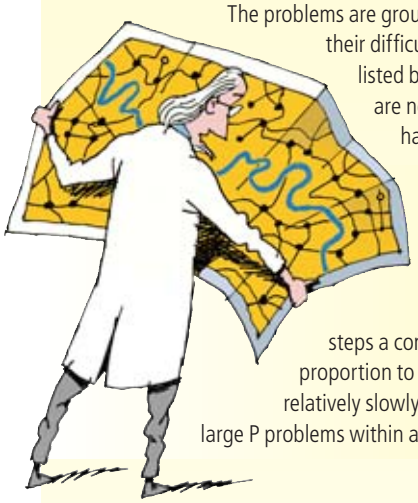
The Good News

If a large, ideal quantum computer would face most of the same limitations as our present-day classical computers do, should the physicists working on the extraordinarily hard task of building even rudimentary quantum computers pack up and go home? I believe the answer is no, for four reasons.

- If quantum computers ever become a reality, the "killer app" for them will most likely not be code breaking but rather something so obvious it is rarely even mentioned: simulating quantum physics. This is a fundamental problem for chemistry, nanotechnology and other fields, important enough that Nobel Prizes have been awarded even for partial progress.
- As transistors in microchips approach the atomic scale, ideas from quantum computing are likely to become relevant for classical computing as well.
- Quantum computing experiments focus attention directly on the most mystifying features of quantum mechanics—and I hope that the less we can sweep those puzzles under the rug, the more we will be forced to understand them.
- Quantum computing can be seen as the most stringent test to which quantum mechanics itself has ever been subjected. In my opinion, the most exciting possible outcome of quantum computing research would be to discover a fundamental reason why quantum computers are *not* possible. Such a failure would overturn our current picture of the physical world, whereas success would merely confirm it. —***S.A.***

What Classical Computers Can and Cannot Do

Computer scientists categorize problems according to how many computational steps it would take to solve a large example of the problem using the best algorithm known. The problems are grouped into broad, overlapping classes based on their difficulty. Three of the most important classes are listed below. Contrary to myth, quantum computers are not known to be able to solve efficiently the very hard class called NP-complete problems.



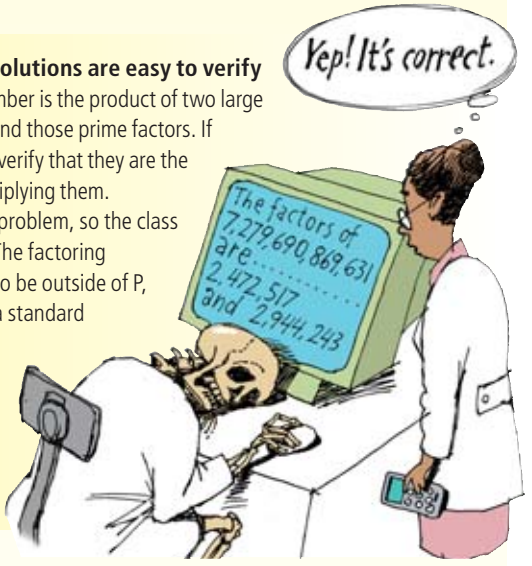
P PROBLEMS: Ones computers can solve efficiently, in polynomial time

Example: Given a road map showing n towns, can you get from any town to every other town? For a large value of n, the number of steps a computer needs to solve this problem increases in proportion to n^2, a polynomial. Because polynomials increase relatively slowly as n increases, computers can solve even very large P problems within a reasonable length of time.

NP PROBLEMS: Ones whose solutions are easy to verify

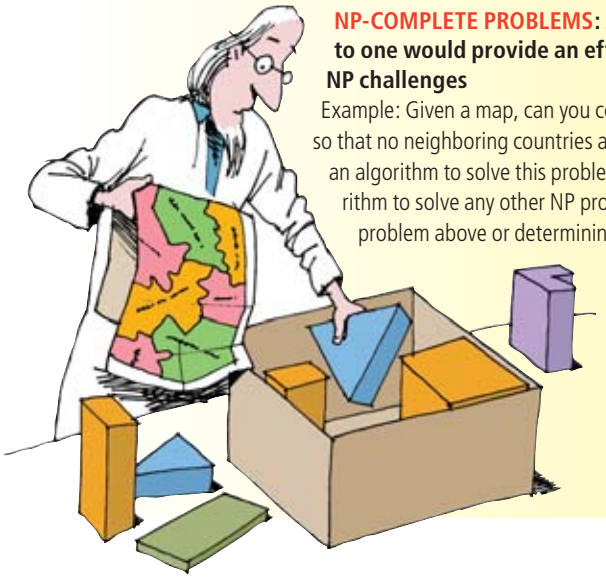
Example: You know an n-digit number is the product of two large prime numbers, and you want to find those prime factors. If you are given the factors, you can verify that they are the answer in polynomial time by multiplying them.

Every P problem is also an NP problem, so the class NP contains the class P within it. The factoring problem is in NP but conjectured to be outside of P, because no known algorithm for a standard computer can solve it in only a polynomial number of steps. Instead the number of steps increases exponentially as n gets bigger.



NP-COMPLETE PROBLEMS: An efficient solution to one would provide an efficient solution to all NP challenges

Example: Given a map, can you color it using only three colors so that no neighboring countries are the same color? If you had an algorithm to solve this problem, you could adapt the algorithm to solve any other NP problem (such as the factoring problem above or determining if you can pack n boxes of various sizes into a trunk of a certain size) in about the same number of steps. In that sense, NP-complete problems are the hardest of the NP problems. No known algorithm can solve an NP-complete problem efficiently.



called destructive interference. So a good quantum computer algorithm would ensure that computational paths leading to a wrong answer would cancel out in this way. It would also ensure that the paths leading to a correct answer would all have amplitudes with the same sign—which yields constructive interference and thereby boosts the probability of finding them when the particles are measured at the end.

For which computational problems can we choreograph this sort of interference, using fewer steps than it would take to solve the problem classically?

In 1994 Peter Shor, now at the Massachusetts Institute of Technology, found the first example of a quantum algorithm that could dramatically speed up the solution of a practical problem. In particular, Shor showed how a quantum computer could factor an n-digit number using a number of steps that increases only as about n^2—in other words, in polynomial time. As mentioned earlier, the best algorithm known for classical computers uses a number of steps that increases exponentially.

Black Boxes

So at least for factoring, one really can get an exponential speedup over known classical algorithms by using quantum methods. But despite a widespread misconception to the contrary, the factoring problem is neither known nor believed to be NP-complete. To create his algorithm, Shor exploited certain mathematical properties of composite numbers and their factors that are particularly well suited to producing the kind of constructive and destructive interference that a quantum computer can thrive on. The NP-complete problems do not seem to share those special properties. To this day, researchers have found only a few other quantum algorithms that appear to provide a speedup from exponential to polynomial time for a problem.

The question thus remains unanswered: Is there an efficient quantum algorithm to solve NP-complete problems? Despite much trying, no such algorithm has been found—though not surprisingly, computer scientists cannot prove that it does not exist. After all, we cannot even prove that there is no polynomial-time classical algorithm to solve NP-complete problems.

What we *can* say is that a quantum algorithm capable of solving NP-complete problems efficiently would, like Shor's algorithm, have to exploit the problems' structure, but in a way that is far beyond present-day techniques. One cannot

achieve an exponential speedup by treating the problems as structureless "black boxes," consisting of an exponential number of solutions to be tested in parallel. Some speedup can nonetheless be wrung out of this black box approach, and computer scientists have determined just how good—and how limited—that speedup is. The algorithm that produces the speedup is the second major quantum algorithm.

The black box approach can be illustrated by pretending that you are searching for the solution to a difficult problem and that the only operation you know how to perform is to guess a solution and see if it works. Let us say there are S possible solutions, where S grows exponentially as the problem size n increases. You might get lucky and guess the solution on your first try, but in the worst case you will need S tries, and on average you will need $S/2$.

Now suppose you can ask about all the possible solutions in quantum superposition. In 1996 Lov Grover of Bell Laboratories developed an algorithm to find the correct solution in such a scenario using only about $\sqrt{S}$ steps. A speedup from $S/2$ to $\sqrt{S}$ is a useful advance for some problems—if there are a million possible solutions, you will need around 1,000 steps instead of 500,000. But the square root does not transform an exponential time into a polynomial time; it just produces a smaller exponential. And Grover's algorithm is as good as it gets for this kind of black box searching: in 1994 researchers had shown that any black box quantum algorithm needs at least $\sqrt{S}$ steps.

Over the past decade, researchers have shown that similar modest speedups are the limit for many other problems besides searching a list, such as counting ballots in an election, finding the shortest route on a map, and playing games of strategy such as chess or Go. One problem that presented particular difficulty was the so-called collision problem, the problem of finding two items that are identical, or that "collide," in a long list. If there were a fast quantum algorithm to solve this problem, many of the basic building blocks of secure electronic commerce would be useless in a world with quantum computers.

Searching a list for an item is like looking for a needle in a haystack, whereas searching for a collision is like looking for two identical pieces of hay, which provides the problem with a kind of structure that a quantum computer could potentially exploit. Nevertheless, I showed in 2002 that within the black box model, any quantum algorithm needs exponential time to solve the collision problem.

Admittedly, these black box limitations do

[THE AUTHOR]

Scott Aaronson is an assistant professor of electrical engineering and computer science at the Massachusetts Institute of Technology. A high school dropout, he went on to receive a bachelor's degree from Cornell University and a Ph.D. in computer science from the University of California, Berkeley, advised by Umesh Vazirani. Outside of research, Aaronson is best known for his widely read blog (www.scottaaronson.com/blog), as well as for creating the Complexity Zoo (www.complexityzoo.com), an online encyclopedia of more than 400 complexity classes.

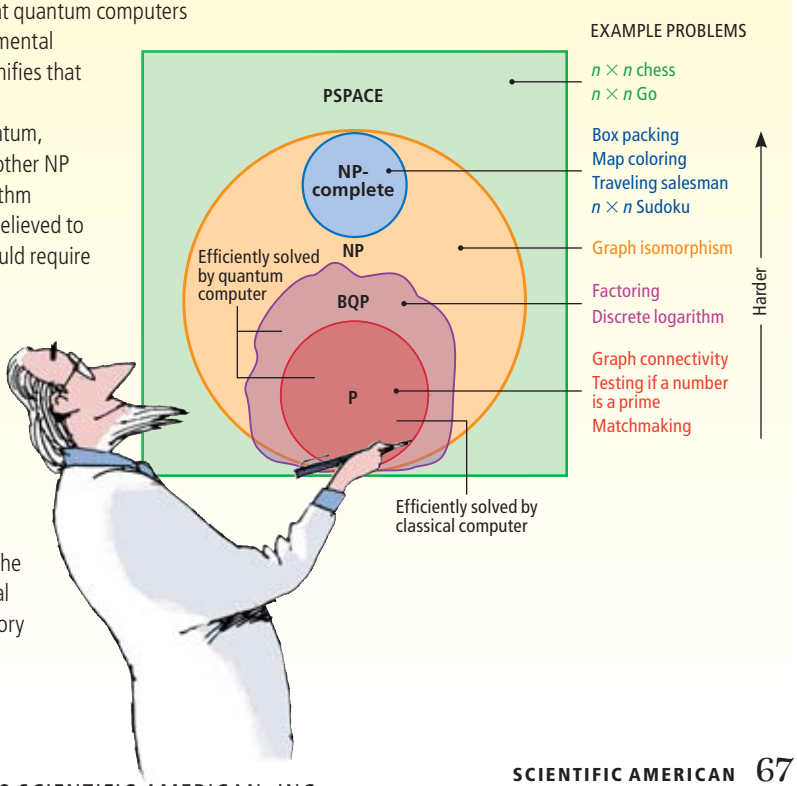
Where Quantum Computers Fit In

The map at the right depicts how the class of problems that quantum computers would solve efficiently (BQP) might relate to other fundamental classes of computational problems. (The irregular border signifies that BQP does not seem to fit neatly with the other classes.)

The BQP class (the letters stand for *b*ounded-error, *q*uantum, *p*olynomial time) includes all the P problems and also a few other NP problems, such as factoring and the so-called discrete logarithm problem. Most other NP and all NP-complete problems are believed to be outside BQP, meaning that even a quantum computer would require more than a polynomial number of steps to solve them.

In addition, BQP might protrude beyond NP, meaning that quantum computers could solve certain problems faster than classical computers could even check the answer. (Recall that a conventional computer can efficiently verify the answer of an NP problem but can efficiently solve only the P problems.) To date, however, no convincing example of such a problem is known.

Computer scientists do know that BQP cannot extend outside the class known as PSPACE, which also contains all the NP problems. PSPACE problems are those that a conventional computer can solve using only a polynomial amount of memory but possibly requiring an exponential number of steps.



JASON DORFMAN (*Aaronson*)

ZONES OF THOUGHT

Unlike the real world, in which computational limits are believed to be the same everywhere, the galaxy in Vernor Vinge's 1992 science-fiction novel *A Fire Upon the Deep* is divided into concentric "zones of thought" having different inherent computational and technological limits.

In the Unthinking Depths, nearest the galactic core, even simple automation fails and IQs plummet.

The Slow Zone contains Earth and is as limited as we know it.

In the Beyond, nearly sentient nanotechnology factories construct wonders such as anti-gravity fabrics, and hypercomputation enables faster-than-light travel.

The Transcend is populated by dangerous, godlike über-intelligences having technologies and thought processes unfathomable to lower beings.

not rule out the possibility that efficient quantum algorithms for NP-complete or even harder problems are waiting to be discovered. If such algorithms existed, however, they would have to exploit the problems' structure in ways that are unlike anything we have seen, in much the same way that efficient classical algorithms for the same problems would have to. Quantum magic by itself is not going to do the job. Based on this insight, many computer scientists now conjecture not only that $P \neq NP$ but also that quantum computers cannot solve NP-complete problems in polynomial time.

Magical Theories

Everything we know is consistent with the possibility that quantum computers are the end of the line—that is, that they are the most general kind of computer compatible with the laws of physics. But physicists do not yet have a final theory of physics, so one cannot rule out the possibility that someday a future theory might reveal a physical means to solve NP-complete problems efficiently. As you would expect, people speculate about yet more powerful kinds of computers, some of which would make quantum computers look as pedestrian as vending machines. All of them, however, would rely on speculative changes to the laws of physics.

One of the central features of quantum mechanics is a mathematical property called linearity. In 1998 Daniel S. Abrams and Seth Lloyd, both then at M.I.T., showed that if a small nonlinear term is added to the equations of quantum mechanics, quantum computers would be able to solve NP-complete problems efficiently. Before you get too excited, you should realize that if such a nonlinear term existed, then one could also violate Heisenberg's uncertainty principle and send signals faster than the speed of light. As Abrams and Lloyd pointed out, perhaps the best interpretation of these results is that they help to explain why quantum mechanics is linear.

Another speculative type of machine would achieve extravagant computational abilities by cramming an infinite number of steps into a finite time. Unfortunately, according to physicists' current understanding, time seems to degenerate into a sea of quantum fluctuations—something like a foam instead of a uniform smooth line—on the scale of 10^{-43} second (the Planck time), which would seem to make this kind of machine impossible.

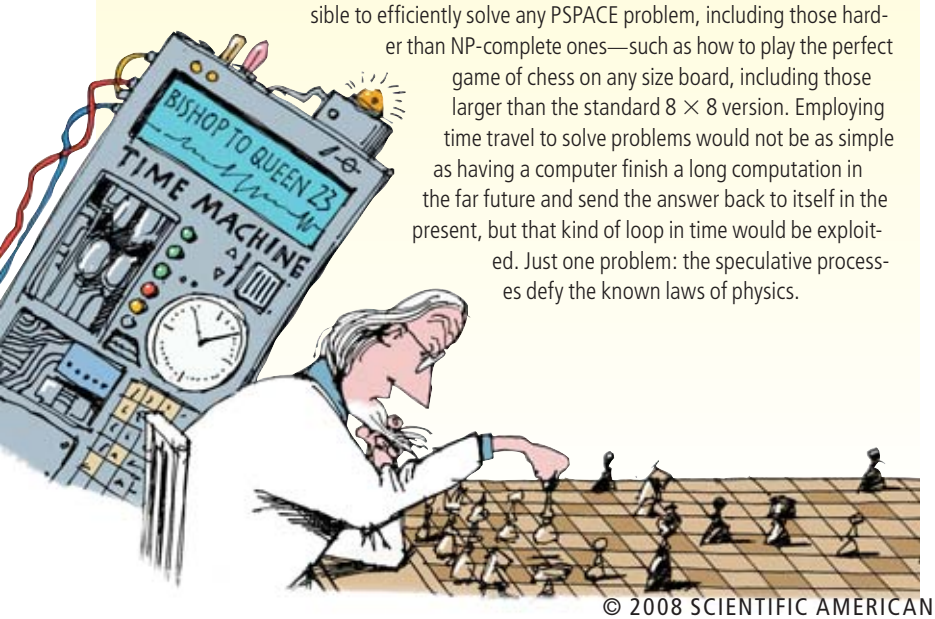
If time cannot be sliced with arbitrary thinness, then perhaps another way to solve NP-complete problems efficiently would be to exploit time travel. Physicists studying the issue talk not about time machines but about closed timelike curves (CTCs). In essence a CTC is a route through space and time that matter or energy could travel along to meet up with itself in the past, forming a closed loop. Current physical theory is inconclusive on whether CTCs can exist, but that need not stop us from asking what the consequences would be for computer science if they did exist.

It seems obvious how one could use a CTC to speed up a computation: program your computer to take however long it needs to solve the problem and then send the answer back in time to yourself at a point before the computer started. Alas, this simple idea does not work, because it ignores the famous grandfather paradox, where you go back in time to kill your own grandfather (but then you are never born, so you never go back in time, and so your grandfather lives to have children after all, and later you *are* born, but then …). In our setting, what would happen if you turned off the computer after you received its answer from the future?

In 1991 physicist David Deutsch of the University of Oxford defined a model of computation with CTCs that avoids this difficulty. In Deutsch's model, nature will ensure that as events unfold along the circular timeline that makes up the CTC, no paradoxes ever arise, a fact that can be exploited to program a computer that loops

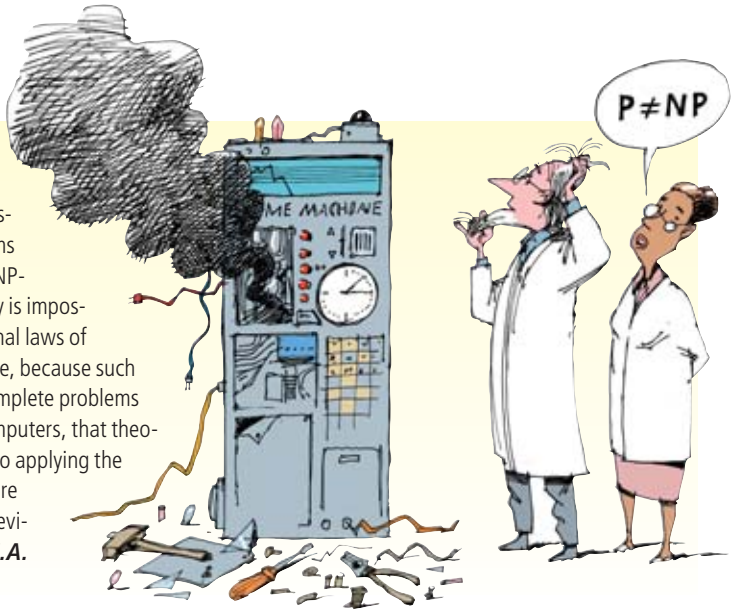
Über-Computers from Exotic Physics?

Although quantum computers seem unlikely to solve NP-complete problems quickly, certain other extraordinary, speculative physical processes would allow construction of computers with that ability and much more. Time travel, for instance, would make it possible to efficiently solve any PSPACE problem, including those harder than NP-complete ones—such as how to play the perfect game of chess on any size board, including those larger than the standard 8×8 version. Employing time travel to solve problems would not be as simple as having a computer finish a long computation in the far future and send the answer back to itself in the present, but that kind of loop in time would be exploited. Just one problem: the speculative processes defy the known laws of physics.



A New Physical Principle?

Because implausible kinds of physics (such as time travel) seem necessary for constructing a computer able to solve NP-complete problems quickly, I predict that scientists might one day adopt a new principle: "NP-complete problems are hard." That is, solving those problems efficiently is impossible on any device that could be built in the real world, whatever the final laws of physics turn out to be. The principle implies that time travel is impossible, because such travel would enable creation of über-computers that could solve NP-complete problems efficiently. Further, if a proposed theory were shown to permit such computers, that theory could be ruled out. Application of the principle would be analogous to applying the laws of thermodynamics to conclude that perpetual-motion machines are impossible (the laws of thermodynamics forbid them) and to deduce previously unknown features of physical processes. —*S.A.*



around inside the CTC to solve hard problems.

Indeed, by using a CTC, we could efficiently solve not only NP problems but even problems in an apparently larger class called PSPACE. PSPACE is the class of problems that could be solved on a conventional computer using a polynomial amount of memory but possibly taking an exponential amount of time. In effect, a CTC would make time and space interchangeable as computational resources. (I did not have to mention the polynomial memory constraint until now, because for P and NP problems it makes no difference if the computer has access to more than polynomial memory.) Recently John Watrous of the University of Waterloo in Ontario and I showed that using a quantum computer in a CTC instead of a conventional one does not enable anything beyond PSPACE to be efficiently solved. In other words, if CTCs exist, then quantum computers are no more powerful than classical ones.

Computational Kryptonite

Physicists do not know if future theories will permit any of these extraordinary machines. Yet without denying our ignorance, we can view that ignorance from a different perspective. Instead of starting from physical theories and then asking about their computational implications, we could start by assuming that NP-complete problems are hard and then study the consequences of that assumption for physics. For instance, if CTCs would let us solve NP-complete problems efficiently, then by starting from the assumption that NP-complete problems are intractable, we could conclude that CTCs cannot exist.

To some, such an approach will seem overly dogmatic. To me, it is no different from assuming the second law of thermodynamics or the impossibility of faster-than-light communication—two earlier limitations on technology that over time earned the status of physical principles. Yes, the second law might be experimentally falsified tomorrow—but until that happens, physicists find it vastly more useful to assume it is correct and then use that assumption for studying everything from car engines to black holes. I predict that the hardness of NP-complete problems will someday be seen the same way: as a fundamental principle that describes part of the essential nature of our universe. There is no way of telling what theoretical enlightenment or what practical consequences might come from future application of this kind of fundamental principle.

In the meantime, we know not to expect magic from quantum computers. To some, the apparent limitations of quantum computers might come as a letdown. One can, however, give those same limitations a more optimistic spin. They mean that although certain cryptographic codes could be broken in a world with quantum computers, other codes would probably remain secure. They increase our confidence that quantum computing will be possible at all—because the more a proposed technology sounds like a science-fiction caricature, the more skeptical we should be. (Who would you be more inclined to believe: the salesperson offering a device that produces unlimited free energy from the quantum vacuum or the one offering a refrigerator that is more efficient than last year's model?) And last, such limitations ensure that computer scientists will continue to have their work cut out for them in designing new quantum algorithms. Like Achilles without his heel or Superman without kryptonite, a computer without any limitations would get boring pretty quickly. ■

SCI AM

More about nonlinear quantum mechanics, hypercomputing, use of time travel, and another scheme called anthropic computing can be found at **www.SciAm.com/ontheweb**

MORE TO EXPLORE

Quantum Computation and Quantum Information. Michael A. Nielsen and Isaac L. Chuang. Cambridge University Press, 2000.

NP-Complete Problems and Physical Reality. Scott Aaronson in *ACM SIGACT News,* Complexity Theory Column, Vol. 36, No. 1, pages 30–52; March 2005. Available at **www.scottaaronson.com/papers/npcomplete.pdf**

Quantum Computer Science: An Introduction. N. David Mermin. Cambridge University Press, 2007.

Shor, I'll Do It. (An explanation of Shor's algorithm for the layperson.) Scott Aaronson. Available at **www.scottaaronson.com/blog/?p=208**

Quantum Computing since Democritus. Lecture notes from course PHYS771, University of Waterloo, Fall 2006. Available at **www.scottaaronson.com/democritus/**

The Bluefin in Peril

The only way to save the bluefin tuna, one of the most marvelous and endangered fish in the ocean, may be to domesticate the species

By Richard Ellis

CHRIS PARK AP Photo

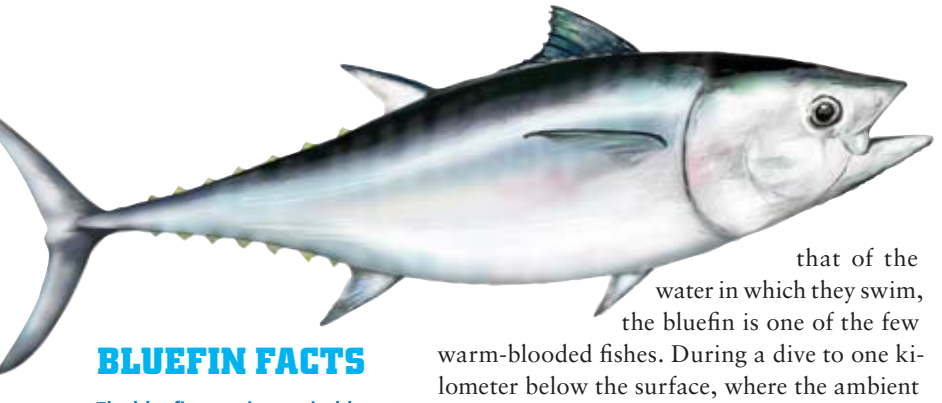
KEY CONCEPTS

- The surging popularity of sushi and sashimi has devastated the bluefin tuna. Overfishing has slashed populations in the Atlantic, Pacific and Indian oceans, pushing the species toward extinction. Regulatory bodies have failed to set sufficiently strict catch quotas, and illegal fishing is rampant.
- Captive breeding of the bluefin could save the species, but the effort will be challenging. Research groups in Japan and Europe have bred the tuna in laboratories, and now an Australian company is attempting to perform the feat on a commercial scale.

—The Editors

All tuna are not alike. The canned tuna fish in sandwiches and salads comes from either skipjack, a meter-long species that is caught in prodigious quantities around the world and served as "light meat tuna," or albacore, another small fish that is marketed as "white meat tuna." The yellowfin and the bigeye tuna are larger species that are also heavily fished, but neither makes for particularly wonderful sushi, and they are usually served grilled. But the bluefin tuna, a giant among fishes, is the premier choice for sushi and sashimi and has become the most desirable food fish in the world. As such, it has vaulted to the top of another, more insidious list: it is probably the most endangered of all large fish species. Heedless overfishing is steadily pushing the bluefin toward extinction, and the species may soon disappear unless entrepreneurial fish farmers can learn how to breed the tuna in captivity.

Reaching a maximum known weight close to three quarters of a ton and a length of four meters, the bluefin is a massive hunk of superheated muscle that cleaves the water by flicking its scimitar-shaped tail. Whereas most of the approximately 20,000 fish species are cold-blooded, possessing a body temperature the same as



BLUEFIN FACTS

The bluefin tuna is a veritable eating machine, superbly adapted to hunting in the cold waters of the temperate oceans.

Size: The largest bluefin ever recorded, caught off Nova Scotia in 1979, weighed **679 kilograms** (1,496 pounds). A typical adult weighs about half as much and measures **two meters** (six feet) long.

Speed: Bluefins can swim as fast as **80 kilometers per hour** in short bursts and can travel across the Atlantic Ocean in less than 60 days.

Reproduction: Females produce up to **10 million eggs a year.** Newborn fish (larvae) are three millimeters long at hatching and grow at a rate of one millimeter a day.

Life span: Bluefin larvae have only a one-in-40-million chance of reaching adulthood, but a mature tuna can live as long as **30 years.**

Price: In 2001 a 200-kilogram bluefin was purchased at a Japanese fish market for **$173,600,** or about $868 per kilogram ($394 per pound).

that of the water in which they swim, the bluefin is one of the few warm-blooded fishes. During a dive to one kilometer below the surface, where the ambient water temperature can be five degrees Celsius (41 degrees Fahrenheit), the bluefin can maintain a body temperature of 27 degrees C (81 degrees F), close to that of a mammal. The bluefin is also among the fastest of all fishes, capable of speeds up to 80 kilometers per hour and able to migrate across entire oceans. It is such a marvelous swimmer that when scientists in the 1990s endeavored to build a mechanical fish, they used the species as a model, designing a robot with a tapered, bullet-shaped body and a rigid, quarter-moon tail fin [see "An Efficient Swimming Machine," by Michael S. Triantafyllou and George S. Triantafyllou; Scientific American, March 1995]. The researchers found that the tail's efficiency lay in the interaction of the vortices created by its rapid flexing, but the hydrodynamics of their electronic models did not even come close to that of a true bluefin. "The more sophisticated our robotic-tuna designs become," the Triantafyllou brothers wrote, "the more admiration we have for the flesh-and-blood model."

Like wolves, bluefins often hunt in packs, forming a high-speed parabola that concentrates the prey, making it easier for the hunters to close in. Tuna are metabolically adapted for high-speed chases, but as opportunistic (and by necessity, compulsive) feeders, they will eat whatever presents itself, whether it is fast-swimming mackerel, bottom-dwelling flounder or sedentary sponge. A study of the stomach contents of New England bluefins by Bradford Chase of the Massachusetts Division of Marine Fisheries revealed that the predominant food item, by weight, was Atlantic herring, followed by sand lance, bluefish and miscellaneous squid. (Other prey included butterfish, silver hake, windowpane flounder, winter flounder, menhaden, sea horses, cod, plaice, pollack, filefish, halfbeak, sculpin, spiny dogfish, skate, octopus, shrimp, lobster, crab, salp and sponges.) Tuna will eat anything they can catch, and they can catch almost anything that swims (or floats, crawls or just sits on the bottom). By and large, they hunt by vision.

From Horse Mackerel to Sushi

The bluefin was not always considered a delicacy. In the early 1900s the fish was known as "horse mackerel," and its red, strong-flavored flesh was considered suitable fare only for dogs and cats. Nevertheless, big-game fishers off New Jersey and Nova Scotia targeted the bluefin because these powerful fish were considered worthy opponents. Zane Grey, the popular author of Western novels such as *Riders of the Purple Sage,* invested most of his not inconsiderable royalties (his books sold more than 13 million copies) on fishing gear, boats and travel to exotic locales in search of tuna, swordfish and marlin. Although swordfish were certainly considered edible, tuna and marlin were thought of as strictly objects of the hunt. The bluefin did not become valuable as a food fish until the latter half of the 20th century, when sushi began to appear on menus around the globe.

One might assume that sushi and sashimi have been staples of the Japanese diet for centuries, but in fact the widespread consumption of raw fish is a relatively modern phenomenon. Dependent on the sea to provide the great majority of their protein, the Japanese could not store fish for any length of time before it spoiled, so they preserved it by smoking or pickling. But when refrigerators were introduced to postwar Japan, fish that once were smoked or pickled could now be stored almost indefinitely. As the fishing industry adopted new technologies such as longlining (using extremely long lines with many baited hooks), purse seining (deploying large nets that can enclose an entire school of fish) and onboard freezers, the circumstances became propitious for an unprecedented modification of Japanese eating habits. The bluefin tuna changed from a fish that samurai would not eat because they believed it was unclean, to *maguro,* a delicacy that can be as expensive as truffles or caviar. *Toro,* the best quality *maguro,* comes from the fatty belly meat of the adult bluefin. Truffles or caviar are expensive because they are rare, but bluefin tuna, previously perceived as inedible, could be found in huge offshore schools and soon became an internationally exalted menu item. In 2001 a single bluefin tuna sold at the Tsukiji fish market in Tokyo for $173,600.

In America—where the thought of eating raw fish was anathema 40 years ago—sushi and sashimi have become commonplace, sold in supermarkets, delis and high-end restaurants. Perhaps the grandest temple to sushi in the U.S. is Masa, a New York City restaurant opened in

LUCY READING-IKKANDA

▲ **OVERFISHING** of the bluefin tuna is particularly severe in the Mediterranean Sea. In southern Spain, a group of tuna fishers raises a net full of bluefins (*above*) and brings the heap of giant fish to shore (*right*).

2004 by Japanese chef Masayoshi Takayama. With a prix fixe of $350 (excluding tax, tips and beverages), Masa immediately became the most expensive restaurant in the city; a lunch or dinner for two can easily exceed $1,000.

It stands to reason that a fish that can be sold for hundreds of dollars a slice is going to attract fishing fleets. The pell-mell rush to provide tuna for the Japanese sushi and sashimi markets has—not surprisingly—intensified tuna fishing around the world. The Japanese tried to fill their larders (and freezers and fish markets) with tuna that they could catch off their own shores (this would have been the Pacific bluefin, *Thunnus orientalis*), but they soon observed that the bluefins were larger and more plentiful in the North Atlantic. The buyers for Japanese fish importers became a familiar sight at the docks of American ports such as Gloucester and Barnstable in Massachusetts, ready to test the tuna for fat content and, if they passed, buy the fish on the spot and ship them to Japan.

At one time, researchers believed that there were two separate populations of North Atlantic bluefins (*Thunnus thynnus*), one that bred in the Gulf of Mexico and stayed in the western Atlantic and another that spawned in the Mediterranean and foraged in the eastern part of the ocean. The International Commission for the Conservation of Atlantic Tunas (ICCAT), a regulatory body established in 1969, based its catch quotas for the bluefin on this two-population concept, setting strict limits in the western Atlantic (where bluefin were becoming scarce as early as the 1970s) while allowing much larger catches in the eastern Atlantic. But tagging experiments—pioneered in the 1950s and 1960s by Frank J. Mather and Francis G. Carey of the Woods Hole Oceanographic Institution and refined in recent years by Barbara A. Block of Stanford University's Hopkins Marine Station—showed that the bluefin confounds the conventional wisdom. The Gulf of Mexico and the Mediterranean are indeed the breeding areas of the North Atlantic bluefins, but individual fishes can migrate across the ocean, and the foraging grounds of the two populations overlap. Because ICCAT has failed to stop overfishing in the eastern Atlantic, bluefin stocks have collapsed throughout the ocean.

If possible, things are worse in the Mediterranean. Employing ideas and technology originally developed in South Australia (with the southern bluefin, *Thunnus maccoyii*), fishers corral schools of half-grown tuna and tow them in floating pens to marine ranches where they are fed and fattened until they can be killed and shipped to Japan. There are rules banning fishing fleets from taking undersize tuna out of the Mediterranean, but none that prevent catching immature tuna and fattening them in floating pens. Every country on the Mediterranean (except Israel) takes advantage of this loophole and maintains tuna ranches offshore. The fishers from Spain, France, Italy, Greece, Turkey, Cyprus, Croatia, Egypt, Libya, Tunisia, Algeria,

[THE AUTHOR]

One of America's leading marine conservationists, **Richard Ellis** is generally recognized as the foremost painter of marine natural history subjects in the world. His paintings of whales have appeared in *Audubon, National Wildlife, Australian Geographic, Encyclopedia Britannica* and many other publications. His books include *The Book of Whales, The Book of Sharks, Imagining Atlantis* and *The Empty Ocean.* Ellis is a special adviser to the American Cetacean Society, a member of the Explorers Club and a research associate at the American Museum of Natural History in New York City. From 1980 to 1990 he was a member of the U.S. delegation to the International Whaling Commission. He is currently working on a book about tuna and serving as co-curator of the "Mythic Creatures" exhibit for the American Museum of Natural History.

AMERICAN MUSEUM OF NATURAL HISTORY (*Ellis*); NEIL LUCAS *Nature Picture Library* (*fishing photographs*)

▲ FISH BUYERS inspect the bluefin carcasses on sale at the Tsukiji wholesale market in Tokyo. This fish market was the site of the record purchase in 2001, when a single bluefin sold for 20.2 million yen, or about $173,600.

SUSHI HISTORY

The insatiable demand for sashimi (sliced raw fish) and sushi (rice topped or rolled with fish or vegetables) is the primary threat to the bluefin tuna.

Fourth Century B.C. Sushi originates in Southeast Asia as a method for preserving fish. Fermentation of the rice prevents the fish from rotting. The cuisine arrives in Japan in the eighth century A.D.

1800s Nigiri sushi, in which the fish is eaten raw rather than preserved, is popular in the market stalls of Edo, now known as Tokyo. But raw fish does not become a Japanese staple until the widespread adoption of refrigeration after World War II.

1970s to the Present Sushi consumption explodes in the U.S., but only high-end restaurants serve bluefin; most tuna sushi in America comes from the yellowfin or bigeye species (which are also threatened). The vast majority of bluefin is still consumed in Japan.

Morocco and Malta are capturing half-grown tuna by the hundreds of thousands. If you had to design a way to guarantee the decimation of a breeding population, this would be it: catch the fish before they are old enough to breed and keep them penned up until they are killed. The tuna ranches, once seen as a solution to the problem, are only making it worse. In 2006 the World Wildlife Fund called for the cessation of all tuna fishing in the Mediterranean, but given the tremendous financial rewards of the status quo, you can imagine how effective this plea was. At its meeting last November, ICCAT ignored the arguments of conservationists and set the 2008 quotas at approximately the same levels as 2007. The organization adopted a plan to scale back Mediterranean tuna fishing by 20 percent by 2010, with further reductions to follow, but the head of the U.S. delegation decried this half-measure, saying that ICCAT had "failed to live up to its founding mission."

Even if lower quotas were in place, however, the bluefin would still be imperiled. The tuna fishery is rife with illegal, unregulated fleets that ignore quotas, restrictions, boundaries, and any other rules and regulations that might threaten their catch. Furthermore, the Japanese market—which devours about 60,000 tons of bluefin every year, or more than three quarters of the global catch—is only too eager to buy the tuna, regardless of where or how it is caught. Japanese fishers have contrived to circumvent even their own country's restrictions, bringing in thousands of tons of illegal tuna every year and then falsifying their records. It would be good for the tuna and, in the end, good for the consumer if tuna fishing was not practiced in such a remorseless manner, but such change would entail nothing less than a modification of the fundamentals of human nature. As the tuna populations continue to fall, the Japanese demand for *toro* is increasing; fewer tuna will mean higher prices, and higher prices will mean intensified fishing. Intensified fishing will, of course, result in fewer tuna. (All bets would be off if the Japanese somehow relaxed their demand for *maguro,* but that seems as likely as Americans giving up hamburgers.) It appears that the only hope for the bluefin is captive breeding.

Cattle of the Sea

In an article entitled "When Will We Tame the Oceans?" that appeared in *Nature* in 2005, John Marra, a biological oceanographer at Columbia University's Lamont-Doherty Earth Observatory, observed that "fishing in the ocean is no longer sustainable. Worldwide, we have failed to manage the ocean's fisheries—in a few decades, there may be no fisheries left to manage." His recommendation? A large-scale domestication of the ocean—with fish farmers breeding, raising and harvesting commercially valuable species. Marra acknowledged that existing fish farms have harmed the environ-

COURTESY OF RICHARD ELLIS (*fish market*); DEX IMAGE/CORBIS (*sashimi*)

ment, polluting coastal ecosystems and putting additional pressure on wild fish populations by spreading disease and toxic chemicals. His solution is to move these so-called mariculture operations farther offshore, to the waters of the outer continental shelves, and to deploy much larger fish pens (closed net structures containing as much as 100,000 cubic meters of water) that could be floated below the surface and towed from one destination to another. This strategy would at least disperse the pollutants generated by fish farming, mitigating the environmental damage.

Marra also suggested taking advantage of the inclination of certain tunas to aggregate under an object that is significantly different from their surroundings. This propensity has already been exploited by fishers in the design and implementation of fish-aggregating devices, which are towed behind boats to attract schools of tuna. Instead of netting all the fish at once, though, fish farmers could create a sustainable business by feeding, maintaining and periodically harvesting some of the tuna in the school, handling the fish in much the same way that ranchers on land manage herds of cattle.

Unless tuna can be raised as if they were domesticated animals, their world populations will continue to crash. Breeding the bluefin in captivity, however, is a major challenge. One company that is attempting this feat is Clean Seas Aquaculture Growout, owned by the Stehr Group in Port Lincoln, South Australia. The Australian government has provided Clean Seas with a grant of 4.1 million Australian dollars ($3.4 million) to assist in the commercialization of southern bluefin breeding. The company has already raised captive-bred yellowtail kingfish (*Seriola lalandi*) and mulloway (*Argyrosomus hololepidotus*), which are now in significant commercial production. In October 2006 Clean Seas airlifted southern bluefin broodstock (sexually mature males and females) from their pens to a three-million-liter (790,000-gallon) tank that had been designed to replicate the optimum conditions for spawning. Hagen Stehr, founder of the company, said in a 2006 interview in *The Australian*, "We've got it all on computer, we can make [the tank] lighter or darker, we can leave the fish in a state of well-being, we've got the sun going up, the sun going down.... This is a world first, the Japanese won't try it at all, the Americans have tried it and failed and the Europeans have failed too."

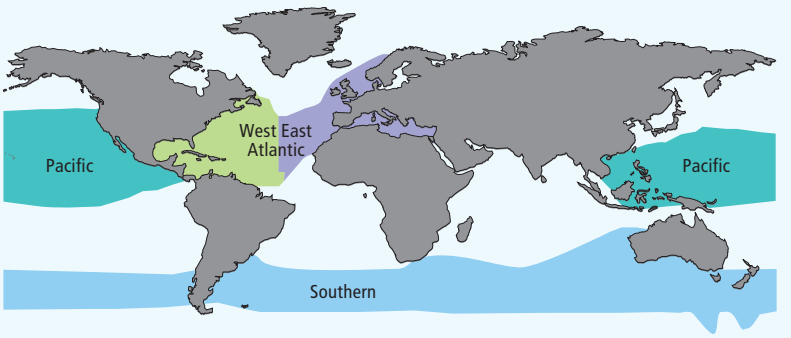
During my February 2007 visit to Port Lincoln, Rob Staunton, the farm manager for the Stehr Group, drove me to Arno Bay, 120 kilometers north of Port Lincoln, on the western shore of Spencer Gulf. I was granted limited entrée into the holy grail of the tuna business, the giant enclosed tuna tank at the Arno Bay hatchery. I say "limited" because my visit, personally sanctioned by Stehr himself, came with severe restrictions,

[THE BIG PICTURE]

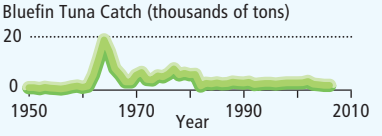
A WORLDWIDE DEPREDATION

Fishing fleets have extracted so many thousands of tons of bluefin tuna from the world's oceans that the species now faces extinction. The bluefin appears to be in the greatest peril in the West Atlantic. Despite the fact that the International Commission for the Conservation of Atlantic Tunas has imposed strict catch quotas in the West Atlantic since 1981, researchers estimate that the amount of sexually mature fish in that region (measured by their total mass) is less than 20 percent of what it was in the mid-1970s. Part of the problem is that bluefins from the West Atlantic migrate to the eastern part of the ocean, where the catch quotas are about 10 times higher. What is more, the reported catch figures (*graphs below*) do not include illegal fishing; some scientists believe the actual bluefin harvest in the East Atlantic and the Mediterranean Sea may be nearly twice as great as the reported data. Similar trends are shrinking the populations of the Pacific bluefin tuna and the Southern bluefins, which spawn in the Indian Ocean.

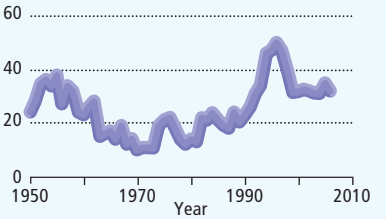
BLUEFIN TUNA GEOGRAPHIC DISTRIBUTION



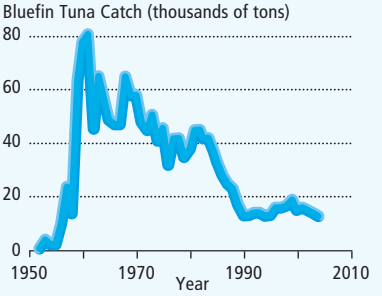
WEST ATLANTIC
Bluefin Tuna Catch (thousands of tons)



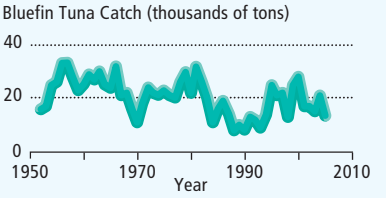
EAST ATLANTIC (and Mediterranean)
Bluefin Tuna Catch (thousands of tons)



SOUTHERN
Bluefin Tuna Catch (thousands of tons)



PACIFIC
Bluefin Tuna Catch (thousands of tons)



LUCY READING-IKKANDA; SOURCE: FISHERIES AND AQUACULTURE DEPARTMENT, FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS

▲ TUNA PENS, which are accelerating the demise of the bluefin, are towed across the Mediterranean Sea to tuna ranches near Sicily (*top*). Each pen contains about 250 bluefins. Off the coast of Ensenada, Mexico, tuna ranchers feed schools of Pacific bluefins until they are fat enough to be sold (*middle*). The slabs of fresh tuna are sent to Los Angeles International Airport, where they embark on nonstop flights to Japan. Inside a tuna pen in the Adriatic Sea off Croatia, a diver swims with the doomed bluefins before they, too, are shipped to Japan (*bottom*).

all of which are perfectly understandable. No photography is allowed in the facility itself, because the engineering, water processing, climate control and every other element in the design of this potential miracle must be carefully safeguarded to prevent corporate theft of the ideas. Along with the grant from the Australian government, the Stehr Group has invested millions in the innovative design of this facility, and it would be a disaster if someone borrowed or modified their designs and somehow beat them to the punch. It is hard to imagine anybody replicating this massive operation without all of Australia knowing about it, but of course, entrepreneurs in other countries—Japan, for instance—are also very interested in the business of captive-breeding bluefin tuna. Indeed, Japanese scientists at Kinki University have already hatched bluefin tuna from eggs and raised them to breeding age in the laboratory but not on the commercial scale being attempted by Clean Seas.

To begin our tour, Staunton and I had to change into special white rubber boots, sterilized to prevent the introduction of alien microbes into the tanks where the bluefin tuna are nurtured. Chaperoned by Thomas Marguritte, the Frenchman-turned-Australian who manages the facility, we exchanged our white boots for blue ones as we entered the sanctum sanctorum of the Arno Bay hatchery, the tuna-breeding tank. In a cavernous room illuminated by a battery of fluorescent lights, with the quiet hum of air-conditioning as the only background noise (the temperature outside was near 38 degrees C, or 100 degrees F), we climbed up to the concrete rim of the vast tank and looked down.

The tank is about 25 meters in diameter and six meters deep, and because the light level was fairly low, we could see very little until Marguritte tossed in a couple of small fish. Suddenly the surface broke with an ultramarine and chrome flash as one of the tuna charged at the baitfish. The tank came alive with froth, pierced by the sicklelike dorsal and tail fins of the tuna, which were anticipating a meal even though, as our docent explained, they had been fed only an hour before. As they circled excitedly underneath us, we could see that these were breeding-size bluefins: 300 kilograms of sleek, polished torpedo, pointed at both ends, with a dotted line of yellow finlets just before the tail, and the startling parentheses that mark the species' horizontal keels, chrome yellow in the southern bluefin and black in the northern varieties. No one can tell a live male from a live female except another tuna.

GAVIN NEWMAN *AFP/Getty Images (top)*; CHRIS PARK *AP Photo (middle)*; AFP/GETTY IMAGES (*bottom*)

Poised on the rim of the tank, we talked about the breeding program. "We can replicate the exact conditions in Indonesian waters where they are known to spawn naturally," Marguritte said. "If they usually spawn in the Southern Hemisphere summer when the days are longest and the water temperature is highest, we can make this tank conform to—pick a date, say, November 20—and set the length of daylight hours, air temperature, water temperature and even currents to conform to that moment in the Indian Ocean, south of the Indonesian archipelago." The only variable they cannot duplicate is the depth of the water, and they are praying that it is not a critical factor in the breeding of the southern bluefin. Just south of the Indonesian arc of islands—Java, Bali, Flores, Sumba, Komodo, Timor—is the Java Trench, which descends to one of the deepest points in the Indian Ocean, nearly eight kilometers down. If depth is a factor, the Clean Seas project is doomed. The broodstock at Clean Seas did not produce offspring in 2007, but they will try again this spring.

Taming the Bluefin

At the Clean Seas conference room in the Port Lincoln headquarters, I met with Marcus Stehr, Hagen's 42-year-old son and the managing director of the company. The day before, Marcus had been onboard one of the company's purse seiners in the Great Australian Bight, the huge open bay off the continent's southern coast, as a net cage containing perhaps 100 tons of tuna started on its journey to the pens off Port Lincoln. Like everyone else associated with this venture, Marcus is enthusiastic and optimistic about the potential for success and believes it is imminent. When I asked him if that success would completely change the way bluefin tuna are perceived in Australia, he said, "It's not a question of *if*, mate—it's *when*."

Although the Aussies appear to be in the lead, it remains to be seen if they, the Japanese, or the Europeans will win the race to breed the bluefin in captivity. In 2005, for example, a research team at the Spanish Institute of Oceanography in Puerto de Mazarrón, Spain, successfully retrieved eggs and sperm from captive Atlantic bluefin broodstock, performed in vitro fertilization and produced larvae. (The hatchlings of bony marine fishes are called larvae because they look so different from the adults.) Somehow or other, it has to happen, because the survival of the species—and the tuna industry—depends on it.

The big-game fisher sees the bluefin tuna as a sleek and powerful opponent; to the harpooner, it is an iridescent shadow below the surface, flicking its scythelike tail to propel it out of range; the purse seiner sees a churning maelstrom of silver and blue bodies to be hauled onboard his boat; the long-liner sees a dead fish, pulled onto the deck along with many other glistening marine creatures; the tuna rancher sees the bluefin as an anonymous creature to be force-fed until it is time to drive a spike into its brain; the auctioneer at the Tsukiji fish market in Tokyo sees row on row of tailless, icy, tuna-shaped blocks; Japanese consumers see it as *toro,* a slice of rich red meat to be eaten with wasabi and soy sauce; to the biologist, the tuna is a marvel of hydrodynamic engineering, its body packed with modifications that enable it to outeat, outgrow, outswim, outdive, and outmigrate any other fish in the sea; and to those who wish to rescue *Thunnus thynnus* from biological oblivion, it has to be seen as a domesticated animal, like a sheep or a cow.

For some, such a shift is almost impossible to contemplate; the bluefin tuna, the quintessential ocean ranger, the wildest, most powerful fish in the sea, cannot be—and probably *should* not be—tamed. But if it remains wild, the future looks bleak for the *maguro* industry—and for the great bluefin tuna. ■

▲ ACTIVISTS from Greenpeace, the global environmental group, staged a protest at a Mediterranean beach in southern Turkey last November to coincide with the annual conference of the International Commission for the Conservation of Atlantic Tunas, which was held in that country. Ignoring the calls for an immediate cessation of bluefin fishing in the Mediterranean, the regulatory body adopted a plan that called for modest reductions of catch quotas over the next few years.

GREENPEACE/HO *AP Photo*

MORE TO EXPLORE

Giant Bluefin. Douglas Whynott. North Point Press, 1995.

Song for the Blue Ocean. Carl Safina. Henry Holt and Company, 1997.

Rapid Worldwide Depletion of Predatory Fish Communities. Ransom A. Myers and Boris Worm in *Nature,* Vol. 423, pages 280–283; May 15, 2003.

Dollars without Sense: The Bait for Big-Money Tuna Ranching around the World. John P. Volpe in *BioScience* Vol. 55, No. 4, pages 301–302; April 2005.

Electronic Tagging and Population Structure of Atlantic Bluefin Tuna. Barbara A. Block et al. in *Nature,* Vol. 434, pages 1121–1127; April 28, 2005.

The Sushi Economy. Sasha Issenberg. Gotham, 2007.

SPACE WEAPONS CONCEPTS include a variety of satellite killers—projectiles, microwave- and laser-beam weapons, and orbital mines—as well as arms launched from space at surface targets, such as the heavy tungsten bunker busters nicknamed "rods from God."

A RECENT SHIFT IN U.S. MILITARY STRATEGY AND PROVOCATIVE ACTIONS BY CHINA THREATEN TO IGNITE A NEW ARMS RACE IN SPACE. BUT WOULD PLACING WEAPONS IN SPACE BE IN ANYONE'S NATIONAL INTEREST?

BY THERESA HITCHENS

In war, do not launch an ascending attack head-on against the enemy who holds the high ground. Do not engage the enemy when he makes a descending attack from high ground. Lure him to level ground to do battle.

—Sun Tzu, Chinese military strategist, *The Art of War,* circa 500 B.C.

"Take the high ground and hold it!" has been standard combat doctrine for armies since ancient times. Now that people and their machines have entered outer space, it is no surprise that generals the world over regard Earth orbit as the key to modern warfare. But until recently, a norm had developed against the weaponization of space—even though there are no international treaties or laws explicitly prohibiting nonnuclear antisatellite systems or weapons placed in orbit. Nations mostly shunned such weapons, fearing the possibility of destabilizing the global balance of power with a costly arms race in space.

That consensus is now in danger of unraveling. In October 2006 the Bush administration adopted a new, rather vaguely worded National Space Policy that asserts the right of the U.S. to conduct "space control" and rejects "new legal regimes or other restrictions that seek to prohibit or limit U.S. access to or use of space." Three months later the People's Republic of China shocked the world by shooting down one of its own aging Fengyun weather satellites, an act that resulted in a hailstorm of dangerous orbital debris and a deluge of international protests, not to mention a good deal of handwringing in American military and political circles. The launch was the first test of a dedicated antisatellite weapon in more than two decades—making China only the third country, after the U.S. and the Russian Federation, to have demonstrated such a technology. Many observers wondered whether the test might be the first shot in an emerging era of space warfare.

Critics maintain it is not at all clear that a nation's security would be enhanced by developing the means to wage space war. After all, satellites and even orbiting weapons, by their very nature, are relatively easy to spot and easy to track, and they are likely to remain highly vulnerable to attack no matter what defense measures are taken. Further, developing antisatellite systems would almost surely lead to a hugely expensive and potentially runaway arms race, as other countries would conclude that they, too, must compete. And even tests of the technology needed to conduct space battles—not to mention a real battle—could generate enormous amounts of wreckage that would continue to orbit Earth. Converging on satellites and crewed space vehicles at speeds approaching several miles a second, such space debris would threaten satellite-based telecommunications, weather forecasting, precision navigation, even military command and control, potentially sending the world's economy back to the 1950s.

KEY CONCEPTS

- Although the "high ground" of outer space seems to offer clear military advantages, nations have so far resisted placing weapons into Earth orbit. That strategic forbearance may be changing.
- The National Space Policy adopted by the U.S. in 2006 seemed to open the way to the further militarization of space. Soon afterward, China tested a ground-based antisatellite missile.
- But space weaponry could trigger a costly international arms race. Satellites and space weapons will remain vulnerable to attack no matter what defenses are mounted.
- And space warfare, or even "live" tests of the weapons, could create so much space debris that Earth orbit would become unnavigable to civilian satellites and crewed spacecraft.

—The Editors

ILLUSTRATIONS BY KENN BROWN

"Star Wars" Redux

Since the dawn of the space age, defense planners have hatched concepts for antisatellite and space-based weaponry—all in the interest of exploiting the military advantages of the ultimate high ground. Perhaps the most notable effort was President Ronald Reagan's Strategic Defense Initiative (SDI)—derided by its critics as "Star Wars." Yet by and large, U.S. military strategy has never embraced such weapons.

Traditionally, space weapons have been defined as destructive systems that operate in outer space after having been launched directly from Earth or parked in orbit. The category includes antisatellite weapons; laser systems that couple ground-based lasers with airship- or satellite-mounted mirrors, which could reflect a laser beam beyond the ground horizon; and orbital platforms that could fire projectiles or energy beams from space. (It is important to note that all nations would presumably avoid using a fourth kind of antisatellite weapon, namely, a high-altitude nuclear explosion. The electromagnetic pulse and cloud of highly charged particles created by such a blast would likely disable or destroy nearly all satellites and manned spacecraft in orbit [see "Nuclear Explosions in Orbit," by Daniel G. Dupont; Scientific American, June 2004].)

But virtually no statement about space weapons goes politically uncontested. Recently some proponents of such weapons have sought to expand the long-held classification I just described to include two existing technologies that depend on passage through space: intercontinental ballistic missiles (ICBMs) and ground-based electronic warfare systems. Their existence, or so the argument goes, renders moot any question about whether to build space weapons systems. By the revised definition, after all, "space weapons" already exist. Whatever the exact meaning of the term, however, the questions such weapons raise are hardly new to think tanks and military-planning circles in Washington: Is it desirable, or even feasible, to incorporate antisatellite weapons and weapons fired from orbit into the nation's military strategy?

The new National Space Policy, coupled with the Chinese test, has brought renewed urgency to that behind-the-scenes debate. Many American military leaders expressed alarm in the wake of the Chinese test, worrying that in any conflict over Taiwan, China could threaten U.S. satellites in low Earth orbit. In April 2007 Michael Moseley, the U.S. Air Force chief of staff, compared China's antisatellite test with the launch of Sputnik by the Soviet Union in 1957, an act that singularly intensified the arms race during the cold war. Moseley also revealed that the Pentagon had begun reviewing the nation's satellite defenses, explaining that outer space was now a "contested domain."

Congressional reaction fell along predictable political lines. Conservative "China hawks" such as Senator Jon Kyl of Arizona immediately called for the development of antisatellite weapons and space-based interceptors to counter Chinese capabilities. Meanwhile more moderate politicians, including Representative Edward Markey of Massachusetts, urged the Bush administration to begin negotiations aimed at banning all space weapons.

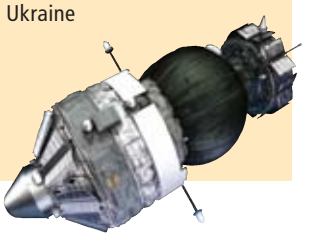
THE PLAYERS

Since the start of the space age, the list of countries, multinational entities and private commercial consortia that have demonstrated an ability to launch satellites into orbit—and thus potentially to shoot one down—has grown long. The chief worry among observers is that any effort by the U.S. to develop orbital weapons would drive the People's Republic of China, the Russian Federation and others to join in a costly arms race in space.

DEMONSTRATED GROUND-BASED ANTISATELLITE WEAPONS
China, Russia, U.S.

ATTAINED GEOSTATIONARY ORBIT (36,000 km above Earth)
European Space Agency (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, U.K.), France, International Launch Services (Russia, U.S.), Japan, Sea Launch (Norway, Russia, U.S.)

ATTAINED ONLY LOW EARTH ORBIT (between 100 and 2,000 km above Earth)
India, Israel, Pakistan, Ukraine



ESA

International Power Plays

Perhaps of even greater concern is that several other nations, including one of China's regional rivals, India, may feel compelled to seek offensive as well as defensive capabilities in space. The U.S. trade journal *Defense News,* for instance, quoted unidentified Indian defense officials as stating that their country had already begun developing its own kinetic-energy (nonexplosive, hit-to-kill) and laser-based antisatellite weapons.

If India goes down that path, its archrival Pakistan will probably follow suit. Like India, Pakistan has a well-developed ballistic missile program, including medium-range missiles that could launch an antisatellite system. Even Japan, the third major Asian power, might join such a space race. In June 2007 the National Diet of Japan began considering a bill backed by the current Fukuda government that would permit the development of satellites for "military and national security" purposes.

As for Russia, in the wake of the Chinese test President Vladimir Putin reiterated Moscow's stance against the weaponization of space. At the same time, though, he refused to criticize Beijing's actions and blamed the U.S. instead. The American efforts to build a missile defense system, Putin charged, and the increasingly aggressive American plans for a military position in space were prompting China's moves. Yet Russia itself, as a major spacefaring power that has incorporated satellites into its national security structure, would be hard-pressed to forgo entering an arms race in space.

Given the proliferation of spacefaring entities [*see box at left*], proponents of a robust space

warfare strategy believe that arming the heavens is inevitable and that it would be best for the U.S. to get there first with firepower. Antisatellite and space-based weapons, they argue, will be necessary not only to defend U.S. military and commercial satellites but also to deny any future adversary the use of space capabilities to enhance the performance of its forces on the battlefield.

Yet any arms race in space would almost inevitably destabilize the balance of power and thereby multiply the risks of global conflict. In such headlong competition—whether in space or elsewhere—equilibrium among the adversaries would be virtually impossible to maintain. Even if the major powers did achieve stability, that reality would still provide no guarantee that both sides would perceive it to be so. The moment one side saw itself to be slipping behind the other, the first side would be strongly tempted to launch a preemptive strike, before things got even worse. Ironically, the same would hold for the side that perceived itself to have gained an advantage. Again, there would be strong temptation to strike first, before the adversary could catch up. Finally, a space weapons race would ratchet up the chances that a mere technological mistake could trigger a battle. After all, in the distant void, reliably distinguishing an intentional act from an accidental one would be highly problematic.

Hit-to-Kill Interceptors

According to assessments by U.S. military and intelligence officials as well as by independent experts, the Chinese probably destroyed their weather satellite with a kinetic-energy vehicle boosted by a two-stage medium-range ballistic missile. Technologically, launching such direct-ascent antisatellite weapons is one of the simplest ways to take out a satellite [*see box at right*]. About a dozen nations and consortia can reach low Earth orbit (between roughly 100 and 2,000 kilometers, or 60 to 1,250 miles, high) with a medium-range missile; eight of those countries can reach geostationary orbit (about 36,000 kilometers, or 22,000 miles, above Earth).

But the real technical hurdle to making a hit-to-kill vehicle is not launch capacity; it is the precision maneuverability and guidance technology needed to steer the vehicle into its target. Just how well China has mastered those techniques is unclear. Because the weather satellite was still operating when it was destroyed, the Chinese operators would have known its exact location at all times.

Ground-Based Lasers

The test of China's direct-ascent antisatellite device came on the heels of press reports in September 2006 that the Chinese had also managed to "paint," or illuminate, U.S. spy satellites with a ground-based laser [*see lower box on page 83*]. Was Beijing actually trying to "blind" or otherwise damage the satellites? No one knows, and no consensus seems to have emerged in official Washington circles about the Chinese intent. Perhaps China was simply testing how well its network of low-power laser-ranging stations could track American orbital observation platforms.

Even so, the test was provocative. Not all satellites have to be electronically "fried" to be put out of commission. A 1997 test of the army's MIRACL system (for *mi*d*i*nfra*r*ed *a*dvanced *c*hemical *l*aser) showed that satellites designed to collect optical images can be temporarily disrupted—dazzled—by low-power beams. It follows that among the satellites vulnerable to such an attack are the orbital spies.

The U.S. and the former Soviet Union began experimenting with laser-based antisatellite weapons in the 1970s. Engineers in both coun-

KINETIC-ENERGY INTERCEPTORS

FEASIBILITY: High

COST ESTIMATES*:

- **Ground-based kinetic-energy interceptor (adapted from existing U.S. ballistic missile defense program): $0–$3 billion**
- **Airborne kinetic-energy interceptor: $3 billion**

Apart from jamming the radio communications or attacking ground-control stations, probably the simplest way to disable a satellite is to launch a missile-borne payload and crash it into an orbital target. Medium-range ballistic missiles fielded by about a dozen nations can directly reach low Earth orbit (between 100 and 2,000 kilometers, or about 60 to 1,250 miles, high). Small air-launched kill vehicles can also attack satellites in low Earth orbit. Assaulting a target in the much higher geostationary orbit (about 36,000 kilometers, or 22,000 miles, high) requires a more powerful launch booster, now possessed by eight countries and space consortia. But the real technical challenge is to guide and maneuver the kill vehicle precisely onto its mark.

**Estimates generally include development and procurement costs associated with building a system and operating it for 20 years.*

SOURCE: Arming the Heavens: A Preliminary Assessment of the Potential Cost and Cost-Effectiveness of Space-Based Weapons, *by Steven Kosiak. Center for Strategic and Budgetary Assessments, 2007.*

tries have focused on the many problems of building high-power laser systems that could reliably destroy low-flying satellites from the ground. Such systems could be guided by "adaptive optics": deformable mirrors that can continuously compensate for atmospheric distortions. But tremendous amounts of energy would be needed to feed high-power lasers, and even then the range and effectiveness of the beams would be severely limited by dispersion, by attenuation as they passed through smoke or clouds, and by the difficulty of keeping the beams on-target long enough to do damage.

During the development of the SDI, the U.S. conducted several laser experiments from Hawaii, including a test in which a beam was bounced off a mirror mounted on a satellite. Laser experiments continue at the Starfire Optical Range at Kirtland Air Force Base in New Mexico. Pentagon budget documents from fiscal years 2004 through 2007 listed antisatellite operations among the goals of the Starfire research, but that language was removed from budget documents in fiscal year 2008 after Congress made inquiries. The Starfire system incorporates adaptive optics that narrow the outgoing laser beam and thus increase the density of its power. That capability is not required for imagery or tracking, further suggesting that Starfire could be used as a weapon.

Yet despite decades of work, battle-ready versions of directed-energy weapons still seem far away. An air force planning document, for instance, predicted in 2003 that a ground-based weapon able to "propagate laser beams through the atmosphere to [stun or kill low Earth orbit] satellites" could be available between 2015 and 2030. Given the current state of research, even those dates seem optimistic.

Co-orbital Satellites

Recent advances in miniaturized sensors, powerful onboard computers and efficient rocket thrusters have made a third kind of antisatellite technology increasingly feasible: the offensive microsatellite [*see upper box on opposite page*]. One example that demonstrates the potential is the air force's experimental satellite series (XSS) project, which is developing microsatellites intended to conduct "autonomous proximity operations" around larger satellites. The first two microsatellites in the program, the XSS-10 and XSS-11, were launched in 2003 and 2005. Though ostensibly intended to inspect larger satellites, such microsatellites could also ram target satellites or carry explosives or directed-energy payloads such as radio-frequency jamming systems or high-powered microwave emitters. Air force budget documents show that the XSS effort is tied to a program called Advanced Weapons Technology, which is dedicated to research on military laser and microwave systems.

During the cold war the Soviet Union developed, tested and even declared operational a co-orbital antisatellite system—a maneuverable interceptor with an explosive payload that was launched by missile into an orbit near a target satellite in low Earth orbit. In effect, the device was a smart "space mine," but it was last demonstrated in 1982 and is probably no longer working. Today such an interceptor would likely be a microsatellite that could be parked in an orbit that would cross the orbits of several of its potential targets. It could then be activated on command during a close encounter.

In 2005 the air force described a program that would provide "localized" space "situational awareness" and "anomaly characterization" for friendly host satellites in geostationary orbit. The program is dubbed ANGELS (for *a*utonomous *n*anosatellite *g*uardian for *e*valuating *l*ocal *s*pace), and the budget line believed to represent it focuses on acquiring "high value space asset defensive capabilities," including a "warning sensor for detection of a direct ascent or co-orbital vehicle." It is clear that such guardian nanosatellites could also serve as offensive weapons if they were maneuvered close to enemy satellites.

And the list goes on. A "parasitic satellite" would shadow or even attach itself to a target in geostationary orbit. Farsat, which was mentioned in an appendix to the [Donald] Rumsfeld Space Commission report in 2001, "would be placed in a 'storage' orbit (perhaps with many microsatellites housed inside) relatively far from its target but ready to be maneuvered in for a kill."

Finally, the air force proposed some time ago a space-based radio-frequency weapon system, which "would be a constellation of satellites containing high-power radio-frequency transmitters that possess the capability to disrupt/destroy/disable a wide variety of electronics and national-level command and control systems."

Air force planning documents from 2003 envisioned that such a technology would emerge after 2015. But outside experts think that orbital radio-frequency and microwave weapons are technically feasible today and could be deployed in the relatively near future.

THE CASE AGAINST

1. All satellites and space-based weapons are likely to remain highly vulnerable to attack.
2. Developing advanced antisatellite weapons will probably trigger a new international arms race.
3. The cost of space weaponry is huge.
4. Testing and using space weapons could leave enormous quantities of debris in orbit that would threaten all satellites and crewed spacecraft.

Space Bombers

Though not by definition a space weapon, the Pentagon's Common Aero Vehicle/Hypersonic Technology Vehicle (often called CAV) enters into this discussion because, like an ICBM, it would travel through space to strike Earth-bound targets [*see top box on next page*]. An unpowered but highly maneuverable hypersonic glide vehicle, the CAV would be deployed from a future hypersonic space plane, swoop down into the atmosphere from orbit and drop conventional bombs on ground targets. Congress recently began funding the project but, to avoid stoking a potential arms race in space, has prohibited any work to place weapons on the CAV. Although engineers are making steady progress on the key technologies for the CAV program, both the vehicle and its space plane mothership are still likely decades off.

Some of the congressional sensitivity to the design of the CAV may have arisen from another, much more controversial space weapons concept with parallel goals: hypervelocity rod bundles that would be dropped to Earth from orbital platforms. For decades air force planners have been thinking about placing weapons in orbit that could strike terrestrial targets, particularly buried, "hardened" bunkers and caches of weapons of mass destruction. Commonly called "rods from God," the bundles would be made up of large tungsten rods, each as long as six meters (20 feet) and 30 centimeters (12 inches) across. Each rod would be hurled downward from an orbiting spacecraft and guided to its target at tremendous speed.

Both high costs and the laws of physics, however, challenge their feasibility. Ensuring that the projectiles do not burn up or deform from reentry friction while sustaining a precise, nearly vertical flight path would be extremely difficult. Calculations indicate that the nonexplosive rods would probably be no more effective than more conventional munitions. Furthermore, the expense of lofting the heavy projectiles into orbit would be exorbitant. Thus, despite continued interest in them, rods from God seem to fall into the realm of science fiction.

Obstacles to Space Weapons

What, then, is holding the U.S. (and other nations) back from a full-bore pursuit of space weapons? The countervailing pressures are threefold: political opposition, technological challenges and high costs.

The American body politic is deeply divided

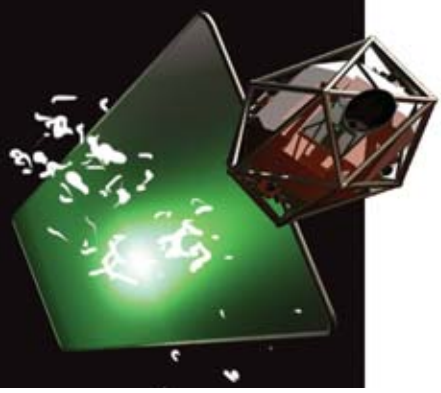
CO-ORBITAL SATELLITES

FEASIBILITY: Medium to high

COST ESTIMATES:

- Space-based (kinetic-energy and other) interceptor: $5 billion–$19 billion
- Space-based radio-frequency jammer: not available
- Space mine: $100 million–$2 billion

Small antisatellite weapons, or microsatellites, could be lofted into the same orbits as their targets, where they could shadow or attach themselves to the targets. Once in place, such "space mines" could attack on command with explosives, small projectiles, radio-frequency jamming systems or high-powered microwave emitters—or they could simply smash into their targets. In one early design the resulting space debris was to be caught in a so-called flyswatter or large net (*right*).



DIRECTED-ENERGY SYSTEMS

FEASIBILITY: Medium

COST ESTIMATES:

- Ground-based laser: $4 billion–$6 billion
- Space-based laser (low- to high-power capability): $3 billion–$60 billion
- Space-based microwave radiator: $200 million–$5 billion

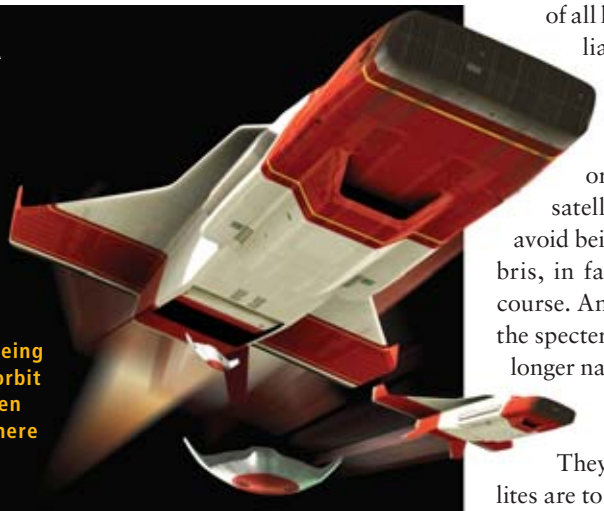
Ground-based laser beams that are precision-guided onto their targets by adaptive optics (deformable mirrors that compensate for atmospheric distortions) could blind, disable or destroy satellites in low Earth orbit. Moderate-power lasers could "dazzle" optical-imaging satellites or damage their sensitive detectors. High-power lasers could "fry" satellites by damaging their electronics or even piercing their skin. Because fast-moving orbital targets lie mostly over Earth's horizon at any one time, ground stations could also direct laser beams at airships or satellite-borne transfer mirrors, which could redirect the beams toward their targets.

SPACE BOMBER

FEASIBILITY: Low

COST ESTIMATE:
- Space bomber: $4 billion

The Pentagon's Common Aero Vehicle/Hypersonic Technology Vehicle is not by definition a space weapon, but it would travel through space to strike terrestrial targets within an hour or two of being deployed. It could be released in orbit from a hypersonic space plane, then glide unpowered into the atmosphere before delivering conventional munitions onto surface targets.



over the wisdom of making space warfare a part of the national military strategy. The risks are manifold. I remarked earlier on the general instabilities of an arms race, but there is a further issue of stability among the nuclear powers. Early-warning and spy satellites have traditionally played a crucial role in reducing fears of a surprise nuclear attack. But if antisatellite weapons disabled those eyes-in-the-sky, the resulting uncertainty and distrust could rapidly lead to catastrophe.

One of the most serious technological challenges posed by space weapons is the proliferation of space debris, to which I alluded earlier. According to investigators at the air force, NASA and Celestrak (an independent space-monitoring Web site), the Chinese antisatellite test left more than 2,000 pieces of junk, baseball-size and larger, orbiting the globe in a cloud that lies between about 200 kilometers (125 miles) and 4,000 kilometers (2,500 miles) above Earth's surface. Perhaps another 150,000 objects that are a centimeter (half an inch) across and larger were released. High orbital velocities make even tiny pieces of space junk dangerous to spacecraft of all kinds. And ground stations cannot reliably monitor or track objects smaller than about five centimeters (two inches) across in low Earth orbit (around a meter in geostationary orbit), a capability that might enable satellites to maneuver out of the way. To avoid being damaged by the Chinese space debris, in fact, two U.S. satellites had to alter course. Any shooting war in space would raise the specter of a polluted space environment no longer navigable by Earth-orbiting satellites.

Basing weapons in orbit also presents difficult technical obstacles. They would be just as vulnerable as satellites are to all kinds of outside agents: space debris, projectiles, electromagnetic signals, even natural micrometeoroids. Shielding space weapons against such threats would also be impractical, mostly because shielding is bulky and adds mass, thereby greatly increasing launch costs. Orbital weapons would be mostly autonomous mechanisms, which would make operational errors and failures likely. The paths of objects in orbit are relatively easy to predict, which would make hiding large weapons problematic. And because satellites in low Earth orbit are overhead for only a few minutes at a time, keeping one of them constantly in range would require many weapons.

Finally, getting into space and operating there is extremely expensive: between $2,000 and $10,000 a pound to reach low Earth orbit and between $15,000 and $20,000 a pound for geostationary orbit. Each space-based weapon would require replacement every seven to 15 years, and in-orbit repairs would not be cheap, either.

Alternatives to Space Warfare

Given the risks of space warfare to national and international security, as well as the technical and financial hurdles that must be overcome, it

OTHER ANTISATELLITE SYSTEMS

Most of the major military powers have probably experimented with ground-based radio-frequency systems that could disable the communications systems of satellites. Moreover, any country with nuclear-tipped ballistic missiles could explode an atomic weapon in orbit, which would wreak havoc on most of the satellites and spacecraft there.

COST ESTIMATES:
- Ground-based radio-frequency jammer: several tens of millions of dollars
- Nuclear weapon (for nations already possessing missiles with nuclear warheads): minimal

CENTER FOR DEFENSE INFORMATION (*Hitchens*)

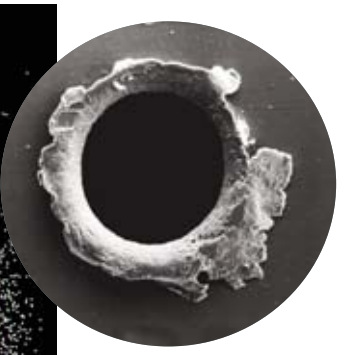
[THE AUTHOR]

Theresa Hitchens directs the Center for Defense Information in Washington, D.C., and leads its Space Security Project, in cooperation with the Secure World Foundation. She is author of *Future Security in Space: Charting a Cooperative Course* (2004) and was editor of *Defense News* from 1998 until 2000. As a journalist, she has focused on the military, the defense industry and North Atlantic Treaty Organization affairs. Most recently, Hitchens served as director of research at the British American Security Information Council, a defense think tank.

[THE AFTERMATH]

WHEN THE DUST WON'T CLEAR

A military conflict in space could release an enveloping cloud of debris that could damage or destroy satellites and crewed spacecraft that circle the globe. At orbital speeds, even minuscule objects could deeply penetrate a vehicle and wreck vital equipment (*far right*). The results of a nuclear detonation in space could be even worse: the electromagnetic pulse and blast of charged particles would degrade all but the most heavily shielded electronics systems in orbit. Space war could push the world economy back into the 1950s, as communications, navigation, weather and other advanced satellite services would be rendered impractical for years to come.



▲ PUNCTURE to NASA's Solar Maximum Mission satellite resulted from a strike by tiny orbital debris.

would seem only prudent for spacefaring nations to find ways to prevent an arms race in space. The U.S. focus has been to reduce the vulnerability of its satellite fleet and explore alternatives to its dependence on satellite services. Most other space-capable countries are instead seeking multilateral diplomatic and legal measures. The options range from treaties that would ban antisatellite and space-based weapons to voluntary measures that would help build transparency and mutual confidence.

The Bush administration has adamantly opposed any form of negotiations regarding space weapons. Opponents of multilateral space weapons agreements contend that others (particularly China) will sign up but build secret arsenals at the same time, because such treaty violations cannot be detected. They argue further that the U.S. cannot sit idly as potential adversaries gain spaceborne resources that could enhance their terrestrial combat capabilities.

Proponents of international treaties counter that failure to negotiate such agreements entails real opportunity costs. An arms race in space may end up compromising the security of all nations, including that of the U.S., while it stretches the economic capacities of the competitors to the breaking point. And whereas many advocates of a space weapons ban concede that it will be difficult to construct a fully verifiable treaty—because space technology can be used for both military and civilian ends—effective treaties already exist that do not require strict verification. A good example is the Biological Weapons Convention. Certainly a prohibition on the testing and use (as opposed to the deployment) of the most dangerous class of near-term space weapons—destructive (as opposed to jamming) antisatellite systems—would be easily verifiable, because earthbound observers can readily detect orbital debris. Furthermore, any party to a treaty would know that all its space launches would be tracked from the ground, and any suspicious object in orbit would promptly be labeled as such. The international outcry that would ensue from such overt treaty violations could deter would-be violators.

Since the mid-1990s, however, progress on establishing a new multilateral space regime has lagged. The U.S. has blocked efforts at the United Nations Conference on Disarmament in Geneva to begin negotiations on a treaty to ban space weapons. China, meanwhile, has refused to accept anything less. Hence, intermediate measures such as voluntary confidence-building, space traffic control or a code of responsible conduct for spacefaring nations have remained stalled.

Space warfare is not inevitable. But the recent policy shift in the U.S. and China's provocative actions have highlighted the fact that the world is approaching a crossroads. Countries must come to grips with their strong self-interest in preventing the testing and use of orbital weapons. The nations of Earth must soon decide whether it is possible to sustain the predominantly peaceful human space exploration that has already lasted half a century. The likely alternative would be unacceptable to all. ■

MORE TO EXPLORE

Report of the Commission to Assess United States National Security Space Management and Organization. Rumsfeld Space Commission report, 2001. Available at **www.fas.org/spp/military/commission/report.htm**

The U.S. Air Force Transformation Flight Plan. Future Concepts and Transformation Division, November 2003. Available at **www.af.mil/library/posture/AF_TRANS_FLIGHT_PLAN-2003.pdf**

The Physics of Space Security: A Reference Manual. David Wright, Laura Grego and Lisbeth Gronlund. American Academy of Arts and Sciences, 2005. Available at **www.ucsusa.org/global_security/space_weapons/the-physics-of-space-security.html**

China's ASAT Test: Motivations and Implications. Phillip C. Saunders and Charles D. Lutes. INSS Special Report, Institute for National Strategic Studies. National Defense University, 2007. Available at **www.ndu.edu/inss/Research/SRjun07.pdf**

The World Security Institute's Center for Defense Information: **www.cdi.org**

NASA ORBITAL DEBRIS PROGRAM OFFICE

SOLVING A MASSIVE WORKER

The largest workplace health study ever conducted is applying cutting-edge techniques to investigating an apparent cancer cluster—and highlighting the reasons why science doesn't always protect us at work • By Carole Bass

KEY CONCEPTS

- A seven-year investigation of brain cancer cases among Pratt & Whitney Aircraft workers in Connecticut has become the largest workplace health study ever performed, covering some quarter of a million employees over a span of 50 years.
- The study's massive scale has made data collection labor-intensive, but the size of the study population should also give the analyses power to detect even subtle patterns that might point to tumor causes, including previously unrecognized brain cancer triggers.
- More industrial epidemiology of this kind could improve worker health protections, many of which are extremely outdated. But funding and political support for such research are lacking.

—The Editors

In John Shea and John Greco's day, the cavernous Pratt & Whitney Aircraft plant was filled with an oily mist that sprayed from the grinding machines, coated the ceiling and covered the workers, who came home drenched in pungent machine oil. Degreasing pits, filled with solvent for cleaning the engine parts, dotted the factory floor; workers used squirt cans of solvent to clean their hands and clothes. Shea spent 34 years grinding engine blades and vanes at the million-square-foot facility in North Haven, Conn. In 1999, at age 56, he was diagnosed with brain cancer. Six months later Shea's friend and co-worker Greco learned he had the same disease: glioblastoma multiforme, the most aggressive type of brain tumor. A year after Shea's diagnosis, both men were dead, but their widows had already begun asking questions about the seemingly unusual number of cases of this deadly form of cancer at one of the world's top jet-engine manufacturers.

What began in 2001 as an investigation into an apparent cluster of brain cancers at North Haven—13 cases of primary malignant brain tumor among the workers, 11 of them glioblastoma, in just the previous decade—has turned into the largest workplace health study ever conducted. A team led by principal investigator Gary Marsh of the University of Pittsburgh and Nurtan Esmen of the University of Illinois at Chicago is engaged in painstaking detective work to solve a complex puzzle: first the researchers must trace an as yet undisclosed number of brain cancer cases among nearly 250,000 employees at eight Pratt & Whitney plants over a span of 50 years and then determine, if possible, what might have caused the tumors by reconstructing workers' exposures to a slew of potentially toxic agents. The group expects to publish preliminary findings in the first half of 2008 and final results in 2009.

Marsh and Esmen's logistically daunting task illustrates the difficulty of such workplace epidemiology, involving multiple exposures at multiple facilities. The researchers' ability to provide concrete answers about what happened to Pratt & Whitney workers in the past will also be limited by incomplete scientific understanding, both of brain tumor triggers and of the toxicity of many chemicals used in industry. The Pratt & Whitney study, using the most sophisticated techniques available, could shed new light on both subjects. The investigation also underscores the fact that determining safe exposures to workplace toxics remains very much a problem of the present.

The National Institute for Occupational Safety and Health estimates that nearly 49,000 Americans die prematurely every year from work-related illnesses—more than eight times the number killed in on-the-job accidents. Most federal workplace exposure limits are based on science from the 1960s, however. As a result, newer Environmental Protection Agency guidelines for safe levels of a chemical in the air outside a factory can be as much as 45,000 times lower than regulations governing air inside the plant set by the Occupational Safety and Health Administration (OSHA). Politics and economics,

PHOTOGRAPH OF RUBIK'S CUBE BY JAMES PORTO; PHOTOILLUSTRATION BY JEN CHRISTIANSEN; BETTMANN/CORBIS (*yellow*); JAMES PORTO (*blue*); UNIVERSITY OF PITTSBURGH (*red and orange*); CONNECTICUT DEPARTMENT OF ENVIRONMENTAL PROTECTION (*green*)

HEALTH PUZZLE

rather than the limitations of science, may be the greatest barriers to updating those worker health protections. The Pratt & Whitney investigation therefore also illustrates how much better occupational epidemiology could be if the political will existed to bring the best modern science to the task of making workplaces safe.

Amassing Evidence

When their husbands were diagnosed with the same rare tumor, Carol Shea and Kate Greco knew nothing about brain cancer or epidemiology. But it seemed an unlikely coincidence, so they started asking Pratt & Whitney for answers: How many other workers had brain cancer?

[EVIDENCE COLLECTION]

BUILDING THE DATABASES

To determine whether an unusual number of cases of brain cancer arose among Pratt & Whitney Aircraft workers and, if so, why, the investigators spent more than five years compiling enormous volumes of information about all of the company's Connecticut workers and manufacturing processes over a 50-year period. One of the two research teams set out to identify the study population, or cohort, and determine how many of that group developed brain cancer by sorting and tracing more than a quarter-million employee names. Meanwhile, the other team was culling a wide variety of sources to reconstruct what substances workers were exposed to between 1952 and 2001 in the course of doing their jobs.

EPIDEMIOLOGY AND BIOSTATISTICS

DATA SOURCES

500,000 pages of employee files, union rosters, National Death Index, Social Security Death Index, national and state cancer registries, and more than 3 million work history entries

IDENTIFY COHORT

All workers employed between 1952 and 2001 at 8 Connecticut plants, roughly equaling 224,000 individuals

MORTALITY ANALYSIS

- Identify all deceased cohort members and causes of death
- Identify living and deceased workers diagnosed with brain cancer between 1976 and 2001
- Compare brain cancer incidence and mortality rates in cohort with general population

NESTED CASE-CONTROL STUDY AND GENETIC ANALYSIS

- Match cases to similar workers without brain cancer as controls
- Contact controls and cases or their family members for interviews and consent to review medical records and to gather and analyze specimens

EXPOSURE RECONSTRUCTION

DATA SOURCES

Operations manuals, purchasing records, time-efficiency studies, air-sampling records, review of 320,000 job titles and employee interviews

DETERMINE EXPOSURES

BY AGENT

- Ionizing radiation (from: thoriated nickel, thoriated tungsten or radioactive particles)
- Metalworking fluids
- Hexavalent chromium
- Electromagnetic fields
- Solvents

BY PROCESS

Grinding, milling, degreasing, etc.

BY PART

Gears, blades, shafts, vanes, etc.

CONSOLIDATED ANALYSIS

JOB DICTIONARY WITH EMPLOYEE-SPECIFIC EXPOSURE PROFILES

MERGE DATA SETS FOR FINAL STUDY RESULTS

Half a million pieces of paper illustrate the number of pages of employee records scanned by University of Pittsburgh researchers to compile worker lists and information.

JAMES PORTO

What might have caused it? By August 2001 a Connecticut Department of Public Health investigation found that the incidence in the preceding 10 years of glioblastoma among workers at the North Haven plant represented between 2.8 and seven times the expected rate, depending on assumptions.

At that point the state health department asked Pratt & Whitney, which declined to comment for this article, to hire an independent epidemiologist to investigate further. The company turned to Marsh, a biostatistician at Pitt's Graduate School of Public Health. Marsh specializes in what he calls "messy and labor-intensive" workplace health investigations, often involving tens of thousands of employees and multiple work sites. He immediately contacted Esmen, an expert in assessing and reconstructing workplace exposures, with whom he frequently collaborates.

The pair initially focused on the North Haven factory, which was shut down in 2002. But when they learned that the company did grinding work similar to the operations at North Haven in its main factory, in East Hartford, and different types of work at other Connecticut facilities, they decided to study all eight existing and defunct plants in the state. Thus, a study initially projected to cover about 100,000 employees grew to a $12-million, seven-year investigation of close to a quarter of a million workers during the years 1952 to 2001.

The bigger project has two scientific advantages, Marsh explains: a higher statistical power, which reduces the chance of false negative results and increases the likelihood of detecting even subtle patterns; and better internal comparisons of work practices, exposures and health outcomes. Occupational epidemiology often suffers from the so-called healthy worker effect—misleadingly comparing disease rates among a group of workers with those of the general population, which includes people too sick to work. Comparing subgroups of Pratt & Whitney workers with one another should produce a more accurate picture.

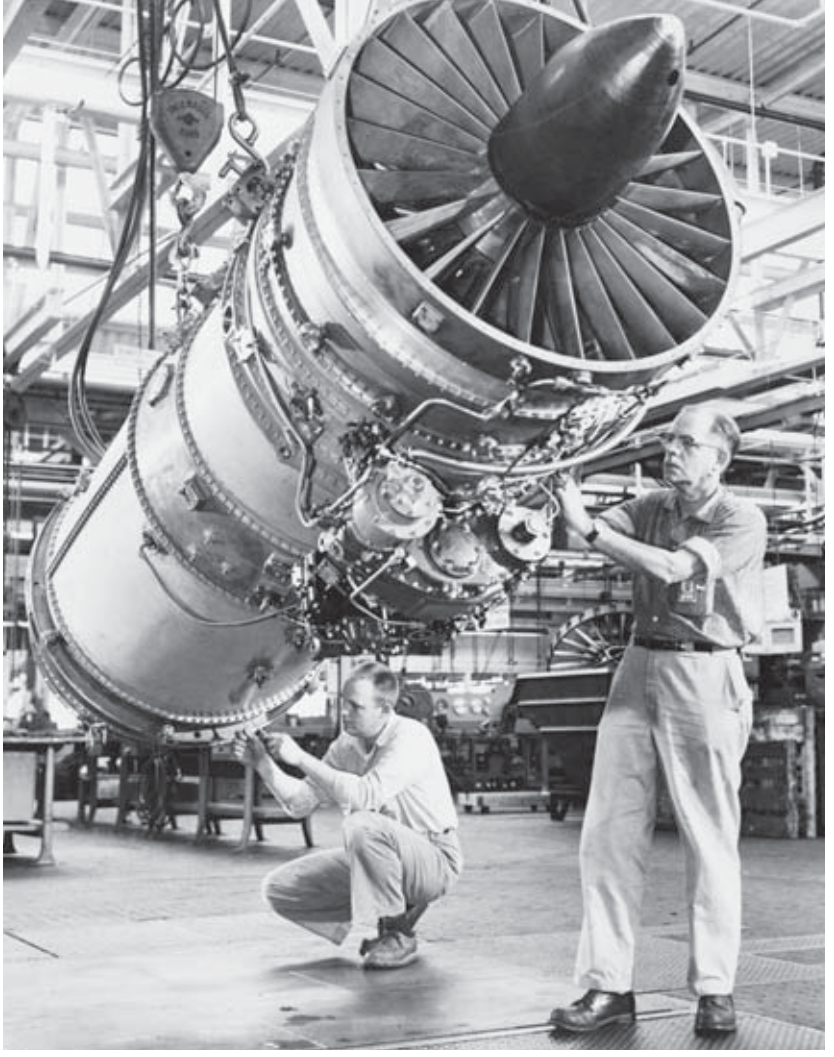
But the study's massive size also represents one of the researchers' greatest challenges. Working under project manager Jeanine Buchanich, Pitt employees and contractors spent a year onsite at Pratt & Whitney, scanning half a million pages of personnel records and abstracting them into a database of employee vital status information. Buchanich then began tracing the roughly 266,000 names—collectively known as the cohort—through national databases to see which employees had died and from what causes. A computer programmer wrote a protocol for sampling names from union membership rosters, which Buchanich checked against the cohort to see if people were missing. She also had to rectify database entries where the dates did not make sense: for example, where an employee was ostensibly hired before he was born or after he died. "The cohort file was fantastically clean," Buchanich says—with an error rate of less than 0.1 percent—"but it was still a couple hundred errors that you have to look up and resolve." After eliminating these and further refining the database, the cohort now comprises about 224,000 workers.

Meanwhile case manager Zb Bornemann has been hunting for brain cancer cases. He ran the entire cohort through the National Death Index and Social Security Death Index. And he continues to contact state cancer registries around the country, asking whether any names from the cohort match records of people with brain tumors. Where a match turns up, Bornemann tries to trace the next of kin through online databases. Some registries are a dead end. Some feel more

[THE AUTHOR]

Carole Bass is an investigative journalist who writes about public health, legal affairs and the environment. She is an Alicia Patterson Fellow for 2008, reporting and writing about toxic chemicals on the job. Bass is a former reporter and editor for the *New Haven Advocate,* where she wrote extensively about brain cancer cases at Pratt & Whitney Aircraft, and for the *Connecticut Law Tribune.* She lives with her husband and two teenage daughters in New Haven, Conn.

FRANK VAN RIPER (*Bass*)



WORKERS at Pratt & Whitney Aircraft's East Hartford plant prepare a jet engine developed for the Boeing 727 to be tested in July 1961. The factory is one of eight in Connecticut included in an investigation of brain cancer cases among the company's employees.

CANCER CLUSTER DEFINED

A seemingly unusual number of cancers occurring during a specific time period among people who live or work together can happen by chance. But an apparent cluster can also indicate that the illnesses have a common source. The mortality analysis portion of the Pratt & Whitney investigation will use statistical techniques "to determine whether the total number of observed malignant and/or benign brain cancer cases and/or deaths is greater than the number expected based on standardized comparisons with the general populations of the total U.S., the state of Connecticut and the local counties from which the workforces are drawn, and to determine whether any observed excesses are likely to be due to chance factors alone."

like *Alice in Wonderland*'s rabbit hole: in Washington State, Bornemann submitted six rounds of application materials for information about cancer patients, including one document that was rejected because it referred to the state cancer registry instead of the state department of health, which houses the cancer registry. Once Bornemann does locate a case's next of kin or, occasionally, a person living with brain cancer, he sends a letter asking for participation in the study: a phone interview, medical records and permission to analyze tissue from the person's brain tumor.

Finding cases factors into the first part of the study—the mortality analysis—which will determine whether a higher-than-expected rate of brain cancer or other diseases existed among Pratt & Whitney workers overall and among various subgroups. The second part is a nested case control study, in which investigators match each brain cancer case with a Pratt & Whitney employee of the same age, sex and year of hire in whom brain cancer did not develop. By comparing their medical and work histories, including the exposure assessment being developed by Esmen at U.I.C., the researchers hope to detect patterns that could explain why brain tumors occurred in some people but not in others.

In a third arm of the study, Pitt neuro-oncologist Frank Lieberman is looking at gene mutations in Pratt & Whitney employees' brain tumor tissue. If he finds a distinctive pattern, it could suggest that the Pratt & Whitney tumors were not random but shared some common causation.

Here some of the science is quite new. Lieberman is working with paraffin-embedded tumor tissue that has been stored for years at hospitals where Pratt & Whitney cancer patients underwent surgery. Until recently, that meant he would be limited to methods that allow only 15 to 20 genes per sample to be examined for changes known to be involved in tumor growth. Now, thanks to improved technology, he can also use microarray techniques previously available only for fresh tissue, which allow him to examine thousands of genes at a time, looking for small mutations as well as duplications or deletions of whole genes. "You can look for changes in patterns of [gene activity]," he says, "not just for changes up or down in specific genes."

Lieberman is comparing those profiles with a National Cancer Institute database and with brain tumor samples from patients at his Pitt clinic, as controls. "This is a very powerful technique," he states, in part because researchers "don't necessarily have to have a hypothesis at the beginning about which genes are important."

One of Lieberman's Pitt colleagues, Emanuela Taioli, uses similar molecular techniques to identify DNA damage caused by specific known carcinogens. The two groups are working together on a pilot effort to collect normal tissue from Pratt & Whitney workers with brain cancer; they hope to detect any molecular changes in that tissue and correlate them with carcinogens for which Esmen's team is finding exposures at Pratt & Whitney. In principle, those fingerprints of toxic substance exposure could represent early steps toward cancer. The science behind these strategies "is still very young," Lieberman cautions. "We're trying to use the opportunity that's presented by the enormous size and the sophistication of the epidemiology to get as much information about possible triggers for brain cancer as we can. But the techniques that are being used are, in a profound sense, still experimental themselves."

BETTMANN/CORBIS

Industrial Archaeology

The work of Esmen's Chicago team is less experimental but just as monumental. The group has spent five years delving into Pratt & Whitney operations from the 1950s through the 1990s, trying to figure out what employees were exposed to and at what levels. "If the data is not there, it has to be reconstructed," says Esmen, a professor of environmental and occupational health sciences at U.I.C.'s School of Public Health. "It's almost like industrial archaeology."

The digging would be easier if the team knew what it was looking for. Scientists have long suspected an occupational source for some brain cancers. But the only proven cause is ionizing radiation, which a few Pratt & Whitney operations did generate. Beyond that, the list of suspects comes from previous studies that found high rates of brain cancer among people who worked with certain metals, machine oils, and solvents but that have not been consistently replicated.

Working from Pratt & Whitney records, Esmen's team is boiling 320,000 job titles down to a manageable number of broader job categories. For each category, the researchers then try to quantify workers' exposure to the suspected agents during various time periods.

The numbers are only relative, though. "The important thing is to get things in the right order," Esmen says. If the researchers estimate a particular exposure was 10 units, "you don't know if it is really 6 or it is 12—but it's definitely

OSHA Lags Behind

"Understand this," says Roger Hancock, a member of the Pratt & Whitney study team: today's OSHA standards "were the latest toxicology data in 1968." He is not joking. OSHA opened shop in 1971 with a statutory mandate: "to assure so far as possible every working man and woman in the Nation safe and healthful working conditions." It adopted permissible exposure limits (PELs) for about 400 chemicals in one swoop, lifting them directly from voluntary industry standards. Those standards were developed by the nonprofit American Conference of Governmental Industrial Hygienists, or ACGIH, in 1968. During the rest of the 1970s, OSHA put together new exposure limits for nine other substances.

But in the next decade, the PEL express turned into a train wreck. In 1980 the U.S. Supreme Court struck down an OSHA standard that reduced the permissible level of benzene by an order of magnitude, saying the agency had to prove that its regulation would prevent a "significant risk of harm." Without defining "significant risk," the court suggested that one additional death for every 1,000 exposed workers was probably significant, whereas one in a billion was not. Ever since then, OSHA has viewed the one-in-1,000 number as the strictest possible standard. Some PELs allow much graver risks: for example, the 2006 limit for hexavalent chromium, issued under court order, corresponds to a cancer risk of 35 to 45 per 1,000, according to OSHA estimates.

In 1987 OSHA undertook a sweeping update of its limits for air contaminants. Less than two years later it issued PELs for 376 chemicals. More than half of those were tightened standards for substances on the original 1971 list; the rest were newly regulated substances. But industry and labor both challenged the law, and in 1992 a federal appeals court threw it out, ruling that OSHA had to carry out separate rulemaking procedures for each substance. It never did.

Despite the scientific difficulties, the ACGIH continues to crank out about 20 to 40 of its voluntary exposure limits, known as TLVs (for threshold limit values), every year. As a result, TLVs now cover more than 700 chemicals, compared with the 400-odd that OSHA regulates. "It's never-ending, because there's always more information," says Terry Gordon, who leads the TLV effort. "We're volunteers doing our best. If OSHA picked up the ball and ran with it, that'd be a great day."

In written responses to questions from *Scientific American,* OSHA emphasized that "updating an exposure limit is not a simple undertaking." Meeting the burdens imposed by Congress and the courts requires "extensive research" and "significant resources to adequately characterize the effect of revised PELs on employee health risks and to evaluate feasibility to ensure that revised standards are necessary, will be effective in protecting employees from health hazards in the workplace, and will be possible for employers to achieve." For hexavalent chromium, OSHA says it adopted "the lowest level that was feasible," based on "the totality of the evidence in the rulemaking record."

But OSHA could do far better, in the view of some who used to work there. Harry Ettinger, an industrial hygienist, led the Reagan-era effort to update the air contaminant limits. "It's an embarrassment" that most PELs still date from 1968, he exclaims. He sees the perfect as the enemy of the good. "I tried to convince labor that they were crazy to sue us. They wanted perfection. Perfection doesn't exist."

To another former OSHA official, Adam Finkel, the problem is priorities. "Most people who know what they're talking about would agree that occupational health [accounts for] 80 to 90 percent" of work-related deaths, says Finkel, who was the agency's director of health standards from 1995 to 2000. Yet "the emphasis has always been on safety" rather than health. He himself was forced from his job after calling attention to on-the-job health hazards facing OSHA's own inspectors; the agency eventually agreed to a six-figure settlement of his whistle-blower claim. Far from taming a perfectionist streak, he believes, "the agency just has to catch up to late 20th-century science. There are so many single agents that we know workers are exposed to at 1,000 times higher than they should be." With so much low-hanging fruit, Finkel says, OSHA's top priority should be "old-style industrial hygiene."

In any case, the "extensive research" and "significant resources" that OSHA states are necessary to update worker health protections have to come from somewhere. More often than not, the industrial interests that would be subject to those regulations are the ones paying for the studies. On the workplace health front, the National Institute for Occupational Safety and Health's research budget has been flat or declining. And manufacturers have not only fought governmental regulation; they have also taken the independent ACGIH to court, trying to block it from releasing its nonbinding exposure limits. —***C.B.***

FAST FACTS

- **On average, nearly 16 workers die from injuries sustained at work and 134 die from work-related diseases in the U.S. every day.**
- **An estimated 11,500 private-sector workers have a nonfatal work-related injury or illness every day.**
- **9,000 workers are treated in emergency rooms because of occupational injuries every day.**

SOURCE: National Institute for Occupational Safety and Health

not 100." Because no measurements exist, the researchers extrapolate from interviews with factory workers and engineers. They also use mounds of company-supplied data, such as purchasing records (for quantities of materials used), 1970s time-efficiency studies (for length of time spent on any given task), internal publications with esoteric titles like *A Versatile Engineering and Manufacturing Capability*, and whatever air sampling Pratt & Whitney may have done over the years.

That last information source might sound like a rich lode for exposure reconstruction. But it is trickier than it seems, Esmen points out. An epidemiologist trying to assess exposures across the entire workforce would take random samples from each work group and document changes from shift to shift or day to day. An industrial hygienist called in to fix a problem—respiratory complaints, for example—would sample only the "problem" area and would consider only the highest levels registered of a suspect substance.

Industrial hygiene textbooks instruct future practitioners to take a full range of samples, says researcher Steve Lacey, who teaches these techniques to U.I.C. graduate students. "But that's not the reality." Roger Hancock, another team member who spent a quarter of a century practicing industrial hygiene in the private sector, knows the reality: "You arrive at a plant with a luggage cart full of [testing] equipment, and you have a week. Maybe they're running that process once that week, so you have one chance to take that sample. If the highest sample is below concern, then you don't take any more samples."

"Below concern," for most companies, means legal compliance. If a workplace meets OSHA standards, that is good enough. But academic researchers and medical textbooks recognize that what is good enough for OSHA is not always good enough to protect workers' health [*see box on preceding page*].

Determining safe exposure limits is not easy, acknowledges Terry Gordon, who chairs the American Conference of Governmental Industrial Hygienists committee that issues voluntary exposure limits for chemicals. Like OSHA, his group of about 20 volunteers does not conduct original research but rather relies on published studies. Animal toxicology investigations, with their controlled lab conditions, are tidier than the murky epidemiology of real people exposed to unknown amounts and combinations of various substances, both on and off the job. But the clarity of animal studies is also a weakness: they measure the effect of one chemical at a time, whereas workplaces typically contain multiple toxicants. "Human data is always preferable to toxicology,"

PHOTOGRAPH OF MONITOR BY JAMES PORTO; PHOTOILLUSTRATION BY JEN CHRISTIANSEN; UNIVERSITY OF PITTSBURGH (*tumor samples*)

[ANALYSIS]

INTERPRETING THE DATA

With the consent of affected workers or their families, the investigators obtained tumor samples and, in some cases, normal tissue, to examine for signs of DNA damage that could have been caused by carcinogens (*left*). When they combine those results with information about what substances employees were exposed to, and when (*right*), they hope to reveal whether on-the-job exposures contributed to the cancers.

GENETIC CLUES

Changes, or mutations, in DNA isolated from tumor samples like those shown on the computer screen can hint at what triggered the tumor.

- A distinctive pattern of mutations in the tumors of different workers could suggest a common causation, for example.
- Such analyses can also determine the duration of cumulative change and even the order in which changes occurred, to help pinpoint when cancer growth might have started.
- Certain known carcinogens damage DNA in specific ways, leaving a fingerprint of exposure in tumor cells and even normal tissue.
- Mutations in genes for carcinogen-metabolizing enzymes are another category of cancer-promoting changes detectable in noncancerous tissue that could help explain how a tumor was initiated.

VISUALIZING EXPOSURES

To manage the enormous amount of data the investigators gathered about the factories' operations between 1952 and 2001, they created a "geographic information system." The database allows them to retrieve and connect data about the locations of employees, parts and processes—and, hence, worker exposures to potentially toxic agents—at different times during the study period.

says Gordon, a professor and researcher in New York University's department of environmental medicine. But "with human data, there's often not enough exposure data, and the exposure information is often not linked to health effects."

The Pratt & Whitney researchers are doing their best to avoid those pitfalls, but they recognize some of the elements that make this study "messy," as Marsh puts it. Because Bornemann has incomplete information from state cancer registries, he will likely miss brain cancer cases. Of those he has located, only 41 percent have agreed to participate; Marsh wants at least 60 percent to assure scientific validity. Marsh also notes that while participants' recollections about medical history and lifestyle are "reasonably good at a broader level, it breaks down when you get to any level of detail." And the U.I.C. team's exposure reconstruction, for all its thoroughness, will still be an estimate of what happened decades ago on shop floors that have since been shuttered or drastically cleaned up.

A Start in the Right Direction

After seven years and $12 million, there is a good chance the Pratt & Whitney study will wind up like so many other workplace health investigations: inconclusive. Researchers say the phenomenon stems from the difficulty of the science. Like all epidemiology studies, this one can at best prove associations between exposures and health effects, not strict causation. It is especially hard to detect causes of diseases like cancer, which typically shows up decades after the offending exposure. And finding a definitively safe level of exposure to carcinogens can be impossible.

Many researchers might love to have $12 million for studies that could produce clearer results in less time, but Pratt & Whitney's millions would probably not otherwise have been spent on workplace health research. Even if the study never provides conclusive answers to all the questions it raises, the effort will hardly have been a waste of time or money.

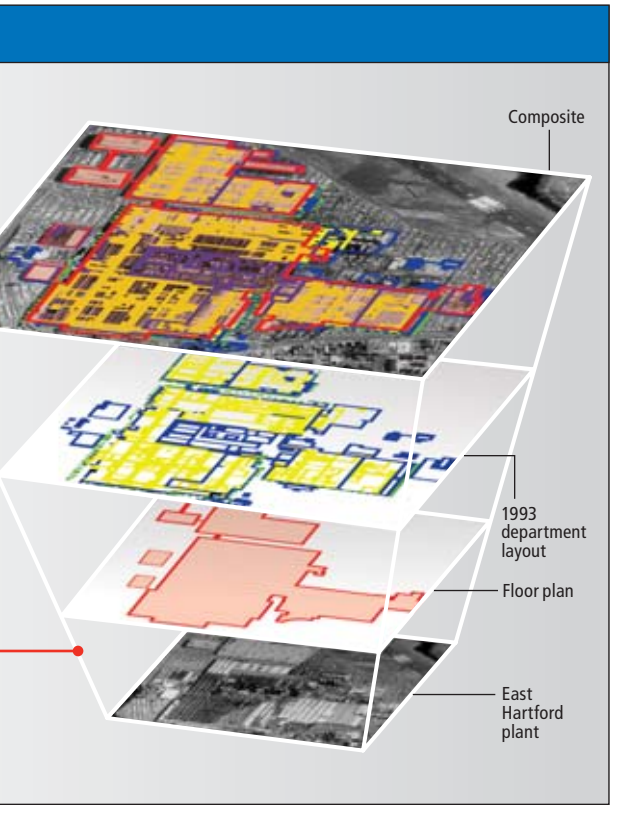
For starters, there is the prospect of some kind of answers for the Pratt & Whitney families. "I've been involved in this project since the day it started," says Pitt project manager Buchanich. "That we're finally going to be able to tell these workers something—it's been a long time."

Regardless of the specific outcomes, the project might be able to tell the rest of us something as well. The mounds of new data it is generating could help unravel the complexities of multiple toxic exposures and could contribute, in Lieberman's words, to understanding "the basic biology of how brain tumors get started." In addition, the study's unprecedented scope is spurring new techniques for managing the massive amount of information. The U.I.C. team, for example, is building a geographic information system, or GIS, database that will let team members map factory operations across time and space. Such technology could be useful for any study in which spatial relations are important, such as archaeology or industrial engineering, according to Esmen.

So this archaeological foray into Connecticut's industrial past may yield tools and information to help workers and brain cancer patients in the future. In doing so, it could help lower the scientific barriers that stand between Americans and healthy workplaces. Reducing the nonscientific barriers will require a different set of tools.

"The largest preventable health and safety risks remaining to be addressed in our society occur disproportionately in U.S. workplaces," wrote Adam Finkel, a former OSHA health standards director, in a letter last May to Representative Lynn Woolsey of California, who chairs the House Subcommittee on Workforce Protections. "The solution is not to complain about the need to do good science but simply to get back to doing good science, like OSHA used to do." ■

CONNECTICUT DEPARTMENT OF ENVIRONMENTAL PROTECTION (*bottom layer*); UNIVERSITY OF ILLINOIS AT CHICAGO (*middle layers*)



MORE TO EXPLORE

Workers at Risk: The Failed Promise of the Occupational Safety and Health Administration. Thomas O. McGarity and Sidney A. Shapiro. Praeger Publishers, 1993.

Occupational and Environmental Health: Recognizing and Preventing Work-Related Disease and Injury. Edited by Barry S. Levy, David H. Wegman, Sherry L. Baron and Rosemary K. Sokas. Fifth edition. Lippincott Williams & Wilkins, 2005.

Doubt Is Their Product: How Industry's Assault on Science Threatens Your Health. David Michaels. Oxford University Press, 2008.

Getting Home Safe and Sound: OSHA at 38. Michael Silverstein in *American Journal of Public Health* (in press).

Cancer Cluster Fact Sheet from the National Cancer Institute: **www.cancer.gov/cancertopics/factsheet/Risk/clusters**

Pratt & Whitney investigation background is available at the Connecticut Department of Public Health Web site: **www.ct.gov/dph/cwp/view.asp?a=3140&q=387474**

Not Tonight, Dear, I Have to Reboot

Is love and marriage with robots an institute you *can* disparage? Computing pioneer David Levy doesn't think so—he expects people to wed droids by midcentury. Is that a good thing? BY CHARLES Q. CHOI

At the Museum of Sex in New York City, artificial-intelligence researcher David Levy projected a mock image on a screen of a smiling bride in a wedding dress holding hands with a short robot groom. "Why not marry a robot? Look at this happy couple," he said to a chuckling crowd.

When Levy was then asked whether anyone who would want to marry a robot was deluded, his face grew serious. "If the alternative is that you are lonely and sad and miserable, is it not better to find a robot that claims to love you and acts like it loves you?" Levy responded. "Does it really matter, if you're a happier person?" In his 2007 book, *Love and Sex with Robots*, Levy contends that sex, love and even marriage between humans and robots are coming soon and, perhaps, are even desirable. "I know some people think the idea is totally outlandish," he says. "But I am totally convinced it's inevitable."

The 62-year-old London native has not reached this conclusion on a whim. Levy's academic love affair with computing began in his last year of university, during the vacuum-tube era. That is when he broadened his horizons beyond his passion for chess. "Back then people wrote chess programs to simulate human thought processes," he recalls. He later became engrossed in writing programs to carry on intelligent conversations with people, and then he explored the way humans interact with computers, a topic for which he earned his doctorate last year from the University of Maastricht in the Netherlands. (Levy was sidetracked from a Ph.D. when he became an international master at chess, which led him to play around the world and to found several computer and chess organizations and businesses.)

DAVID LEVY

I DO: Predicts the first place to legalize marriage with robots will be Massachusetts, where liberal jurisdiction and high-tech research meet.

NO FANTASY: Despite his thinking about robot love, he is not a fan of science fiction: "The only sci-fi book I ever read was as a favor to a publisher who wanted a quote from me on the back cover, but the book was so dreadful that I couldn't support it."

Over the decades, Levy notes, interactions between humans and robots have become increasingly personal. Whereas robots initially found work, say, building cars in a factory, they have now moved into the home in the form of Roomba the robotic vacuum cleaner and digital pets such as Tamagotchis and the Sony Aibo.

And the machines can adopt a decidedly humanoid look: the robot Repliee from Hiroshi Ishiguro, director of Osaka University's Intelligent Robotics Laboratory, can fool people into believing that it is a real person for about 10 seconds from a few feet away. And "it's just a matter of time before someone takes parts from a

SIGN TOUHIG

vibrator, puts it into a doll, and maybe adds some basic speech electronics, and then you'll have a fairly primitive sex robot," Levy remarks.

Science-fiction fans have witnessed plenty of action between humans and characters portraying artificial life-forms, such as with Data from the *Star Trek* franchise or the Cylons from the reimagined *Battlestar Galactica*. And Levy is betting that a lot of people will fall in love with such devices. Programmers can tailor the machines to match a person's interests or render them somewhat disagreeable to create a desirable level of friction in a relationship. "It's not that people will fall in love with an algorithm but that people will fall in love with a convincing simulation of a human being, and convincing simulations can have a remarkable effect on people," he says.

Indeed, a 2007 study from the University of California, San Diego, found that toddlers grew to accept a two-foot-tall humanoid robot named QRIO after it responded to the children who touched it. Eventually the kids considered QRIO as a near equal, even covering it with a blanket and telling it "night night" when its batteries ran out. "People who grow up with all sorts of electronic gizmos will find android robots to be fairly normal as friends, partners, lovers," Levy speculates. He also cites 2005 research from Stanford University that showed people grew to like and trust computer personalities that cared about their wins and losses in blackjack and were generally supportive, much as they would respond to being cared about by other people.

The modern age of telecommunications has already made it possible to fall in love without ever having met face to face, Levy adds. "So many people nowadays are developing strong emotional attachments across the Internet, even agreeing to marry, that I think it doesn't matter what's on the other end of the line," he says. "It just matters what you experience and perceive."

Based on what researchers know about how humans fall in love, human-robot connections may not be all that surprising. Rutgers University biological anthropologist Helen Fisher, renowned for her studies on romantic love, suggests that love seems dependent on three key components: sex, romance and deep attachments. These components, she remarks, "can be triggered by all kinds of things. One can trig-

ROBO NUPTIALS: David Levy thinks that human-robot marriages are inevitable. Others find the prospect ludicrous.

ger the sex drive just by reading a book or seeing a movie—it doesn't have to be triggered by a human being. You can feel a deep attachment to your land, your house, an idea, a desk, alcohol or whatever, so it seems logical that you can feel deeply attached to a robot. And when it comes to romantic love, you can fall madly in love with someone who doesn't know you exist. It shows how much we want to love."

Still, both Fisher and Levy agree that many if not most humans will continue to love and have sex the old-fashioned way. "But I think there are people who feel a void in their emotional and sex lives for any number of reasons who could benefit from robots," Levy states. He cites a Massachusetts Institute of Technology student dubbed "Anthony" in M.I.T. psychologist Sherry Turkle's book *The Second Self*, which explores human-computer interactions. Anthony tried having human girlfriends but preferred relationships with computers. Levy says that he dedicated his book "to 'Anthony' and all the other 'Anthonys' before and since of both sexes,

ROY MCMAHON *Corbis (wedding)*; VOLKER MOEHRKE *zefa/Corbis (robot)*; EMILY HARRISON (*photoillustration*)

to all those who feel lost and hopeless without relationships, to let them know there will come a time when they can form relationships with robots."

Whether those bonds are emotionally healthy, however, is debatable. As Turkle puts it: "If you are lonely but afraid of intimacy, relationships with machines can enable you to be a loner yet never alone, give you the illusion of companionship without the demands of friendship. There is nothing to celebrate here. To me, the seductiveness of relationships with robots speaks to what we are not getting from people."

Instead of throwing robots at social problems, Turkle feels humans should do the job. "What people like Anthony need are experiences that will increase their repertoire for dealing with the complexity and challenges of relationships with people," she explains. Levy contends that there are not going to be enough people to handle social concerns such as loneliness or care for the elderly, but Turkle dismisses the idea: "If we paid people to take care of the elderly in the way we invested in other things, this wouldn't be an issue."

Both Fisher and Turkle find the idea of legal human-robot marriages ridiculous. But Levy counters that "if you went back 100 years, if you proposed the idea that men would be marrying men, you'd be locked up in the loony bin. And it was only in the second half of the 20th century that you had the U.S. federal government repealing laws in about 12 states that said marriage across racial boundaries was illegal. That's how much the nature of marriage has changed."

As to what Levy's wife thinks, he laughs: "She was totally skeptical of the idea that humans would fall in love with robots. She's still fairly skeptical." A reasonable reaction—then again, a Stepford wife with contrariness programmed into her would say that, too. ■

Charles Q. Choi is a frequent contributor. A Q&A version of his interview with Levy is at **www.SciAm.com/ontheweb**

Page Intentionally Blank

The Switch Is On

By Mark Fischetti

Incandescent lightbulbs may be history. As of 2007, compact fluorescent lightbulbs (CFLs), which are more energy-efficient, had made only modest inroads because they were more expensive. But in December the U.S. Congress passed a major energy bill that included a new lighting standard. By 2012 manufacturers selling any 100-watt (W) bulb must make it 30 percent more efficient than today's 100-W incandescent bulb. Similar requirements will phase in for 75-W bulbs in 2013 and 60-W and 40-W in 2014. Europe has passed its own rules, too. CFLs already meet the specs, and although makers will try to improve the old hardware, the new bulbs clearly have the edge and continue to improve.

The technology inside a lightbulb is quite advanced [*see illustrations*]. The tungsten filament in an incandescent glows at more than 2,200 degrees Celsius and must be made perfectly uniform because any tiny imperfection will cause it to rapidly burn out. Yet only about 10 percent of the electricity entering the bulb is emitted as visible light; 90 percent or so is radiated as heat. A CFL is about four times as efficient as an incandescent bulb. A 26-W CFL can therefore shine as brightly as a 100-W incandescent, requiring only one quarter of the energy. The tubular fluorescent bulbs common in overhead lighting are slightly more efficient still but do not fit standard light sockets, as CFLs do.

CFLs still present some problems, which manufacturers are solving. For example, some consumers find the light too harsh. "The human eye wants to see all color wavelengths," says James Dakin, senior consulting engineer at GE Lighting in Cleveland, but the phosphor coating inside the bulbs fails to emit certain wavelengths. Phosphor improvements are filling in those holes, Dakin explains.

Early bulbs took several minutes to reach full output and may have hummed or flickered, but the electronic ballasts that have replaced the original magnetic ballasts have ended those shortcomings and also allowed smaller tubular shapes. "We have a huge effort under way to create CFLs that are suitable for more applications," says James Meyer, general manager of GE Lighting. Lowering cost further, he says, is now mostly a matter of even greater mass production.

Light-emitting diodes (LEDs) could also pose competition in years to come; the solid-state lights are about as efficient and last three times longer than CFLs. But the screw-in versions struggle to outshine even a 25-W incandescent and cost far more. For now, CFLs have the brightest future.

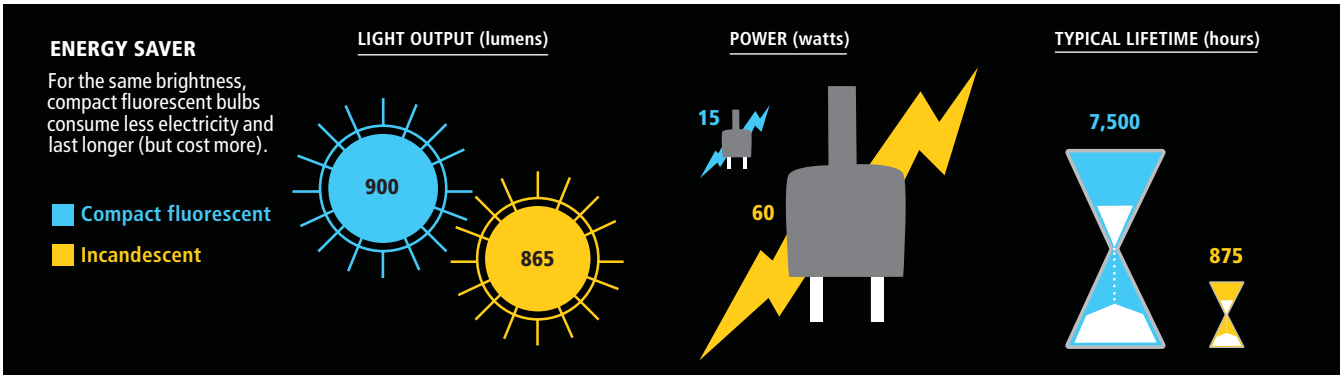
DID YOU KNOW ...

AC/DC: Unlike virtually all appliances, the common incandescent bulb used in standard alternating-current sockets would work equally well in a direct-current socket—in case anyone ever runs across one.

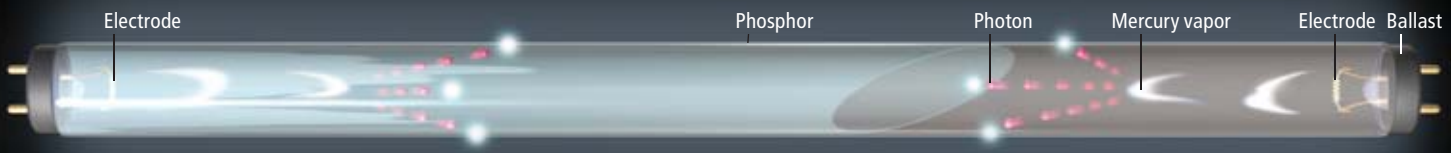
COOL REFRIGERATOR: Recent improvement in white light–emitting diodes (LEDs) has prompted Wal-Mart to replace the hot, incandescent bulbs inside refrigerated food cases and freezers in all its stores. The LEDs run much cooler, reducing the refrigeration load, and consume less electricity, too.

MERCURY: As their packaging indicates, regular and compact fluorescents (CFLs) contain small amounts of mercury. Dead bulbs must be properly disposed; regulations vary by state. Manufacturers are trying to lessen the mercury needed. They also counter CFL critics by noting that the extra electricity required to power an equivalent incandescent sends mercury (and other contaminants) into the atmosphere—that is, if the power is produced by burning coal.

BROWN SPOT: Tungsten atoms that gradually evaporate from a filament may collect in one area on a glass bulb, creating a brown spot. If the bulb is poorly sealed and air leaks in, tungsten may react with it to form brown, purple or yellow oxide or trioxide deposits.

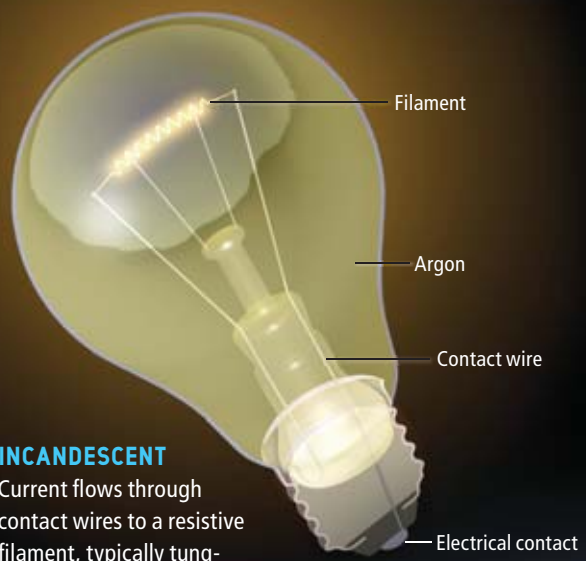


DANIELA NAOMI MOLNAR



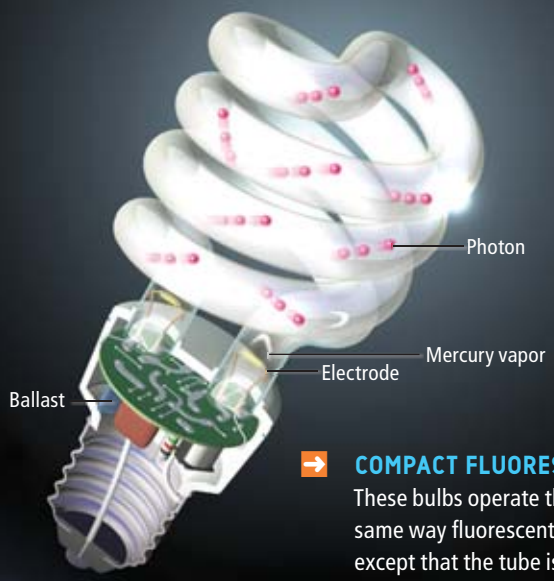
FLUORESCENT

A ballast (*far right*) provides a high voltage that creates an electric arc across tungsten electrodes. The arc excites mercury vapor atoms, which discharge ultraviolet photons. The photons strike phosphor coatings on the glass, causing them to emit visible light (fluoresce). The ballast then regulates voltage and power at a lower level to sustain the arc. Argon (*not shown*) in the tube quickens start-up and enhances brightness.



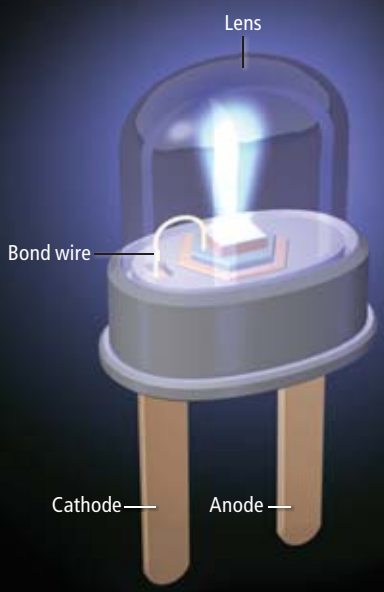
INCANDESCENT

Current flows through contact wires to a resistive filament, typically tungsten, causing it to heat so much it glows (incandesces). Over time tungsten atoms evaporate inside the evacuated chamber, thinning the filament until a threadbare spot breaks. In bulbs higher than 25 watts, manufacturers fill the chamber with an inert gas such as argon or xenon to slow evaporation.



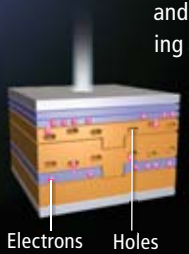
COMPACT FLUORESCENT

These bulbs operate the same way fluorescents do, except that the tube is bent; both ends terminate in one base that can fit a conventional incandescent socket.



LIGHT-EMITTING DIODE

Current flows through a semiconductor diode, causing electrons and holes to move; when these entities meet, they emit a photon of a certain color. Different semiconductors emit different colors. The appearance of white light can be created by combining red, blue and green diodes within one bulb or by coating a blue diode with a yellowish phosphor.



DIODE emits light when electrons and holes collide.

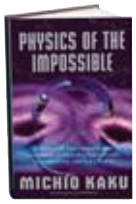
GEORGE RETSECK

SEND TOPIC IDEAS *to workingknowledge@SciAm.com*

"Shields Up!" ▪ The Whale Trial ▪ Antimatter

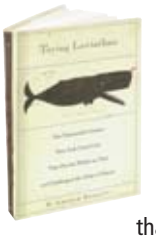
BY MICHELLE PRESS

PHYSICS OF THE IMPOSSIBLE: A Scientific Exploration into the World of Phasers, Force Fields, Teleportation, and Time Travel
by Michio Kaku. Doubleday, 2008 ($26.95)



"If at first an idea does not sound absurd, then there is no hope for it." Einstein's words make an apt motif for Kaku's premise: just because something is impossible today doesn't mean it will be impossible in the future. Kaku—a well-known theoretical physicist at the City University of New York and an accomplished writer—reminds us of how fantastic the World Wide Web would have seemed in 1908. He goes on to discuss a number of currently impossible technologies, categorizing them into technologies that are not doable today but may be so in the foreseeable future; those that are impossible in the foreseeable future but do not violate the laws of physics; and those that violate the laws of physics as we know them today. In the first category, for example, a combination of several technologies—a supercharged plasma window, a curtain of high-energy laser beams and a screen of carbon nanotubes—might one day create a realistic force field, not unlike the one Captain Kirk summoned with "Shields up!" in countless episodes of *Star Trek.*

TRYING LEVIATHAN: The Nineteenth-Century New York Court Case That Put the Whale on Trial and Challenged the Order of Nature
by D. Graham Burnett. Princeton University Press, 2007 ($29.95)



Burnett, who teaches history at Princeton, tells the strange story of an 1818 trial that set in opposition the new science of taxonomy and the biblically sanctioned view that the whale is a fish. The immediate dispute was whether whale oil is fish oil and therefore subject to various regulations, but the grander implications did indeed challenge the order of nature. And just as in the 2005 trial in Dover, Pa., that pitted evolution against intelligent design, the whale trial fueled a sensational public debate in which a parade of colorful experts took the witness stand. This wonderfully detailed book is aimed at scholars, and the story awaits a more popular telling, maybe even its own *Inherit the Wind.*

EXCERPT

THE MYSTERY OF THE MISSING ANTIMATTER
by Helen R. Quinn and Yossi Nir. Princeton University Press, 2008 ($29.95)

In addition to the stuff we call matter, physicists have identified another category of material called antimatter, but they can find very little of it in nature. In this elegantly written book, two physicists guide readers through cutting-edge scientific developments that aim to figure out what happened to antimatter:

"This much we do know: The fate of antimatter to disappear was sealed by the time the Universe was no older than a millionth of a second. At that time, matter particles and antiparticles were both still very abundant, but there must have been a tiny edge for particles over antiparticles, about one extra particle for every ten billion particle-antiparticle pairs. This tiny excess is all that matter needed for a total victory over antimatter in the present Universe. All the visible structures in the Universe that we observe today—planets, stars, galaxies, clusters of galaxies—are made from that surplus of particles over antiparticles....

"What laws of nature, not yet manifest in experiments and not part of our current Standard Model, were active in the early Universe, allowing the observed amount of matter to persist while all antimatter disappeared from the Universe?"

NEW BOOKS ABOUT THE NATURAL WORLD

1 **The Last Flight of the Scarlet Macaw: One Woman's Fight to Save the World's Most Beautiful Bird**
by Bruce Barcott. Random House, 2008 ($26)
Global trade, the world's demand for energy, and the realities of economic survival come together in Belize, with bad news for the macaw.

2 **Macachiavellian Intelligence: How Rhesus Macaques and Humans Have Conquered the World**
by Dario Maestripieri. University of Chicago Press, 2007 ($25)
A graceful and witty examination of power struggles among this widespread, contentious group of monkeys reveals unexpected parallels with human society.

3 **Natural Acts: A Sidelong View of Science and Nature**
by David Quammen. W. W. Norton, 2007 ($24.95)
If you missed the first edition of this exemplar of nature writing 20 years ago, don't pass up the new expanded edition.

4 **How and Why Species Multiply: The Radiation of Darwin's Finches**
by Peter R. Grant and B. Rosemary Grant. Princeton University Press, 2008 ($35)
Thirty-four years of studying Darwin's finches in the Galápagos Islands.

5 **No Way Home: The Decline of the World's Great Animal Migrations**
by David S. Wilcove. Island Press, 2007 ($24.95)
Fences, farms, skyscrapers, cell towers, pavement and global warming threaten nature's long-distance travelers.

6 **Life in the Valley of Death: The Fight to Save Tigers in a Land of Guns, Gold, and Greed**
by Alan Rabinowitz. Island Press, 2007 ($25.95)
Combines adventure, wildlife biology, personal crises and good prose into a gripping story, culminating in the establishment of an 8,500-square-mile tiger reserve in troubled Myanmar.

BIOSPHOTO/FISCHER BERNDT/PETER ARNOLD, INC. (*monkey*); TOM BRAKEFIELD *Getty Images (tiger*); FRANS LANTING *Minden Pictures (parrot*);

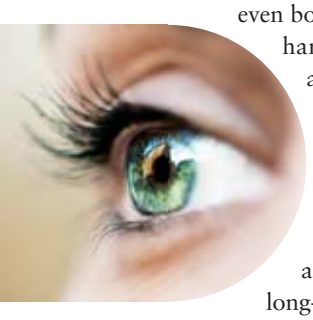
*If mutations **occur at random** over a species' entire genome, how can an organ as complex as an **eye evolve**?*

—V. Rautenbach, London

University of Utah biology professor **Jon Seger** explains:

Although it is highly unlikely that such an intricate and useful organ would arise spontaneously from random hereditary accidents, an eye can easily evolve through the same ongoing interaction between mutation and selection that drives the evolution of other adaptations.

Within a population, each individual mutation is extremely rare when it first occurs; often just one copy of it exists in the gene pool of an entire species. But huge numbers of mutations may occur every generation in the species as a whole. Some of these mutations are so harmful that they are eliminated before their carriers are even born. But the vast majority of mutations are harmless or at least tolerable, and a very few are actually helpful. These surviving mutations enter the population as exceedingly rare alternative versions of the genes in which they occur.



Most are then lost just because they are rare; however, very small effects on survival and reproduction may significantly affect the long-term rates at which different mutations accumulate in particular genes and at particular sites within genes. The resulting pattern of evolutionary change looks nonrandom and, in fact, really is nonrandom: some sites almost never change, some change occasionally and others change relatively often.

Nevertheless, that does not mean that the mutations themselves occurred nonrandomly. In retrospect, it is as if the mutations occurred where needed, but appearances can be deceiving, and selection is a great illusionist. In actuality they just *accumulated* where needed—first one, then another and another, over very many generations. Although getting two or more new "cooperating" mutations together in the same genome may take time, they will eventually find one another in a sexual species, assuming they are not lost from the population.

Visually oriented vertebrates such as humans have stunningly intricate eyes, but there is great variation in most aspects of the organ even within vertebrates, and several fundamentally different designs have emerged in animals as a whole. This diversity shows with living examples how simple, nonfocusing light sensors could gradually be elaborated and refined to become the complex, sophisticated imaging eyes of the kinds we know today.

To an organism that needs information about fast-changing aspects of its environment, a crude light-sensing organ may be much better than none at all. Given such a primitive eye, there may be thousands of different mutations that would slightly improve its functioning in various ways. When one such mutation occurs, is lucky enough not to be immediately lost and then rises in frequency under the force of natural selection, it sets the stage for others. Given enough time and continued selection, this process will readily improve the functioning of the eye, often by making it more complex.

*Why do parrots have the **ability to mimic**?*

—H. Messing, Westwood, Mass.

Michael Schindlinger, an avian biologist at Lesley University who maintains the Web site www.freeparrots.net, replies:

Parrots most likely imitate for reasons that seem to vary from species to species. In some cases, the regional dialects that arise from this mimicry may help males and females from similar areas find—or perhaps avoid—one another. Similarly, song learning in some species allows territorial neighbors to know one another and helps to distinguish drifters. And one study on small Australian parrots known as budgerigars reports that they seem to use call similarity in judging mates.

Imitative vocal learning is also a reliable display of neural functions that may be under consideration by a potential mate or ally, including hearing, memory and muscle control for sound production.

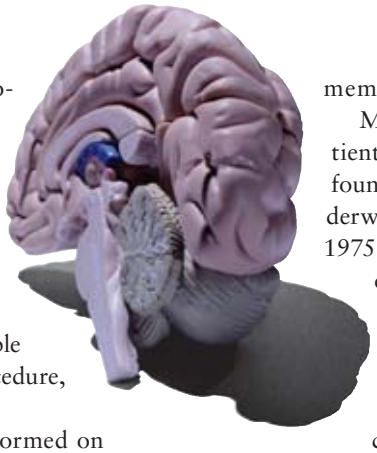
Playback studies of geographic dialects from wild parrot populations have shown that birds react more strongly to their local tongue. This phenomenon hints at another important benefit of imitation: to better command the attention of a potential listener by producing sounds for which the listener already has a memory, or a so-called neural template. The existence of this preformed perceptual template makes another parrot's vocalizations easier to perceive in a noisy environment. ■

HAVE A QUESTION?... *Send it to experts@SciAm.com or go to* **www.SciAm.com/asktheexperts**

JULIET WHITE *Getty Images (eye)*; P. WEGNER *Peter Arnold, Inc. (birds)*

Do You Need Only Half Your Brain?

BY CHARLES Q. CHOI



The operation known as a hemispherectomy—the removal of half the brain—sounds too radical to ever consider, much less perform. In the past century, however, surgeons have done it hundreds of times for disorders that cannot be controlled any other way. Perhaps surprisingly, the surgery has no apparent effect on personality or memory. Does that mean a person needs only half a brain? Yes and no. People can survive and function pretty well after the procedure, but they will have some physical disabilities.

The first known hemispherectomy was performed on a dog in 1888 by German physiologist Friedrich Goltz. Neurosurgeon Walter Dandy pioneered the use of the procedure on humans at Johns Hopkins University in 1923, operating on a patient who had a brain tumor. (That man lived for more than three years before ultimately succumbing to cancer.)

In 1938, after performing a hemispherectomy on a 16-year-old girl, Canadian neurosurgeon Kenneth McKenzie reported that it could stop debilitating seizures. And today brain surgeons perform hemispherectomies on patients who undergo dozens of seizures daily that resist all medication and stem from conditions that primarily afflict one hemisphere. "These disorders are often progressive and damage the rest of the brain if not treated," explains neurosurgeon Gary W. Mathern of the University of California, Los Angeles.

The surgery takes two forms. Anatomical hemispherectomies involve the removal of an entire hemisphere, whereas functional hemispherectomies take out only parts of a hemisphere—as well as severing the corpus callosum, the fiber bundle that connects the two halves of the brain. The evacuated cavity is left empty, filling up with cerebrospinal fluid over time.

Doctors often prefer anatomical hemispherectomies because "leaving even a little bit of brain behind can lead seizures to return," says neurologist John Freeman of Johns Hopkins, which specializes in the procedure. On the other hand, functional hemispherectomies, which U.C.L.A. surgeons usually perform, lead to less blood loss. "Our patients are usually under two years of age, so they have less blood to lose," Mathern says. Neurosurgeons have performed the functional operation on children as young as three months old. In these tiny patients, memory and personality develop normally.

Most Johns Hopkins hemispherectomy patients are older than five years. A recent study found that 86 percent of the 111 children who underwent the procedure at Johns Hopkins between 1975 and 2001 are either seizure-free or have nondisabling seizures that do not require medication. Another study found that children who underwent a hemispherectomy often improved academically once their seizures stopped. "One was champion bowler of her class, one was chess champion of his state, and others are in college doing very nicely," Freeman notes.

But Can You Dance?

Of course, removal of half the brain does have its downside—notably significant loss of function on one side of the body. "You can walk, run—some dance or skip—but you lose use of the hand opposite of the hemisphere that was removed," Freeman says. Vision is impaired as well.

Also, if the left side of the brain is taken out, most people have problems with their speech. But, Freeman notes, the younger a person is when he or she undergoes hemispherectomy, the less speech disability the person is likely to have.

Mathern and his colleagues have recently conducted the first functional magnetic resonance imaging study into hemispherectomy patients, investigating how their brain changes with physical rehabilitation. Probing how the remaining cerebral hemispheres of these patients acquire language, sensory, motor and other functions "could shed a great deal of light on the brain's plasticity, or ability to change," Freeman observes. Still, he says, hemispherectomy is among the most drastic kinds of brain surgery and is "something that one does only when the alternatives are worse." ■

Charles Q. Choi is a freelance science writer based in New York City.

MORE ONLINE . . . *Find more investigations into popular myths at* **www.SciAm.com/factorfiction**
To suggest a topic, write to factorfiction@SciAm.com

MICHAEL MATISSE *Getty Images*