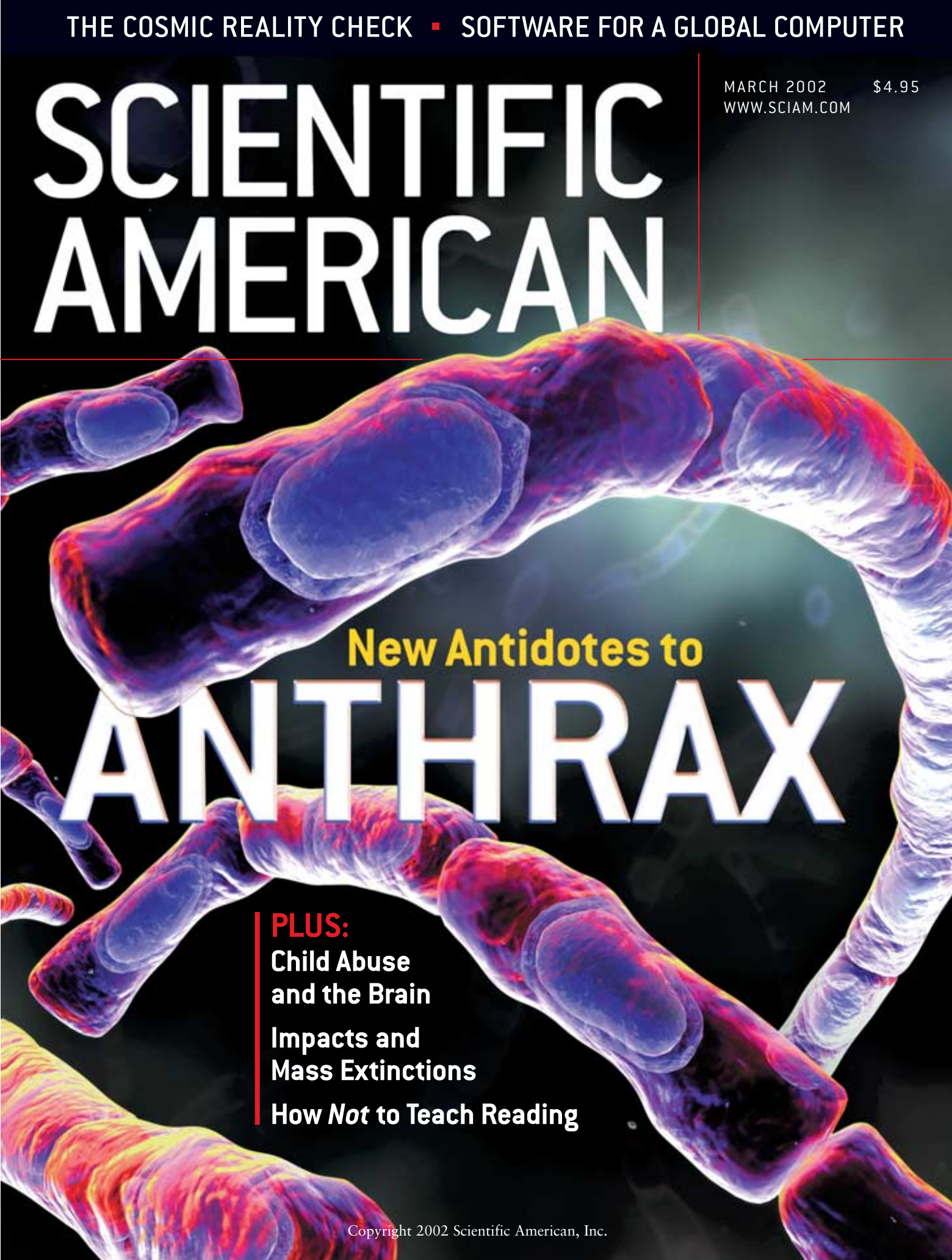
THE COSMIC REALITY CHECK ▪ SOFTWARE FOR A GLOBAL COMPUTER

SCIENTIFIC AMERICAN

MARCH 2002 $4.95
WWW.SCIAM.COM



New Antidotes to ANTHRAX

PLUS:

Child Abuse and the Brain

Impacts and Mass Extinctions

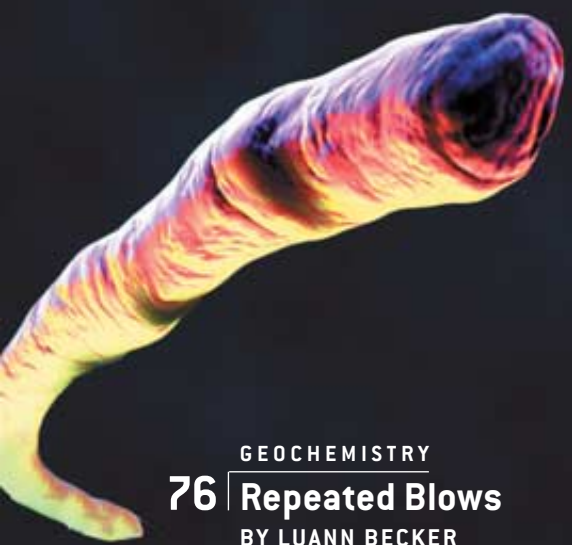
How *Not* to Teach Reading

march 2002

contents

features

SCIENTIFIC AMERICAN Volume 286 Number 3



48 Anthrax bacteria

SCIENTIFIC AMERICAN Volume 286 Number 3

departments

columns

Cover photograph and page 3: Jeff Johnson; this page, clockwise from top left: John McFaul; courtesy of the University of Cambridge; AMEC Border Wind

Treat AIDS Globally

An ounce of prevention is worth a pound of cure, or so the saying goes. That line of thinking has driven AIDS policy in the developing world for almost a generation now: clinics around the globe have dispensed untold millions of free condoms and have counseled hordes of people about how to change their behavior to reduce the risk of HIV infection.

Such prevention efforts—which tend to be less expensive than offering life-prolonging drugs to already infected individuals—have indeed helped stabilize or reduce the incidence of new HIV infections in various countries, most notably Uganda and Thailand. But the time has come for the developed world to acknowledge that prevention alone is not enough to battle AIDS in developing nations.

CAREGIVERS hold boxes containing ashes of South African babies who died of AIDS.

DENIS FARRELL *AP Photo*

United Nations Secretary-General Kofi A. Annan ruffled some public health officials' feathers last June at the U.N. Special Assembly on HIV/AIDS when he included treatment as a priority in the newly established Global Fund to Fight AIDS, Tuberculosis and Malaria. Some thought the money amassed by the fund (only $1.6 billion so far toward the stated goal of more than $7 billion a year) would be better spent concentrating on prevention.

But prevention and treatment go hand in hand when it comes to HIV. People are more likely to inquire about testing—during which they are offered prevention counseling—when they know they will receive treatment if they turn out to be positive. Without hope of therapy, many are afraid to learn their HIV status or may succumb to a fatalism that discourages them from practicing safer sex. And so they unwittingly spread the virus further. Moreover, access to treatment reduces the stigma of HIV infection, which has been a barrier to prevention efforts in places such as sub-Saharan Africa.

The link between treatment and prevention is most apparent in the use of antiretroviral drugs to prevent mothers from passing on HIV to their babies during birth. A single dose of nevirapine to a laboring mother and one to the newborn, a regimen that costs roughly $8, can reduce the likelihood of HIV transmission by roughly 50 percent.

Treating HIV-infected adults might also decrease the transmission of the virus, because such drugs slash the amount of virus in the body. A study reported last year in the *Lancet* found that individuals in Uganda who had lower concentrations of HIV in their blood were less likely to infect their spouses.

The high cost of most antiretroviral drugs and a dearth of doctors, clinics and hospitals block the use of AIDS drugs in many developing countries; political obstacles prevent their employment in others. South African president Thabo Mbeki's refusal to acknowledge that HIV causes AIDS, for example, puts treatment out of reach for his nation, which has the highest overall number of HIV-infected people in the world.

But such hurdles can be surmounted. This year one of South Africa's neighbors, Botswana, initiated a program to offer free antiretroviral treatment to its population. With $50 million each from the Bill and Melinda Gates Foundation and from Merck (which is also providing free drugs), Botswana is providing medicines to citizens in the four hardest-hit areas of the country, which has the highest percentage of HIV-infected adults (40 percent). We can only hope it is not too late elsewhere to broaden the definition of prevention to include using treatment to prevent death.

THE EDITORS editors@sciam.com

How to Contact Us

EDITORIAL

For Letters to the Editors:
Letters to the Editors
Scientific American
415 Madison Ave.
New York, NY 10017-1111
or
editors@sciam.com
Please include your name and mailing address, and cite the article and the issue in which it appeared. Letters may be edited for length and clarity. We regret that we cannot answer all correspondence.

For general inquiries:
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-754-0550
fax: 212-755-1976
or
editors@sciam.com

SUBSCRIPTIONS

For new subscriptions, renewals, gifts, payments, and changes of address:
U.S. and Canada
800-333-1199
Outside North America
515-247-7631
or
www.sciam.com
or
Scientific American
Box 3187
Harlan, IA 51537

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 from *SA*:
permissions@sciam.com
or
212-451-8546 for procedures
or
Permissions Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
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.

New York
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8893
fax: 212-754-1138

Los Angeles
310-234-2699
fax: 310-234-2670

San Francisco
415-403-9030
fax: 415-403-9033

Chicago
Christiaan Rizy
212-451-8228
fax: 212-754-1138

Dallas
The Griffith Group
972-931-9001
fax: 972-931-9074

Detroit
Karen Teegarden & Associates
248-642-1773
fax: 248-642-6138

Canada
Fenn Company, Inc.
905-833-6200
fax: 905-833-2116

U.K.
The Powers Turner Group
+44-207-592-8331
fax: +44-207-630-6999

France and Switzerland
PEM-PEMA
+33-1-4143-8300
fax: +33-1-4143-8330

Germany
Publicitas Germany GmbH
+49-69-71-91-49-0
fax: +49-69-71-91-49-30

Sweden
Andrew Karnig & Associates
+46-8-442-7050
fax: +49-8-442-7059

Belgium
Publicitas Media S.A.
+32-2-639-8445
fax: +32-2-639-8456

Middle East and India
Peter Smith Media & Marketing
+44-140-484-1321
fax: +44-140-484-1320

Japan
Pacific Business, Inc.
+813-3661-6138
fax: +813-3661-6139

Korea
Biscom, Inc.
+822-739-7840
fax: +822-732-3662

Hong Kong
Hutton Media Limited
+852-2528-9135
fax: +852-2528-9281

On the Web

WWW.SCIENTIFICAMERICAN.COM

FEATURED STORY

Teaching Aibo New Tricks

Part of the fun of owning a pet is teaching it tricks. The same holds true for Sony's popular robotic dog, Aibo. Enthusiastic owners have devised a number of software additions to grant the mechanical mutts new abilities—among them how to dance, speak, obey wireless commands and share the color video that serves as their vision.

Last fall, though, one experimenter was temporarily forced to dismantle Web sites offering the software packages when Sony threatened to sue. For now, the robotics hobbyists have prevailed and Sony has backed down—but the whole event has raised interesting questions about the U.S. Digital Millennium Copyright Act of 1998.

www.sciam.com/explorations/2002/012102aibo/

NEW TO THE SITE

Daily Trivia

- How long is one cosmic year?
- Where does the smallest lizard on earth live?
- What does the word "algebra" really mean?

Think you know? Now you can visit ScientificAmerican.com every day and add a new fact to your file of science and technology trivia.

www.sciam.com

ASK THE EXPERTS

How do you get laryngitis?

Rebecca Gaughan of the American Academy of Otolaryngology provides an answer to this question.

www.sciam.com/askexpert/

SCIENTIFIC AMERICAN JOBS

www.scientificamerican.com/jobs

Looking to make a career change in 2002?

Visit Scientific American Jobs for positions in the science and technology sectors. Find opportunities in computers, sales and marketing, research, and more.

SCIENTIFIC AMERICAN®
Established 1845

Letters

EDITORS@SCIAM.COM

"SOCIETY SHOULD CONSIDER what sort of legacy we would be leaving future generations if there were a wholesale move away from paper and into electronic books and texts," writes Chris Rohrs of Novato, Calif., about "The Electronic Paper Chase," by Steve Ditlea [November 2001]. "Paper will last a long time (the Dead Sea scrolls, for example, are almost 2,000 years old), whereas electronic texts are ephemeral and last as long as the current flows. Paper will hold notes and scribbles in the margins, so we can see Mark Twain's or Fermat's edits and musings. How will future literary researchers know about the ebb and flow of ideas of an author who was published electronically?"

Scroll down for additional letters from November and December 2001.

CHEMICAL WEAPONS ON THE QT

Publishing "Better Killing through Chemistry," by George Musser [News Scan, December], on how easy it is to make sarin, seems foolhardy, despite your reassurances. You can't possibly know what every potential terrorist already knows or has read, and you will never know whether you haven't just given someone the idea to try it.

I think you base your reassurances on probabilities, which I would normally accept readily. But before September 11, who among us would have considered possible what happened?

DAN BAGNELL
Cleguer, France

MUSSER REPLIES: The information we provide is extremely minimal; to use it to mount an attack, someone would have to undertake the chemical syntheses, devise a dispersal apparatus and handle the other logistics. Anyone who could do all that would find the explanation we give such a small fraction of the endeavor that it hardly bears mentioning. The government does not monitor, let alone regulate, the domestic purchase of chemical-weapons precursors. It is this lack of oversight that truly facilitates such terrorist actions.

THE CASE AGAINST GROWTH

The truth about biodiversity ["On the Termination of Species," by W. Wayt Gibbs, November] has nothing to do with accurately measuring extinction rates or numbers of species. As long as the extinction rate exceeds that of species generation, biodiversity will decrease, eventually destroying the ecosystems on which people depend—and us along with them. It is not a question of if, but when.

Our real impact is based on the ever growing overall human economy, which in turn is based not only on our global population size but on our average resource usage rate per capita. Even at a stable population level, the economy can continue to grow and threaten our ecological underpinnings.

Growth is limited here on Earth—and thus, by definition, unsustainable. As long as we pursue growth, we will simply be deepening the hole out of which we must climb. Land and species cannot be successfully set aside and kept pristine, because no place is immune to the flow of toxic substances through the air and the water. Such attempts at conservation, along with improved measuring or modeling, will always fail to help us out of our hole.

But as soon as we give up growth in favor of dynamic equilibrium as the hallmark of economic strength, everything from biodiversity to humanity's social ills will come to take care of itself. People, businesses and the nonhuman world will work synergistically, and all will be the better for it.

MARK S. MERITT
Red Hook, N.Y.

Letters

A FACE IN THE CROWD

"Facing a New Menace" [News Scan, by Gary Stix and Philip Yam, November] raises objections to new security measures that could be used to combat terrorism. Face-recognition technology, in particular, has "strong deleterious side effects," according to Deborah Hurley, and the authors cite an "inevitable … loss of some personal liberty."

Nowhere does the article state exactly what liberties are being lost or what deleterious side effects exist in the use of face-recognition technology. In public places, especially transportation centers, we expect, and even demand, increased scrutiny of passengers for the protection of everyone. Since when is showing one's face in public a private act? And since when do we oppose what a machine can do simply because it does it faster and more efficiently than humans?

KEN SIMON
Detroit

SMALLER CLASSES are often better, even if it means converting closets into classrooms.

FOR THE CLASSROOM, A JEFFERSONIAN APPROACH

As a parent and teacher, I am infuriated by the application of cost-benefit analysis to one of our species' basic tasks—the nurturing of the young ["Does Class Size Matter?" by Ronald G. Ehrenberg, Dominic J. Brewer, Adam Gamoran and J. Douglas Willms, November]. If small classes and higher salaries for more competent teachers are effective, why not fund them both? Perhaps the $30 billion–plus that we annually squander on an intelligence establishment that cannot connect terrorist A with plot B and consistently hires personnel who sell our deepest secrets to the highest bidder could be used for this purpose.

In an era when human stupidity seems to be threatening our existence, we might discover, with a relatively minor investment in education, that Jefferson's well-informed citizenry is not only our best defense but also our only hope.

ALEX PIRIE
Somerville, Mass.

If student motivation and academic ability correlate positively with neighborhood income and if smaller class size does improve student achievement, why not have larger classes in upper-income neighborhoods and smaller ones in lower-income areas? That way there would be little, if any, extra cost.

M. G. BLAKESLEE
Portland, Ore.

HUNTING—OR NOT—IN ALASKA

Ned Ford of the Sierra Club alleged in the November Letters column that "one reason the caribou are increasing near the Trans Alaska Pipeline is because pipeline workers were encouraged to kill all the wolves in the area during their off-hours hunting." This statement is absolutely false. The oil companies have maintained a strict ban against hunting of any species, wolves included, by their employees in the oil field and pipeline areas. Some workers caught violating this restriction were summarily fired.

JOHN A. MORRISON
Alaska Outdoor Information Services
Anchorage, Alaska

EDITOR W. WAYT GIBBS REPLIES: *The readers who have criticized Ford's letter are, to the best of my knowledge, correct: neither firearms nor hunting has been allowed by workers or visitors on the oil fields for many years. The construction of the Alaska pipeline did, however, require building a parallel haul road that transects the state. That road opened hundreds of miles of forest lands to much readier access for recreational hunters and fishermen. Caribou are easier and more valuable for hunters to take than wolves, so it seems a bit implausible that hunting has provided a significant net boost in caribou populations in areas remote from human settlements. There may be instances in which the destruction of a pack of wolves for safety reasons has led to surges in local caribou populations.*

ERRATA The first spacecraft to escape our solar system was Pioneer 10, launched in March 1972, not the Voyagers ["A Short Stroll through the Solar System," by W. Wayt Gibbs, November], which were launched five years later.

"Beyond Chicken Soup," by William A. Haseltine [November], implies that Frederick Sanger of the University of Cambridge and his colleagues were the first to decipher the full sequence of a viral genome. Sanger's group was the first to report the sequence of a DNA virus, in 1977. But Walter Fiers of the University of Ghent in Belgium and his colleagues reported the sequence of an RNA virus in 1976.

BRUCE STRACHAN

50, 100 & 150 Years Ago

FROM SCIENTIFIC AMERICAN

Logic Circuits ▪ Simian Brains ▪ Steam Evolution

MARCH 1952

LOGIC MACHINES—"First formulated in the 19th century by the English mathematician George Boole, symbolic logic has been developed into a powerful tool for dealing with complex problems in mathematics and in business. At the moment, logic machines have very limited value, due to the fact that science is seldom confronted with problems of a strictly logical nature which are complex enough to require mechanical aid. Logic networks may become increasingly useful in the operation of the giant electronic computers. Problems frequently arise in deciding the best way to set the machine for a given task, and often these problems are purely logical in character. Computers of the future may have logic circuits built into them so that such decisions will be made automatically. —Martin Gardner" [*Editors' note: This article was the first of many illustrious contributions to the magazine by this author.*]

MARCH 1902

TURBINES TAKE OVER—"Unquestionably, in the development of the steam engine, we are just now entering upon a new era, when steam has ceased to be used as a prime mover. With the demands of locomotive service, the turbine is never likely to displace the reciprocating engine in this class of work. As an electrical drive, however, it is pre-eminently qualified, and since electrical power seems destined to indefinitely enlarge its field of application, the growth of the steam turbine is destined to be rapid and widespread."

MAPPING THE SIMIAN BRAIN—"Sufferers from nervous complaints, especially such as cause interruption of muscular action, may have reason to bless the memory of certain great apes who have co-operated unselfishly with, and without being consulted by, some British scientists and surgeons in privately conducted experiments. Studies of the brains of the higher apes have shown that their composition was sufficiently like that of a man to justify the belief that investigations would furnish knowledge of the human brain."

SPECIAL ISSUE, 1902

AUTOMOBILE CRAZE—"The development of the automobile industry has been absolutely without a parallel. Also in a remarkably short space of time, the automobile has grown from the first crude conception to its present highly developed condition. The year 1901–1902 is likely to rank as one of the most important in the history of the automobile."

IT'S NOT HOGWARTS—"Children of some wealthy parents are to be the subjects of food experiments by scientists in a splendidly equipped home known as the Chicago Hospital School for nervous and delicate children, says the New York Medical Journal. Only the well-to-do can afford to send their children to the school. The home can accommodate only fifteen children, and has more applications than it can fill."

MARCH 1852

SMALLPOX WARNING—"A work by Dr. T. H. Buckler, physician to the Baltimore Almshouse, alludes to the propagation of disease by means of banknotes: 'The inmate of a small-pox hospital, if he wants a lemon, sends a note saturated with the poison (and having, perhaps, the very sea-sick odor of small-pox) to a confectioner, who takes it, of course. It would be impossible to conceive of any better mode of distributing the poison of a disease known to be so contagious and infectious.'"

NEBULAR SKEPTICISM—"Pierre-Simon LaPlace thought the solar system was, at first, one vast nebula, in a high state of heat from chemical action. We [the editors] object entirely to the Nebular hypothesists. If this world were originally in a state of gas, just imagine a mass of gases in chaotic confusion, of more than thirteen million miles in diameter, and this tossing away through space like a ship without sail or rudder. These philosophers have strange ideas of the Divine Government."

SECURITY

I Seek You

ARE NEW SECURITY TECHNOLOGIES WORTH THE INTRUSION AND THE COST? BY WENDY M. GROSSMAN

Within hours of the September 11 attacks, even rabid civil libertarians were talking about the need for national identification systems, giant linked databases, face-recognition technology, closed-circuit television (CCTV) monitors, biometric authentication, profiling and increased government wiretapping powers. Some of these measures—particularly, more latitude in wiretapping—have already been enacted as law, as security services around the world have seemingly dusted off every plan once deemed too invasive and presented it to legislatures. If to gain security in the U.S. we must compromise some of the rights that have been considered essential, at least we should be reasonably sure that such measures will be worth the money and lost liberty. Yet based on current uses of the security technology, there is reason to remain skeptical.

READY FOR YOUR CLOSE-UP?
Closed-circuit televisions, common in the U.K., are supposed to deter crime, but data suggest they don't.

ALISTAIR GRANT *AP Photo*

Most of the proposed technologies are not only controversial but also expensive, slow and complicated to deploy. Most are either untried or untested on the necessary scale and carry risks that are not well understood. Solid scientific data are frequently lacking—few studies exist detailing the success rate of psychological profiling, for example. One rare exception is a January/February 2001 study published in *Australasian Science* that tentatively concluded that the few profilers who agreed to be tested (only five did) performed only slightly better than competing groups of psychologists, science students, detectives and, pulling up the rear, civilians and psychics.

Media hype and overblown claims by firms selling the technology—several companies involved in biometrics, the field that attempts to identify people through their biological traits, hired lobbyists in October—don't help. Take, for example, the idea of combining face recognition with CCTV systems to scan airport terminals for suspected terrorists. In the camera-filled U.K., the London borough of Newham claimed its pilot scheme produced a 21 percent drop in crimes "against the person" and unprecedented decreases in criminal property damage, vehicle-related crime, and burglary. In August 2001 the U.K. approved a further £79 million (about $114 million) for 250 new CCTV systems. Simon Davies, a fellow at the London School of Economics and the founder and director of Privacy International, estimates that the country has at least 1.5 million CCTV cameras now in place.

Jason Ditton, professor of law at the University of Sheffield in England and director of the Scottish Center for Criminology in Glasgow, is one of the few academic sources of

BIOMETRICS includes fingerprint analyses. The system shown, developed by NTT in Japan, recognizes a print in 0.5 second.

CCTV information. His research, funded by the government's Scottish Office, shows that the cameras are not cost-effective and that they reduce neither crime nor the fear of crime. His 1999 study of CCTV in Glasgow's city center revealed that although crime fell in the areas covered by the cameras, the drop was insignificant once general crime trends were taken into account. Even worse results were in Sydney, Australia, where a $1-million system accounted for an average of one arrest every 160 days—a quarter of the Glasgow rate, which Ditton thought was poor.

Moreover, it is not clear how much of a role the displacement effect—the shifting of crime from one area to another—plays. A Sydney city council's report indicates that the cameras probably displaced some crime to areas outside the lens's view. And therein lies a fundamental design conflict. For the cameras to be an effective deterrent, everyone has to know they're there; however, to be effective in spotting criminals they need to be covert.

Trying to add face recognition to the camera system leads to an even more fundamental problem: you can only catch people you're already looking for. James L. Wayman, director of the U.S. National Biometric Test Center at San Jose State University, says flatly: "You cannot hang a camera on a pole and expect to ever find anybody. Even the vendors say that." Indeed, the American Civil Liberties Union reported in January that such a system in Tampa, Fla., failed to identify any individuals in the police database of photos and misidentified some innocents as suspects.

Even if it worked, the difficulty remains of predicting what people will do. Wayman is a strong proponent of the Immigration and Naturalization Service Passenger Accelerated Service System (INSPASS), which lets frequent travelers register handprints and speed through immigration checks. But "how do you know someone's going to be a terrorist when they get on an airplane?" Wayman asks. "It's beyond what science is capable of predicting." Besides, as the September 11 events showed, terrorists could patiently build up seemingly legitimate travel logs and apparently innocent lives before committing their acts.

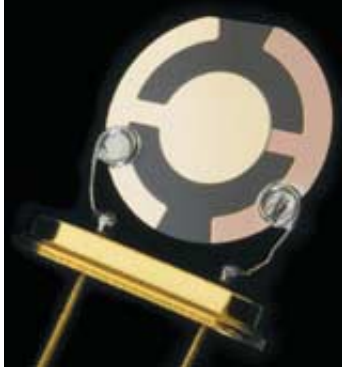
Much of the debate about new security technologies is framed around the assumption that they will work and that personal privacy is a necessary sacrifice, when in fact the effectiveness of such technologies is questionable. An alternative solution, notes Philip E. Agre, associate professor of information studies at the University of California at Los Angeles, is to spend the money to bolster existing security practices: improving authentication for airport staff, training flight attendants in martial arts, improving luggage searches and finding ways to prevent identity theft. These and other measures might eliminate the possibility of trading security for dearly held freedoms.

Wendy M. Grossman, based in London, is a frequent contributor who specializes in computer and information technology.

WHEN POWER TRUMPS PRIVACY

Fearing that power, once handed over, is not likely to be rescinded, privacy advocates are concerned about granting law enforcement greater latitude for surveillance. Currently European privacy laws require that all communications data (telephone records, e-mail, Web logs) be destroyed once they are no longer needed by the service provider for billing purposes. Most closed-circuit TV systems follow a similar principle, so that tapes are typically retained for 31 days.

President George W. Bush is asking a reluctant European Union to loosen these rules in the interests of fighting terrorism, even though such data retention is not required under U.S. law. Meanwhile the U.K.'s Anti-Terrorism, Crime and Security Act whizzed through Parliament to become law in December; it includes a confusing clause allowing the retention of data in the interests of national security.



SHAKE, RATTLE AND POP: Quartz resonator vibrates so quickly, it audibly separates a virus from antibodies.

DIAGNOSTICS

Hears to Your Health

A SENSOR LETS RESEARCHERS LISTEN FOR GERMS **BY MICHAEL BEHAR**

When University of Cambridge scientists first heard a virus wresting itself from the tenacious clutch of an antibody, the sound should have elicited a collective sigh of relief from fretting patients everywhere. The researchers were testing a new device that can hear the presence of a virus in a blood sample. For many patients, who sometimes wait days to get test results, the invention could mean on-the-spot detection of HIV, hepatitis and dozens of other pathogens, including anthrax and smallpox.

ITSUO INOUYE *AP Photo (top)*; MATTHEW COOPER AND PAUL LARHAM *University of Cambridge (bottom)*

The Cambridge experiment involved a tiny slice of quartz crystal layered with antibodies. A virus—in the first case, herpes simplex—was introduced and subsequently bound to an antibody on the crystal. The scientists then slowly increased the frequency of an electric current flowing into the quartz. As the quartz oscillated, it whipped the virus and antibody back and forth. Eventually the herpesvirus tore away from the antibody, emitting a faint pop.

"If you apply enough force to a stick, it will snap and you hear a sound," explains Matthew Cooper, one of six researchers involved in the project. "Likewise, we can hear the sound of the bonds snapping when we break apart a virus and an antibody." The quartz acts like a piezoelectric microphone, converting mechanical vibrations into electrical impulses. Similarly, when a virus breaks from an antibody, the quartz changes the vibrations into detectable electrical signals.

The entire process, termed rupture event scanning, is far better than current enzyme- or biochemical-based viral tests, which reveal the existence of antibodies but can't determine whether or not a subject is carrying the associated virus. "We are directly detecting the virus," Cooper points out, "which gives you a much more accurate prognosis."

Using targeted antibodies, the quartz microphone could be fashioned to recognize the sounds of a multitude of viruses. "It could even detect bioterrorist germs," Cooper says: add a microthin film of anthrax or smallpox antibodies to the crystal, then douse it with a sample of infected blood for an instant diagnosis. He is quick to add that the technology is at least three years from its commercial debut. To that end, the Cambridge team has formed a company called Akubio. With $1.7 million in funding (mostly from venture capitalists), Cooper wants to engineer a cell phone–size tool that can eavesdrop on "cells, bacteria and a variety of different substances in the body."

Michael Behar is a Washington, D.C.–based science and technology journalist and a former senior editor at Wired *magazine.*

RIDING THE WHIPSAW

To shake the virus loose from the tightly clinging antibodies, researchers had to snap them back and forth 14 million times a second. As a result, the virus and antibodies experience a force roughly 10 million times that of gravity.

HEALTH

Throw the Switch?

NEW VACCINES MAY NOT BE A REASON TO KEEP SMALLPOX STOCKS

BY DANIEL GROSSMAN

In a statement last November, U.S. Health and Human Services secretary Tommy G. Thompson announced his opposition to the execution of one of the world's most infamous mass murderers. The killer is variola, the virus responsible for smallpox, which took more than 300 million lives in the 20th century.

After the World Health Organization eradicated smallpox in 1977, all known cultures were consolidated in two repositories, one at the Centers for Disease Control in Atlanta and one at the State Research Center of Virology and Biotechnology in Koltsovo, Russia. Since eradication, health officials and scientists have been debating whether to destroy these stocks and, if so, when. Some argue that the variola could be the basis for novel vaccines or a smallpox cure should anyone release any secret stashes of the virus. Others think that there are no good scientific or public health reasons to believe that workable drugs could be created from the existing stocks.

Though very effective for preventing smallpox, today's vaccine is not suitable for everyone. It contains live vaccinia virus (a distant cousin of variola), which causes severe complications in people with impaired immune systems, including chemotherapy and AIDS patients, and is not considered adequately tested to use on pregnant women. Certain otherwise healthy individuals also develop serious side effects, among them, in rare cases, permanent neurological damage. A new vaccine free of live virus might be safer.

The Bush administration also wants drugs to treat smallpox after it has been contracted. "No one wants to keep this virus forever," confesses one high-level government official familiar with smallpox deliberations. "We just want to get rid of all of it or have the tools to handle it if someone has it in a freezer." Pursuing these goals requires further research with live variola virus.

Frank Fenner, an eradication program

EXISTING SMALLPOX VACCINE is good enough, some researchers argue.

FRITZ GORO *Timepix*

news SCAN

A POX OF MICE AND MEN

Last year's report of an experiment conducted on mice in Australia has increased the intensity of the debate over what to do with variola (smallpox) virus stocks. Researchers hoping to control that continent's wild mouse population added a single gene to the relatively benign virus that causes mousepox. The addition made the virus devastatingly lethal even in mice vaccinated against mousepox.

Some scientists say that if someone were sinister enough to make a similar change to variola, then existing cultures might be helpful in developing countermeasures. Others argue that samples of the newly altered virus, not the stocks from which it was produced, would be the critical foundation of a treatment.

alumnus and a longtime WHO adviser on variola research, says new drugs are not needed. The existing vaccine, he points out, already works as a treatment if administered within several days of exposure. He predicts that efforts to find a cure that could treat smallpox in its later stages will prove "fruitless." And if you do have a new smallpox drug, Fenner asks, "How on earth do you test it?" There are no longer any smallpox victims.

Laboratory animals could be the answer. Peter B. Jahrling, a biologist at the U.S. Army Medical Research Institute of Infectious Diseases, succeeded in infecting monkeys—an important development because animals don't naturally contract smallpox. The monkeys had symptoms and tissue and organ damage similar to those in humans and so might pave the way for new drugs. With continued access to the virus, Jahrling thinks he could have a treatment for smallpox ready within 10 years: "With clenched teeth, I could do it in five."

Critics say it is premature to conclude that Jahrling's monkeys are a valid analogue of human subjects. The animals received the variola virus intravenously, at doses far in excess of what it takes to produce smallpox in humans. In fact, many compounds that work well in lab animals fail miserably in humans. Rather than gambling on a drug tested only on animals, Fenner argues that researchers should look to improve existing vaccines. Figuring out how to treat the complications from the smallpox vaccine would be cheaper and easier to accomplish, and such work does not require the variola virus.

Underlying the debate over the variola repositories is a disagreement about human nature. Those who want to use the virus for more research say September 11 proves that bad people don't necessarily feel bound by international laws or accepted standards of behavior. Those who would like to destroy the stockpiles—heavily represented by veterans of the eradication campaign—insist that civilized nations of the world should set an example and send a message to would-be bioterrorists. The WHO World Health Assembly is expected to consider these conflicting views in May when it convenes in Geneva for its annual meeting.

Daniel Grossman is a writer and radio producer based in Watertown, Mass.

TRAINING

Astronaut Boot Camp

NASA FINDS A NEW WAY TO IMBUE RECRUITS WITH THE RIGHT STUFF **BY PHIL SCOTT**

Back in the good old days, going on a space mission meant training, training and more training—in simulators. But these days NASA makes sure astronauts also spend time at sleep-away camp with a few fellow astronauts, dining outdoors and sleeping under the stars. Okay, it's a little rougher than roasting marshmallows and telling ghost stories. In fact, it makes TV's *Survivor* look like a day at Six Flags. The campsite: Cold Lake in Alberta, Canada. "It's really cold, –30 degrees Fahrenheit. It gets your attention," says NASA astronaut Andrew Thomas.

BRUTAL COLD of Canadian camping builds character for would-be astronauts—here in January 2000.

CPL. L. A. WHITE *4 Wing Imaging*

Thomas put the program together in 1999, after pitching his tent for four and half months on Russia's Mir space station. Like the six other Americans sent to Mir, Thomas felt culturally isolated. "So I thought it wise to develop a program to prepare astronauts for interpersonal issues on long space flights," he explains.

The experience breaks down into three main topics: leadership, self-management and teamwork. Thomas teaches the first workshop, which consists of classroom lectures on the behavior of astronauts and on leadership in close quarters and in isolation. He draws comparisons between Norwegian Roald Amundsen and Englishman Robert Scott's Antarctic race. "Amundsen had extraordinary capacity to lead and to give attention to details," Thomas says. Amundsen successfully reached and returned from the South Pole

because he planned ahead, adapted skis and sled dogs from his studies of the Inuits up north, and adopted a democratic style in everyday decision making. But when the tough calls had to be made, he would do it.

Scott, who perished with his team on the return trip, "made his decisions in an autocratic, hierarchical style," Thomas continues. "He then made infamous blunders"—such as setting out with inadequate food and fuel reserves—and "having his men drag back sleds filled with rocks in the name of science while they died in their tracks." (Recent analyses suggest that unusually cold weather, more than poor leadership, doomed Scott's expedition [see "Thawing Scott's Legacy," Profile; SCIENTIFIC AMERICAN, December 2001].)

After five half-days in the classroom, a group of six astronauts take it outside: to the National Outdoor Leadership School, conducted in Utah and Wyoming. Next comes the true and final stress test: Cold Lake. There the group receives a couple of days of training with its cold-weather equipment and then is dropped by helicopter into the middle of the Canadian military base. Assigned to map an unfamiliar area, trainees set up a central base and receive commands by radio, just as they would from Mission Control. "This may be at two in the morning," Thomas says.

Each astronaut takes a turn as leader for few days. "The leader has to decide who's best to go, who's been working hardest and needs a rest. The risks are real in the sense of providing stress," Thomas adds. "It's a good analogue for when they end up in space."

Although some campers have griped that long cold-weather outings are just NASA's latest big new idea, response has been positive overall. Soon, however, the astronauts might contend with even more claustrophobic togetherness. NASA has contracted with the National Oceanic and Atmospheric Administration to use its underwater lab Aquarius, off Florida's Key Largo. That could mean that NASA plans training in addition to Cold Lake—or that the space agency has moved on to its latest big new idea.

Phil Scott is based in New York City.

A FEW COLD FACTS

- **Cold Lake training lasts 11 days.**
- **The International Space Station orbits at an average altitude of 247 statute miles (397 kilometers).**
- **Right now the average crew of three—the most that can be accommodated for an emergency exit on the attached Soyuz capsule—stays on the ISS for three months.**

ENERGY

Blowing Out to Sea

OFFSHORE WIND FARMS MAY FINALLY REACH THE U.S. **BY WENDY WILLIAMS**

With little alteration to the national power grid, the U.S. could quickly get at least 12 percent of its electricity from wind. Yet currently, wind generators supply only about 0.5 percent, in part because people don't want to live underneath the tall turbines. In Europe one solution to the people problem is to place the wind machines out at sea, where the winds are stronger anyway.

Acknowledging this potential, a Yarmouth, Mass., company plans to build America's first offshore wind farm by the end of 2005. Cape Wind Associates has slated construction of a 420-megawatt wind project on a shallow sandbar known as Horseshoe Shoal, located five miles south of Cape Cod between the islands of Nantucket and Martha's Vineyard. It would be the world's second largest, after Ireland's recently proposed 520-megawatt farm.

Each of the 170 ultra-high-tech wind turbines will stand 260 feet tall at the turbine hub, and each blade will be up to 150 feet long. The turbines, which should be visible in the distance from the Hyannisport Kennedy enclave, will be laid out in a grid pattern over 25 square miles of saltwater. An underwater cable will run from the turbine complex to a Cape Cod substation. Project developers claim that at peak operation the farm will satisfy almost all the electricity needs of Cape residents—a critical selling point in a region that suffers increasingly from air inversions and smog.

Less than a decade old, offshore wind technology has been virtually ignored by U.S. companies until now. In Europe, though, it's the hot idea in "green" energy. Denmark, for example, trumpets the fact that 50 percent of its energy will come from wind by 2030. If successful, offshore wind farms could solve many problems encountered with land-based wind technology in densely populated regions. Ocean winds are stronger and steadier. Land

OFFSHORE WIND FARMS, such as these in the North Sea off the coast of Blyth in the U.K., are less likely to draw complaints of noise and unsightliness.

AMEC BORDER WIND

acquisition is unnecessary. And, perhaps most important, the huge turbines are out of sight and earshot of most people. Initially fishermen worried about their catch volume decreasing, but several European studies suggest that the heavily anchored turbines act like shipwrecks and in fact improve fish numbers.

On the flip side, investment costs are mammoth. Cape Wind, having already invested several million dollars in planning studies, expects to spend a total of $600 million. James S. Gordon, president of Cape Wind, is confident that the whole package can be financed through private sources. Under his 27-year leadership, Energy Management, a partner in Cape Wind, has built a number of natural gas–fired plants in New England. Says Gordon: "We're creating a national model for America's energy and environmental future."

The U.S. Department of Energy is "watching the Cape project very closely," remarks Brian Parsons, a scientist with the DOE's National Renewable Energy Laboratory. But the size of the undertaking has raised some eyebrows. "I'd be a little skeptical about starting with something that big," warns wind-farm engineering expert Tim Cockerill, a research fellow at the University of Sunderland in England. Others in Europe, however, are thinking along the same lines as Cape Wind. Researchers at the Dutch Offshore Wind Energy Converter project are aiming for a single six-megawatt offshore turbine by 2008. Continued interest may prove within the decade whether this alternative to fossil fuels is more than just a passing gust.

Wendy Williams, based in Mashpee, Mass., is studying technologies that reduce carbon emissions through a grant from the Fund for Investigative Journalism.

news SCAN

PASSING THE CARBON BUCK

The Kyoto Protocol, an international agreement to curb emissions of global warming gases, allows countries to trade emissions through a commodity called a CO_2 equivalent, which equals the amount of industrial greenhouse gases that have the heat-trapping ability of one metric ton of carbon dioxide. The supply of CO_2 equivalents is severely limited.

According to a trading group formed by the financial-services firm Cantor Fitzgerald, the price of one CO_2 equivalent ranges from $1 to $2 a year, although in the future it may reach $5 to $9. Cape Wind claims that the 420-megawatt wind farm will displace a plant that would have annually spewed 1.37 million metric tons of carbon dioxide.

COSMOLOGY

Been There, Done That

THE BIG BANG MAY NOT HAVE BEEN A SINGULAR EVENT **BY GEORGE MUSSER**

Singularities are the toxic waste of cosmology. Theories, let alone children, are well advised not to touch anything with an infinite density or temperature: the zero time of the big bang, say, or the very center of a black hole. At such places, physics dissolves into metaphysics. These mathematical points admit of no explanation; they just are. To dispose of them, cosmologists usually have opted for burial. For instance, cosmic inflation—the favored mechanism for how our universe expanded from the big bang—does not eliminate the primeval singularity but simply isolates it from today's universe.

Lately, though, a more thorough decontamination is becoming a viable option, especially with the maturing of string theory, physicists' best candidate for a theory of everything. Last fall cosmologists Paul Steinhardt of Princeton University and Neil Turok of the University of Cambridge, building on earlier work with Steinhardt's graduate student Justin Khoury and string theorists Burt A. Ovrut of the University of Pennsylvania and Nathan Seiberg of the Institute for Advanced Study in Princeton, proposed that the big bang is not a one-of-a-kind event but part of a recurring cycle. "What we're motivated by string theory to believe is that the big bang is not what we've always thought—a beginning of space and time, where temperature and energy diverge," Steinhardt says. "Rather

THE GREAT CYCLE OF BEING

The idea of a cycling universe seems to cycle around every now and then. Its last appearance was motivated by the possibility that the universe has enough matter to reverse its expansion and collapse in a big crunch. Observations have since ruled that out.

it is a transition between the current expanding phase and a preexisting contracting phase."

Detoxifying singularities has long been one of string theory's major goals. According to the theory, elementary particles ultimately consist of wriggling strings, which have multidimensional counterparts known as branes. The intrinsic size of strings and branes prevents them from collapsing into points of infinite density. The theory has already had some success in explaining black holes as a novel type of particle, and over the past decade it has inspired several alternatives to the standard picture of inflation.

Like some of those alternatives, the cyclic model is based on the idea that our universe is a three-dimensional brane that bounds a four-dimensional space. Another brane—a parallel universe—resides a subsubatomic distance away. That universe is closer to you than your own skin, yet you can never see or touch it.

These two branes act as if connected by a spring, which pulls the branes together when they are far apart and pushes them apart when they are close. Thus, they oscillate to and fro. Periodically the branes hit and rebound like cymbals. To those of us stuck inside one of the branes, the collision looks exactly like a big bang. The hot primordial soup was the energy dumped into the branes when they hit. The density fluctuations that seeded galaxies began as wrinkles in the branes.

Many a cosmological model has foundered on the question of these density fluctuations. Observations indicate that the fluctuations had the same amplitude no matter what their size. The cyclic model predicts exactly that—the only model besides inflation to do so. "Without any notion of inflation whatsoever, we are able to account for that near-scale-invariant spectrum," Ovrut says. "That really was a remarkable discovery."

And unlike inflation, the cyclic model naturally incorporates the dark energy that is now causing cosmic expansion to accelerate: it is none other than the spring energy.

Like a bicycle pump, the back-and-forth motion of the fourth dimension puffs up the volume of our three dimensions. The pump allows a little backflow, so just before each collision, the branes contract slightly. But the density never becomes infinite. "The only thing that is singular is that one dimension shrinks to zero for one moment," Turok says. "This is the mildest of all possible singularities."

Unfortunately, a mild singularity is still a singularity. String theory is too provisional to detoxify it fully, so researchers can't be sure that some unsuspected effect won't undo each cycle's careful preparation for the next. "How do small perturbations come through the big crunch and go out of it?" asks cosmologist Andrei Linde of Stanford University, a leading critic of the model. "It is like throwing a chair into a black hole and expecting it to rematerialize later." And that is not the only problem; the precise behavior of the spring-like force, for instance, seems rather ad hoc.

New observations of the cosmic microwave background radiation should be able to confirm or dispel these misgivings. Whatever becomes of the model, it has encouraged cosmologists to question conventional wisdom. Gabriele Veneziano of CERN, a pioneer of both string theory and its application to cosmology, says, "Thanks partly to the work of Turok, Steinhardt and colleagues, our community is much more ready to accept that the big bang was the outcome of something rather than the cause of everything."

A longer discussion appears at www.sciam.com/explorations/2002/021102cyclic/

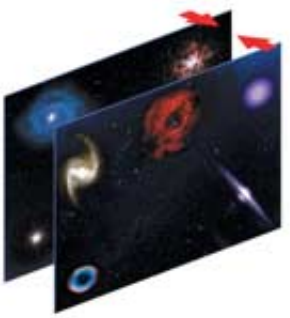
A CYCLE BUILT FOR TWO BRANES

Earlier models of the cycling universe had a fatal flaw: the big crunch is not a perfect mirror image of the big bang. As space contracts, photons gain energy at the expense of the gravitational field, so the universe ends up hotter than when it started. No true cycle could develop; the model requires an ultimate beginning as surely as the one-time big bang does.

The new cyclic model solves that problem. The accelerating expansion wrought by dark energy dilutes the photons, so each bang begins afresh. (This acceleration fulfills the same role as inflation in the standard big bang theory but occurs at a different point in cosmic history.) The universe can be infinitely old, thereby eliminating the puzzle of what came "before."

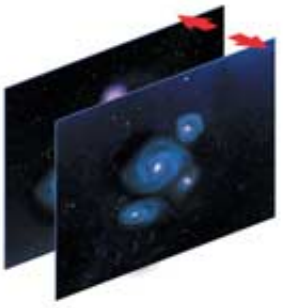
CYCLIC COSMOLOGY posits that our universe and a twin—shown here as planes, but actually three-dimensional—periodically bounce off each other.

1. The universes stop moving apart and start to approach each other.



2. Even as they do so, each universe continues to expand.

3. They collide. A new big bang commences.



4. The collision refills each universe with matter.

SAMUEL VELASCO

DATA POINTS: TAKING STOCK

Since the time of Carolus Linnaeus, who devised the modern species classification system 250 years ago, scientists have categorized only a small fraction of life on the earth. Proponents of the ambitious All-Species Inventory (www.all-species.org) hope to finish the job, which would include bacteria and fungi.

Estimated number of species: 7 million to 100 million

Estimated number identified so far: 1.8 million

Target time of completion: 25 years

Cost, lower estimate: $3 billion

Cost, upper estimate: $50 billion

Total raised so far: $1 million

SOURCES: www.all-species.org; The Scientist, July 23, 2001; the New York Times, December 9, 2001.

ASTRONOMY

Space Rock Candy

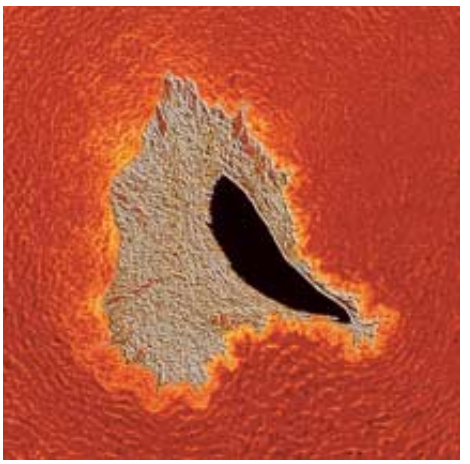
Rocks from space have always posed a threat to life, so how ironic that life's building blocks keep showing up on meteorites. First it was amino acids; now it's sugar. NASA researchers analyzed sugar molecules coating two kinds of carbon-rich meteorite leftover from the solar system's first days and found that the abundance of the compounds decreased with their size and that the sugars were present in many different molecular arrangements. Both characteristics suggest an extraterrestrial origin, because biological sugars tend to be larger and of particular shapes. The isotope concentrations of the meteoric confection were also unlike those of earthly sweets. The authors suggested that the simple sugars could have arisen when starlight bombarded dense clouds of dust floating between stars, which were later caught up in the solar system as asteroids. Their research appeared in the December 20/27, 2001, *Nature*. —*JR Minkel*

HEART DISEASE

Inflamed Blame Game

Some scientists believe that past infections may increase the chance of an inflammatory immune response to plaque-filled arteries. In possible support, a recent study found a correlation between exposure to multiple infectious organisms and the extent of atherosclerosis and the risk of death from it. German researchers tested 572 people suffering from heart disease for antibodies to eight organisms, from herpes and Epstein-Barr viruses to the bacteria that cause pneumonia and stomach ulcers. Participants with the most exposures were up to five times more likely than those least exposed to have advanced atherosclerosis.

After three years, the death rate for patients with advanced heart disease who tested positive for few or no pathogens was 7 percent, whereas 20 percent of those who tested positive for most or all of the infections died. Increasing pathogen exposure also correlated with higher mortality in limited atherosclerosis. The study was published in the January 1 *Circulation*. —*JR Minkel*



ARTERY clogged with plaque (*gray*) only has a little opening left (*black*).

PHYSICS

Superfluid Freeze

Atoms in a Bose-Einstein condensate, the strange gaseous superfluid that forms near absolute zero, do not have definite locations. Instead each atom is "smeared out" across the whole cloud of atoms, and the cloud behaves a lot like a single entity. Now physicists in Germany have used lasers to "freeze" individual atoms in the condensate. Laser beams bathe the cloud from six directions, and the interfering light waves form an optical egg crate for the atoms. At high intensities, the egg crate's pits are deep, and each one captures an atom and holds it in place. The characteristic quantum properties of the condensate are lost. Turning down the lasers restores the condensate, like ice melting to water. The frozen state, called a Mott insulator, may provide yet another route to building a quantum computer by using each atom in the lattice as one quantum bit. The January 3 *Nature* contains the results.

—*Graham P. Collins*

ILLUSTRATION BY MATT COLLINS; A. PASIEKA *Photo Researchers, Inc.* (*bottom*)

AGRICULTURE

Microwaves of Grain

The moisture content of grains such as corn, wheat, barley and soybeans is a crucial factor in determining the proper time to harvest them. If moisture levels are too high, grains may be damaged during threshing and shelling; low levels increase the risk of grains being shattered and kernels breaking. Currently samples are collected and tested by hand, and each type of grain requires a separate set of measurements. A new technique, developed by the Agricultural Research Service of the U.S. Department of Agriculture, can significantly improve moisture measurements. The system sends microwaves through the grain to a receiving antenna, which measures changes in the waves that reveal the moisture content. Just as important, the same technique can be used on all grains. *—Steve Mirsky*

HARVEST TIME through moisture

ECOLOGY

Bubble Bath of Death

Deoxygenating ballast water could help prevent stowaway species from spreading around the world. Current approaches to killing off invaders rely on heat, poisons and filtration, which are costly and may harm local waters. Researchers found that simply bubbling nitrogen into ballast water depletes it of oxygen, spelling doom for the larvae of tubeworms, crabs and zebra mussels in a matter of days. The idea first originated as a means to minimize rust; ship owners spend about $100,000 per vessel every year for the paint needed to protect against corrosion. Though effective against many species, deoxygenating ballast water isn't a panacea—it won't work against anaerobic bacteria or organisms in certain life stages that require no oxygen. The study appears in the January *Biological Conservation*. *—Philip Yam*

ZEBRA MUSSELS and other invasive species can overrun native fauna.

FARRELL GREHAN *Corbis (top)*; SCOTT CAMAZINE *Photo Researchers, Inc. (bottom)*

STENCH WARFARE

Blows to the Nose

Of all the repulsive smells you've whiffed in your life, which ones are apt to clear a room the fastest? Experts at the U.S. Department of Defense want to know so they might use them in a nonlethal "odor bomb." Such a nasty device could be useful for quelling demonstrations or repelling enemy troops. Researchers at the Monell Chemical Senses Center in Philadelphia received a DOD grant three years ago to find the stinkiest stenches. "We focused on biological odors because we thought those had the best chance of being recognized universally," explains Pam Dalton, the cognitive psychologist who led the study.

In recently completed tests, subjects reacted most profoundly to the potent reeks of human fecal waste and rotting food. The former packs foul-smelling skatole compounds; the latter emits rancid-smelling butyric acid and various sulfurous decay byproducts. Do the champion malodors work as planned? "Well, one time I managed to evacuate the building," Dalton reports. "And the people around here are used to offensive smells." *—Steven Ashley*

WWW.SCIAM.COM/NEWS

BRIEF BITS

- In a step toward xenotransplantation, researchers have made genetically modified pig clones that lack a copy of a gene that causes immune system rejection. /010402/2.html
- Challenging conventional wisdom, an experiment shows that a language learned in adulthood is processed the same way as the primary language learned in childhood. /010202/2.html
- In mouse studies, gene therapy cured sickle-cell anemia. After a virus delivered a modified gene into the bone marrow's stem cells, the mice began churning out mostly normal red blood corpuscles. /121401/1.html

BY THE NUMBERS

Down with Evolution!

CREATIONISTS ARE CHANGING STATE EDUCATIONAL STANDARDS **BY RODGER DOYLE**

EVOLUTIONARY STANDARDS

Criteria used in rating state standards on the teaching of evolution:

- **Is the word "evolution" used?**
- **Is biological evolution treated?**
- **Is human evolution treated?**
- **Is geological evolution treated?**
- **Is cosmology treated?**
- **Are connections among historical sciences treated (for example, discussion of the essential role of living things in the transition to an oxidizing atmosphere)?**
- **Is creationist jargon used?**
- **Is there a disclaimer that subverts the sound treatment of evolution?**

SOURCE: Good Science, Bad Science: Teaching Evolution in the States, *by Lawrence S. Lerner. Thomas B. Fordham Foundation, Washington, D.C., 2000 (www.edexcellence.net). Notations on the map are quotations from the book.*

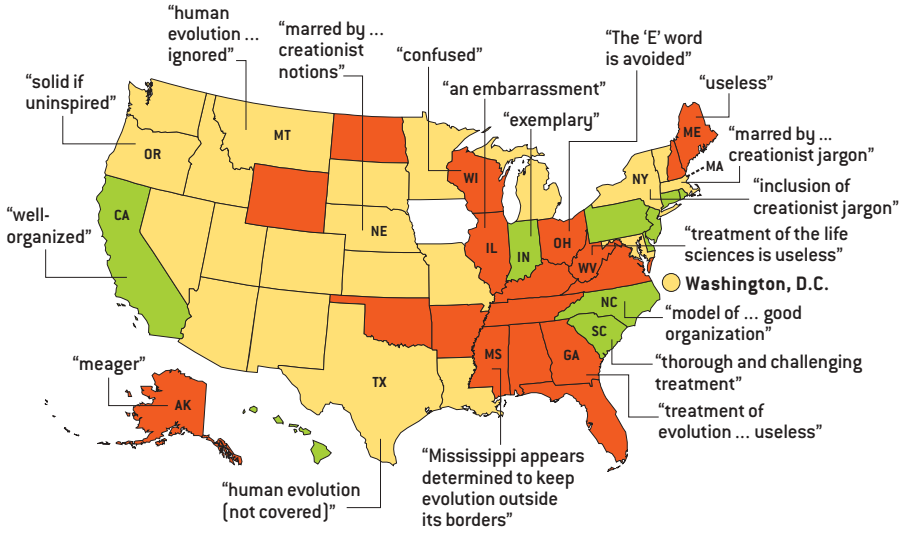
Since 1920 creationists have been successful in persuading legislatures in five Southern states to pass laws favorable to their views, but the courts consistently struck them down, saying that they violated the establishment clause of the Constitution. In the 1990s creationists began focusing instead on changing state educational standards. The most famous attempt to do so in recent years—the decision of the Kansas Board of Education to eliminate evolution from the state's science standards—was not a success: the decision was reversed in 2001 when antievolution board members were defeated for reelection.

Still, creationists have been victorious in many other states, a trend catalogued by Lawrence S. Lerner of California State University at Long Beach. His evaluation, summarized and updated in the map below, is valuable in part because it points up the widespread sway of creationists in Northern states, such as Illinois, Ohio and Wisconsin, that have a liberal or moderate tradition. Furthermore, it highlights the fact that certain Southern states—North and South Carolina—have more rigorous educational standards than some Northern states, such as New York and Massachusetts.

There is little information on what is actually taught in individual classrooms and school districts, so it is not clear what effect state standards have on the quality of evolution teaching. The influence of the standards is, however, potentially great because they are likely to affect the content of textbooks and lesson plans. Standards set the tone under which teachers and administrators work and, if written well, make it easier for science-oriented educators to insist that all teachers, including the one third who advocate equal time for creationism, observe proper guidelines.

Creationists have been able to alter state education standards despite being a fairly small minority. According to a 1999 poll by the People for the American Way Foundation, a Washington, D.C.–based organization opposed to the teaching of creationism in science classes, only 16 percent of Americans support the teaching of creationism to the exclusion of evolution. A huge majority—83 percent—favor teaching evolution, but most of them maintain that creationism should be discussed in science classes with evolution. Only 37 percent expressed strong support for evolution—that is, teaching it to the exclusion of all religious doctrine in science classes.

In the absence of a majority favoring strict standards for evolution teaching, it is easy to see why creationists have been able to make headway even outside the circle of evangelical Christianity. In 1996 Pope John Paul II reaffirmed the Catholic Church's commitment to evolution, first stated in 1950, saying that his inspiration for doing so came from the Bible. Despite this, 40 percent of American Catholics in a 2001 Gallup poll said they believed that God created human life in the past 10,000 years. Indeed, fully 45 percent of all Americans subscribe to this creationist view. Many who are indifferent to conservative theology give creationism some support, perhaps because, as mathematician Norman Levitt of Rutgers University suggests, the subject of evolution provokes anxiety about the nature of human existence, an anxiety that antievolutionists use to promote creationist ideas.



Rating of Evolution Treatment in State Public School Science Standards

Unsatisfactory, useless or absent | Satisfactory/ good | Very good/ excellent | No state standard

Rodger Doyle's e-mail is rdoyle2@delphi.com

RODGER DOYLE

Defying Gravity

A small Swiss firm develops an innovative G suit for fighter pilots BY MICHAEL BEHAR

Col. Hank Morrow, commander of the 149th Fighter Wing of the Texas Air National Guard, has been flying for more than two decades. In that time, he has seen aircraft push the high-performance envelope: planes today are so fast and nimble that standard evasive maneuvers can add nine times the weight of gravity, or nine g's, to the mass of a pilot's body. That amount of force causes fatigue, blackouts, even death as gravity drives blood and oxygen from the brain, lungs and heart. G suits are supposed to protect pilots by filling with compressed air and squeezing the lower extremities to shove bodily fluids upward. Yet G-suit technology has stagnated for almost half a century, while rapid innovations in aircraft design have put many pilots at the mercy of their machines. All that could change if the air force chooses to outfit its aviators with a revolutionary liquid-filled G suit called the Libelle.

PILOTS sing the praises of the Libelle G suit.

The suit is the brainchild of Andreas Reinhard, a former Swiss Air Force fighter pilot turned inventor and founder of Life Support Systems, a company he launched in 1996 to develop the Libelle. Instead of using air, the Libelle forms a liquid barrier around the pilot, much like a baby is protected in the womb. Morrow recently tested the suit at Edwards Air Force Base in California and was so ecstatic with the results that he told the members of the Libelle team he would write them a personal check on the spot if they would sell him one.

Reinhard says he first got the idea for the Libelle—the German word for "dragonfly"—in 1987, when he was still in the Swiss Air Force. He was inspired by the dragonfly because it is the only animal that can withstand 30 g's of force, because its cardiac system is encased in liquid. "After a dogfight training session, I was extremely exhausted," he recalls. "I imagined filling the whole cockpit with a fluid that had the same viscosity and density of blood."

In crafting the Libelle, Reinhard revived a concept developed in the 1940s, when antigravity suits first appeared. The first suit was developed in Canada by Wilbur Franks of the University of Toronto. Franks found that when he suspended glass test tubes in water, they didn't break in the centrifuge. He applied this observation to a crude prototype suit by sandwiching a layer of water between two rubber panels. Later he devised a workable suit with air bladders; his basic design is more or less identical to what pilots currently use.

In seeking to improve today's suits, one of the first challenges for Reinhard and his engineers was to find a liquid that could absorb g forces but that was nontoxic and nonflammable. After making several prototype suits, including one filled with silicone ("like what's in breast implants," Reinhard notes), the team settled on distilled water spiked with a secret "special material" that prevents the Libelle from freezing should the pilot eject at high altitudes. The liquid—housed in two-inch-wide channels that run the length of the arms, legs and torso—is harmless enough to drink, even serving as an emergency ration for a downed pilot, Reinhard says.

Another task was finding a fabric that could dynamically respond to sudden changes in gravity. "We had to cover the whole body with a material that wouldn't stretch under pressure," Reinhard remarks. "At the same time, the suit had to be flexible so the

COURTESY OF LIFE SUPPORT SYSTEMS

pilots could move." After unsuccessfully searching fabric mills, the Libelle's engineers decided to make their own material. They devised a hybrid weave, blending Du Pont's flame-resistant Nomex fabric with tough Kevlar aramid fibers, that is "rigid on the horizontal axis but flexible vertically."

As g forces intensify during a hypersonic turn or downward spiral, the Libelle's liquid tubes compress, pulling with them the surrounding fabric. Imagine a self-contained hydrostatic vise in which water progressively squeezes the pilot as he hits the afterburner. A conventional G suit takes a few seconds to respond, because air must be pumped into various bladders from the plane's onboard pneumatic system. "With the Libelle, you didn't feel it working," says Lt. Col. Christian Ledet, a senior flight surgeon with the Iowa Air National Guard who also tested the suit at Edwards. "It was just kind of there, responding to the laws of physics, doing its job before you even knew it."

The Libelle suit is the first of its kind to reach production and has proved a worthy competitor to the most advanced air-filled suits, such as the U.S. Air Force's Combat Edge system. During several test flights at Edwards, one pilot wore the Libelle, while another, seated in the rear cockpit, donned the Combat Edge. "We went up to 18,000 feet, hit the afterburner and started 9-g spirals until the guy in the backseat said uncle," Morrow remembers. "It was easier to breathe and easier to communicate" in the Libelle suit, Ledet confirms. "When I came back, I didn't feel like a wrung-out wet rag. With a regular G suit I can't even get out of the cockpit."

"When I came back, I didn't feel like a wrung-out wet rag. With a regular G suit I can't even get out of the cockpit."

Two years after starting the Swiss-based Life Support Systems, Reinhard formed an alliance with Autoflug, a German producer of aircraft rescue and safety equipment, to help handle the final development and marketing of the Libelle. Besides the U.S. tests, a number of test pilots in the German and Netherlands air forces have flown in the suit, trials that proved a success. The company's aim now is to convince Britain, Germany, Italy and Spain to put the Libelle on the Eurofighter, a next-generation aircraft that the

four countries are building to eventually replace the Mig-29, the Phantom and the Tornado.

In October the U.S. Department of Defense chose the Libelle for its foreign comparative testing (FCT) program. Reinhard contends that "the Pentagon process is too slow." But in this case, patience could pay off: the yearlong evaluation will begin this spring. Assuming it passes with high marks, the "Libelle could be recommended for an air force–wide buy," says Maj. John Ryan, program manager at the air force FCT office. There are no guarantees, but air force procurement of the Libelle would most likely make Reinhard, Autoflug and the Libelle's private investors a bundle of cash.

Although first-round trials at Edwards and in Europe were promising, Ulf Balldin, a senior scientist at the Wyle Laboratories unit in Houston, Tex., and president of the International Academy of Aviation and Space Medicine, feels that medical evidence establishing the Libelle's outright superiority is lacking. "I have been working with G suits for many years, and as far as I'm concerned it has not been tested and proven properly," he says. But Reinhard does not agree: "After a few hundred centrifuge rides and hundreds of test flights with more than 80 different subjects, we believe we are ready to go to market."

Any doubts about the Libelle should be sorted out by the rigorous FCT program, when pilots, physicians and engineers will poke and prod the suit to ensure that it is the best technology available. "We'll do sortie after sortie to see where the suit has merit, where it's better and where it's not as good," explains Col. Peter Demitry, chief of the U.S. Air Combat Command's Human Systems Integration Division.

In the meantime, Reinhard and teams at Life Support Systems and Autoflug are tweaking the Libelle design to satisfy the demanding specs mandated by potential customers (U.S. pilots want more pockets) while continuing to make their pitch to the Eurofighter contingent. As for a windfall sale to the Pentagon, Reinhard and company will have to wait until at least 2003. But if the exalted testimony from test pilots is any measure, it's hard to believe that the Libelle won't come out on top. SA

Michael Behar is a Washington, D.C.–based science and technology journalist and a former senior editor at Wired *magazine.*

Who Owns You?

A mock trial explores the intersection of patents and genetic-property rights

By GARY STIX

A man named Salvador Dolly gives blood for a routine genetic test to determine his fitness to father a child. The testing company, Advanced Genetic Testing Company (AGTC), then sells the remains of the sample to NuGenEra, a biotechnology company. NuGenEra discovers that Dolly's genes make him resistant to HIV.

The company responds to this discovery by taking out a patent on both Dolly's genome and a series of gene sequences that confer resistance. When NuGenEra informs Dolly that his genes guard against the deadly virus, he decides to set up a business to market his blood to research institutions. To protect its patent, NuGenEra sues Dolly for patent infringement, saying that it owns his genome.

JOHN McFAUL

Does the patent mean that Dolly must forgo any rights to his own genome? Does it violate his privacy or property rights? Should these rights be balanced against society's need for the tests and therapies for HIV that might be derived from NuGenEra's research on Dolly's genome? These issues were highlighted last November in a mock trial at the California Institute of Technology as part of the school's Program for Law and Technology, in collaboration with Loyola Law School.

During arguments made by students from both schools, Judge Marilyn Hall Patel, who presided over the Napster copyright case, had to decide whether to invalidate the NuGenEra patent and throw out the company's suit against Dolly for violating the patent on his own genes. Many of the arguments centered on the usefulness of Dolly's genes—utility being one of the principal criteria for granting a patent. In its patent, NuGenEra claimed that both Dolly's entire genome and 10 genes within it, called the P sequences, could be employed to create diagnostic tests for determining resistance to HIV and to produce gene therapies to cure the disease.

Dolly's attorneys argued that the genome—and even the P sequences—consisted of DNA for which the specific genes that conferred resistance had not yet been identified, a lack of utility that meant the patent should be declared invalid. They also contended that the patent violated Dolly's rights to privacy, property and personal autonomy.

In her decision, Patel allowed the mock case to move forward to a jury trial (see http://techlaw.lls.edu/atc3/order.pdf). In doing so, she affirmed that the P sequences had a legitimate use as a diagnostic tool to ascertain HIV resistance. But she invalidated the part of NuGenEra's patent that covered Dolly's whole genome because of a lack of any clear-cut applications.

Acknowledging an aversion to judge-made law, Patel would not embrace privacy or other public policy arguments made by Dolly's attorneys, citing the absence of legislation and case law to guide her. But she did seem inclined to find some means of suggesting protection for genetic property within the bounds of existing law. The judge noted that genetic material is unique to each individual. Thus, Dolly may have the right to sue in California for misuse of his likeness for commercial purposes.

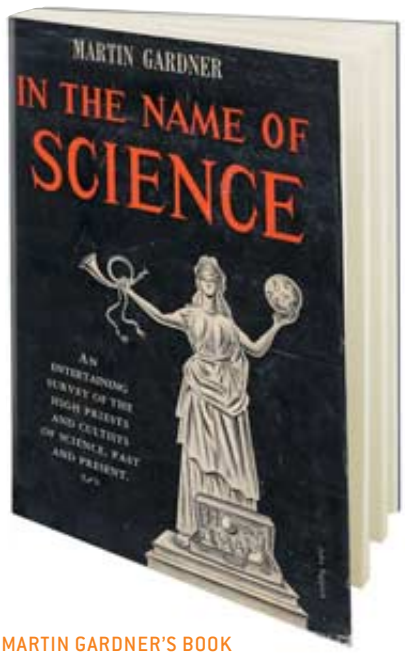
The case illustrates how the genomics era may affect existing patent law. "I think that if this were a real opinion and it carried weight, it would mean that the patent laws are going to be aggressively pursued irrespective of these countervailing social policy issues," says Karl Manheim, who directs the law and technology program at Loyola. So if *NuGenEra v. Salvador Dolly* is any portent, whatever part of one's self that is locked up in the genetic code may be eligible to be owned and bottled by someone else. SA

Please let us know about interesting and unusual patents. Send suggestions to: patents@sciam.com

Hermits and Cranks

Fifty years ago Martin Gardner launched the modern skeptical movement. Unfortunately, much of what he wrote about is still current today By MICHAEL SHERMER

In 1950 Martin Gardner published an article in the *Antioch Review* entitled "The Hermit Scientist," about what we would today call pseudoscientists. It was Gardner's first publication of a skeptical nature (he was the math games columnist for *Scientific American* for more than a quarter of a century). In 1952 he expanded it into a book called *In the Name of Science,* with the descriptive subtitle "An entertaining survey of the high priests and cultists of science, past and present." Published by Putnam, the book sold so poorly that it was quickly remaindered and lay dormant until 1957, when it was republished by Dover. It has come down to us as *Fads and Fallacies in the Name of Science,* which is still in print and is arguably *the* skeptic classic of the past half a century.



MARTIN GARDNER'S BOOK
In the Name of Science is the bible of the modern skeptical movement.

Thankfully, there has been some progress since Gardner offered his first criticisms of pseudoscience. Now largely antiquated are his chapters on believers in a flat earth, a hollow earth, Atlantis and Lemuria, Alfred William Lawson, Roger Babson, Trofim Lysenko, Wilhelm Reich and Alfred Korzybski. But disturbingly, a good two thirds of the book's contents are relevant today, including Gardner's discussions of homeopathy, naturopathy, osteopathy, iridiagnosis (reading the iris of the eye to determine bodily malfunctions), food faddists, cancer cures and other forms of medical quackery, Edgar Cayce, the Great Pyramid's alleged mystical powers, handwriting analysis, ESP and PK (psychokinesis), reincarnation, dowsing rods, eccentric sexual theories, and theories of group racial differences.

The "hermit scientist," a youthful Gardner wrote, works alone and is ignored by mainstream scientists. "Such neglect, of course, only strengthens the convictions of the self-declared genius." But Gardner was wrong by half in his prognostications: "The current flurry of discussion about Velikovsky and Hubbard will soon subside, and their books will begin to gather dust on library shelves." Adherents to Immanuel Velikovsky's views on how celestially caused global catastrophes shaped the beliefs of ancient humans are a quaint few surviving in the interstices of fringe culture. L. Ron Hubbard, however, has been canonized by the Church of Scientology as the founding saint of a world religion.

In 1952 Gardner could not have known that the nascent flying saucer craze would turn into an alien industry: "Since flying saucers were first reported in 1947, countless individuals have been convinced that the earth is under observation by visitors from another planet." Absence of evidence then was no more a barrier to belief than it is today, and ufologists proffered the same conspiratorial explanations for the dearth of proof: "I have heard many readers of the saucer books upbraid the government in no uncertain terms for its stubborn refusal to release the 'truth' about the elusive platters. The administration's 'hush hush policy' is angrily cited as proof that our military and political leaders have lost all faith in the wisdom of the American people."

Even then Gardner was bemoaning that some beliefs never seem to go out of vogue, as he recalled an H. L. Mencken quip from the 1920s: "Heave an egg out of a Pullman window, and you will hit a Fundamentalist almost anywhere in the U.S. today." Gardner cautions that when religious superstition should be on the wane, it is easy "to forget that thousands of

BRAD HINES

high school teachers of biology, in many of our southern states, are still afraid to teach the theory of evolution for fear of losing their jobs." Today creationism has spread northward and mutated into the oxymoronic form of "creation science."

And the motives of the hermit scientists have not changed either. Gardner recounts the day that Groucho Marx interviewed Louisiana state senator Dudley J. LeBlanc about a "miracle" cure-all vitamin-and-mineral tonic called Hadacol that the senator had invented. When Groucho asked the senator what it was good for, LeBlanc answered with surprising honesty: "It was good for five and a half million for me last year."

What I find especially valuable about Gardner's views are his insights into the differences between science and pseudoscience. On the one extreme we have ideas that are most certainly false, "such as the dianetic view that a one-day-old embryo can make sound recordings of its mother's conversation." In the borderlands between the two "are theories advanced as working hypotheses, but highly debatable because of the lack of sufficient data." Of these Gardner selects a most propitious example: "the theory that the universe is expanding." That theory would now fall at the other extreme end of the spectrum, where lie "theories almost certainly true, such as the belief that the earth is round or that men and beasts are distant cousins."

Even then Gardner was bemoaning that some beliefs never seem to go out of vogue, as he recalled H. L. Mencken's quip: "Heave an egg out of a Pullman window, and you will hit a Fundamentalist almost anywhere in the U.S. today."

How can we tell if someone is a scientific crank? Gardner offers this advice: (1) "First and most important of these traits is that cranks work in almost total isolation from their colleagues." Cranks typically do not understand how the scientific process operates—that they need to try out their ideas on colleagues, attend conferences and publish their hypotheses in peer-reviewed journals before announcing to the world their startling discovery. Of course, when you explain this to them they say that their ideas are too radical for the conservative scientific establishment to accept. (2) "A second characteristic of the pseudo-scientist, which greatly strengthens his isolation, is a tendency toward paranoia," which manifests itself in several ways:

> (1) He considers himself a genius. (2) He regards his colleagues, without exception, as ignorant blockheads.... (3) He believes himself unjustly persecuted and discriminated against. The recognized societies refuse to let him lecture. The journals reject his papers and either ignore his books or assign them to "enemies" for review. It is all part of a dastardly plot. It never occurs to the crank that this opposition may be due to error in his work.... (4) He has strong compulsions to focus his attacks on the greatest scientists and the best-established theories. When Newton was the outstanding name in physics, eccentric works in that science were violently anti-Newton. Today, with Einstein the father-symbol of authority, a crank theory of physics is likely to attack Einstein.... (5) He often has a tendency to write in a complex jargon, in many cases making use of terms and phrases he himself has coined.

We should keep these criteria in mind when we explore controversial ideas on the borderlands of science. "If the present trend continues," Gardner concludes, "we can expect a wide variety of these men, with theories yet unimaginable, to put in their appearance in the years immediately ahead. They will write impressive books, give inspiring lectures, organize exciting cults. They may achieve a following of one—or one million. In any case, it will be well for ourselves and for society if we are on our guard against them." So we still are, Martin. That is what skeptics do, and in tribute for all you have done, we shall continue to honor your founding command. SA

Michael Shermer is founding publisher of Skeptic *magazine (www.skeptic.com) and author of* How We Believe *and* The Borderlands of Science.

Aspirations in Science and Civics

From the carbon-nanotube lab to the corridors of Washington power, Mildred S. Dresselhaus has followed a career that combines scientific research with public service By DAVID APPELL

Standing in a well-worn hallway of Building 13 of the Massachusetts Institute of Technology, Mildred S. Dresselhaus is quietly but firmly directing the show. She answers questions from a member of her lab group and, in the next sentence, asks another if he is free to pick up a visitor at the airport that afternoon.

KATHLEEN DOOHER

MILDRED S. DRESSELHAUS: INDEFATIGABLE

- Married to M.I.T. physicist Gene F. Dresselhaus; four children; son Peter is a physicist at the National Institute of Standards and Technology.
- Avid violinist and violist; member of Amateur Chamber Music Society in New York City.
- National Medal of Science, 1990; president, American Association for the Advancement of Science, 1997–1998; 17 honorary doctorates.

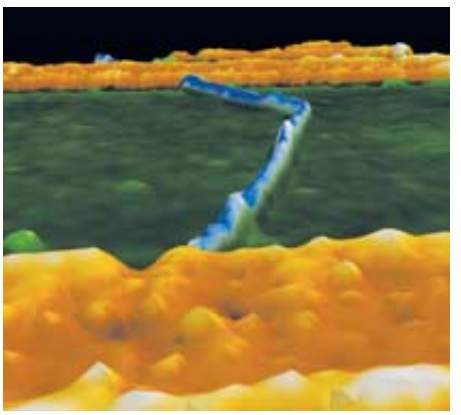
Next, she moves on to me and says, correctly, "You look like you're looking for me." We find an empty conference room "away from the phone," and immediately I have her full attention. It is, I learn later from her friends and colleagues, "typical Millie." "She has these fantastic personal skills and inexhaustible energy," says M.I.T. colleague physicist Daniel Kleppner. "She manages to do two or three things at once and do them well. She's never sitting idle."

Indeed, the scope of Dresselhaus's career is imposingly impressive: a leader in carbon research for 40 years; author or co-author of nearly 1,000 scientific papers, articles and reviews; adviser to more than 60 doctoral students; national officeholder in several professional science associations; and past director in the U.S. Department of Energy's Office of Science, one of the largest funders of basic research. To top it off, she has 17 honorary doctorates and was awarded the National Medal of Science in 1990.

All this, begun at a time when professional currents pushed much harder against her gender than they do today. "We didn't think we had a career in physics," she says of women physicists of her generation. "We were just doing it because we were interested and hoping we could do some kind of research." But she got lucky and found herself under the tutelage of past and future Nobel Prize winners. In 1951 she graduated with a physics degree from Hunter College, where one of her mentors was medical physicist Rosalyn S. Yalow. After a year at the University of Cambridge on a Fulbright fellowship, she received her master's degree from Harvard University, where her adviser was physicist Norman F. Ramsey. She completed her Ph.D. in solid-state physics at the University of Chicago, where she took classes from Enrico Fermi. "I learned a lot from him about teaching methods, how it's important to get things simple," she recalls. "If you can't explain it simply, Fermi wasn't really that interested in it."

At Chicago in 1956, she met her husband, solid-state physicist Gene F. Dresselhaus. Four years later the Dresselhauses had found themselves posts at M.I.T.'s Lincoln Labs. It was a great place to work, she recounts, "but it had only one drawback—they expected me to start the job at eight o'clock in the morning," which wasn't always possible once they'd had four children. After many reprimands, she found refuge in a visiting professorship in the electrical department, which turned her post into a tenured position a year later. In 1985 M.I.T. granted her an Institute Professorship, its highest academic honor; only about a dozen such posts exist.

Dresselhaus has made her name in the field of carbon physics. Early in her career she unveiled the mysteries of the way electrons are organized in graphite (technically, its electronic band structure). She discovered that the identification of electrons and holes (positive charges created by the absence of electrons) in the material had been interchanged, thus answering a variety of previously open questions. She was among the first to utilize lasers for magneto-optics experiments, and she was a pioneer in determining how certain semimetals, including vanadium and niobium, transport heat and electricity. Dresselhaus also did founding work in intercalation physics, the determination of the properties of materials that are interlaced with other materials, such as graphite with alkali metal layers.



CARBON NANOTUBES are promising in part because they can have various electrical properties, Dresselhaus found. Here a tube (*blue*) acts as a nanometer-wide wire connecting two electrodes (*yellow*).

COURTESY OF C. DEKKER *TU Delft*, © Nature, *1999*

Over the past decade Dresselhaus has focused on the burgeoning field of carbon nanotubes, which in the last year has seen a doubling of published papers, to an annual output of more than 1,500. In 1992 she predicted with her husband and Riichiro Saito from the University of Electro-Communications in Tokyo that carbon nanotubes could be either semiconducting or metallic depending on their geometric characteristics—an extraordinary hypothesis borne out by experiment in 1998.

"The science is booming right now," Dresselhaus says, and the potential applications for it are diverse, including flat-panel displays, hydrogen energy storage, building construction and drug delivery. Nanotubes hold great promise as well for electronics, portending a drastic shrinkage of wires and electronic devices, with a concomitant increase in speed [see "Nanotubes for Electronics," by Philip G. Collins and Phaedon Avouris, SCIENTIFIC AMERICAN; December 2000].

Even as she was helping to lay the foundation for a nanotube future, Dresselhaus kept in mind the service side of science. "My own undergraduate education at Hunter College cost me $5 per semester, which covered tuition, laboratory charges and textbooks on loan," she noted in 1997 while president of the American Association for the Advancement of Science (AAAS). "The taxpayers invested in me." Her decision to tackle the administration of science was prompted, she elaborates now, by her election to the National Academy of Engineering in 1974. She has been president of the American Physical Society and treasurer of the National Academy of Sciences, just two of the many posts from which she has been able to give back to taxpayers, to her discipline and to science at large. Name a committee, and Dresselhaus has been on it—and was probably its chair.

Her last public appointment began in August 2000 and effectively ended in November, with the presidential election. Energy Secretary Bill Richardson had asked her to direct the Department of Energy's Office of Science, overseeing a $3.1-billion budget (placing it among the largest federal supporters of science) and five national laboratories. Dresselhaus agreed, even though it was at the tail end of the Clinton administration. "Then came a confirmation process that was very, very discouraging," she says: it was drawn out over 10 months while legislators played political football. She can't pinpoint the reasons for the delay. "It had nothing to do with me, nothing to do with my post, nothing to do with science. It's the way politics is done," she observes.

Late in his term President Bill Clinton became particularly interested in science, she says: "He saw the connection between science and the economy." But Dresselhaus is disappointed in the present administration's attitude. "The president personally has not much interest in it, and he doesn't have people around him with much interest." The Bush administration did not even respond to her offer to stay on at the DOE until the next appointment was made, and ultimately she simply packed up and left. (In December, President George W. Bush announced his intention to nominate Raymond L. Orbach, now chancellor of the University of California at Riverside.)

Noted ecologist Peter H. Raven calls Dresselhaus a "foremost role model." Director of the Missouri Botanical Garden and current president of the AAAS, Raven was secretary of the National Academy when Dresselhaus was treasurer. He praises her for being firmly anchored in family life while forging ahead with contributions as a scientist. In a scientific life that has been remarkably broad in scope, Mildred Dresselhaus has not forgotten to go deep as well. SA

David Appell is a science writer based in Gilford, N.H.

The Worldwide Computer

By David P. Anderson
and John Kubiatowicz

An operating system spanning the Internet would bring the power of millions of the world's Internet-connected PCs to everyone's fingertips

When Mary gets home

from work and goes to her PC to check e-mail, the PC isn't just sitting there. It's working for a biotech company, matching gene sequences to a library of protein molecules. Its DSL connection is busy downloading a block of radio telescope data to be analyzed later. Its disk contains, in addition to Mary's own files, encrypted fragments of thousands of other files. Occasionally one of these fragments is read and transmitted; it's part of a movie that someone is watching in Helsinki. Then Mary moves the mouse, and this activity abruptly stops. Now the PC and its network connection are all hers.

This sharing of resources doesn't stop at her desktop computer. The laptop computer in her satchel is turned off, but its disk is filled with bits and pieces of other people's files, as part of a distributed backup system. Mary's critical files are backed up in the same way, saved on dozens of disks around the world.

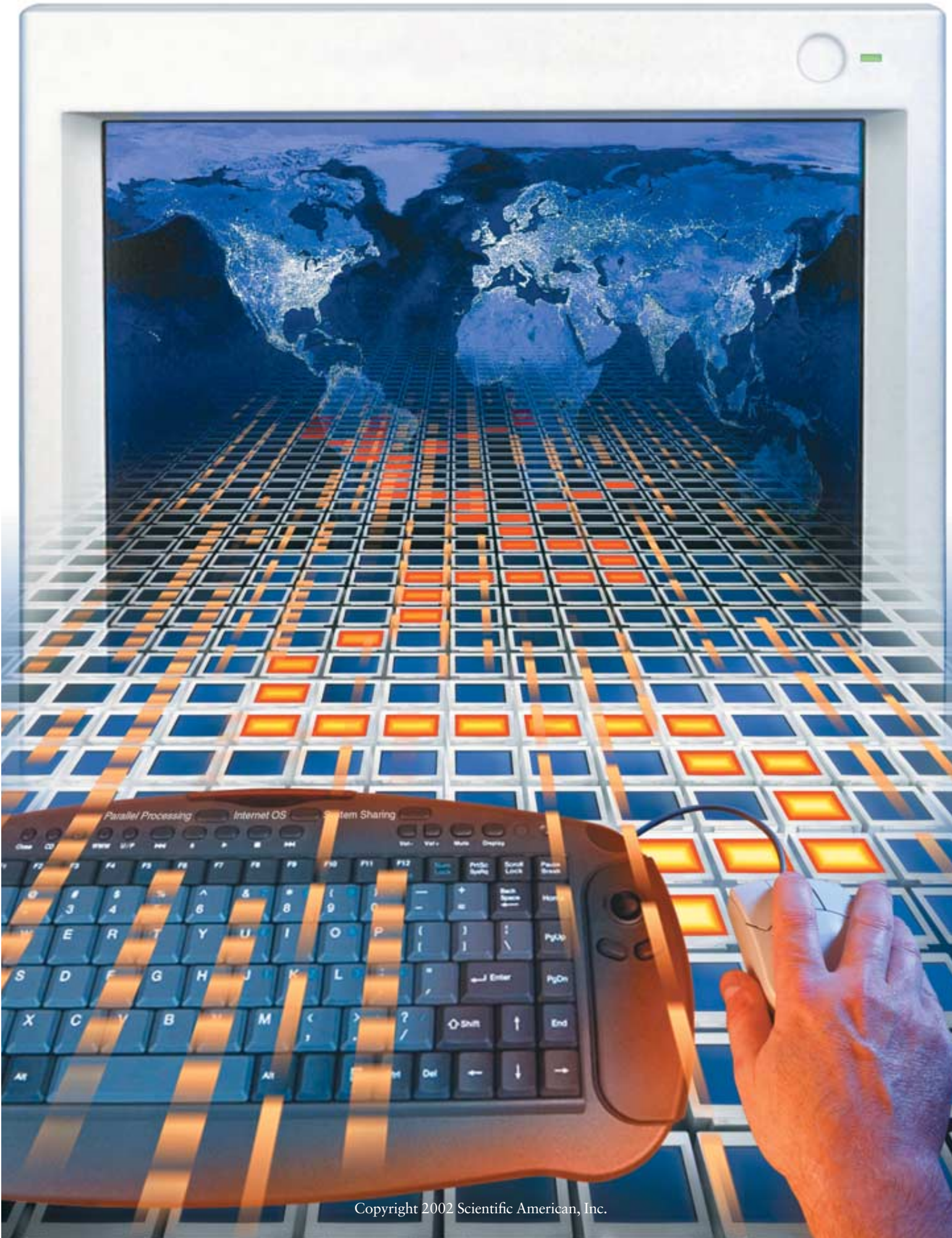
Later, Mary watches an independent film on her Internet-connected digital television, using a pay-per-view system. The movie is assembled on the fly from fragments on several hundred computers belonging to people like her.

Mary's computers are moonlighting for other people. But they're not giving anything away for free. As her PC works, pennies trickle into her virtual bank account. The payments come from the biotech company, the movie system and the backup service. Instead of buying expensive "server farms," these companies are renting time and space, not just on Mary's two computers but on millions of others as well. It's a win-win situation. The companies save money on hardware, which enables, for instance, the movie-viewing service to offer obscure movies. Mary earns a little cash, her files are backed up, and she gets to watch an indie film. All this could happen with an Internet-scale operating system (ISOS) to provide the necessary "glue" to link the processing and storage capabilities of millions of independent computers.

Internet-Scale Applications

ALTHOUGH MARY'S WORLD is fictional—and an Internet-scale operating system does not yet exist—developers have already produced a number of Internet-scale, or peer-to-peer, applications that attempt to tap the vast array of underutilized machines available through the Internet [*see box on page 42*].

PHILIP HOWE



These applications accomplish goals that would be difficult, unaffordable or impossible to attain using dedicated computers. Further, today's systems are just the beginning: we can easily conceive of archival services that could be relied on for hundreds of years and intelligent search engines for tomorrow's Semantic Web [see "The Semantic Web," by Tim Berners-Lee, James Hendler and Ora Lassila; SCIENTIFIC AMERICAN, May 2001].

Unfortunately, the creation of Internet-scale applications remains an imposing challenge. Developers must build each new application from the ground up, with much effort spent on technical matters, such as maintaining a database of users, that have little to do with the application itself. If Internet-scale applications are to become mainstream, these infrastructure issues must be dealt with once and for all.

We can gain inspiration for eliminating this duplicate effort from operating systems such as Unix and Microsoft Windows. An operating system provides a virtual computing environment in which programs operate as if they were in sole possession of the computer. It shields programmers from the painful details of memory and disk allocation, communication protocols, scheduling of myriad processes, and interfaces to devices for data input and output. An operating system greatly simplifies the development of new computer programs. Similarly, an Internet-scale operating system would simplify the development of new distributed applications.

More than 150 MILLION hosts are connected to the Internet, and the number is GROWING exponentially.

Existing Distributed Systems

COMPUTING

GIMPS (Great Internet Mersenne Prime Search): www.mersenne.org/
Searches for large prime numbers. About 130,000 people are signed up, and five new primes have been found, including the largest prime known, which has four million digits.

distributed.net: www.distributed.net/
Has decrypted several messages by using brute-force searches through the space of possible encryption keys. More than 100 billion keys are tried each second on its current decryption project. Also searches for sets of numbers called optimal Golomb rulers, which have applications in coding and communications.

SETI@home (Search for Extraterrestrial Intelligence): http://setiathome.berkeley.edu/
Analyzes radio telescope data, searching for signals of extraterrestrial origin. A total of 3.4 million users have devoted more than 800,000 years of processor time to the task.

folding@home: http://folding.stanford.edu/
Run by Vijay Pande's group in the chemistry department at Stanford University, this project has about 20,000 computers performing molecular-dynamics simulations of how proteins fold, including the folding of Alzheimer amyloid-beta protein.

Intel/United Devices cancer research project: http://members.ud.com/projects/cancer/
Searches for possible cancer drugs by testing which of 3.5 billion molecules are best shaped to bind to any one of eight proteins that cancers need to grow.

STORAGE

Napster: www.napster.com/
Allowed users to share digital music. A central database stored the locations of all files, but data were transferred directly between user systems. Songwriters and music publishers brought a class-action lawsuit against Napster. The parties reached an agreement whereby rights to the music would be licensed to Napster and artists would be paid, but the new fee-based service had not started as of January 2002.

Gnutella: www.gnutella.com/
Provides a private, secure shared file system. There is no central server; instead a request for a file is passed from each computer to all its neighbors.

Freenet: http://freenetproject.org/
Offers a similar service to Gnutella but uses a better file-location protocol. Designed to keep file requesters and suppliers anonymous and to make it difficult for a host owner to determine or be held responsible for the Freenet files stored on his computer.

Mojo Nation: www.mojonation.net/
Also similar to Gnutella, but files are broken into small pieces that are stored on different computers to improve the rate at which data can be uploaded to the network. A virtual payment system encourages users to provide resources.

Fasttrack P2P Stack: www.fasttrack.nu/
A peer-to-peer system in which more powerful computers become search hubs as needed. This software underlies the Grokster, MusicCity ("Morpheus") and KaZaA file-sharing services.

An ISOS consists of a thin layer of software (an ISOS agent) that runs on each "host" computer (such as Mary's) and a central coordinating system that runs on one or more ISOS server complexes. This veneer of software would provide only the core functions of allocating and scheduling resources for each task, handling communication among host computers and determining the reimbursement required for each machine. This type of operating system, called a microkernel, relegates higher-level functions to programs that make use of the operating system but are not a part of it. For instance, Mary would not use the ISOS directly to save her files as pieces distributed across the Internet. She might run a backup application that used ISOS functions to do that for her. The ISOS would use principles borrowed from economics to apportion computing resources to different users efficiently and fairly and to compensate the owners of the resources.

Two broad types of applications might benefit from an ISOS. The first is distributed data processing, such as physical simulations, radio signal analysis, genetic analysis, computer graphics rendering and financial modeling. The second is distributed online services, such as file storage systems, databases, hosting of Web sites, streaming media (such as online video) and advanced Web search engines.

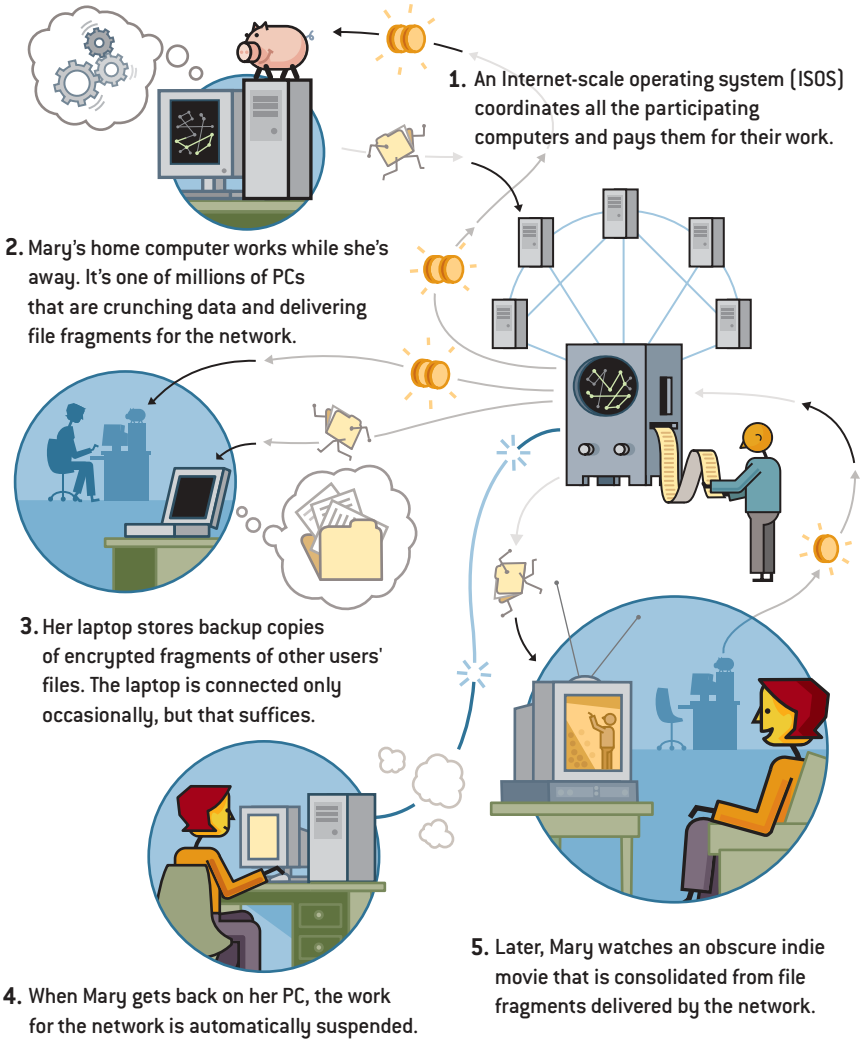
What's Mine Is Yours

COMPUTING TODAY operates predominantly as a private resource; organizations and individuals own the systems that they use. An ISOS would facilitate a new paradigm in which it would be routine to make use of resources all across the Internet. The resource pool—hosts able to compute or store data and networks able to transfer data between hosts—would still be individually owned, but they could work for anyone. Hosts would include desktops, laptops, server computers, network-attached storage devices and maybe handheld devices.

The Internet resource pool differs from private resource pools in several important ways. More than 150 million hosts are connected to the Internet, and the number is growing exponentially. Consequently, an ISOS could provide a virtual computer with potentially 150 million times the processing speed and storage capacity of a typical single computer. Even when this virtual computer is divided up among many users, and after one allows for the overhead of running the network, the result is a bigger, faster and cheaper computer than the users could own privately. Continual upgrading of the resource pool's hardware causes the total speed and capacity of this über-computer to increase even faster than the number of connected hosts. Also, the pool is self-maintaining: when

MOONLIGHTING COMPUTERS

With Internet-scale applications, PCs around the world can work during times when they would otherwise sit idle. Here's how it works:

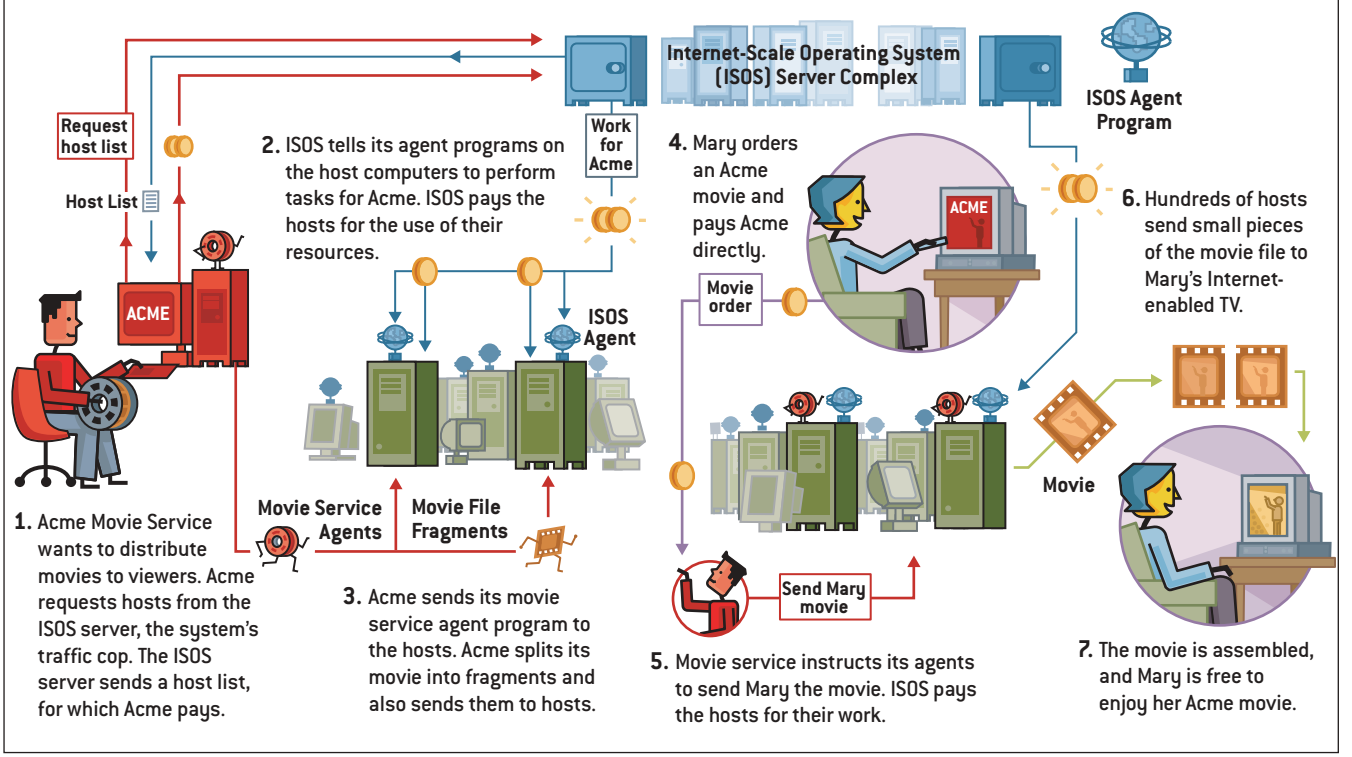


THE AUTHORS

DAVID P. ANDERSON and *JOHN KUBIATOWICZ* are both associated with the University of California, Berkeley. Anderson was on the faculty of the computer science department from 1985 to 1991. He is now director of the SETI@home project and chief science officer of United Devices, a provider of distributed computing software that is allied with the distributed.net project. Kubiatowicz is an assistant professor of computer science at Berkeley and is chief architect of OceanStore, a distributed storage system under development with many of the properties required for an ISOS.

XPLANE

HOW A DISTRIBUTED SERVICE WOULD OPERATE



a computer breaks down, its owner eventually fixes or replaces it.

Extraordinary parallel data transmission is possible with the Internet resource pool. Consider Mary's movie, being uploaded in fragments from perhaps 200 hosts. Each host may be a PC connected to the Internet by an antiquated 56k modem—far too slow to show a high-quality video—but combined they could deliver 10 megabits a second, better than a cable modem. Data stored in a distributed system are available from any location (with appropriate security safeguards) and can survive disasters that knock out sections of the resource pool. Great security is also possible, with systems that could not be compromised without breaking into, say, 10,000 computers.

In this way, the Internet-resource paradigm can increase the bounds of what is possible (such as higher speeds or larger data sets) for some applications, whereas for others it can lower the cost. For certain applications it may do neither—it's a paradigm, not a panacea. And designing an ISOS also presents a number of obstacles.

Some characteristics of the resource pool create difficulties that an ISOS must deal with. The resource pool is heterogeneous: Hosts have different processor types and operating systems. They have varying amounts of memory and disk space and a wide range of Internet connection speeds. Some hosts are behind firewalls or other similar layers of software that prohibit or hinder incoming connections. Many hosts in the pool are available only sporadically; desktop PCs are turned off at night, and laptops and systems using modems are frequently not connected. Hosts disappear unpredictably—sometimes permanently—and new hosts appear.

The ISOS must also take care not to antagonize the owners of hosts. It must have a minimal impact on the non-ISOS uses of the hosts, and it must respect limitations that owners may impose, such as allowing a host to be used only at night or only for specific types of applications. Yet the ISOS cannot trust every host to play by the rules in return for its own good behavior. Owners can inspect and modify the activities of their hosts. Curious and malicious users may attempt to disrupt, cheat or spoof the system. All these problems have a major influence on the design of an ISOS.

Who Gets What?

AN INTERNET-SCALE operating system must address two fundamental issues—how to allocate resources and how to compensate resource suppliers. A model based on economic principles in which suppliers lease resources to consumers can deal with both issues at once. In the 1980s researchers at Xerox PARC proposed and analyzed economic approaches to apportioning computer resources. More recently, Mojo Nation developed a file-sharing system in which users are paid in a virtual currency ("mojo") for use of their resources and they in turn must pay mojo to use the system. Such economic models encourage owners to allow their resources to be used by other organizations, and theory shows that they lead to optimal allocation of resources.

Even with 150 million hosts at its disposal, the ISOS will be dealing in "scarce"

XPLANE

resources, because some tasks will request and be capable of using essentially unlimited resources. As it constantly decides where to run data-processing jobs and how to allocate storage space, the ISOS must try to perform tasks as cheaply as possible. It must also be fair, not allowing one task to run efficiently at the expense of another. Making these criteria precise—and devising scheduling algorithms to achieve them, even approximately—are areas of active research.

The economic system for a shared network must define the basic units of a resource, such as the use of a megabyte of disk space for a day, and assign values that take into account properties such as the rate, or bandwidth, at which the storage can be accessed and how frequently it is available to the network. The system must also define how resources are bought and sold (whether they are paid for in advance, for instance) and how prices are determined (by auction or by a price-setting middleman).

Within this framework, the ISOS must accurately and securely keep track of resource usage. The ISOS would have an internal bank with accounts for suppliers and consumers that it must credit or debit according to resource usage. Participants can convert between ISOS currency and real money. The ISOS must also ensure that any guarantees of resource availability can be met: Mary doesn't want her movie to grind to a halt partway through. The economic system lets resource suppliers control how their resources are used. For example, a PC owner might specify that her computer's processor can't be used between 9 A.M. and 5 P.M. unless a very high price is paid.

Money, of course, encourages fraud, and ISOS participants have many ways to try to defraud one another. For instance, resource sellers, by modifying or fooling the ISOS agent program running on their computer, may return fictitious results without doing any computation. Researchers have explored statistical methods for detecting malicious or malfunctioning hosts. A recent idea for preventing unearned computation credit is to ensure that each work unit has a number of intermediate results that the server can quickly check and that can be obtained only by performing the entire computation. Other approaches are needed to prevent fraud in data storage and service provision.

The cost of ISOS resources to end users will converge to a fraction of the cost of owning the hardware. Ideally, this fraction will be large enough to encourage owners to participate and small enough to make many Internet-scale applications economically feasible. A typical PC owner might see the system as a barter economy in which he gets free services, such as file backup and Web hosting, in exchange for the use of his otherwise idle processor time and disk space.

A Basic Architecture

WE ADVOCATE two basic principles in our ISOS design: a minimal core operating system and control by central servers.

Curious and malicious USERS may attempt to DISRUPT, CHEAT or spoof the system.

Primes and Crimes

By Graham P. Collins

NO ONE HAS SEEN signs of extraterrestrials using a distributed computation project (yet), but people have found the largest-known prime numbers, five-figure reward money—and legal trouble.

The Great Internet Mersenne Prime Search (GIMPS), operating since 1996, has turned up five extremely large prime numbers so far. The fifth and largest was discovered in November 2001 by 20-year-old Michael Cameron of Owen Sound, Ontario. Mersenne primes can be expressed as $2^P - 1$, where P is itself a prime number. Cameron's is $2^{13,466,917} - 1$, which would take four million digits to write out. His computer spent 45 days discovering that his number is a prime; altogether the GIMPS network expended 13,000 years of computer time eliminating other numbers that could have been the 39th Mersenne.

The 38th Mersenne prime, a mere two million digits long, earned its discoverer (Nayan Hajratwala of Plymouth, Mich.) a $50,000 reward for being the first prime with more than a million digits. A prime with 10 million digits will win someone $100,000.

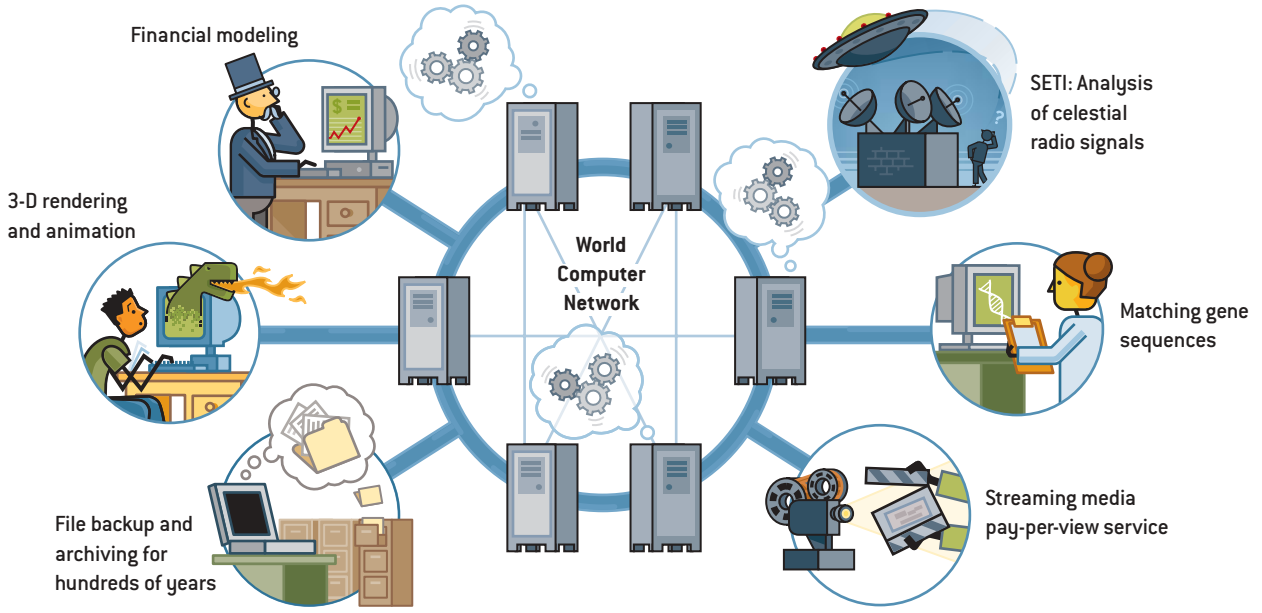
A Georgia computer technician, on the other hand, has found nothing but trouble through distributed computation. In 1999 David McOwen installed the client program for the "distributed.net" decryption project on computers in seven offices of the DeKalb Technical Institute, along with Y2K upgrades. During the Christmas holidays, the computers' activity was noticed, including small data uploads and downloads each day. In January 2000 McOwen was suspended, and he resigned soon thereafter.

Case closed? Case barely opened: The Georgia Bureau of Investigation spent 18 months investigating McOwen as a computer criminal, and in October 2001 he was charged with eight felonies under Georgia's computer crime law. The one count of computer theft and seven counts of computer trespass each carry a $50,000 fine and up to 15 years in prison. On January 17, a deal was announced whereby McOwen will serve one year of probation, pay $2,100 in restitution and perform 80 hours of community service unrelated to computers or technology.

Graham P. Collins is a staff writer and editor.

WHAT AN INTERNET-SCALE OPERATING SYSTEM COULD DO

By harnessing the massive unused computing resources of the global network, an ISOS would make short work of daunting number-crunching tasks and data storage. Here are just a few of the possibilities:



A computer operating system that provides only core functions is called a microkernel. Higher-level functions are built on top of it as user programs, allowing them to be debugged and replaced more easily. This approach was pioneered in academic research systems and has influenced some commercial systems, such as Windows NT. Most well-known operating systems, however, are not microkernels.

The core facilities of an ISOS include resource allocation (long-term assignment of hosts' processing power and storage), scheduling (putting jobs into queues, both across the system and within individual hosts), accounting of resource usage, and the basic mechanisms for distributing and executing application programs. The ISOS should not duplicate features of local operating systems running on hosts.

The system should be coordinated by servers operated by the ISOS provider, which could be a government-funded organization or a consortium of companies that are major resource sellers and buyers. (One can imagine competing ISOS providers, but we will keep things simple and assume a unique provider.) Centralization runs against the egalitarian approach popular in some peer-to-peer systems, but central servers are needed to ensure privacy of sensitive data, such as accounting data and other information about the resource hosts. Centralization might seem to require a control system that will become excessively large and unwieldy as the number of ISOS-connected hosts increases, and it appears to introduce a bottleneck that will choke the system anytime it is unavailable. These fears are unfounded: a reasonable number of servers can easily store information about every Internet-connected host and communicate with them regularly. Napster, for example, handled almost 60 million clients using a central server. Redundancy can be built into the server complex, and most ISOS online services can continue operating even with the servers temporarily unavailable.

The ISOS server complex would maintain databases of resource descriptions, usage policies and task descriptions. The resource descriptions include, for example, the host's operating system, processor type and speed, total and free disk space, memory space, performance statistics of its network connections, and statistical descriptions of when it is powered on and connected to the network. Usage policies spell out the rules an owner has dictated for using her resources. Task descriptions include the resources assigned to an online service and the queued jobs of a data-processing task.

To make their computers available to the network, resource sellers contact the server complex (for instance, through a Web site) to download and install an ISOS agent program, to link resources to their ISOS account, and so on. The ISOS agent manages the host's resource usage. Periodically it obtains from the ISOS server complex a list of tasks to perform.

Resource buyers send the servers task requests and application agent programs (to be run on hosts). An online service provider can ask the ISOS for a set of hosts on which to run, specifying its resource requirements (for example, a distributed backup service could use sporadically connected resource hosts—Mary's laptop—which would cost less than constantly connected hosts). The ISOS supplies the service with addresses and descriptions of the granted hosts and allows the application agent program to communicate directly between hosts on which it is running. The service can re-

XPLANE

quest new hosts when some become unavailable. The ISOS does not dictate how clients make use of an online service, how the service responds or how clients are charged by the service (unlike the ISOS-controlled payments flowing from resource users to host owners).

An Application Toolkit

IN PRINCIPLE, the basic facilities of the ISOS—resource allocation, scheduling and communication—are sufficient to construct a wide variety of applications. Most applications, however, will have important subcomponents in common. It is useful, therefore, to have a software toolkit to further assist programmers in building new applications. Code for these facilities will be incorporated into applications on resource hosts. Examples of these facilities include:

Location independent routing. Applications running with the ISOS can spread copies of information and instances of computation among millions of resource hosts. They have to be able to access them again. To facilitate this, applications name objects under their purview with Globally Unique Identifiers (GUIDs). These names enable "location independent routing," which is the ability to send queries to objects without knowing their location. A simplistic approach to location independent routing could involve a database of GUIDs on a single machine, but that system is not amenable to handling queries from millions of hosts. Instead the ISOS toolkit distributes the database of GUIDs among resource hosts. This kind of distributed system is being explored in research projects such as the OceanStore persistent data storage project at the University of California at Berkeley.

Persistent data storage. Information stored by the ISOS must be able to survive a variety of mishaps. The persistent data facility aids in this task with mechanisms for encoding, reconstructing and repairing data. For maximum survivability, data are encoded with an "*m*-of-*n*" code. An *m*-of-*n* code is similar in principle to a hologram, from which a small piece suffices for reconstructing the whole image. The encoding spreads information over *n* fragments (on *n* resource hosts), any *m* of which are sufficient to reconstruct the data. For instance, the facility might encode a document into 64 fragments, any 16 of which suffice to reconstruct it. Continuous repair is also important. As fragments fail, the repair facility would regenerate them. If properly constructed, a persistent data facility could preserve information for hundreds of years.

Secure update. New problems arise when applications need to update stored information. For example, all copies of the information must be updated, and the object's GUID must point to its latest copy. An access control mechanism must prevent unauthorized persons from updating information. The secure update facility relies on Byzantine agreement protocols, in which a set of resource hosts come to a correct decision, even if a third of them are trying to lead the process astray.

Other facilities. The toolkit also assists by providing additional facilities, such as format conversion (to handle the heterogeneous nature of hosts) and synchronization libraries (to aid in cooperation among hosts).

An ISOS suffers from a familiar catch-22 that slows the adoption of many new technologies: Until a wide user base exists, only a limited set of applications will be feasible on the ISOS. Conversely, as long as the applications are few, the user base will remain small. But if a critical mass can be achieved by convincing enough developers and users of the intrinsic usefulness of an ISOS, the system should grow rapidly.

The Internet remains an immense untapped resource. The revolutionary rise in popularity of the World Wide Web has not changed that—it has made the resource pool all the larger. An Internet-scale operating system would free programmers to create applications that could run on this World Wide Computer without worrying about the underlying hardware. Who knows what will result? Mary and her computers will be doing things we haven't even imagined. SA

A typical PC owner might see the system as a BARTER ECONOMY that provides free services in exchange for PROCESSOR TIME and DISK SPACE.

MORE TO EXPLORE

The Ecology of Computation. B. A. Huberman. North-Holland, 1988.

The Grid: Blueprint for a New Computing Infrastructure. Edited by Ian Foster and Carl Kesselman. Morgan Kaufmann Publishers, 1998.

Peer-to-Peer: Harnessing the Power of Disruptive Technologies. Edited by Andy Oram. O'Reilly & Associates, 2001.

Many research projects are working toward an Internet-scale operating system, including:

Chord: **www.pdos.lcs.mit.edu/chord/**

Cosm: **www.mithral.com/projects/cosm/**

Eurogrid: **www.eurogrid.org/**

Farsite: **http://research.microsoft.com/sn/farsite/**

Grid Physics Network (Griphyn): **www.griphyn.org/**

OceanStore: **http://oceanstore.cs.berkeley.edu/**

Particle Physics Data Grid: **www.ppdg.net/**

Pastry: **www.research.microsoft.com/~antr/pastry/**

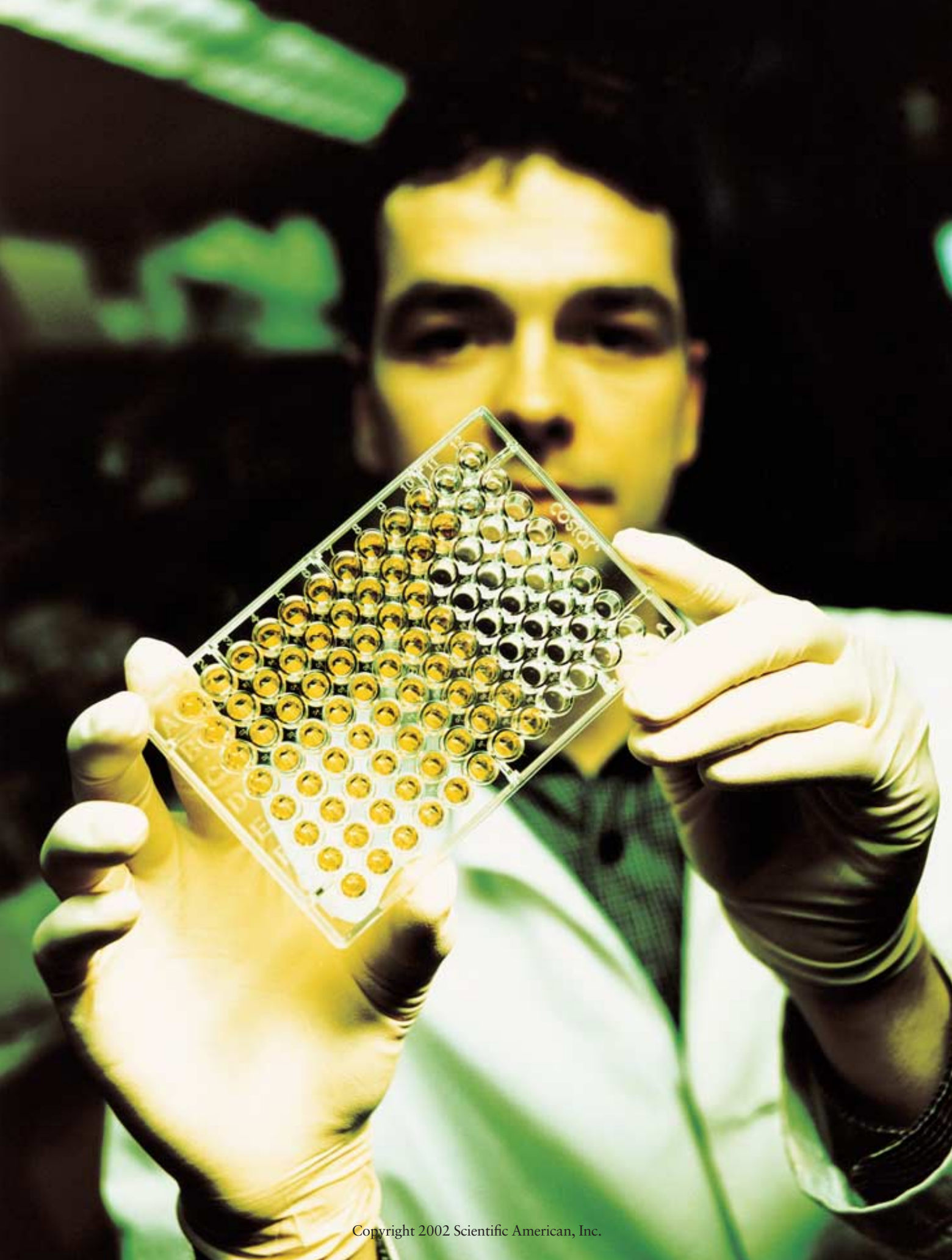
Tapestry: **www.cs.berkeley.edu/~ravenben/tapestry/**

ATTACKING ANTHRAX

Recent discoveries are suggesting much-needed strategies for improving prevention and treatment. High on the list: ways to neutralize the anthrax bacterium's fiendish toxin

by John A. T. Young and R. John Collier

CULTURES OF CELLS survived exposure to the anthrax toxin after being treated with a potential antitoxin. Michael Mourez of Harvard University holds a plate containing the cultures.



The need for new anthrax therapies became all too clear last fall

when five people died of inhalation anthrax, victims of the first purposeful release of anthrax spores in the U.S. Within days of showing initially unalarming symptoms, the patients were gone, despite intensive treatment with antibiotics. Six others became seriously ill as well before pulling through.

Fortunately, our laboratories and others began studying the causative bacterium, *Bacillus anthracis,* and seeking antidotes long before fall 2001. Recent findings are now pointing the way to novel medicines and improved vaccines. Indeed, in the past year alone, the two of us and our collaborators have reported on three promising drug prototypes.

An Elusive Killer

THE NEW IDEAS for fighting anthrax have emerged from ongoing research into how *B. anthracis* causes disease and death. Anthrax does not spread from individual to individual. A person (or animal) gets sick only after incredibly hardy spores enter the body through a cut in the skin, through contaminated food or through spore-laden air. Inside the body the spores molt into "vegetative," or actively dividing, cells.

Anthrax bacteria that colonize the skin or digestive tract initially do damage locally and may cause self-limited ailments: black sores and swelling in the first instance; possibly vomiting and abdominal pain and bleeding in the second. If bacterial growth persists unchecked in the skin or gastrointestinal tract, however, the microbes may eventually invade the bloodstream and thereby cause systemic disease.

Inhaled spores that reach deep into the lungs tend to waste little time where they land. They typically convert to the vegetative form and travel quickly to lymph nodes in the middle of the chest, where many of the cells find ready access to the blood. (Meanwhile bacteria that remain in the chest set the stage for a breath-robbing buildup of fluid around the lungs.)

Extensive replication in the blood is generally what kills patients who succumb to anthrax. *B. anthracis*'s ability to expand so successfully derives from its secretion of two substances, known as virulence factors, that can profoundly derail the immune defenses meant to keep bacterial growth in check. One of these factors encases the vegetative cells in a polymer capsule that inhibits ingestion by the immune system's macrophages and neutrophils—the scavenger cells that normally degrade disease-causing bacteria. The capsule's partner in crime is an extraordinary toxin that works its way into those scavenger cells, or phagocytes, and interferes with their usual bacteria-killing actions.

The anthrax toxin, which also enters other cells, is thought to contribute to mortal illness not only by dampening immune responses but also by playing a direct role. Evidence for this view includes the observation that the toxin alone, in the absence of bacteria, can kill animals. Conversely, inducing the immune system to neutralize the toxin prevents *B. anthracis* from causing disease.

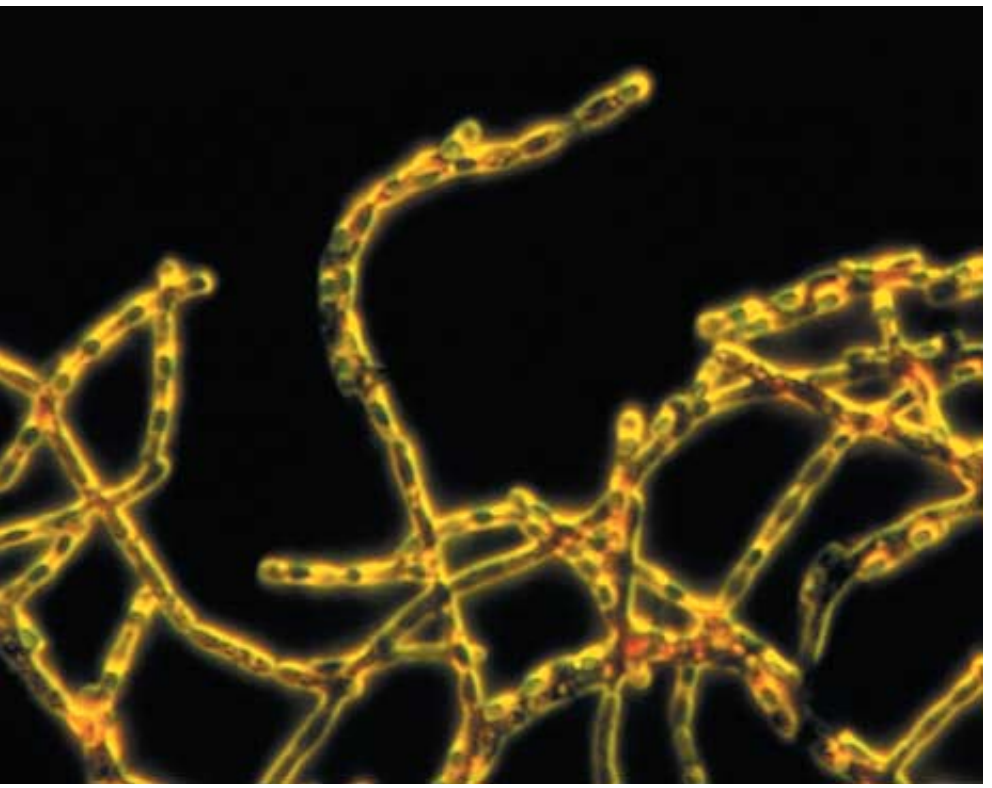
A Terrible Toxin

HARRY SMITH and his co-workers at the Microbiological Research Establishment in Wiltshire, England, discovered the toxin in the 1950s. Aware of its central part in anthrax's lethality, many researchers have since focused on learning how the substance "intoxicates" cells—gets into them and disrupts their activities. Such details offer essential clues to blocking its effects. Stephen H. Leppla and Arthur M. Friedlander, while at the U.S. Army Medical Research Institute of Infectious Diseases, initiated that effort with their colleagues in the 1980s; the two

Overview/*Anthrax*

- A three-part toxin produced by the anthrax bacterium, *Bacillus anthracis,* contributes profoundly to the symptoms and lethality of anthrax.
- The toxin causes trouble only when it gets into the cytosol of cells, the material that bathes the cell's internal compartments.
- Drugs that prevented the toxin from reaching the cytosol would probably go a long way toward limiting illness and saving the lives of people infected by the anthrax bacterium.
- Analyses of how the toxin enters cells have recently led to the discovery of several potential antitoxins.

JASON GROW *(preceding page)*



ACTIVELY DIVIDING CELLS of the anthrax bacterium arrange themselves into chains that resemble linked boxcars.

of us and others took up the task somewhat later.

The toxin turns out to consist of three proteins: protective antigen, edema factor and lethal factor. These proteins cooperate but are not always joined together physically. They are harmless individually until they attach to and enter cells, which they accomplish in a highly orchestrated fashion.

First, protective antigen binds to the surface of a cell, where an enzyme trims off its outermost tip. Next, seven of those trimmed molecules combine to form a ring-shaped structure, or heptamer, that captures the two factors and is transported to an internal membrane-bound compartment called an endosome. Mild acidity in this compartment causes the heptamer to change shape in a way that leads to the transport of edema factor and lethal factor across the endosomal membrane into the cytosol (the internal matrix of cells), where they do their mischief. In essence, the heptamer is like a syringe loaded with edema factor and lethal factor, and the slight acidity of the endosome causes the syringe to pierce the membrane of the endosome and inject the toxic factors into the cytosol.

Edema factor and lethal factor catalyze different molecular reactions in cells. Edema factor upsets the controls on ion and water flow across cell membranes and thereby promotes the swelling of tissues. In phagocytes, it also saps energy that would otherwise be used to engulf bacteria.

The precise behavior of lethal factor, which could be more important in causing patient deaths, is less clear. Scientists do know that it is a protease (a protein-cutting enzyme) and that it cleaves enzymes in a family known as MAPKKs. Now they are trying to tease out the molecular events that follow such cleavage and to uncover the factor's specific contributions to disease and death.

Therapeutic Tactics

CERTAINLY DRUGS able to neutralize the anthrax toxin would help the immune system fight bacterial multiplication and would probably reduce a patient's risk of dying. At the moment, antibiotics given to victims of inhalation anthrax may control microbial expansion but leave the toxin free to wreak havoc.

In principle, toxin activity could be halted by interfering with any of the steps in the intoxication process. An attractive approach would stop the sequence almost before it starts, by preventing protective antigen from attaching to cells. Scientists realized almost 10 years ago that this protein initiated toxin entry by binding to some specific protein on the surface of cells; when cells were treated with enzymes that removed all their surface proteins, protective antigen found no footing. Until very recently, though, no one knew which of the countless proteins on cells served as the crucial receptor.

The two of us, with our colleagues Kenneth Bradley, Jeremy Mogridge and Michael Mourez, found the receptor last summer. Detailed analysis of this molecule (now named ATR, for anthrax toxin receptor) then revealed that it spans the cell membrane and protrudes from it. The protruding part contains an area resembling a region that serves in other receptors as an attachment site for particular proteins. This discovery suggested that the area was the place where protective antigen latched onto ATR, and indeed it is.

We have not yet learned the normal function of the receptor, which surely did not evolve specifically to allow the anthrax toxin into cells. Nevertheless, knowledge of the molecule's makeup is enabling us to begin testing inhibitors of its activity. We have had success, for instance, with a compound called sATR, which is a soluble form of the receptor domain that binds to protective antigen. When sATR molecules are mixed into the medium surrounding cells, they serve as effective decoys, tricking protective antigen into binding to them instead of to its true receptor on cells.

THE AUTHORS

JOHN A. T. YOUNG and *R. JOHN COLLIER* have collaborated for several years on investigating the anthrax toxin. Young is Howard M. Temin Professor of Cancer Research in the McArdle Laboratory for Cancer Research at the University of Wisconsin–Madison. Collier, who has studied anthrax for more than 14 years, is Maude and Lillian Presley Professor of Microbiology and Molecular Genetics at Harvard Medical School.

Detecting Anthrax

Rapid sensing would save lives

By Rocco Casagrande

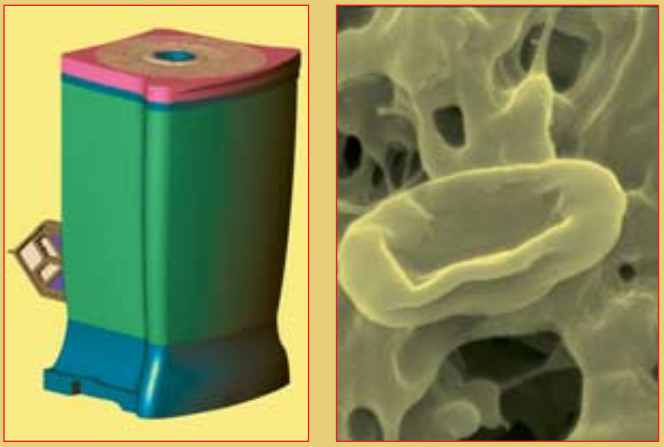
IF A TERRORIST GROUP spread anthrax spores into the open air, the release could affect large numbers of people but would probably go unnoticed until victims showed up at hospitals. Many would undoubtedly seek help too late to be saved by current therapies. Much illness could be prevented, however, if future defenses against anthrax attacks included sensors that raised an alarm soon after spores appeared in the environment. The needed instruments are not yet ready for deployment, but various designs that incorporate cutting-edge technology are being developed.

Environmental sensors must discriminate between disease-causing agents (pathogens) and the thousands of similar but harmless microorganisms that colonize air, water and soil. Most of the tools being investigated work by detecting unique molecules on the surface of the pathogens of interest or by picking out stretches of DNA found only in those organisms.

The Canary, which is being developed at the Massachusetts Institute of Technology Lincoln Laboratory, is an innovative example of the devices that detect pathogens based on unique surface molecules. The sensors of the Canary consist of living cells—B cells of the immune system—that have been genetically altered to emit light when their calcium levels change. Protruding from these cells are receptors that will bind only to a unique part of a surface molecule on a particular pathogen. When the cells in the sensor bind to their target, that binding triggers the release of calcium ions from stores within the cells, which in turn causes the cells to give off light. The Canary can discern more than one type of pathogen by running a sample through several cell-filled modules, each of which reacts to a selected microorganism.

The GeneXpert system, developed by Cepheid, in Sunnyvale, Calif., is an example of a gene-centered approach. It begins its work by extracting DNA from microorganisms in a sample. Then, if a pathogen of concern is present, small primers (strips of genetic material able to recognize specific short sequences of DNA) latch onto the ends of DNA fragments unique to the pathogen. Next, through a procedure called the polymerase chain reaction (PCR), the system makes many copies of the bound DNA, adding fluorescent labels to the new copies along the way. Within about 30 minutes GeneXpert can make enough DNA to reveal whether even a small amount of the worrisome organism inhabited the original sample.

This system contains multiple PCR reaction chambers with distinct primer sets to allow the detection of different pathogens simultaneously. Furthermore, the GeneXpert system could be used to determine whether the anthrax bacterium is present in a nasal swab taken from a patient in as little as half an hour, significantly faster than the time it takes for conventional microbiological techniques to yield results.



CARTRIDGE used in the experimental GeneXpert system is about as tall as an adult's thumb (*left*). Inside, sound waves bombard material to be tested, causing any cells to break open and release their DNA. If a pathogen of interest is present, its DNA will be amplified in the arrow-shaped reaction tube (*protrusion*), and the edges of the arrowhead will fluoresce. The micrograph (*right*) shows the remains of a cell that has disgorged its contents.

Instruments designed specifically to detect spores of the anthrax bacterium or of closely related microbes (such as the one that causes botulism) can exploit the fact that such spores are packed full of dipicolinic acid (DPA)—a compound, rarely found elsewhere in nature, that helps them to survive harsh environmental conditions. Molecules that fluoresce when bound to DPA have shown promise in chemically based anthrax detectors. "Electronic noses," such as the Cyranose detection system made by Cyrano Sciences in Pasadena, Calif., could possibly "smell" the presence of DPA in an air sample laced with anthrax spores.

The true danger of an anthrax release lies in its secrecy. If an attack is discovered soon after it occurs and if exposed individuals receive treatment promptly, victims have an excellent chance of surviving. By enhancing early detection, sensors based on the systems discussed above or on entirely different technologies could effectively remove a horrible weapon from a terrorist's arsenal.

ROCCO CASAGRANDE is a scientist at Surface Logix in Brighton, Mass., where he is developing methods and devices for detecting biological weapons.

We are now trying to produce sATR in the amounts needed for evaluating its ability to combat anthrax in rodents and nonhuman primates—experiments that must be done before any new drug can be considered for fighting anthrax in people. Other groups are examining whether carefully engineered antibodies (highly specific molecules of the immune system) might bind tightly to protective antigen in ways that will keep it from coupling with its receptor.

More Targets

SCIENTISTS ARE ALSO seeking ways to forestall later steps in the intoxication pathway. For example, a team from Harvard has constructed a drug able to clog the regions of the heptamer that grasp edema and lethal factors. The group—from the laboratories of one of us (Collier) and George M. Whitesides—reasoned that a plugged heptamer would be unable to draw the factors into cells.

We began by screening randomly constructed peptides (short chains of amino acids) to see if any of them bound to the

COURTESY OF CEPHEID

ANTHRAX IN ACTION

Physicians classify anthrax according to the tissues that are initially infected. The disease turns deadly when the causative bacterium, *Bacillus anthracis,* reaches the bloodstream and proliferates there, producing large amounts of a dangerous toxin. Much research is now focused on neutralizing the toxin.

THREE TYPES

INHALATION ANTHRAX
Spores are breathed in

GASTROINTESTINAL ANTHRAX
Spores are ingested by eating contaminated meat

CUTANEOUS ANTHRAX
Spores penetrate the skin through a break

HOW INHALATION ANTHRAX ARISES

Inhalation anthrax is the most dangerous form, probably because bacteria that land in the lungs are more likely to reach the bloodstream and thus disseminate their toxin through the body.

1 Immune system cells called macrophages ingest *B. anthracis* spores and carry them to lymph nodes in the chest. En route, or in the macrophages, the spores transform into actively dividing cells

2 Proliferating *B. anthracis* cells erupt from macrophages and infiltrate the blood readily

3 In the blood, the active bacteria evade destruction by macrophages and other cells of the immune system by producing a capsule (*detail*) that blocks the immune cells from ingesting them and by producing a toxin that enters immune cells and impairs their functioning

4 Protected from immune destruction, the bacteria multiply freely and spread through the body

MACROPHAGE
SPORE
REPLICATING BACTERIAL CELLS
BACTERIUM
CAPSULE
BACTERIA IN BLOOD
CELL
MACROPHAGE FILLED WITH TOXIN
TOXIN MOLECULES

HOW THE TOXIN INVADES CELLS ... AND HOW TO STOP IT

THE ANTHRAX TOXIN must enter cells to hurt the body. It consists of three collaborating proteins: protective antigen (PA), edema factor (EF) and lethal factor (LF). The last two disrupt cellular activities, but only after protective antigen delivers them to the cytosol—the matrix surrounding the cell's intracellular compartments. Molecular understanding of how the factors reach the cytosol has led to ideas for blocking that journey and thus for neutralizing the toxin and saving lives. The antitoxins depicted in the boxes have shown promise in laboratory studies.

LF

EF

HEPTAMER

PA

1 PA binds to its receptor on a cell

ANTHRAX TOXIN RECEPTOR (ATR)

2 PA gets cleaved

3 Seven copies combine, forming a heptamer

4 Up to three copies of EF or LF or a combination of the two bind to the heptamer

HEPTAMER COMPLEX

5 The heptamer complexed with EF and LF is delivered to a membrane-bound compartment called an endosome

ENDOSOME

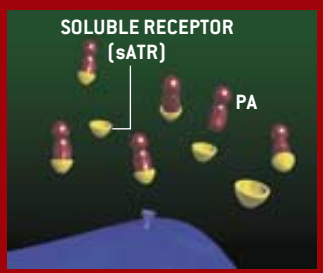
6 Mild acidity in the endosome causes the heptamer to inject EF and LF into the cytosol

CYTOSOL

7 EF causes tissues to swell and prevents immune system cells from ingesting and degrading bacteria

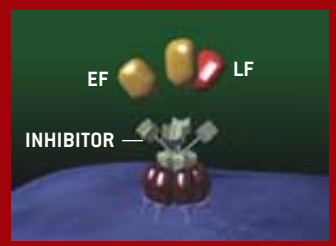
8 LF is believed to be important in causing disease and death, but exactly how it does so is in question

TREATMENT IDEA



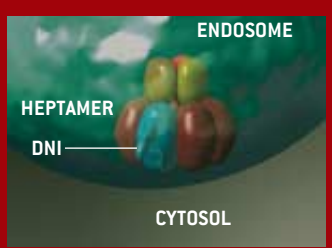
Prevent PA from linking to its receptor on cells. Induce it to bind instead to decoys, such as soluble copies of the toxin receptor's PA binding site.

TREATMENT IDEA



Keep EF and LF from attaching to their binding sites on PA heptamers. Plug those sites with linked copies of a molecule that also has affinity for the sites.

TREATMENT IDEA



Block transport of EF and LF from the endosome into the cytosol by causing newly forming heptamers to incorporate a version of PA known as a dominant negative inhibitor (DNI). DNI-containing heptamers cannot move EF and LF across the endosome's membrane.

heptamer. One did, so we examined its ability to block toxin activity. It worked, but weakly. Assuming that fitting many plugs into the heptamer's binding domains for edema and lethal factor would be more effective, we took advantage of chemical procedures devised by Whitesides's group and linked an average of 22 copies of the peptide to a flexible polymer. That construction showed itself to be a strong inhibitor of toxin action—more than 7,000 times better than the free peptide—both in cell cultures and in rats.

Another exciting agent, and the one probably closest to human testing, would alter the heptamer itself. This compound was discovered after Bret R. Sellman in Collier's group noted that when certain mutant forms of protective antigen were mixed with normal forms, the heptamers formed on cells as usual but were unable to inject edema and lethal factors into the cytosol. Remarkably, some of these mutants were so disruptive that a single copy in a heptamer completely prevented injection.

In a study reported last April, these mutants—known as dominant negative inhibitors, or DNIs—proved to be potent blockers of the anthrax toxin in cell cultures and in rats. Relatively small amounts of selected DNIs neutralized an amount of protective antigen and lethal factor that would otherwise kill a rat in 90 minutes. These findings suggest that each mutant copy of protective antigen is capable of inactivating six normal copies in the bloodstream and that it would probably reduce toxin activity in patients dramatically.

Of course, as more and more questions about the toxin are answered, scientists should discover further treatment ideas. Now that the receptor for protective antigen has been identified, researchers can use it as a target in screening tests aimed at finding drugs able to bar the receptor from binding to protective antigen. And understanding of the receptor's three-dimensional structure would reveal the precise contact points between protective antigen and the receptor, enabling drugmakers to custom-design receptor blocking agents.

Scientists would also like to uncover the molecular interactions that enable protective antigen heptamers to move from the cell surface into endosomes inside the cell. Impeding that migration should be very useful. And what happens after lethal factor cleaves MAPKK enzymes? How do those subsequent events affect cells? Although the latter question remains a vexing challenge, recent study of lethal factor has brightened the prospects for finding drugs able to inactivate it. Last November, Robert C. Liddington of the Burnham Institute in La Jolla, Calif., and his colleagues in several laboratories published the three-dimensional structure of the part of lethal factor that acts on MAPKK molecules. That site can now become a target for drug screening or design.

New leads for drugs should also emerge from the recent sequencing of the code letters composing the *B. anthracis* genome. By finding genes that resemble those of known functions in other organisms, biologists are likely to discover additional information about how the anthrax bacterium causes disease and how to stop it.

The continuing research should yield several antitoxins. To be most effective, such drugs will probably be used with antibiotics, much as cocktails of antiviral drugs are recommended for treating HIV infection.

To be most effective, antitoxins will probably be USED WITH ANTIBIOTICS, much as cocktails of antiviral drugs are recommended for treating HIV infection.

Promising Preventives

AS PLANS TO IMPROVE therapies proceed, so does work on better vaccines. Vaccines against toxin-producing bacteria often prime the immune system to neutralize the toxin of concern as soon as it appears in the body, thus preventing disease. Livestock in parts of the U.S. receive preparations consisting of *B. anthracis* cells that lack the protective capsule and thus replicate poorly. A similar vaccine for humans has been used in the former Soviet Union. But preparations that contain whole microbes often cause side effects, and they raise the specter that renegade cells might at times give rise to the very diseases they were meant to prevent.

The only anthrax vaccine approved for human use in the U.S. takes a different form. It consists primarily of toxin molecules that have been chemically treated to prevent them from making people ill. It is produced by growing the weakened strain of *B. anthracis* in culture, filtering the bacterial cells from the culture medium, adsorbing the toxin proteins in the remaining filtrate onto an adjuvant (a substance that enhances immune responses) and treating the mixture with formaldehyde to inactivate the proteins. Injection of this preparation, known as AVA (for anthrax vaccine adsorbed), stimulates the immune system to produce antibodies that specifically bind to and inactivate the toxin's components. Most of the antibodies act on protective antigen, however, which explains the protein's name: it is the component that best elicits protective immunity.

AVA is given to soldiers and certain civilians but is problematic as a tool for shielding the general public against biological warfare. Supplies are limited. And even if AVA were available in abundance, it would be cumbersome to deliver on a large scale; the standard protocol calls for six shots delivered over 18 months followed by annual boosters. The vaccine has not been licensed for use in people already exposed to anthrax spores. But late last year officials, worried that spores

BRYAN CHRISTIE DESIGN *(preceding two pages)*

Medical Lessons

Doctors now have a changed view of inhalation anthrax

By Ricki L. Rusting

NORMA WALLACE of Willingboro, N.J., is one of the six patients who survived inhalation anthrax last autumn.

THE RECENT CASES of inhalation anthrax in the U.S. have upended some old assumptions about that disease. When contaminated letters started appearing in September 2001, public health authorities initially believed that only those who received the letters, and perhaps individuals nearby, were in danger. But spores clearly seeped out through the weave of the envelopes, contaminating postal facilities and jumping to other mail. Such "cross contamination" is a leading explanation for the deaths of two of the 11 people confirmed to have contracted inhalation anthrax last year. Also contrary to expectations, spores do not remain sedentary once they land. They can become airborne again as people walk around in a tainted room.

One surprise was positive. Before October 2001 common wisdom held that inhalation anthrax was almost always incurable after symptoms appeared. But doctors beat those odds last fall, saving six of the victims. What made the difference? Researchers cannot draw firm conclusions from so few cases. But some intriguing patterns emerged when John A. Jernigan of the Centers for Disease Control and Prevention (CDC) and a team of others reviewed the medical records of the first 10 patients. Their findings appear in the November/December 2001 *Emerging Infectious Diseases* and online at www.cdc.gov/ncidod/eid/vol7no6/jernigan.htm

Relatively prompt diagnosis may have helped, the researchers report. Inhalation anthrax has two symptomatic phases—an early period marked by maladies common to a variety of ailments (such as fatigue, fever, aches and cough) and a later phase in which patients become critically ill with high fever, labored breathing and shock. Six of the 10 patients received antibiotics active against the anthrax bacterium, *Bacillus anthracis,* while they were still showing early symptoms of infection, and only they survived.

The types of antibiotics prescribed and the use of combinations of drugs might also have had a hand in the unexpectedly high survival rate. Nine of the people discussed in the review sought care before the CDC published what it called "interim" guidelines for treating inhalation anthrax on October 26, but most patients received therapy consistent with those guidelines: ciprofloxacin (the now famous Cipro) or doxycycline plus one or two other agents known to inhibit replication of *B. anthracis* (such as rifampin, vancomycin, penicillin, ampicillin, chloramphenicol, imipenem, clindamycin and clarithromycin). Aggressive "supportive" care—including draining dangerous fluid from around the lungs—probably helped as well, scientists say.

Even the survivors were very sick, however. Jernigan says they are still being observed to see whether long-term complications will develop, although as of mid-January no obvious signs of such problems had emerged. Researchers suspect that anthrax antitoxins would ease the course of many people afflicted with anthrax and might also rescue patients who could not be saved with current therapies.

Ricki L. Rusting is a staff editor and writer.

might sometimes survive in the lungs for a long time, began offering an abbreviated, three-course dose on an experimental basis to postal workers and others who had already taken 60 days of precautionary antibiotics. People who accepted the offer were obliged to take antibiotics for an additional 40 days, after which the immunity stimulated by the vaccine would presumably be strong enough to provide adequate protection on its own.

In hopes of producing a more powerful, less cumbersome and faster-acting vaccine, many investigators are focusing on developing inoculants composed primarily of protective antigen produced by recombinant DNA technology. By coupling the recombinant protein with a potent new-generation adjuvant, scientists may be able to evoke good protective immunity relatively quickly with only one or two injections. The dominant negative inhibitors discussed earlier as possible treatments could be useful forms of protective antigen to choose. Those molecules retain their ability to elicit immune responses. Hence, they could do double duty: disarming the anthrax toxin in the short run while building up immunity that will persist later on.

We have no doubt that the expanding research on the biology of *B. anthracis* and on possible therapies and vaccines will one day provide a range of effective anthrax treatments. We fervently hope that these efforts will mean that nobody will have to die from anthrax acquired either naturally or as a result of biological terrorism. SA

MORE TO EXPLORE

Anthrax as a Biological Weapon: Medical and Public Health Management. Thomas V. Inglesby et al. in *Journal of the American Medical Association,* Vol. 281, No. 18, pages 1735–1745; May 12, 1999.

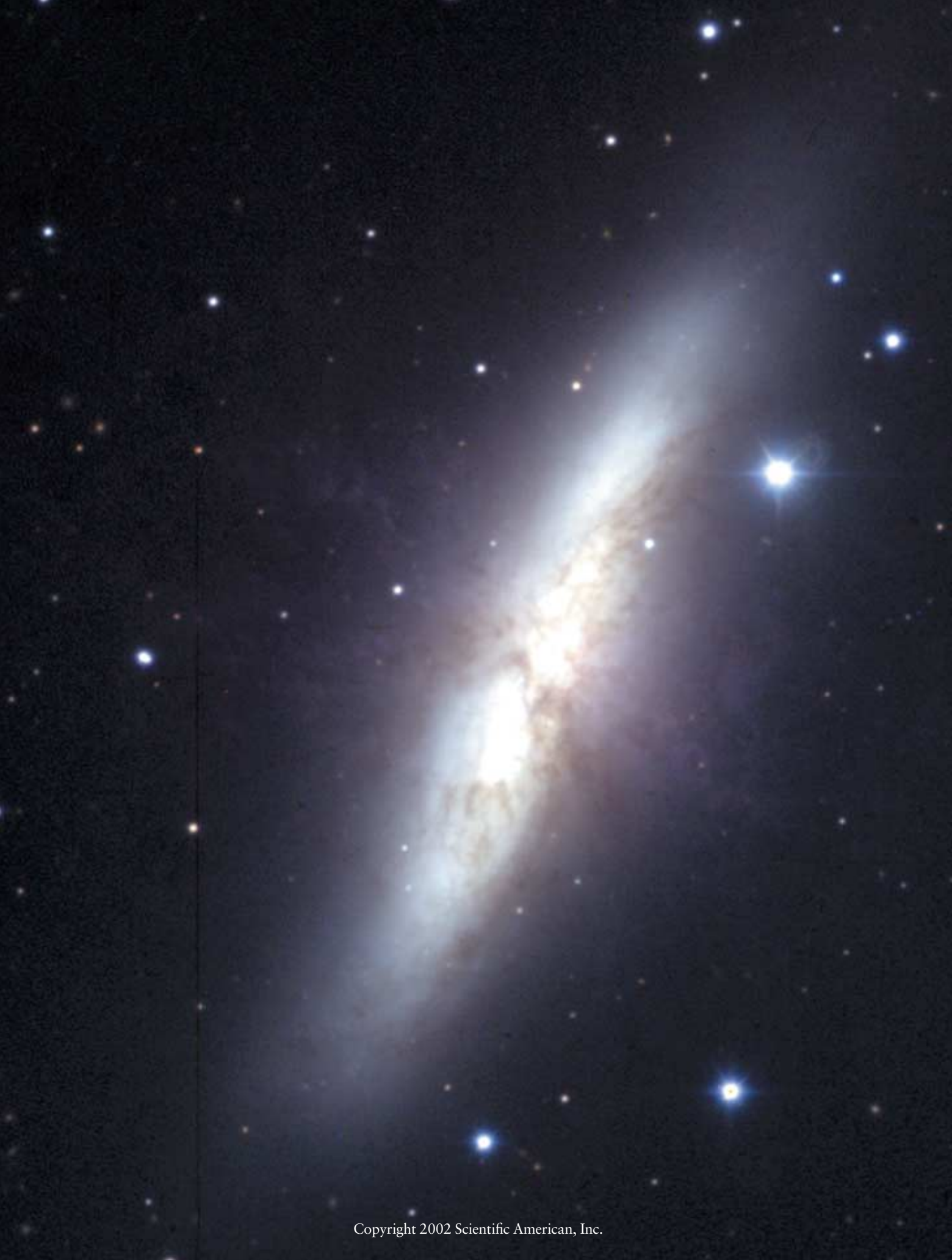
Dominant-Negative Mutants of a Toxin Subunit: An Approach to Therapy of Anthrax. Bret R. Sellman, Michael Mourez and R. John Collier in *Science,* Vol. 292, pages 695–697; April 27, 2001.

Designing a Polyvalent Inhibitor of Anthrax Toxin. Michael Mourez et al. in *Nature Biotechnology,* Vol. 19, pages 958–961; October 2001.

Identification of the Cellular Receptor for Anthrax Toxin. Kenneth A. Bradley, Jeremy Mogridge, Michael Mourez, R. John Collier and John A. T. Young in *Nature,* Vol. 414, pages 225–229; November 8, 2001.

The U.S. Centers for Disease Control and Prevention maintain a Web site devoted to anthrax at **www.cdc.gov/ncidod/dbmd/diseaseinfo/anthrax_g.htm**

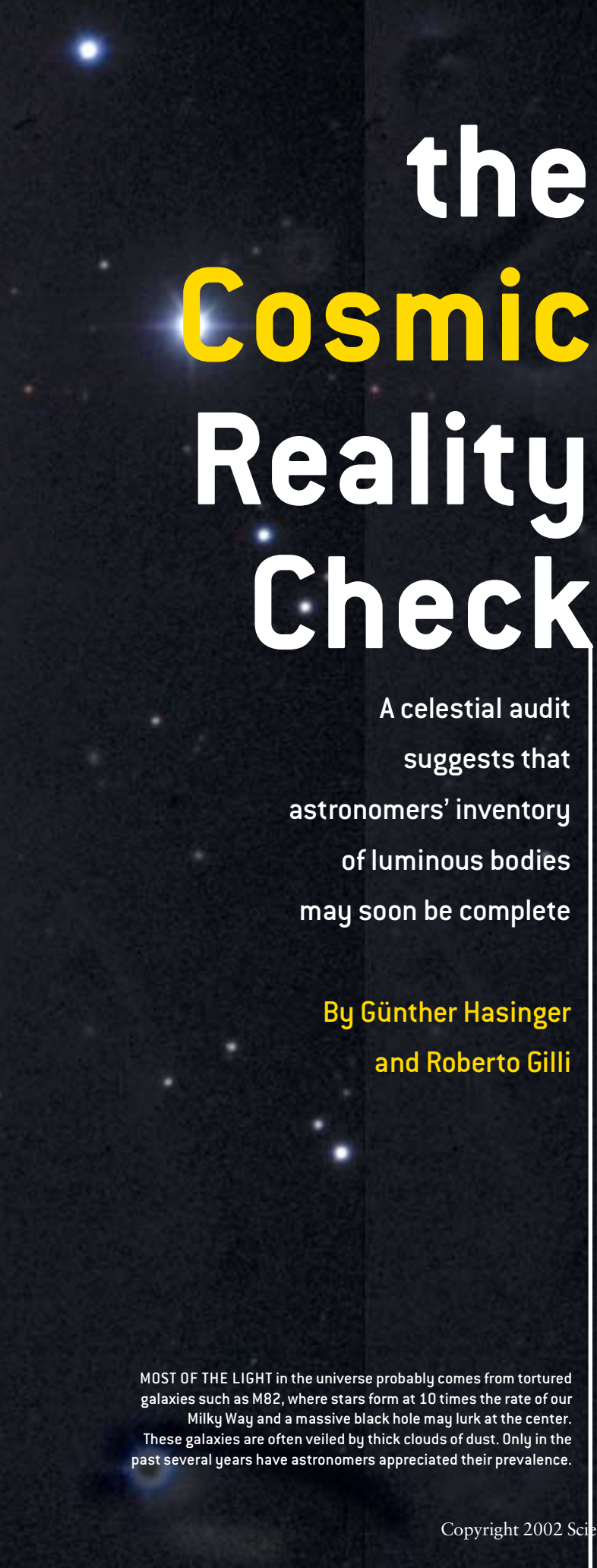
BRIAN BRANCH-PRICE *AP Photo*



the Cosmic Reality Check

A celestial audit suggests that astronomers' inventory of luminous bodies may soon be complete

By Günther Hasinger and Roberto Gilli



MOST OF THE LIGHT in the universe probably comes from tortured galaxies such as M82, where stars form at 10 times the rate of our Milky Way and a massive black hole may lurk at the center. These galaxies are often veiled by thick clouds of dust. Only in the past several years have astronomers appreciated their prevalence.

With each day's newspaper seems to arrive a new astronomical discovery: a new celestial body, a new physical process, a new form of matter. Will the revelations ever end? Will there ever come a day when astronomers feel confident that they have made a complete inventory of the universe? If the question is put so broadly, the answer is clearly no: astronomers already know that not everything in the universe can be seen directly, and additional surprises are inevitable. But a somewhat narrower question—will astronomers ever finish their head count of stars, galaxies and other luminous matter?—has a rather different answer. A day will indeed come when astronomers have accounted for the bulk of the light in the universe, and that day is fast approaching.

Over the years, astronomers have developed a type of quality-control check that can signal whether they have missed any important source of light. The idea is to study a phenomenon that most observers consider a nuisance: the so-called background radiation. When scientists in any discipline talk about a "background," they usually mean everything except what they are interested in. A telescope capturing the radiation from a star cannot avoid collecting light from other bodies near and far. This extraneous light serves only to reduce the precision of the desired measurement.

Those of us who study the background radiation focus our attention precisely on what our colleagues try to ignore. We first add up all the light coming from a given region of space. Then we systematically subtract the contributions from known objects such as stars, galaxies and gas clouds—collectively, the "foreground." If something is left over, some diffuse glow of indeterminate origin, we know that our census of heavenly objects must still be incomplete.

Sometimes a diffuse glow is observed when objects are very

closely spaced and the telescope lacks sufficient angular resolution to pick them apart. Take, for example, the Milky Way, which is a blur to the naked eye. With a simple pair of binoculars, you can see that the blur consists of millions of individual points of light. At other times, a diffuse glow comes from a source that truly is diffuse, such as the zodiacal dust of our own solar system or the gaseous supernova remnants of our galaxy. Many (but by no means all) of these sources within our galaxy and nearby galaxies have been identified, so they can be considered part of the foreground. The radiation that comes from far outside our galaxy, filling the whole universe, is the cosmic background.

In the past half a decade, as the sensitivity and resolution of telescopes have improved dramatically, astronomers have accounted for more and more of the background glow. In so doing, we have discovered that our previous inventories of the universe were incomplete: for instance, we had badly underestimated the prevalence of supermassive black holes. Far from being isolated oddities, as was once thought, they are everywhere. Earlier studies had missed them because they are cloaked by prodigious quantities of dust. With these holes now unveiled, we may soon explain the background fully.

That is not to say we will have seen everything there is to see. We can no more catalogue every celestial body than a biologist can count every beetle. But just as biologists can fairly claim to know all the major types of, say, land mammals, astronomers are on the verge of identifying all the major classes of light-emitting objects.

Not a Whisper Be Lost

WHEN ASTRONOMY AFICIONADOS hear the word "background," they immediately think of the famous cosmic microwave background (CMB). This pervasive radio emission appears to have a truly diffuse origin—namely, a hot plasma that filled the universe when it was only 400,000 years old. Through the expansion of the universe, this radiation is today observed at a peak wavelength of about one millimeter, corresponding to a temperature of 2.7 kelvins. The study of the spectrum and distribution of the CMB has provided compelling evidence for the big bang theory.

Yet the CMB is only part of the story. The whole electromagnetic background is actually a mixture of components, each of which dominates a particular range of wavelengths. Besides the CMB are the lesser-known cosmic x-ray background (CXB), cosmic infrared background (CIB) and cosmic optical background (COB).

The precise measurement of these components is one of the most trying tasks in observational astronomy. Conceptually it seems so simple: look at the sky to measure the total signal and then subtract all the known sources between Earth and the deep universe (the foreground): the noise of the detectors, the signals

LITTLE MORE THAN A BLUR to the naked eye, the Milky Way dissolves into 100 billion stars when viewed through a telescope. The inset shows fine detail in the constellation Ara. Similarly, the cosmic background—a subtle glow that fills the sky between the stars—used to look like a blur. In recent years, telescopes have become advanced enough to pick out individual sources of light, many of which represent new classes of galaxies.

preceding pages: PETER CHALLIS *Harvard-Smithsonian Center for Astrophysics; this page:* ROGER RESSMEYER *Corbis;* WFI/EUROPEAN SOUTHERN OBSERVATORY *(inset)*

from within our solar system, the emission from the rest of the galaxy, and so forth. In addition, one has to correct for any foreground attenuation of the background signal.

Performing all these subtractions with sufficient precision, though, is tricky; subtraction is an operation that amplifies error. In certain wavelength bands, observers are lucky that the background is the brightest emission in the sky, but in other bands they have to extract a cosmic whisper from a foreground roar. Most often the limiting factor is the accuracy with which astronomers know the foreground emission. They try to skirt this problem by concentrating on regions of the sky that are utterly devoid of stars and other known foregrounds—the more boring, the better. Despite the obstacles, observers have now determined the cosmic background spectrum with quite high precision over a broad range of the spectrum [*see illustration on page 67*].

The x-ray component, discovered in 1962, has a characteristic hump at about 30 kiloelectron-volts—corresponding roughly to the wavelength used for medical x-rays—and a long tail toward higher energies, including gamma rays. Below about 1 keV, and superposed on this continuum, are a number of atomic emission lines that appear to be the fingerprint of a gas heated to several million kelvins and most likely located inside or around our galaxy.

In the 1970s the first x-ray satellites, such as UHURU, ARIEL V and HEAO-1, showed the higher-energy x-radiation to be spread uniformly over the sky. Thus, its origin has to be mainly extragalactic: if it came from our solar system or galaxy, the brightness would be strongly skewed in certain directions corresponding to the plane of the planets or galactic disk. Gamma-ray satellites such as SAS-3, COS-B and the Compton Gamma Ray Observatory have found a similar uniformity at still higher energies.

Whereas the CMB and CXB dominate the sky in their respective bands, the other cosmic background components account for only a small fraction of the radiation in their respective wavelength bands. A few years ago several groups independently detected the far-infrared background signal in the high-frequency tail of the CMB [see "Glow in the Dark," by George Musser; Scientific American, March 1998]. In the near- to mid-infrared range, the bright zodiacal light obscures the background, so astronomers have generally resorted to interpolating measurements from other wavelength bands. They have also derived upper limits from observations of high-energy gamma rays: too thick a haze of infrared photons would interfere with the propagation of gamma rays. Only in the past two years have observers made direct measurements at infrared wavelengths.

In the optical and ultraviolet, the first direct background measurements were announced last December by Rebecca A. Bernstein of the University of Michigan and her colleagues. Before their work, astronomers had relied on constraints derived by summing up the light from the faintest galaxies seen by the Hubble Space Telescope. In the extreme ultraviolet, the background is obscured by the interstellar medium, so the background level can be estimated only by interpolating between the ultraviolet and x-ray measurements.

Hidden in the Background

TO USE THE BACKGROUND radiation as a quality-control check, astronomers have had to develop ways to compare what is measured with what is expected. That is no easy task. The background represents a tangled mixture of light from various classes of astronomical objects. Starlight, which is produced by thermonuclear fusion, is mainly confined to near-infrared, optical and ultraviolet wavelengths. Quasars and other active galactic nuclei (AGN), whose black holes suck in matter and efficiently convert its gravitational energy into radiation, shine in a very broad band, from radio to gamma wavelengths. Clouds of dust absorb optical, ultraviolet and x-ray light and reradiate the energy in the far-infrared. To complicate matters further, the background blends together light from objects at vastly different cosmic distances and evolutionary stages.

One strategy is to conduct intensive surveys of the sky—to make observations with the highest possible resolution and sensitivity and thereby get a fix on the specific sources of the background. By comparing the findings made at different wavelengths, we can determine what kind of objects these sources are. This direct approach, however, can achieve the requisite precision only for relatively bright objects in very limited areas of the sky. For the broader picture, we turn to a second technique known as population synthesis: calculate the expected emission from possible combinations of objects, compare this prediction with the background measurements, and continue trying different combinations until one seems to fit.

Because the CXB was the first known background, it has been studied more than the other background components. The most basic question—does the CXB come from unresolved sources or a hitherto unknown type of diffuse gas?—was debated for three decades [see "The Origin of the Cosmic X-ray Background," by Bruce Margon; Scientific American, Jan-

THE AUTHORS

GÜNTHER HASINGER and *ROBERTO GILLI* began to work together on the x-ray background radiation at the Astrophysical Institute in Potsdam, Germany, where Hasinger was the director and Gilli did some of the work for his Ph.D. Hasinger, now a director at the Max Planck Institute for Extraterrestrial Physics in Garching, Germany, is primarily an observer. Gilli, now at Arcetri Observatory in Florence, Italy, is primarily a modeler. Hasinger very nearly became a rock musician instead of an astronomer. He played bass guitar and traverse flute and was recording his first album when his mother insisted that he enroll at the University of Munich, lest she lose child-support benefits. He still jams occasionally. Gilli, for his part, was steered away from a career in soccer.

uary 1983]. In the 1990s an indirect line of argument finally settled the issue. If the CXB comes from hot intergalactic gas, the gas should also act as a screen that distorts our view of the cosmic microwave background. The spectrum of the CMB would then deviate from that of a perfect blackbody. Yet CMB observations, notably by the Cosmic Background Explorer satellite, saw no such deviation. Therefore, only a small fraction of the x-ray background can come from such gas; cooler gas might contribute, but for the most part, the CXB must represent unidentified discrete sources.

But what could these sources be? The first intensive surveys to answer this question were performed in the early 1980s with the Einstein x-ray satellite (HEAO-2) by Riccardo Giacconi, the discoverer of the CXB, and others. They resolved about a fifth of the x-ray background into discrete sources, including quasars. The ROSAT satellite followed up this work. In 1984 a group of scientists consisting of Giacconi, Maarten Schmidt (the discoverer of quasars), Joachim Trümper (the father of ROSAT) and one of us (Hasinger) met at the Max Planck Institute for Extraterrestrial Physics in Garching, Germany, to start planning for deep surveys with that satellite. After the ROSAT launch in 1990, the surveys became a major enterprise lasting over a decade and involving a large number of co-workers, more than we could possibly list here.

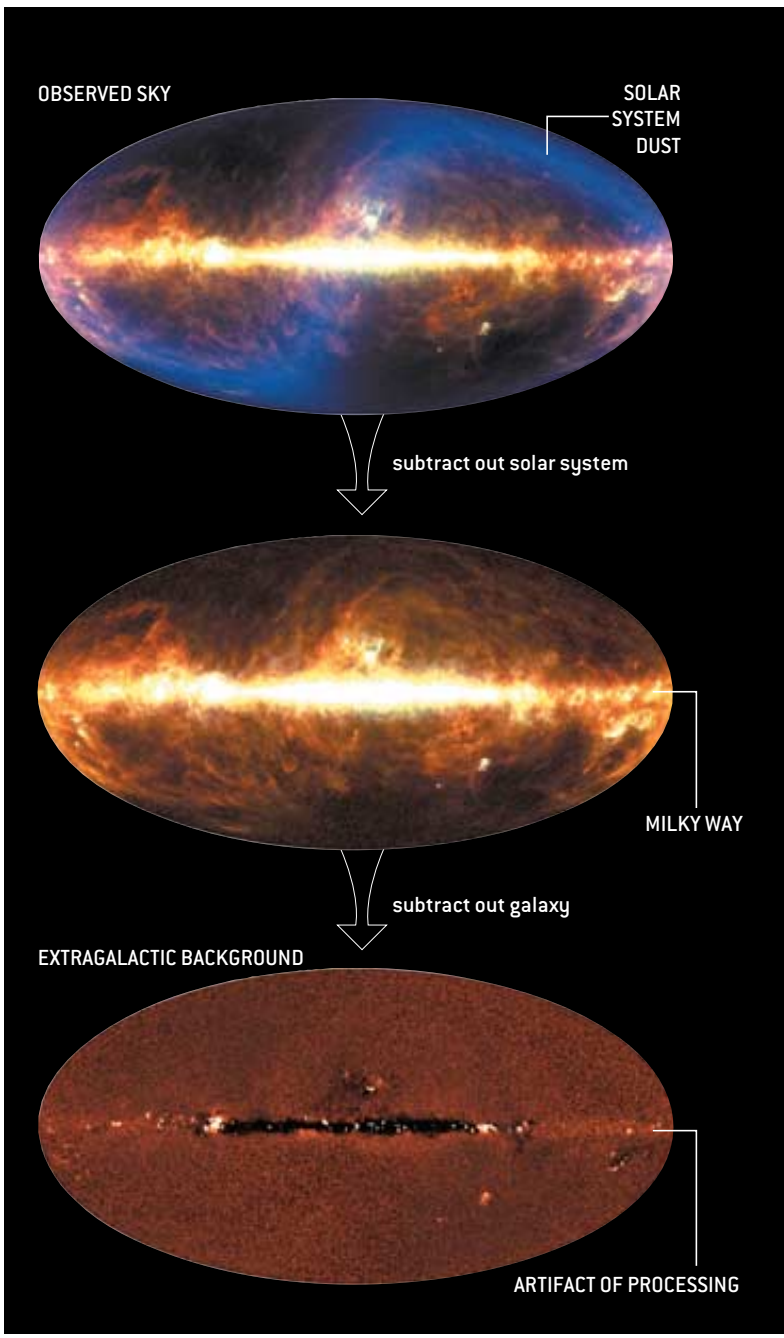
The ROSAT Deep Surveys of the so-called Lockman Hole—an area close to the Big Dipper that is almost free from foreground absorption—are among the longest and deepest x-ray-plus-optical observations ever performed. They have resolved 80 percent of the x-ray background at energies of less than 2 keV, a range that astronomers call soft x-rays. The main bottleneck has been making the optical identifications. We have to look for counterparts of the x-ray sources on deep optical images, and often they are extremely faint. Then we have to obtain their spectra, which reveal the properties of the objects as well as their redshift, a measure of distance. This work would not be possible without the giant Keck telescope, but even its 10-meter mirror has trouble collecting enough light to measure the spectra of the faintest optical counterparts.

About 80 percent of the ROSAT sources have turned out to be active galactic nuclei of various kinds—mostly luminous quasars and so-called Seyfert-1 galaxies. The broad emission lines in the spectra of these AGNs indicate that we have a clear view into their innermost regions, where the monstrous black holes are gorging themselves.

Wallowing in Dust

THE REST OF THE AGNs, however, show only narrow emission lines or no emission lines at all—suggesting that gas and dust block our view of their central black holes. They are classified as type 2 quasars or Seyfert-2 galaxies. The existence of a second type makes sense in the framework of the "unified model" for AGNs. Proposed in the mid-1980s, the unified model posits that all AGNs contain not only a central black hole but also a torus of gas and dust. Depending on how this torus is oriented, it can hide the black hole. The model has since been updated, but the basic prediction has stayed the same: we perceive either an unobscured (type 1) or an obscured (type 2) AGN.

Although these soft-x-ray surveys showed that AGNs are the dominant sources of the x-ray background, an apparent paradox emerged as astronomers began to employ their second strategy to understand the background—namely, population synthesis. When astronomers added together the spectra of dif-



LIKE STRIPPING OFF LAYERS OF PAINT to expose the wall underneath, measuring the cosmic background involves taking an image of the sky, then subtracting light from known objects and seeing what is left. In these images, which combine the far-infrared wavelengths of 60 (*blue*), 100 (*green*) and 240 microns (*red*), the sky is projected so that the Milky Way runs across its center; the plane of the planets has an S shape.

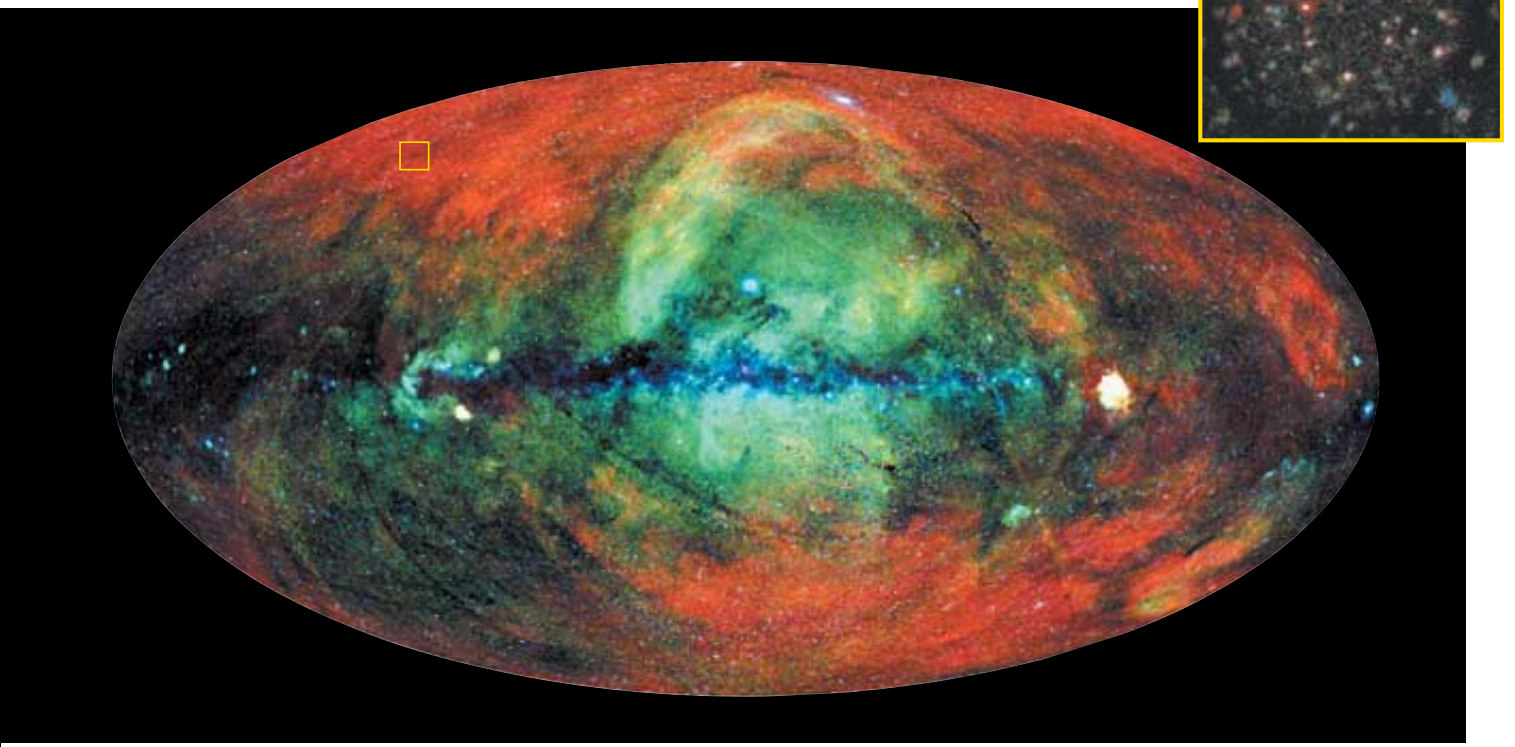
ferent types of AGNs according to their observed proportions, the result should have equaled the spectrum of the CXB. It did not. AGN spectra have a flat or bowl-like shape, whereas the CXB spectrum has a peak at 30 keV.

A solution to this discrepancy was proposed in 1989 by Giancarlo Setti of Bologna University in Italy and Lo Woltjer of Haute-Provence Observatory in France, who at that time were working together at the European Southern Observatory in Garching. They hypothesized that population-synthesis modeling had not added the AGNs in their correct proportions. Contrary to what people had thought, most sources of the x-ray background could be type 2 AGNs. Higher-energy (so-called hard) x-rays can penetrate the dust and gas around these black holes, whereas the soft x-rays are absorbed. In this way, the overall CXB spectrum would differ from that of bright AGNs.

Picking up on this idea, population-synthesis modelers sought the right mixture of type 1 and type 2 AGNs that would explain the CXB spectrum, taking into account how these objects might evolve over time. As shown in 1995 by Andrea Comastri, then at the Max Planck Institute for Extraterrestrial Physics, and his co-workers, such models can reproduce the spectrum up to about 300 keV if the vast majority—80 to 90 percent—of the energy produced by black holes is veiled by thick clouds of gas and dust. If so, these beasts were 100 times more abundant in the early universe than today—a figure consistent with their forming in almost all galaxies. They might have gone unnoticed were it not for the cosmic x-ray background.

A related paradox concerns the optical and infrared backgrounds (the COB and CIB, respectively). The COB is most likely the summed emission of stars, redshifted as the universe expands. The CIB, on the other hand, has the spectrum of dust at a temperature of 10 to 100 kelvins, also redshifted. The energy represented by the dust emission must ultimately originate in stars and AGNs. Yet the CIB is as bright as or brighter than the COB. It is as though the moon (which merely reflects sunlight) were brighter than the sun (the source of that light). The logical resolution of this paradox, like that of the x-ray paradox, is that a substantial fraction of radiation sources in the universe is shrouded by gas and dust.

To confirm these inferences, astronomers have been studying the background radiation at wavelengths that would be unaffected by any obscuring material—namely, hard x-rays. This potent radiation passes through dust as though the dust were not even there. The two great new x-ray observatories now in orbit, the Chandra X-ray Observatory (with superb angular resolution) and XMM-Newton (with a large telescope area), have extended the band covered by ROSAT to substantially higher energies, up to 10 keV, though not yet to the peak of the x-ray background. The most sensitive x-ray surveys to date have been performed with Chandra in two sky areas, the Chandra Deep Field South and the Hubble Deep Field North, by the groups led by Giacconi, who is now at Johns Hopkins University, and by Gordon



IN AN X-RAY IMAGE OF THE SKY, the cosmic background radiation is easy to see: it is the haze in areas away from the Milky Way (*horizontal band*). The ROSAT satellite, which made this image, also took a detailed look at the Lockman Hole, a patch of sky where the background is especially visible (*inset*). The box shows its approximate location. The colors represent high- (*blue*), medium- (*green*) and low-energy (*red*) x-rays.

opposite page: MICHAEL HAUSER *Space Telescope Science Institute* AND NASA; *this page:* MAX PLANCK INSTITUTE FOR EXTRATERRESTRIAL PHYSICS

DISTANT AND DUSTY GALAXIES, the missing link in understanding the cosmic background, show up in the famous Hubble Deep Field (*above*) and a Chandra x-ray image of the same region (*inset*). Comparing the two images allows astronomers to identify these objects. Many are powered by supermassive black holes.

NASA/PSU/G. GARMIRE, N. BRANDT ET AL.

P. Garmire of Pennsylvania State University. These surveys have resolved at least 80 percent of the hard x-ray background.

The optical matchup work has just started. So far the sources are a mixture of type 1 and type 2 AGNs, in excellent agreement with the models. Interestingly, about 10 percent of the x-ray sources discovered by Chandra are very faint galaxies—presumably normal galaxies that contain no AGNs. Their x-ray emission is associated mainly with gas heated by star formation.

Your Friendly Neighborhood ULIRG

THE TWO MAIN STRATEGIES used to study the background leave something to be desired. The intensive surveys push technology to and beyond its limits, and population synthesis is rather abstract. Astronomers have therefore developed a third strategy: scour the nearby universe for counterparts to the distant type 2 galaxies.

They have found their answer in the galaxy NGC 6240. It is one of the black sheep of the Milky Way's neighborhood—a member of an exotic class known as ultraluminous infrared galaxies (ULIRGs). Such galaxies emit most of their total energy output in the far-infrared, a telltale sign that they are saturated with dust. Because dust consists of heavy chemical elements that are synthesized in stars and scattered through space when those stars die, prodigious amounts of dust imply prodigious star formation. Whereas the Milky Way is making a few new stars a year, NGC 6240 must be churning out hundreds. Not only is NGC 6240 wracked by star formation, it is cursed with one of the most voracious black holes in the nearby universe.

The overall spectrum of NGC 6240 has the same shape as that of the cosmic background. It contains all the ingredients we need to explain the background, although we still need to mix them in the right proportions.

Seeing what NGC 6240 looks like, astronomers have realized that the unexpected prevalence of type 2 AGNs in the early universe has a natural explanation: the AGNs were accompanied by bursts of star formation. Stars spewed dust, which hides the holes from our view. Indeed, an accumulating body of evidence indicates that star formation and black-hole feeding were much more common in the past than today. The two processes seem to have hit their peak at roughly the same era in cosmic history.

Why do AGNs and starbursts occur hand in hand? No one yet knows. It seems quite likely that the two processes have the same underlying cause: galaxy collisions, which cause gas to spiral toward the center of the galaxy, whereupon it either forms stars or falls into the hole. Nearly all ULIRGs, including NGC 6240, show signs of having undergone a collision with another galaxy. On the other hand, not all AGNs seem to be associated with major collisions.

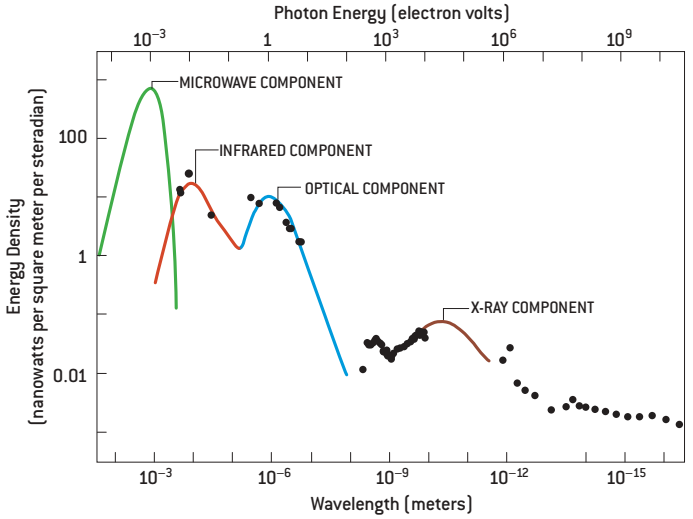
Many researchers think the connection between AGNs and starbursts may be much tighter than merely having a common source of fuel. Black holes could directly stoke the fires of star formation, or stars could help funnel material into the hole. Stars and supermassive holes might even be symbiotic, unable to exist without one another. Such connections might account for the correlation between the properties of galaxies and their central holes [see "The Hole Shebang," by George Musser; SCIENTIFIC AMERICAN, October 2000].

Fortified by studies of NGC 6240 and its ilk, astronomers have used population synthesis to see whether AGNs and starbursts could explain not just the x-ray background but also the optical and infrared backgrounds. The answer appears to be no.

Joint observations by Chandra and the SCUBA instrument, which observes at submillimeter wavelengths between the far-infrared and radio, have failed to note much overlap. Omar Almaini of the Royal Observatory in Edinburgh, Scotland, and his collaborators estimate that up to 30 percent of the cosmic infrared background is ultimately generated by AGNs. Hasinger and his colleagues have combined XMM and Infrared Space Observatory measurements of the Lockman Hole, putting a lower limit—15 percent—on the AGN contribution to the infrared background.

Elese N. Archibald of the Joint Astronomy Center in Hilo, Hawaii, and her co-workers have explained these findings as a natural sequence of galaxy formation. In their scenario, each galaxy forms around a seed black hole of relatively low mass (10 to 1,000 solar masses). At first, starlight dominates the total output of the galaxy, because the little hole still has to grow. The hole does so exponentially by swallowing material as fast as it can. After about 500 million years, the hole is so fat—a billion solar masses—that infalling material outshines the stars. A quasar is born. After a while, this quasar has eaten all the available fuel and falls asleep until new gas falls into the center, waking it up. The hole may also merge with another of like size.

To be sure, some researchers think we may still be missing



COBBLING TOGETHER observations (*dots*) at different wavelengths, astronomers have prepared a spectrum of the cosmic background. Four components (*solid curves*) are obvious; they peak at different wavelengths and represent different ways to generate light. Nearly all the background energy can now be accounted for. (A steradian is about $^{1}/_{13}$ of a sphere.)

JOHNNY JOHNSON; SOURCE: GÜNTHER HASINGER AND ROBERTO GILLI

some crucial piece of the puzzle, such as galaxies that are too spread out to see directly or stars that formed before galaxies did [see "The First Stars in the Universe," by Richard B. Larson and Volker Bromm; SCIENTIFIC AMERICAN, December 2001]. Sources other than AGNs have been proposed for the very high-energy tail of the CXB. For example, a significant fraction of the gamma rays could be produced by electrons catapulted to immense speeds during the formation of the large-scale structure of the universe.

Further intensive surveys are needed to disentangle the various processes that contribute to the background, and future observatories—such as the Space Infrared Telescope Facility, the Herschel Far-Infrared Telescope, the Next Generation Space Telescope and the Atacama Large Millimeter Array—will be required to study some of the objects that x-ray satellites have detected. X-ray spectrometry by the planned XEUS mission could be crucial because it might be able to estimate redshifts from x-ray data alone, thereby allowing observation of objects too heavily obscured to be visible in the optical at all. Such work might finally explain the mysterious link between galaxies and the black holes at their centers, deduce which formed first, and describe how star formation relates to black-hole activity.

The Bright Night Sky

THE STUDY OF THE BACKGROUND is a classic example of how nothing in astronomy is quite what it appears to be. The mere presence of the background indicates that, despite first appearances, the night sky is not completely dark. For most of human history, the darkness of the night sky was taken for granted, and the question was why it was so. In an infinite universe filled with stars, every line of sight should eventually meet the surface of a star. The dimming of starlight with distance should be exactly canceled by the increase in the number of stars you see as you look farther out, so the night sky should appear as bright as the surface of the sun. Day and night should blend into one.

This puzzle, known as Olber's paradox, was solved in 1848 by Edgar Allan Poe. In his prose poem *Eureka,* he argued that the stars must not have had enough time to fill the universe with light. The darkness of the night sky, then, tells us that the universe has not existed forever. Not only has that hypothesis stood the test of time, it eventually proved crucial to formulating the big bang theory.

Still, the night is not pitch-black; it is pervaded by the cosmic background. Although we have made much progress in explaining it, we have much left to do. Whereas 19th-century thinkers had to explain why the night sky isn't bright, modern cosmologists must figure out why it isn't completely dark. SA

MORE TO EXPLORE

The ROSAT Deep Survey I: X-ray Sources in the Lockman Field. G. Hasinger, R. Burg, R. Giacconi, M. Schmidt, J. Trümper and G. Zamorani in *Astronomy & Astrophysics,* Vol. 329, No. 2, pages 482–494; January 1998. Preprint available at **arXiv.org/abs/astro-ph/9709142**

Resolving the Extragalactic Hard X-ray Background. R. F. Mushotzky, L. L. Cowie, A. J. Barger and K. A. Arnaud in *Nature,* Vol. 404, pages 459–464; March 30, 2000. **astro-ph/0002313**

The Cosmic Infrared Background at 1.25 μm and 2.2 μm Using DIRBE and 2MASS: A Contribution Not Due to Galaxies? L. Cambrésy et al. in *Astrophysical Journal,* Vol. 555, No. 2, pages 563–571; July 10, 2001. **astro-ph/0103078**

Wondering in the Dark. Ken Croswell in *Sky & Telescope,* Vol. 102, No. 6, pages 45–50; December 2001.

Testing Current Synthesis Models of the X-ray Background. R. Gilli, M. Salvati and G. Hasinger in *Astronomy & Astrophysics,* Vol. 366, No. 2, pages 407–417; February 2001. **astro-ph/0011341**

For other papers on the background radiation, visit **http://nedwww.ipac.caltech.edu/level5/bcg_radiation.html**

SCARS THAT WON'T HEAL: THE NEUROBIOLOGY OF CHILD ABUSE

Maltreatment at an early age can have enduring negative effects on a child's brain development and function

By Martin H. Teicher

In 1994 Boston police were shocked to discover a malnourished four-year-old locked away in a filthy Roxbury apartment, where he lived in dreadfully squalid conditions. Worse, the boy's tiny hands were found to have been horrendously burned. It emerged that his drug-abusing mother had held the child's hands under a steaming-hot faucet to punish him for eating her boyfriend's food, despite her instructions not to do so. The ailing youngster had been given no medical care at all. The disturbing story quickly made national headlines. Later placed in foster care, the boy received skin grafts to help his scarred hands regain their function. But even though the victim's physical wounds were treated, recent research findings indicate that any injuries inflicted to his developing mind may never truly heal.

Though an extreme example, the notorious case is unfortunately not all that uncommon. Every year child welfare agencies in the U.S. receive more than three million allegations of childhood abuse and neglect and collect sufficient evidence to substantiate more than a million instances.

It is hardly surprising to us that research reveals a strong link between physical, sexual and emotional mistreatment of children and the development of psychiatric problems. But in the early 1990s mental health professionals believed that emotional and social difficulties occurred mainly through psychological means. Childhood maltreatment was understood either to foster the development of intrapsychic defense mechanisms that proved to be self-defeating in adulthood or to arrest psychosocial development, leaving a "wounded child" within. Researchers thought of the damage as basically a software problem amenable to reprogramming via therapy or simply erasable through the exhortation "Get over it."

New investigations into the consequences of early maltreatment, including work my colleagues and I have done at McLean Hospital in Belmont, Mass., and at Harvard Medical School, appear to tell a different story. Because childhood abuse occurs during the critical formative

STUART BRADFORD

More than THREE MILLION allegations of childhood ABUSE and NEGLECT are received every year.

time when the brain is being physically sculpted by experience, the impact of severe stress can leave an indelible imprint on its structure and function. Such abuse, it seems, induces a cascade of molecular and neurobiological effects that irreversibly alter neural development.

Extreme Personalities

THE AFTERMATH of childhood abuse can manifest itself at any age in a variety of ways. Internally it can appear as depression, anxiety, suicidal thoughts or posttraumatic stress; it can also be expressed outwardly as aggression, impulsiveness, delinquency, hyperactivity or substance abuse. One of the more perplexing psychiatric conditions that is strongly associated with early ill-treatment is borderline personality disorder. Someone with this dysfunction characteristically sees others in black-and-white terms, often first putting a person on a pedestal, then vilifying the same person after some perceived slight or betrayal. Those afflicted are also prone to volcanic outbursts of anger and transient episodes of paranoia or psychosis. They typically have a history of intense, unstable relationships, feel empty or unsure of their identity, commonly try to escape through substance abuse, and experience self-destructive or suicidal impulses.

While treating three patients with borderline personality disorder in 1984, I began to suspect that their early exposure to various forms of maltreatment had altered the development of their limbic systems. The limbic system is a collection of interconnected brain nuclei (neural centers) that play a pivotal role in the regulation of emotion and memory. Two critically important limbic regions are the hippocampus and the amygdala, which lie below the cortex in the temporal lobe [*see illustration on opposite page*]. The hippocampus is thought to be important in the formation and retrieval of both verbal and emotional memories, whereas the amygdala is concerned with creating the emotional content of memory—for example, feelings relating to fear conditioning and aggressive responses.

My McLean colleagues Yutaka Ito and Carol A. Glod and I wondered whether childhood abuse might disrupt the healthy maturation of these brain regions. Could early maltreatment stimulate the amygdala into a state of heightened electrical irritability or damage the developing hippocampus through excessive exposure to stress hormones? We reasoned further that hippocampal harm or amygdaloid overexcitation could produce symptoms similar to those experienced by patients with temporal lobe epilepsy (TLE), which sporadically disrupts the function of these brain nuclei. During TLE seizures, patients remain conscious while experiencing a range of psychomotor symptoms brought on by electrical storms within these regions. Associated effects include the abrupt onset of tingling, numbness or vertigo; motor-related manifestations such as uncontrollable staring or twitching; and autonomic symptoms such as flushing, nausea or the "pit in your stomach" feeling one gets in a fast-rising elevator. TLE can also cause hallucinations or illusions in any of the five senses. It is not unusual, for instance, for one afflicted with this condition to experience Alice-in-Wonderland-like distortions of the sizes or shapes of objects. Disconnected feelings of déjà vu and mind-body dissociation are also common.

Abuse-Driven Brain Changes

TO EXPLORE the relation between early abuse and dysfunction of the limbic system, in 1984 I devised a checklist of questions that assess the frequency with which patients experience TLE-related symptoms. In 1993 my co-workers and I reported results from 253 adults who came to an outpatient mental health clinic for psychiatric evaluation. Slightly more than half reported having been abused physically or sexually, or both, as children. Compared with patients who reported no ill-treatment, average checklist scores were 38 percent greater in the patients with physical (but not sexual) abuse and 49 percent higher in the patients with sexual (but not other physical) mistreatment.

Overview/*Insight into Child Abuse*

- Until recently, psychologists believed that mistreatment during childhood led to arrested psychosocial development and self-defeating psychic defense mechanisms in adults. New brain imaging surveys and other experiments have shown that child abuse can cause permanent damage to the neural structure and function of the developing brain itself.
- This grim result suggests that much more effort must be made to prevent childhood abuse and neglect before it does irrevocable harm to millions of young victims. New approaches to therapy may also be indicated.

Patients who acknowledged both physical and sexual abuse had average scores 113 percent higher than patients reporting none. Maltreatment before age 18 had more impact than later abuse, and males and females were similarly affected.

In 1994 our McLean research team sought to ascertain whether childhood physical, sexual or psychological abuse was associated with brain-wave abnormalities in electroencephalograms (EEGs), which provide a more direct measure of limbic irritability than our checklist. We reviewed the records of 115 consecutive admissions to a child and adolescent psychiatric hospital to search for a link. We found clinically significant brain-wave abnormalities in 54 percent of patients with a history of early trauma but in only 27 percent of nonabused patients. We observed EEG anomalies in 72 percent of those who had documented histories of serious physical and sexual abuse. The irregularities arose in frontal and temporal brain regions and, to our surprise, specifically involved the left hemisphere rather than both sides, as one would expect.

Our findings dovetailed with a 1978 EEG study of adults who were victims of incest. The study's author, Robert W. Davies of the Yale University School of Medicine, and his team had found that 77 percent exhibited EEG abnormalities and 27 percent experienced seizures.

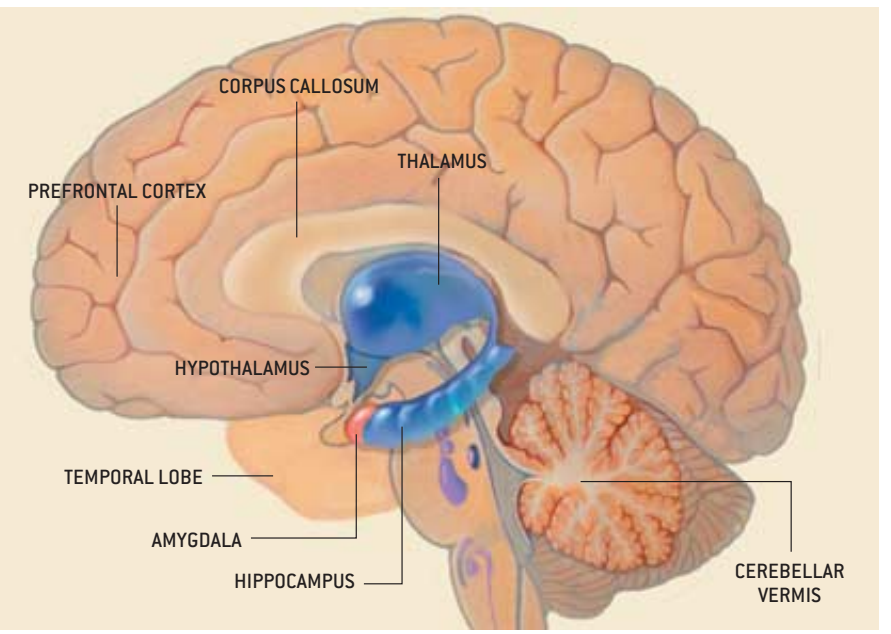
Subsequent work by other investigators using magnetic resonance imaging (MRI) technology has confirmed an association between early maltreatment and reductions in the size of the adult hippocampus. The amygdala may be smaller as well. In 1997 J. Douglas Bremner, then at the Yale University School of Medicine, and his colleagues compared MRI scans of 17 adult survivors of childhood physical or sexual abuse, all of whom had posttraumatic stress disorder (PTSD), with 17 healthy subjects matched for age, sex, race, handedness, years of education, and years of alcohol abuse. The left hippocampus of abused patients with PTSD was, on average, 12 percent smaller than the hippocampus of the healthy control subjects, but the right hippocampus was of normal size. Not surprisingly, given the important role of the hippocampus in memory function, these patients also scored lower on verbal memory tests than the nonabused group.

In 1997 Murray B. Stein of the University of California at San Diego also found left hippocampal abnormalities in 21 adult women who had been sexually abused as children and who had PTSD or dissociative identity disorder (also called multiple personality disorder, a condition thought by some researchers to be common in abused females). Stein determined that in these women the volume of the left hippocampus was significantly reduced but that the right hippocampus was relatively unaffected. In addition, he found a clear correspondence between the degree of reduction in hippocampus size and the severity of the patients' dissociative symptoms. In 2001 Martin Driessen of Gilead Hospital in Bielefeld, Germany, and his colleagues reported a 16 percent reduction in hippocampus size and an 8 percent reduction in amygdala size in adult women with borderline personality disorder and a history of childhood maltreatment.

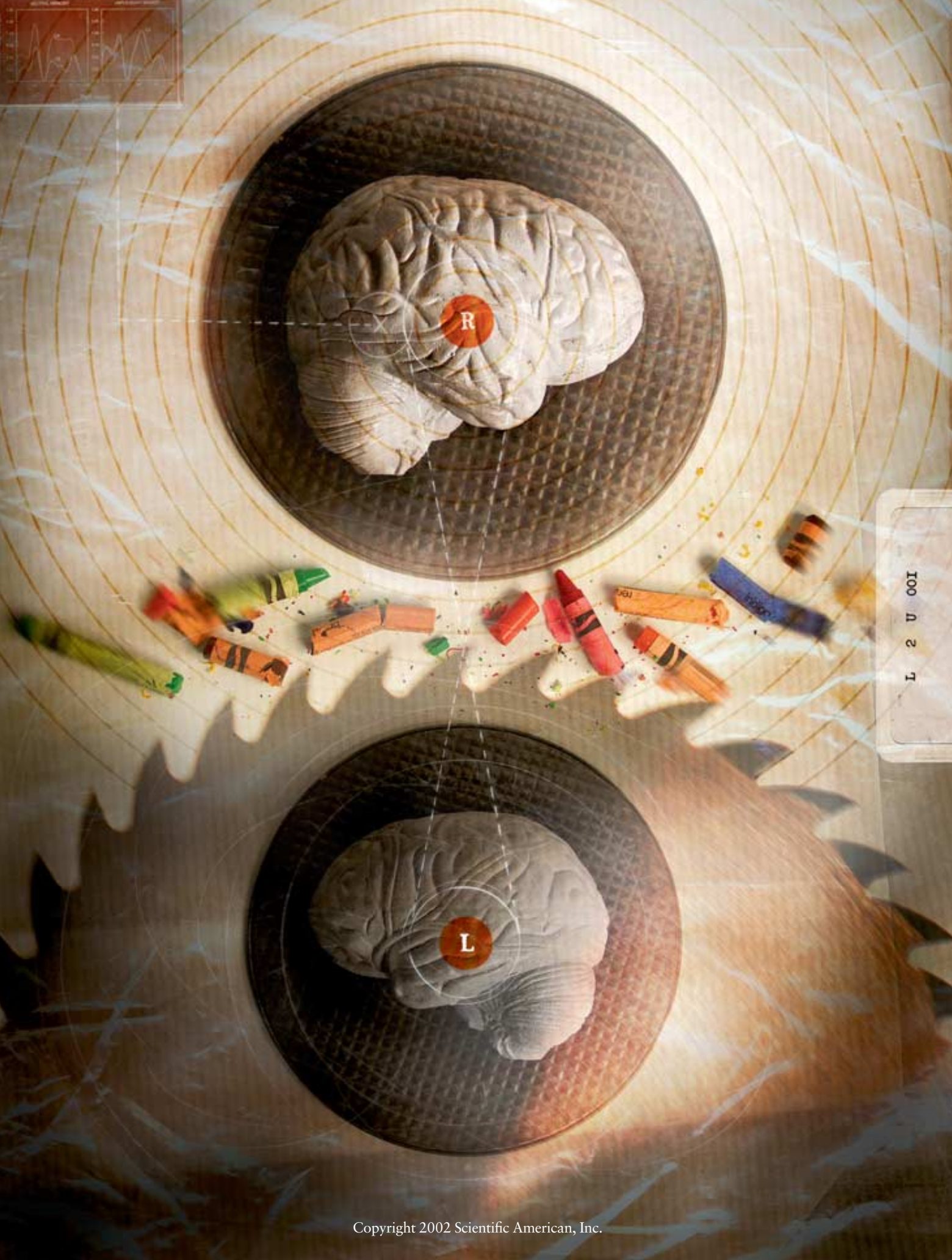
On the other hand, when Michael D. De Bellis and his colleagues at the University of Pittsburgh School of Medicine carefully measured MRI images of the hippocampus in 44 maltreated children with PTSD and 61 healthy control subjects in 1999, they failed to observe a significant difference in volume.

My McLean colleagues Susan Andersen and Ann Polcari and I obtained similar results in our recently completed volumetric analysis of the hippocampus in 18 young adults (18 to 22 years of age) with a history of repeated forced sexual abuse accompanied by fear or terror, who were compared with 19 healthy age-matched controls. Unlike in previous studies, the control subjects were not patients but were recruited from the general public and had fewer mental health problems. We observed no differences in hippocampal volume. Like Driessen's group, however, we did find a 9.8 percent average reduction in the size of the left amygdala, which correlated with feelings of depression and irritability or hostility. We asked ourselves why the hippocampus was smaller in abused subjects in

ANTISOCIAL BEHAVIOR resulting from childhood abuse appears to be caused by overexcitation of the limbic system, the primitive midbrain region that regulates memory and emotion. Two relatively small, deep-lying brain structures—the hippocampus and the amygdala—are thought to play prominent roles in generating this kind of interpersonal dysfunction. The hippocampus is important in determining what incoming information will be stored in long-term memory. The principal task of the amygdala is to filter and interpret incoming sensory information in the context of the individual's survival and emotional needs and then to help initiate appropriate responses.



CAROL DONNER



Researchers thought of the damage as a SOFTWARE PROBLEM amenable to reprogramming via therapy.

studies from Bremner's, Stein's and Dreissen's groups but normal in De Bellis's and in our own investigations. Of the several possible answers, the most likely is that stress exerts a very gradual influence on the hippocampus, so adverse effects may not be discernible at a gross anatomical level until people get older.

Moreover, animal studies by Bruce S. McEwen of the Rockefeller University and Robert M. Sapolsky of Stanford University had previously demonstrated the marked vulnerability of the hippocampus to the ravages of stress. Not only is the hippocampus particularly susceptible because it develops slowly, it also is one of the few brain regions that continues to grow new neurons after birth. Further, it has a higher density of receptors for the stress hormone cortisol than almost any other area of the brain. Exposure to stress hormones can significantly change the shape of the largest neurons in the hippocampus and can even kill them. Stress also suppresses production of the new granule cells (small neurons), which normally continue to develop after birth.

Experiments with rats by Christian Caldji, Michael J. Meaney of McGill University and Paul M. Plotsky of Emory University have shown that early stress reconfigures the molecular organization of these regions. One major result is the alteration of the protein subunit structure of GABA receptors in the amygdala [*see illustration on next page*]. These receptors respond to gamma aminobutyric acid, the brain's primary inhibitory neurotransmitter, and GABA attenuates the electrical excitability of neurons. Reduced function of this neurotransmitter produces excessive electrical activity and can trigger seizures. This discovery provides an elegant molecular explanation for our findings of EEG abnormalities and limbic irritability in patients with childhood abuse.

Left-Side Problems

THE EFFECT on the limbic system was only the most expected consequence of childhood trauma. We were intrigued, however, by our earlier observation that ill-treatment was associated with EEG abnormalities in the left hemisphere. This inspired us to examine the effect of early abuse on the development of the left and right hemispheres. We chose to use EEG coherence, a sophisticated quantitative analysis method that provides evidence about the brain's microstructure—its wiring and circuitry. Conventional EEG, in contrast, reveals brain function. The EEG coherence technique accomplishes its task by generating a mathematical measure of the degree of cross-correlation among the elaborate neuronal interconnections in the cortex that process and modify the brain's electrical signals. In general, abnormally high levels of EEG coherence are evidence of diminished development among these neuron interchanges.

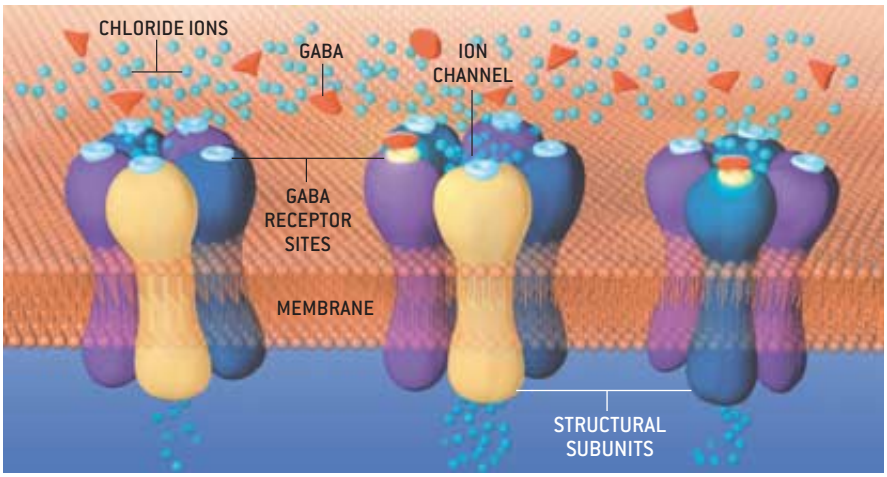
Our research team used this technique in 1997 to compare 15 healthy volunteers with 15 child and adolescent psychiatric patients who had a confirmed history of intense physical or sexual abuse. Coherence measures showed that the left cortices of the healthy control subjects were more developed than the right cortices, a result that is consistent with what is known about dominant hemisphere anatomy—that is, right-handed people tend to be left-cortex dominant. The maltreated patients, however, were notably more developed in the right cortex than the left, even though all were right-handed and hence left-dominant. The right hemispheres of abused patients had developed as much as the right hemispheres of the control subjects, but their left hemispheres lagged substantially behind. This anomalous result showed up regardless of the patient's primary diagnosis. And although the effect extended throughout the entire left hemisphere, the temporal regions were most affected, which supported our original hypothesis.

The left hemisphere is specialized for perceiving and expressing language, whereas the right hemisphere specializes in processing spatial information and in processing and expressing emotions—particularly negative emotions. We had wondered whether mistreated children might store their disturbing memories in the right hemisphere and whether recollecting these memories might preferentially activate the right hemisphere.

To test this hypothesis, Fred Schiffer worked in my laboratory at McLean in 1995 to measure hemispheric activity in adults during recall of a neutral memory and then during recall of an upsetting early memory. Those with a history of abuse appeared to use predominantly their left hemispheres when thinking about neutral memories and their right when recalling an early disturbing memory. Subjects in the control group used both hemispheres to a comparable degree for either task, suggesting that their responses were more integrated between the two hemispheres.

Because Schiffer's research indicated that childhood trauma was associated with diminished right-left hemisphere integration, we decided to look for some deficiency in the primary pathway for infor-

STUART BRADFORD



FEWER INHIBITIONS: Stress causes changes to normal postsynaptic receptors (*left*) for gamma aminobutyric acid (GABA), the major inhibitory neurotransmitter in the central nervous system. It may lead to overstimulation of neurons, resulting in limbic system irritability. The presence of GABA lowers the electrical excitability of neurons by allowing greater flow of chloride ions (*center*). Loss of one of the GABA receptor's key structural subunits impairs its ability to moderate neural activity (*right*).

mation exchange between the two hemispheres, the corpus callosum. In 1997 Andersen and I collaborated with Jay Giedd of the National Institute of Mental Health to search for the posited effect. Together we found that in boys who had been abused or neglected, the middle parts of the corpus callosum were significantly smaller than in the control groups. Furthermore, in boys, neglect exerted a far greater effect than any other kind of maltreatment. In girls, however, sexual abuse was a more powerful factor, associated with a major reduction in size of the middle parts of the corpus callosum. These results were replicated and extended in 1999 by De Bellis. Likewise, the effects of early experience on the development of the corpus callosum have been confirmed by research in primates by Mara M. Sanchez of Emory.

Our latest finding had its roots in the seminal studies of Harry F. Harlow of the University of Wisconsin–Madison. In the 1950s Harlow compared monkeys raised by their mothers with monkeys reared by wire or terrycloth surrogate mothers. Monkeys raised with the surrogates became socially deviant and highly aggressive adults. Working with Harlow, W. A. Mason of the Delta Primate Center in Louisiana discovered that these consequences were less severe if the surrogate mother was swung from side to side. J. W. Prescott of the National Institute of Child Health and Human Development hypothesized that this movement would be conveyed to the cerebellum, particularly the middle part, called the cerebellar vermis, located at the back of the brain just above the brain stem. Among other functions, the vermis modulates the brain-stem nuclei that control the production and release of the neurotransmitters norepinephrine and dopamine. Like the hippocampus, this part of the brain develops gradually and continues to create neurons after birth. It has an even higher density of receptors for stress hormones than the hippocampus, so exposure to such hormones can strongly affect its development.

Abnormalities in the cerebellar vermis have recently been reported to be associated with various psychiatric disorders, including manic-depressive illness, schizophrenia, autism and attention-deficit/hyperactivity disorder. These maladies emerge from genetic and prenatal factors, not childhood mistreatment, but the fact that vermal anomalies seem to sit at the core of so many psychiatric conditions suggests that this region plays a critical role in mental health.

Dysregulation of the vermis-controlled neurotransmitters norepinephrine and dopamine can produce symptoms of depression, psychosis and hyperactivity as well as impair attention. Activation of the dopamine system has been associated with a shift to a more left hemisphere–biased (verbal) attentional state, whereas activation of the norepinephrine system shifts attention to a more right hemisphere–biased (emotional) state. Perhaps most curiously, the vermis also helps to regulate electrical activity in the limbic system, and vermal stimulation can suppress seizure activity in the hippocampus and amygdala.

R. G. Heath, working at Tulane University in the 1950s, found that Harlow's monkeys had seizure foci in their fastigial nuclei and hippocampus. In later work with humans, he found that electrical stimulation of the vermis reduced the frequency of seizures and improved the mental health in a small number of patients with intractable neuropsychiatric disorders. This result led my colleagues and me to speculate whether childhood abuse could produce abnormalities in the cerebellar vermis that contributed to psychiatric symptoms, limbic irritability and gradual hippocampal degeneration.

To begin to test this hypothesis, Carl M. Anderson recently worked in tandem with me and with Perry Renshaw at the Brain Imaging Center at McLean. Anderson used T2-relaxometry methods, a new MRI-based functional imaging technique we developed. For the first time, we can monitor regional cerebral blood flow at rest without the use of radioactive tracers or contrast dyes.

When the brain is resting, the neuronal activity of a region closely matches the amount of blood that area receives to sustain this activity. Anderson found a striking correlation between the activity in the cerebellar vermis and the degree of limbic irritability indicated by my TLE-related question checklist in both healthy young adult controls and young adults

CAROL DONNER

THE AUTHOR

MARTIN H. TEICHER is an associate professor of psychiatry at Harvard Medical School, director of the Developmental Biopsychiatry Research Program at McLean Hospital in Belmont, Mass., and chief of the Developmental Psychopharmacology Laboratory at the Mailman Research Center at McLean.

STRESS sculpts the brain to exhibit various ANTISOCIAL, though adaptive, behaviors.

with a history of repeated sexual abuse.

At any level of limbic symptomatology, however, the amount of blood flow in the vermis was markedly decreased in the individuals with a history of trauma. Low blood flow points to a functional impairment in the activity of the cerebellar vermis. On average, abused patients had higher checklist scores presumably because their vermis could not activate sufficiently to quell higher levels of limbic irritability.

Together these findings suggest an intriguing model that explains one way in which borderline personality disorder can emerge. Reduced integration between the right and left hemispheres and a smaller corpus callosum may predispose these patients to shift abruptly from left- to right-dominated states with very different emotional perceptions and memories. Such polarized hemispheric dominance could cause a person to see friends, family and co-workers in an overly positive way in one state and in a resoundingly negative way in another—which is the hallmark of this disorder. Moreover, limbic electrical irritability can produce symptoms of aggression, exasperation and anxiety. Abnormal EEG activity in the temporal lobe is also often seen in people with a greatly increased risk for suicide and self-destructive behavior.

Adaptive Detriment

OUR TEAM INITIATED this research with the hypothesis that early stress was a toxic agent that interfered with the normal, smoothly orchestrated progression of brain development, leading to enduring psychiatric problems. Frank W. Putnam of Children's Hospital Medical Center of Cincinnati and Bruce D. Perry of the Alberta Mental Health Board in Canada have now articulated the same hypothesis. I have come to question and reevaluate our starting premise, however. Human brains evolved to be molded by experience, and early difficulties were routine during our ancestral development. Is it plausible that the developing brain never evolved to cope with exposure to maltreatment and so is damaged in a nonadaptive manner? This seems most unlikely. The logical alternative is that exposure to early stress generates molecular and neurobiological effects that alter neural development in an adaptive way that prepares the adult brain to survive and reproduce in a dangerous world.

What traits or capacities might be beneficial for survival in the harsh conditions of earlier times? Some of the more obvious are the potential to mobilize an intense fight-or-flight response, to react aggressively to challenge without undue hesitation, to be at heightened alert for danger and to produce robust stress responses that facilitate recovery from injury. In this sense, we can reframe the brain changes we observed as adaptations to an adverse environment.

Although this adaptive state helps to take the affected individual safely through the reproductive years (and is even likely to enhance sexual promiscuity), which are critical for evolutionary success, it comes at a high price. McEwen has recently theorized that overactivation of stress response systems, a reaction that may be necessary for short-term survival, increases the risk for obesity, type II diabetes and hypertension; leads to a host of psychiatric problems, including a heightened risk of suicide; and accelerates the aging and degeneration of brain structures, including the hippocampus.

We hypothesize that adequate nurturing and the absence of intense early stress permits our brains to develop in a manner that is less aggressive and more emotionally stable, social, empathic and hemispherically integrated. We believe that this process enhances the ability of social animals to build more complex interpersonal structures and enables humans to better realize their creative potential.

Society reaps what it sows in the way it nurtures its children. Stress sculpts the brain to exhibit various antisocial, though adaptive, behaviors. Whether it comes in the form of physical, emotional or sexual trauma or through exposure to warfare, famine or pestilence, stress can set off a ripple of hormonal changes that permanently wire a child's brain to cope with a malevolent world. Through this chain of events, violence and abuse pass from generation to generation as well as from one society to the next. Our stark conclusion is that we see the need to do much more to ensure that child abuse does not happen in the first place, because once these key brain alterations occur, there may be no going back. SA

MORE TO EXPLORE

Developmental Traumatology, Part 2: Brain Development. M. D. De Bellis, M. S. Keshavan, D. B. Clark, B. J. Casey, J. N. Giedd, A. M. Boring, K. Frustaci and N. D. Ryan in *Biological Psychiatry,* Vol. 45, No. 10, pages 1271–1284; May 15, 1999.

Wounds That Time Won't Heal: The Neurobiology of Child Abuse. Martin H. Teicher in *Cerebrum* (Dana Press), Vol. 2, No. 4, pages 50–67; Fall 2000.

McLean Hospital: **www.mcleanhospital.org/**

Did extraterrestrial collisions capable of causing widespread extinctions pound the earth not once, but twice— or even several times?

Repeated Blows

By Luann Becker

Most people are unaware of it, but our planet is under a constant barrage by the cosmos. Our galactic neighborhood is littered with comets, asteroids and other debris left over from the birth of the solar system. Most of the space detritus that strikes the earth is interplanetary dust, but a few of these cosmic projectiles have measured five kilometers (about 3.1 miles) or more across. Based on the number of craters on the moon, astronomers estimate that about 60 such giant space rocks slammed into the earth during the past 600 million years. Even the smallest of those collisions would have left a scar 95 kilometers (about 60 miles) wide and would have released a blast of kinetic energy equivalent to detonating 10 million megatons of TNT.

Such massive impacts are no doubt capable of triggering drastic and abrupt changes to the planet and its inhabitants. Indeed, over the same time period the fossil record reveals five great biological crises in which, on average, more than half of all living species ceased to exist. After a period of heated controversy, scientists began to accept that an asteroid impact precipitated one of these catastrophes: the demise of the dinosaurs 65 million years ago. With that one exception, however, compelling evidence for large impacts coincident with severe mass extinctions remained elusive—until recently.

During the past two years, researchers have discovered new methods for assessing where and when impacts occurred, and the evidence connecting them to other widespread die-offs is getting stronger. New tracers of impacts are cropping up, for instance, in rocks laid down at the end of the Permian period—the time 250 million years ago when a mysterious event known as the Great Dying wiped out 90 percent of the planet's species. Evidence for impacts associated with other extinctions is tenuous but growing stronger as well.

Scientists find such hints of multiple life-altering impacts in a variety of forms. Craters and shattered or shocked rocks—the best evidence of an ancient impact—are turning up at key time intervals that suggest a link with extinction. But more often than not, this kind of physical evidence is buried under thick layers of sediment or is obscured by erosion. Researchers now understand that the biggest blows also leave other direct, as well as indirect, clues hidden in the rock record. The first direct tracers included tiny mineral crystals that had been fractured or melted by the blast. Also found in fallout layers have been elements known to form in space but not on the earth. Indeed, my colleagues and I have discovered extraterrestrial gases trapped inside carbon molecules called fullerenes in several suspected impact-related sediments and craters.

Equally intriguing are the indirect tracers that paleontologists have recognized: rapid die-offs of terrestrial vegetation and abrupt declines in the productivity of marine organisms coincident with at least three of the five great extinctions. Such severe and rapid perturbations in the earth's ecosystem are rare, and some scientists suspect that only a catastrophe as abrupt as an impact could trigger them.

The evidence for impacts acting as culprits in widespread die-offs is getting stronger.

Overview/*Deadly Barrage?*

- About 60 meteorites five or more kilometers across have hit the earth in the past 600 million years. The smallest ones would have carved craters some 95 kilometers wide.
- Most scientists agree that one such impact did in the dinosaurs, but evidence for large collisions coincident with other mass extinctions remained elusive—until recently.
- Researchers are now discovering hints of ancient impacts at sites marking history's top five mass extinctions, the worst of which eliminated 90 percent of all living species.

Dinosaur Killer

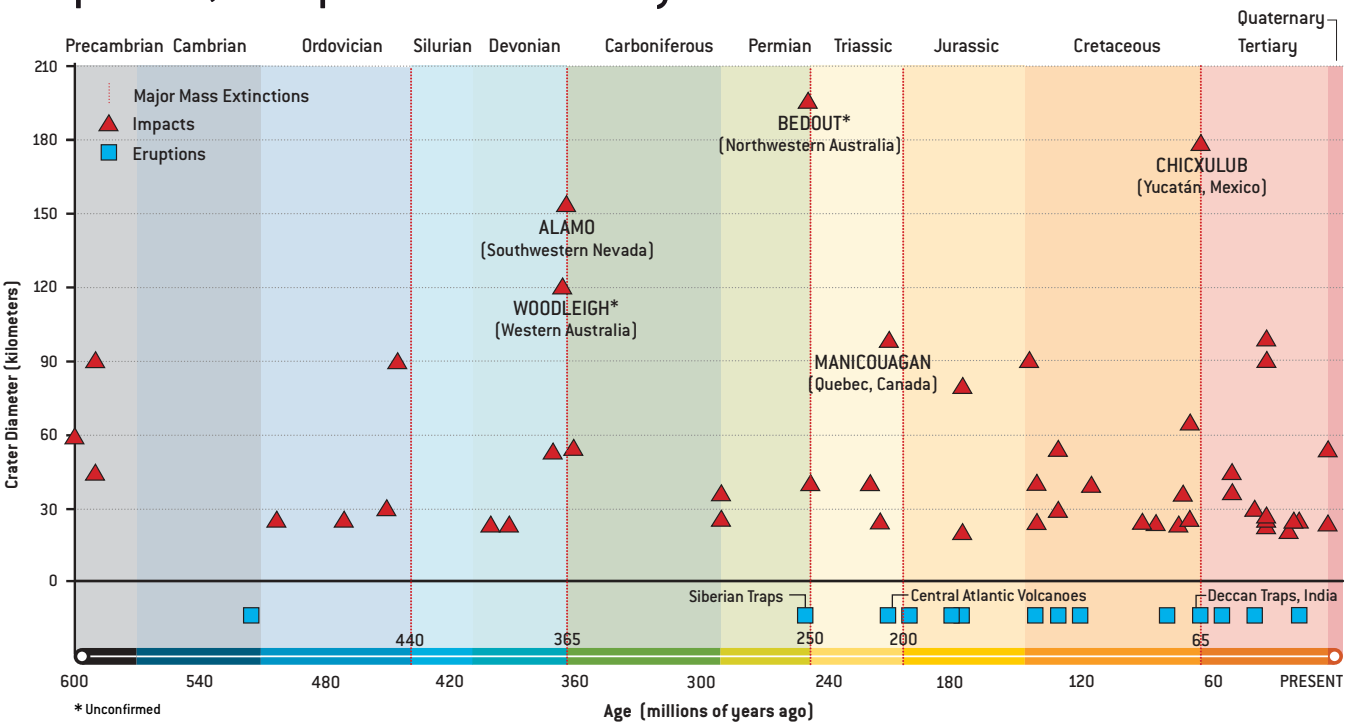
THE FIRST IMPACT TRACER linked to a severe mass extinction was an unearthly concentration of iridium, an element that is rare in rocks on our planet's surface but abundant in many meteorites. In 1980 a team from the University of California at Berkeley—led by Nobel Prize–winning physicist Luis Alvarez and his son, geologist Walter Alvarez—reported a surprisingly high concentration of this element within a centimeter-thick layer of clay exposed near Gubbio, Italy. The Berkeley team calculated that the average daily delivery of cosmic dust could not account for the amount of iridium it measured. Based on these findings, the scientists hypothesized that it was fallout from a blast created when an asteroid, some 10 to 14 kilometers (six to nine miles) across, collided with the earth.

Even more fascinating, the clay layer had been dated to 65 million years ago, the end of the Cretaceous period. From this iridium discovery came the landmark hypothesis that a giant impact ended the reign of the dinosaurs—and that such events may well be associated with other severe mass extinctions over the past 600 million years. Twenty years ago this bold and sweeping claim stunned scientists, most of whom had been content to assume that the dinosaur extinction was a gradual process initiated by a contemporaneous increase in global volcanic activity. The announcement led to intense debates and reexaminations of end Cretaceous rocks around the world.

Out of this scrutiny emerged three additional impact tracers: dramatic disfigurations of the earthly rocks and plant life in the form of microspherules, shocked quartz and high concentrations of soot. In 1981 Jan Smit, now at the Free University in Amsterdam, uncovered microscopic droplets of glass, called microspherules, which he argued were products of the

KAMIL VOJNAR (*preceding two pages*)

Impacts, Eruptions and Major Mass Extinctions



rapid cooling of molten rock that splashed into the atmosphere during the impact. Three years later Bruce Bohor and his colleagues at the U.S. Geological Survey were among the first researchers to explain the formation of shocked quartz. Few earthly circumstances have the power to disfigure quartz, which is a highly stable mineral even at high temperatures and pressures deep inside the earth's crust.

At the time microspherules and shocked quartz were introduced as impact tracers, some still attributed them to extreme volcanic activity. Powerful eruptions can indeed fracture quartz grains—but only in one direction, not in the multiple directions displayed in Bohor's samples. The microspherules contained trace elements that were markedly distinct from those formed in volcanic blasts. Scientists subsequently found enhanced iridium levels at more than 100 end Cretaceous sites worldwide and shocked quartz at more than 30 sites.

Least contentious of the four primary impact tracers to come out of the 1980s were soot and ash, which measured tens of thousands of times higher than normal levels, from impact-triggered fires. The most convincing evidence to support the impact scenario, however, was the recognition of the crater itself, known today as Chicxulub, in Yucatán, Mexico. Shortly after the Alvarez announcement in 1980, geophysicists Tony Camargo and Glen Penfield of the Mexican national oil company, PEMEX, reported an immense circular pattern—later estimated to be some 180 kilometers (about 110 miles) across—while surveying for new oil and gas prospects buried in the Gulf of Mexico. Other researchers confirmed the crater's existence in 1991.

Finding a reasonable candidate for an impact crater marked a turning point in the search for the causes of extreme climate perturbations and mass extinctions—away from earthly sources such as volcanism and toward a singular, catastrophic event. Both volcanoes and impacts eject enormous quantities of toxic pollutants such as ash, sulfur and carbon dioxide into the atmosphere, triggering severe climate change and environmental degradation. The difference is in the timing. The instantaneous release from an impact would potentially kill off species in a few thousand years. Massive volcanism, on the other hand, continues to release its pollutants over millions of years, drawing out its effects on life and its habitats.

While geologists were searching for craters and other impact tracers, paleontologists were adding their own momentum to the impact scenario. Fossil experts had long been inclined to agree with the volcanism theory because the disappearance of species in the fossil record appeared to be gradual. A convincing counterargument came from paleontologists Philip Signor of the University of California at Davis and Jere Lipps,

THE AUTHOR

LUANN BECKER has studied impact tracers since she began her career as a geochemist at the Scripps Institution of Oceanography in La Jolla, Calif., in 1990. In 1998 Becker participated in a meteorite-collecting expedition in Antarctica and in July 2001 was awarded the National Science Foundation Antarctic Service Medal. The following month she joined the faculty at the University of California, Santa Barbara, where she continues to study fullerenes and exotic gases trapped within them as impact tracers. This summer she and her colleagues will conduct fieldwork at end Permian extinction sites in South Africa and Australia. Part of this expedition will be included in a television documentary, scheduled to air this fall, about mass extinctions and their causes.

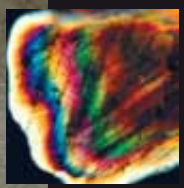
AARON FIRTH (BASED ON GRAPHIC BY MICHAEL PAINE)

Enduring Traces

Craters are the best evidence for an impact, but ejecta from the affiliated blast contains other clues that can settle to the earth and persist in the rock record for millions of years. Such impact tracers are especially prevalent with large, devastating collisions like the hypothetical one illustrated here: an asteroid 10 kilometers (six miles) wide slams into a coastline, transmitting temperatures of several thousand degrees and pressures a million times greater than the weight of the earth's atmosphere.

IMPACT TRACER

SHOCKED MINERALS

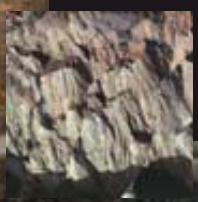


Extreme pressure and heat fracture quartz crystals (*left*) and metamorphose iron-nickel-silica grains.

IMPACT TRACER

DISFIGURED ROCKS

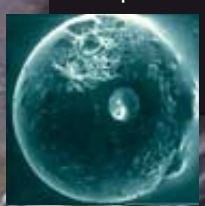
Shock waves are captured in rock as shattercones (*left*). Bedrock fractures; some ejected debris resettles as breccia.



IMPACT TRACER

MICROSPHERULES

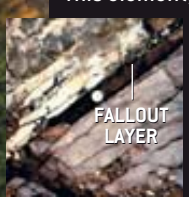
Tiny glass droplets form during the rapid cooling of molten rock that splashes into the atmosphere.



IMPACT TRACER

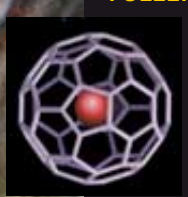
IRIDIUM

This element, which is rare in earthly rocks but abundant in some meteorites, may be preserved in a fallout layer of clay.



IMPACT TRACER

EXTRATERRESTRIAL FULLERENES

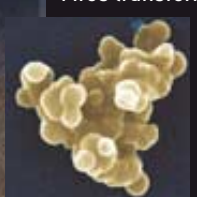


Caged carbon molecules trap extraterrestrial noble gases in space and travel to the earth in the impactor.

IMPACT TRACER

SOOT AND ASH

Fires transform vegetation into soot that accumulates to levels tens of thousands of times higher than normal.



INITIAL DEVASTATION

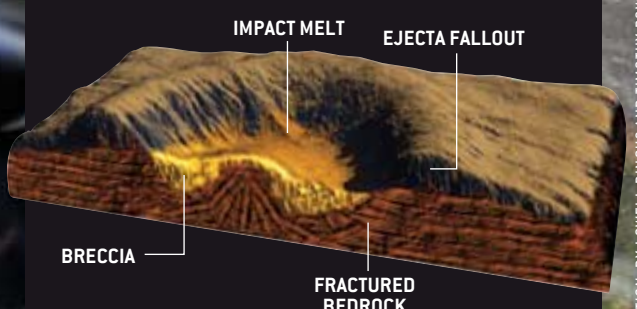
INTO ORBIT
The explosion ejects some 21,000 cubic kilometers (5,000 cubic miles) of debris, about 1,700 cubic kilometers of which is launched into orbit at 50 times the speed of sound.

CHOKED SKY
Little sunlight can penetrate to the ground for several months as ejected debris rains through the atmosphere, and temperatures drop below freezing for up to half a year.

KILLER WAVES
Tsunamis as high as 90 meters (300 feet) destroy coastal ecosystems within hundreds or even thousands of kilometers of the impact.

TERRIBLE TREMOR
A magnitude 13 earthquake—a million times greater than the strongest tremor recorded in human history—courses through the planet.

DISMAL AFTERMATH



This hypothetical catastrophe excavates a crater up to 100 kilometers (60 miles) across and 40 kilometers (25 miles) deep. The nearly instantaneous release of climate-changing pollutants such as ash, sulfur and carbon dioxide kills off species and degrades environments in a few thousand years or less. This geologically rapid timing is reflected in recent scientific studies indicating that species disappear quickly during the worst mass extinctions. Massive volcanism ejects similar pollutants, but its damaging effects are prolonged over millions of years.

illustrations: DON FOLEY; SOURCE: *THE MISTAKEN EXTINCTION*, BY LOWELL DINGUS AND TIMOTHY ROWE, W. H. FREEMAN, 1998. *photographs:* ALAN HILDEBRAND (*quartz*); WALTER PEREDERY (*shattercones*); TIM CULLER *University of California, Berkeley*/APOLLO 11 CREW/NASA (*microspherules*); W. ALVAREZ/SPL/PHOTO RESEARCHERS, INC. (*fallout layer*); WENDY S. WOLBACH *DePaul University* (*soot*)

now at Berkeley. In 1982 they recognized that the typical approach for defining the last occurrence of a given species did not take into account the incompleteness of the fossil record or the biases introduced in the way the fossils were collected.

Many researchers subsequently conducted high-resolution studies of multiple species. These statistically more reliable assessments indicate that the actual extinction time periods at the end of the Cretaceous—and at the end of the Permian—were abrupt (thousands of years) rather than gradual (millions of years). Although volcanically induced climate change no doubt contributed to the demise of some species, life was well on its way to recovery before the volcanism ceased—making the case for an impact trigger more compelling.

Extraterrestrial Hitchhikers

THE RECOGNITION of a shorter time frame for the Great Dying prompted several scientists to search for associated impact tracers and craters. By the early 1990s scientific papers were citing evidence of iridium and shocked quartz from end Permian rocks; however, the reported concentrations were 10- to 100-fold lower than those in the end Cretaceous clay. This finding prompted some paleontologists to claim that the impact that marked the end of the age of dinosaurs was as singular and unique as the animals themselves.

Other scientists reasoned that perhaps an impact had occurred but the rocks simply did not preserve the same clues that were so obvious in end Cretaceous samples. At the end of the Permian period the earth's landmasses were configured into one supercontinent, Pangea, and a superocean, Panthalassa. An asteroid or comet that hit the deep ocean would not generate shocked quartz, because quartz is rare in ocean crust. Nor would it necessarily lead to the spread of iridium worldwide, because not as much debris would be ejected into the atmosphere. Supporting an ocean-impact hypothesis for more ancient extinctions such as the Great Dying, it turned out, would require new tracers.

One of the next impact tracers to hit the scene—and one that would eventually turn up in meteorites and at least two impact craters—evolved out of the accidental discovery of a new form of carbon. In the second year of my doctoral studies at the Scripps Institution of Oceanography in La Jolla, Calif., my adviser, geochemist Jeffrey Bada, showed me an article that had appeared in a recent issue of *Scientific American* [see "Fullerenes," by Robert F. Curl and Richard E. Smalley; October 1991]. It outlined the discovery of a new form of carbon, closed-cage structures called fullerenes (also referred to as buckminsterfullerenes or "buckyballs," after the inventor of the geodesic domes that they resemble). A group of astrochemists and physical chemists had inadvertently created fullerenes in 1985 during laboratory experiments designed to mimic the formation of carbon clusters, or stardust, in some stars. Additional experiments revealed that fullerenes, unlike the other solid forms of carbon, diamond and graphite, were soluble in some organic solvents, a property that would prove their existence and lead to a Nobel Prize in Chemistry for Curl, Smalley and Harold W. Kroto in 1996.

Knowing that stardust, like iridium, is delivered to our plan-

Rough Neighborhood

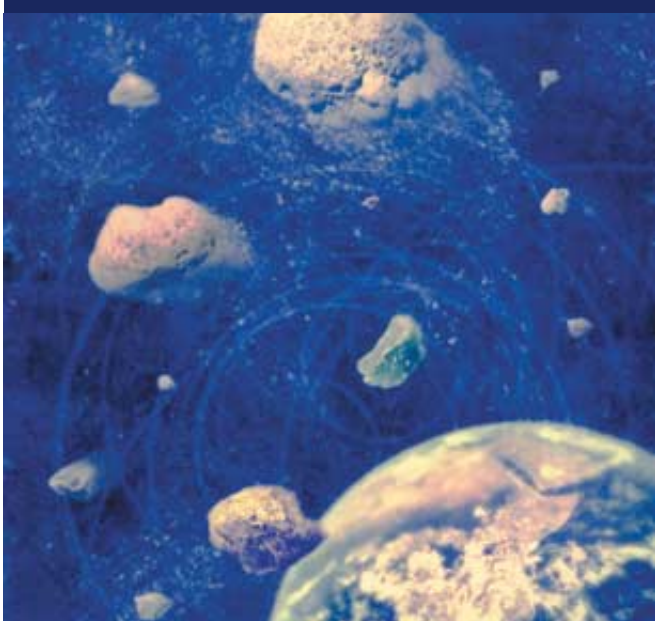
The search for Earth-crossing asteroids expands

ON JANUARY 7 a shopping mall–size rock reminded everyone just how cluttered the solar system really is. Roughly 300 meters in diameter, asteroid 2001 YB5 was small enough to escape notice until late December but big enough to carve a crater the size of a small city had it struck land. Fortunately, its closest approach to Earth was 830,000 kilometers (about twice the distance to the moon), and we are in no danger of a YB5 collision for at least the next several centuries.

But what about the 1,500 other known near-Earth asteroids? (They are so dubbed because they have broken away from the main asteroid belt between Mars and Jupiter and now pose a potential impact risk.) YB5-size space rocks fly this close nearly every year, says David Morrison of the NASA Ames Research Center, but they strike Earth only about every 20,000 to 30,000 years.

Finding hazardous objects long before they become a threat is the aim of the U.K.'s new information center on near-Earth objects, which is scheduled to debut in early April at the National Space Science Center in Leicester. Asteroid hunters at the U.K. center and a handful of other institutions worldwide are especially concerned with objects one kilometer (six tenths of a mile) in diameter, the low-end estimate for the size required to wreak global havoc. The odds of such a catastrophe occurring in the next 100 years range between one in 4,000 and one in 8,600, according to recent calculations by Alan Harris of the Jet Propulsion Laboratory in Pasadena, Calif. NASA's ongoing Spaceguard Survey, which aims to find 90 percent of the Earth-crossing asteroids this size or larger by 2008, will help sharpen this prediction.

—Sarah Simpson, contributing editor



et in the form of cosmic dust, asteroids and comets, we decided to search for these exotic carbon molecules in earthly sediments. We chose a known impact site—the 1.85-billion-year-old Sudbury crater in Ontario, Canada—because of its unique lining of carbon-rich breccia, a mixture of shattered target rocks and other fallout from the blast. (Not unlike the Chicxulub controversy, it took the discovery of shocked quartz and shatter-cones, features described as shock waves captured in the rock, to convince most scientists that the crater was an impact scar rather than volcanic in origin.)

Because fullerene is a pure-carbon molecule, the Sudbury breccia offered a prime location for collecting promising samples, which we did in 1993. By exploiting the unique solubility properties of fullerene, I was able to isolate the most stable molecules—those built from 60 or 70 carbon atoms each—in the laboratory. The next critical questions were: Did the fullerenes hitch a ride to the earth on the impactor, surviving the catastrophic blast? Or were they somehow generated in the intense heat and pressures of the event?

Meanwhile organic chemist Martin Saunders and his colleagues at Yale University and geochemist Robert Poreda of the University of Rochester were discovering a way to resolve this question. In 1993 Saunders and Poreda demonstrated that fullerenes have the unusual ability to capture noble gases—such as helium, neon and argon—within their caged structures. As soon as Bada and I became aware of this discovery, in 1994, we asked Poreda to examine our Sudbury fullerenes. We knew that the isotopic compositions of noble gases observed in space (like those measured in meteorites and cosmic dust) were clearly distinct from those found on the earth. That meant we had a simple way to test where our exotic carbon originated: measure the isotopic signatures of the gases within them.

What we found astounds us to this day. The Sudbury fullerenes contained helium with compositions similar to some meteorites and cosmic dust. We reasoned that the molecules must have survived the catastrophic impact, but how? Geologists agree that the Sudbury impactor was at least eight kilometers (about five miles) across. Computer simulations predicted that all organic compounds in an asteroid or comet of this size would be vaporized on impact. Perhaps even more troubling was the initial lack of compelling evidence for fullerenes in meteorites.

We, too , were surprised that the fullerenes survived. But as for their apparent absence in meteorites, we suspected that previous workers had not looked for all the known types. In the original experiment designed to simulate stardust, a family of large fullerenes formed in addition to the 60- and 70-atom molecules. Indeed, on a whim, I attempted to isolate larger fullerenes in some carbon-rich meteorites, and a whole series of cages with up to 400 carbon atoms were present. Like their smaller counterparts from the Sudbury crater, these larger structures contained extraterrestrial helium, neon and argon.

With the discovery of the giant fullerenes in meteorites, Poreda and I decided to test our new method on sediments associated with mass extinctions. We first revisited fullerene samples that other researchers had discovered at end Cretaceous

KAMIL VOJNAR

sites. One group, led by Dieter Heymann of Rice University, had proposed that the exotic carbon was part of the soot that accumulated in the wake of the massive, impact-ignited fires. The heat of such a fire may have been intense enough to transform plant carbon into fullerenes, but it could not account for the extraterrestrial helium that we found inside them.

Inspired by this success, we wondered whether fullerenes would be a reliable tracer of large impacts elsewhere in the fossil record. Sediments associated with the Great Dying became our next focus. In February 2001 we reported extraterrestrial helium and argon in fullerenes from end Permian locations in China and Japan. In the past several months we have also begun to look at end Permian sites in Antarctica. Preliminary investigations of samples from Graphite Peak indicate that fullerenes are present and contain extraterrestrial helium and argon. These end Permian fullerenes are also associated with shocked quartz, another direct indicator of impact.

As exciting as these new impact tracers linked to the Great Dying have been, it would be misleading to suggest that fullerenes are the smoking gun for a giant impact. Many scientists still argue that volcanism is the more likely cause. Some have suggested that cosmic dust is a better indicator of an impact event than fullerenes are. Others are asking why evidence such as shocked quartz and iridium are so rare in rocks associated with the Great Dying and will remain skeptical if an impact crater cannot be found.

Forging Ahead

UNDAUNTED BY SKEPTICISM, a handful of scientists continues to look for potential impact craters and tracers. Recently geologist John Gorter of Agip Petroleum in Perth, Australia, described a potential, enormous end Permian impact crater buried under a thick pile of sediments offshore of northwestern Australia. Gorter interpreted a seismic line over the region that suggests a circular structure, called the Bedout, some 200 kilometers (about 125 miles) across. If a future discovery of shocked quartz or other impact tracers proves this structure to be ground zero for a life-altering impact, its location could explain why extraterrestrial fullerenes are found in China, Japan and Antarctica—regions close to the proposed impact—but not in more distant sites, such as Hungary and Israel.

Also encouraging are the recent discoveries of other tracers proposed as direct products of an impact. In September 2001 geochemist Kunio Kaiho of Tohoku University in Japan and his colleagues reported the presence of impact-metamorphosed iron-silica-nickel grains in the same end Permian rocks in Meishan, China, where evidence for abrupt extinctions and extraterrestrial fullerenes has cropped up. Such grains have been reported in several end Cretaceous impact sites around the world as well.

In the absence of craters or other direct evidence, it still may be possible to determine the occurrence of an impact by noting symptoms of rapid environmental or biological changes. In 2000, in fact, Peter Ward of the University of Washington and his colleagues reported evidence of abrupt die-offs of rooted plants in end Permian rocks of the Karoo Basin in South Africa. Several groups have also described a sharp drop in productivity in marine species associated with the Great Dying—and with the third of the five big mass extinctions, in some 200-million-year-old end Triassic rocks. These productivity crashes, marked by a shift in the values of carbon isotopes, correlate to a similar record at the end of the Cretaceous, a time when few scientists doubt a violent impact occurred.

Only more careful investigation will determine if new impact tracers—both direct products of a collision and indirect evidence for abrupt ecological change—will prove themselves reliable in the long run. So far researchers have demonstrated that several lines of evidence for impacts are present in rocks that record three of our planet's five most devastating biological crises. For the two other largest extinctions—one about 440 million years ago and the other about 365 million years ago—iridium, shocked quartz, microspherules, potential craters and productivity collapse have been reported, but the causal link between impact and extinction is still tenuous at best. It is important to note, however, that the impact tracers that typify the end of the Cretaceous will not be as robust in rocks linked to older mass extinctions.

Whatever stimulated these mass extinctions made possible our own existence.

The idea that giant collisions may have occurred multiple times is intriguing in its own right. But perhaps even more compelling is the growing indication that these destructive events may be necessary to promote evolutionary change. Most paleontologists believe that the Great Dying, for instance, enabled dinosaurs to thrive by opening niches previously occupied by other animals. Likewise, the demise of the dinosaurs allowed mammals to flourish. Whatever stimulated these mass extinctions, then, also made possible our own existence. As researchers continue to detect impact tracers around the world, it's looking more like impacts are the culprits of the greatest unresolved murder mysteries in the history of life on earth. SA

MORE TO EXPLORE

Impact Event at the Permian-Triassic Boundary: Evidence from Extraterrestrial Noble Gases in Fullerene. Luann Becker, Robert J. Poreda, Andrew G. Hunt, Theodore E. Bunch and Michael Rampino in *Science*, Vol. 291, pages 1530–1533; February 23, 2001.

Accretion of Extraterrestrial Matter throughout Earth's History. Edited by Bernhard Peucker-Ehrenbrink and Birger Schmitz. Kluwer Academic/Plenum Publishers, 2001.

LEARNING TO READ is a joy for many children, but others struggle mightily.

HOW SHOULD READING BE TAUGHT?

Educators have long argued over the best way to teach reading to children. The research, however, indicates that a highly popular method is inadequate on its own

By Keith Rayner, Barbara R. Foorman, Charles A. Perfetti, David Pesetsky and Mark S. Seidenberg

Photographs by Tina West

Most of us are a little fuzzy on how we learned to read, much as we cannot recall anything special about learning to talk. Although these skills are related, the ways we acquire them differ profoundly. Learning to speak is automatic for almost all children brought up in normal circumstances, but learning to read requires elaborate instruction and conscious effort. Remember how hard it once was? Reading this page with the magazine turned upside down should bring back some of the struggles of early childhood, when working through even a simple passage was a slog.

Well aware of the difficulties, educators have given a great deal of thought to how they can best help children learn to read. No single method has triumphed. Indeed, heated arguments about the most appropriate form of reading instruction continue to polarize the teaching community. To help forge a consensus, we recently came together under the aegis of the American Psychological Society to review the voluminous research on the mental processing that underlies skilled reading and on how reading should be taught. The results point strongly in directions that may disturb some parents.

Three general approaches have been tried. In one, called whole-word instruction (also known as the "look-say" method), children learn by rote how to recognize at a glance a vocabulary of 50 to 100 words. Then they gradually acquire other words, often through seeing them used over and over in the context of a story. ("Run, Spot, run," from the well-known Dick and Jane series of readers, is a classic example of a sentence designed to aid whole-word instruction.) This procedure could just as well be used to learn Chinese, in which each character in the written language corresponds to a word or word root.

Actually, for the past half a century, youngsters in China have followed a different prescription: as a first step toward literacy,

they are taught to read Chinese words using the Roman alphabet. Similarly, speakers of most other languages learn the relationship between letters and the sounds associated with them (phonemes). That is, children are taught how to use their knowledge of the alphabet to sound out words. This procedure constitutes a second approach to teaching reading—the phonics so familiar to baby boomers.

The connections between letters and phonemes would appear simple enough. For example, the letter "b" almost always sounds the same as it does in the word "bat." Or consider the silent "e," which denotes that the preceding vowel has a long sound, as in the words "pave," "save" and "gave." Although the final "e" is not voiced, its role is straightforward. English, however, offers plenty of exceptions—take the word "have." There are, in fact, hundreds of deviations from the normal patterns, including "give," "said," "is," "was," "were," "done" and "some." Such problematic yet common words are among the first a child has to learn.

Clearly, the lack of perfect correspondence between letters and sounds is a source of confusion and a potential roadblock for the beginning reader. As a result, many schools have adopted a different approach: the whole-language method (also called literature-based instruction or guided reading). The strategy here is similar to whole-word instruction, but it relies more heavily on the child's experience with language. For example, students are offered engaging books and are encouraged to guess the words that they do not know by considering the context of the sentence or by looking for clues in the story line and illustrations, rather than trying to sound them out. Often children are given the opportunity to write stories of their own, in an effort to instill a love of words and reading.

The whole-language approach aims to make reading instruction enjoyable. One of its key principles is that the rules of phonics should *not* be taught directly. Rather the connection between letters and sounds should be learned incidentally through exposure to text. This methodology stipulates that students should not be corrected when they make errors reading words. The philosophical rationale is that learning to read, like learning to speak, is a natural act that children can essentially teach themselves how to do. Just how well that assumption holds up in practice often depends on the individual.

Although many parents might think that innate INTELLIGENCE WILL GOVERN how well their kids learn to read, the EVIDENCE SUGGESTS OTHERWISE.

Overview/*Teaching Reading*

- Learning to read is a crucial step in children's education because those who fare poorly in the early grades are unlikely to catch up with their more skilled classmates, even after years of further schooling.
- During the 1990s many educators in America abandoned the traditional "phonics" method of reading instruction: teaching children directly the correspondences between spoken sounds and letters that represent them. Instead elementary school teachers turned to various "whole-language" methods, by which students learn the connections between letters and sounds incidentally in the course of literature-based activities.
- Evaluations of the effectiveness of the two methods have shown that children become skilled readers much more readily when their instruction includes phonics. Modern research in psychology and linguistics helps to explain why this is so.

How Beginners Learn to Read

ALTHOUGH MANY PARENTS might think that innate intelligence will govern how well their kids learn to read no matter what type of instruction is given, the evidence suggests otherwise. Two separate studies from the 1960s and 1970s have shown that, in general, IQ has very little bearing on early reading ability. More recently, researchers have found that children who have difficulty learning to read often have above-average IQs.

It might also be tempting to believe that the differences in early reading ability wash out over time, but that, too, is a misconception. Keith E. Stanovich of the University of Toronto has, for example, shown that children's facility with reading in the first grade usually provides a good indication of what their 11th-grade reading proficiency will turn out to be. Why? Because reading requires practice, and those who excel end up practicing the most. Hence, the gap between more and less able readers in the first few grades generally grows over the years.

Teaching children to read well early on obviously helps to develop a valuable lifetime habit; thus, it is no wonder that educators have placed enormous emphasis on finding the best way to teach these skills. At one time, a great deal of debate in educational circles centered on whether whole-word or phonics instruction was the most effective. But over the past decade or so, arguments have revolved around the relative merits of phonics and whole-word's successor, whole-language.

Many teachers adopted the whole-language approach because of its intuitive appeal. After all, making reading fun promises to keep children motivated, and learning to read depends more on what the student does than on what the teacher does. But the prospect of keeping kids interested would not

LARGE BOOKS that children can read together are a common feature of whole-language instruction. This approach emphasizes engaging literature and attempts to motivate youngsters by keeping reading fun.

have been enough by itself to convince teachers to use the whole-language method. What really sold it was an educational philosophy that empowered teachers to compose their own curricula and encouraged them to treat children as active participants, an enticing combination that was promoted with flair by some educator celebrities. The presumed benefits of whole-language instruction—and the stark contrast to the perceived dullness of phonics—led to its growing acceptance across America during the 1990s.

In Massachusetts, for example, whole-language almost became the official state method of instruction with passage of the Massachusetts Education Reform Act of 1993. That legislation changed what had been a tradition of little state involvement in school curriculum. The law promised to increase state funding for public education, and in exchange local school systems were required to meet new state standards.

Despite the previous lack of central control, the reading curricula in Massachusetts public schools were rather uniform—and it is not difficult to understand why. As in other places, teachers and administrators took the same courses at the same handful of universities, attended the same workshops, bought the same textbooks and responded to the same educational fashions. Hence, the committee of educators charged by the state government with framing a statement about how reading should be taught were heavily influenced by the whole-language approach. And naturally enough, the document they produced highlighted the idea that children could learn to read the same way they learned to talk. It presented a vision of language acquisition that attributed the process to curiosity and enthusiasm alone, and it seemed authoritative, claiming support from research.

As it happens, Massachusetts is home to hubs of research in linguistics and the psychology of reading—at the Massachusetts Institute of Technology and the University of Massachusetts at Amherst. After the content of the proposed curriculum document became known, a number of scholars in these places (including two of us) reacted strongly. Dozens of linguists and psychologists signed a letter taking issue with the document's assertion that research supported whole-language instruction. They sent it to the state commissioner of education, who eventually saw to it that corrections were made and that state standards reflected the actual research results.

How Phonics Is Taught

IN TEACHING PHONICS, instructors present the spellings for different sounds in a specific order, introducing the simplest (or most useful) patterns early on. They then practice these patterns with their students using engaging stories. Shown below are 20 of the 120 or so patterns presented to first graders in one modern phonics program, *Open Court Reading,* from SRA/McGraw-Hill. Choosing another published system of phonics instruction would provide the students with a somewhat different scope and sequence, but the general strategy would be the same.

Some teachers prefer to dispense with such structured programs and to create phonics lessons on their own. Doing so is no small chore, because they have so many decisions to make. Should rules be taught for all the ways to spell each of the approximately 40 distinct sounds (phonemes) of American English? For the long "a" alone, there are eight spelling patterns, as in "make," "rain," "say," "they," "baby," "eight," "vein" and "great." And do all the phonemes need attention? For example, do the vowel sounds in "book" and "moon" both need to be taught?

Although some teachers can tackle these questions and create phonics lessons that are every bit as effective as those provided in a published program, most probably have too many demands on their time to take on that task. Just how much latitude phonics instructors should be given and how effectively they can make use of the flexibility remain points of debate in a number of school districts.

—K.R., B.R.F., C.A.P., D.P. and M.S.S.

Letter Pattern to Be Mastered (first 10 patterns taught)	One of the Words Used as Examples	Letter Pattern to Be Mastered (final 10 patterns taught)	One of the Words Used as Examples
m	monkey	ture	nature
a	lamb	ear	earn
t	time	or	worm
h	hound	ar	carry
p	popcorn	er	berry
n	nose	tion	nation
c	camera	ion	million
d	dinosaur	re	reheat
(contractions)	can't	ure	measure
s	sausages	ous	dangerous

ALPHABET CARDS and books that highlight selected sounds are among the tools that teachers use to help children learn the correspondences between letters and sounds.

By chance, this incident took place just as debate about how to teach reading was heating up in other states (most notably, in California and Texas). Sides were often divided along political lines, with conservatives backing phonics and liberals favoring whole-language instruction. Consequently, the Massachusetts dispute drew national attention. In particular, conservative newsletters and Web sites created considerable publicity for the researchers' letter—an ironic twist, given that the list of professors who signed it included several well-known leftists.

Why Phonics?

WHY DID SO MANY LINGUISTS and psychologists object strongly to the abandonment of phonics? In short, because research had clearly demonstrated that understanding how letters relate to the component sounds of words is critically important in reading. Our recent review of the topic shows that there is no doubt about it: teaching that makes the rules of phonics clear will ultimately be more successful than teaching that does not. Admittedly, some children can infer these principles on their own, but most need explicit instruction in phonics, or their reading skills will suffer.

This conclusion rests, in part, on knowledge of how experienced readers make sense of words on a page—an understanding that psychologists have developed over many decades. One of the first researchers to investigate the nature of reading was James M. Cattell, an American psychologist of the Victorian era. To test whether proficient readers were taking in words letter by letter or all at once, he performed a pioneering experiment, exposing subjects very briefly to whole words or to individual letters and asking them what they saw. He found that they were

THE AUTHORS

KEITH RAYNER, BARBARA R. FOORMAN, CHARLES A. PERFETTI, DAVID PESETSKY and *MARK S. SEIDENBERG* collaborated on a paper surveying the teaching of reading for the November 2001 issue of *Psychological Science in the Public Interest* [see "More to Explore," on page 91]. Rayner, Distinguished Professor of Psychology at the University of Massachusetts at Amherst, is currently on sabbatical in England at the University of Durham. Foorman is a professor of pediatrics at the University of Texas–Houston Health Science Center, where she directs the Center for Academic and Reading Skills. Perfetti is University Professor of Psychology and Linguistics at the University of Pittsburgh, where he is associate director of the Learning Research and Development Center. Pesetsky is Ferrari P. Ward Professor of Linguistics at the Massachusetts Institute of Technology. Seidenberg is a professor of psychology at the University of Wisconsin–Madison.

better able to report words than letters. Thus, it seemed apparent to him that people do not absorb printed words one letter at a time. (Such findings helped to motivate the creation of the whole-word method later on.) More recent research has refined our knowledge of this phenomenon. For example, studies that track eye movements during reading show that although people register each letter in a word as a separate symbol, they normally perceive all the letters in a word simultaneously.

Computer programs that simulate how children read suggest that gaining a command of PHONICS IS EASIER than learning to ASSOCIATE WHOLE WORDS with their meanings.

The question of whether accomplished readers mentally sound out words took longer to answer. Advocates of whole-language instruction have argued forcefully for more than 20 years that people often derive meanings directly from print without ever determining the sound of the word. Some psychologists today accept this view, but most believe that reading is typically a process of rapidly sounding out words mentally, even for the highly skilled.

The most compelling evidence for this last contention comes from clever experiments by Guy Van Orden of Arizona State University wherein a subject is first asked a question, such as "Is it a flower?" He or she is then presented with a target word (for example, "rose") and asked whether the word fits the category. Sometimes the subject is offered a word that sounds the same as a correct answer (called a homophone—say, "rows" instead of "rose"). Subjects often mistakenly identify such words as fitting the category, and these incorrect responses show that readers routinely convert strings of letters to sounds (or rather, to their unvoiced mental equivalents), which they then use to ascertain meanings.

Some eye-movement studies have used homophones to demonstrate that the process of sounding out words mentally begins very rapidly after a reader's gaze first fixes on a particular word. And recent brain studies show that the primary motor cortex is active during reading, presumably because it is involved with mouth movements used in reading aloud.

Consequently, psychologists now know that the process of mentally sounding out words is an integral part of silent reading, even for the highly skilled. This understanding suggests that learning the correspondences between letters and sounds—that is to say, phonics—is keenly important for beginners. Further support for phonics instruction comes from experiments designed to mimic the way people learn to read.

Investigators have, for example, trained English-speaking college students to read using unfamiliar symbols such as Arabic letters. One group learned the phonemes associated with individual Arabic letters (the phonics approach), while another group learned entire words associated with certain strings of

U.S. Government Studies Supporting Phonics Instruction

TITLE	ORGANIZER	SCOPE	SUMMARY STATEMENT
Preventing Reading Difficulties in Young Children	National Academy of Sciences/National Research Council (sponsored by the Department of Education); 1998	Literature review covering more than 700 publications	"Failure to grasp that written spellings systematically represent the sounds of spoken words makes it difficult not only to recognize printed words but also to understand how to learn and to profit from instruction. If a child cannot rely on the alphabetic principle, word recognition is inaccurate or laborious and comprehension of connected text will be impeded."
Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction	National Reading Panel (convened by the National Institute of Child Health and Human Development, in consultation with the secretary of education); 2000	Includes a meta-analysis of 38 controlled studies of phonics instruction published since 1970	"The meta-analysis indicated that systematic phonics instruction enhances children's success in learning to read and that systematic phonics instruction is significantly more effective than instruction that teaches little or no phonics."

Arabic letters (whole-word). Then both groups were required to read a new set of words constructed from the original characters. In general, readers who were taught the rules of phonics could read many more new words than those trained with a whole-word procedure. Research using computer programs that simulate how children read also indicates that gaining a command of phonics is easier than learning to associate whole words with their meanings.

Classroom studies comparing phonics with either whole-word or whole-language instruction are also quite illuminating. The late Jeanne S. Chall of Harvard University carried out a comprehensive review of such work, as subsequently did Marilyn J. Adams, who was also affiliated with Harvard. In a nutshell, their reviews, as well as our own, show that systematic phonics instruction produces higher achievement for beginning readers. The differences are greatest for students at risk of failing to learn to read, such as those living in homes where the value of literacy is not emphasized.

One particularly persuasive study was undertaken as long ago as 1985. Mary Ann Evans of the University of Guelph in Canada and Thomas H. Carr of Michigan State University compared two programs used in 20 first-grade classrooms. Half the students were offered traditional reading instruction, which included the use of specially designed readers, phonics drills and applications. The other half were taught using an individualized method that drew from their experiences with language; these children produced their own booklets of stories and developed sets of words to be recognized (common components of the whole-language approach). The two groups spent the same amount of time on reading, had similar socioeconomic profiles and were virtually identical on measures of intelligence and language maturity. Yet this study found that the first group scored higher at year's end on tests of reading and comprehension.

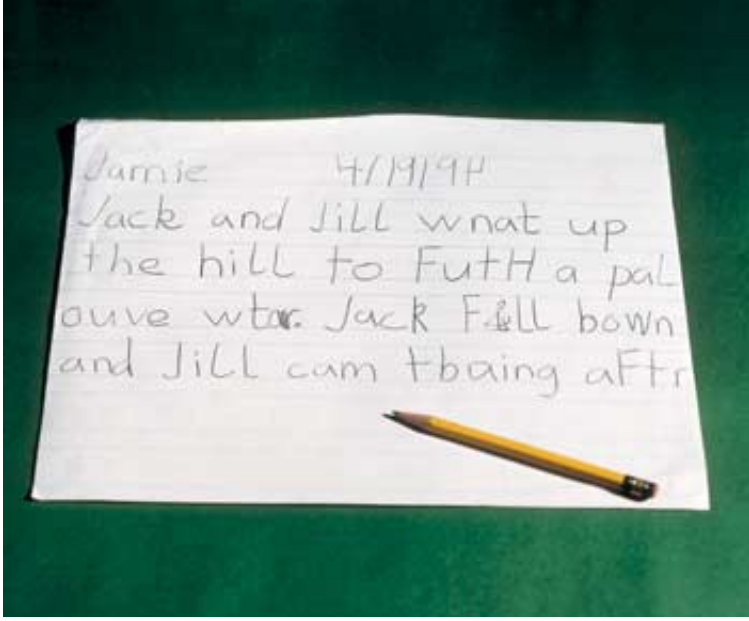
More recent investigations (namely, authoritative evaluations by the National Reading Panel and the National Research Council) examining all the available studies echo these results. Influenced by such findings, the Bush administration is now promoting the inclusion of phonics in reading programs nationwide.

A Delicate Balance

IF RESEARCHERS ARE SO CONVINCED about the need for phonics instruction, why does the debate continue? Because the controversy is enmeshed in the philosophical differences between traditional and progressive approaches, differences that have divided American educators for years. The progressives challenge the results of laboratory tests and classroom studies on the basis of a broad philosophical skepticism about the value of such research. They champion student-centered learning and teacher empowerment. Sadly, they fail to realize that these very admirable educational values are equally consistent with the teaching of phonics.

If schools of education insisted that would-be reading teachers learned something about the vast research in linguistics and psychology that bears on reading, and if these institutions regularly included a modern, high-quality course on phonics, their graduates would be more eager to use phonics and would be prepared to do so effectively. They would not have to follow scripted programs or rely on formulaic workbooks and could allow their pupils to apply the principles of phonics while reading for pleasure. Using whole-language activities to supplement phonics instruction certainly helps to make reading fun and meaningful for children, so no one would want to see such tools discarded. Indeed, recent work has indicated—and many teachers have discovered—that the combination of literature-based instruction and phonics is more powerful than either method used alone.

Teachers obviously need to strike a balance. But in doing so, we urge them to remember that reading must be grounded in a firm understanding of the connections between letters and sounds. Instructors should recognize the ample evidence that youngsters who are directly taught phonics become better at reading, spelling and comprehension than those who must pick up all the confusing rules of English on their own. Educators who deny this reality are neglecting decades of research. They are also neglecting the needs of their students. SA



SPELLING ERRORS abound when early readers take to their pencils. Teachers of phonics routinely correct mistakes. Strict devotees of whole-language instruction tend to be more tolerant of invented spellings, choosing to concentrate on the student's meaning.

MORE TO EXPLORE

Beginning to Read: Thinking and Learning about Print. Marilyn J. Adams. MIT Press, 1990.

Learning to Read: The Great Debate. Jeanne S. Chall. Harcourt Brace, 1996.

Preventing Reading Difficulties in Young Children. Edited by C. E. Snow et al. National Academy Press, 1998. Available at **books.nap.edu/books/030906418X/html/index.html**

Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction. National Reading Panel. National Institute of Child Health and Human Development, 2000. Available at **www.nationalreadingpanel.org/Publications/publications.htm**

How Psychological Science Informs the Teaching of Reading. Keith Rayner, Barbara R. Foorman, Charles A. Perfetti, David Pesetsky and Mark S. Seidenberg in *Psychological Science in the Public Interest,* Vol. 2, No. 2, pages 31–74; November 2001. Available at **www.psychologicalscience.org/newsresearch/publications/journals/pspi2_2.html**

COMBINATION LOCKS

Secret of Spin

For several generations now, people have experienced the same daily mystery as *Scientific American* reader Billy Ellis. "When I was opening my sports locker today," he writes, "it occurred to me that I have no idea how a standard combination lock works. I have been using them for decades. How does spinning the dial 'to the right, to the left, then to the right again' open the lock?"

The answer lies in ingenious yet simple design. Rotating the dial back and forth turns three internal disks. Each disk has a notch. When the notches align, the lock's shackle is freed. The combination for a standard lock from Master Lock consists of three numerals, one for each disk. A disk has 40 possible notch positions, so the dial has 40 numbers, which results in 64,000 potential combinations: 40 times 40 times 40.

Combination locks dominate the market for school, gym and workplace lockers. Although they might keep foes—and friends—out of your stuff, don't be too sanguine. Many institutions keep a chart of locker combinations or use locks that have a master-key override in the back so administrators can get in, according to Steve Fantl, business technical leader at Master Lock in Milwaukee, the largest U.S. supplier of combination locks.

Like nostalgia, a good lock can last for years. The shackle of a Master Lock is made of hardened steel. The dial, certain internal disks and the locking lever, however, are made of plated zinc because the metal can be die-cast. These parts can corrode over long periods if exposed to moisture. If your old bicycle lock is stuck, try squirting some rust-busting oil inside it.

Combination locks that secure wall safes and bank vaults are much more sophisticated than standard combination locks. So why do filmmakers often show crooks with a stethoscope or computerized gizmo that helps them to magically decipher the combination? "Because," Fantl says with a shrug, "it's the movies." And why do most locks have a three-number combination? "Two numbers don't provide enough combinations," Fantl notes, "and studies show four numbers are too many for most people to remember." —*Mark Fischetti*

ILLUSTRATIONS BY BRYAN CHRISTIE DESIGN; SANDIA NATIONAL LABORATORIES (*micrograph*)

IN A LOCK
a spring holds two rotating disks in place. A third disk is attached to the dial. Each disk has a notch, as well as a nub that corresponds to a number on the dial. Aligning all three notches allows the shackle to be released.

DID YOU KNOW . . .

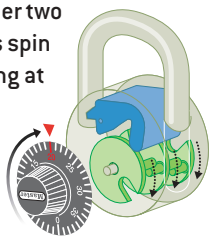
➤ THE MASTER: In 1921 locksmith Harry Soref founded Master Lock in Milwaukee. He built a padlock from laminated layers of steel that was stronger than typical hollow locks. His break came in 1928, when federal agents in New York City ordered 147,600 padlocks to shutter establishments selling booze during Prohibition. Harry Houdini conferred with Soref about trick keys that could help the magician escape from handcuffs. (Soref is also said to be the first to use a beaded lamp pull chain as a key chain.)

➤ SMALL BUT SECURE: Engineers at Sandia National Laboratories have crafted a tiny microelectromechanical combination lock (*micrograph below*). It has six pinwheels stacked in series. Each is less than 300 microns across. From a keyboard, a user enters six digits in succession, from 0 to 9. A correct keystroke sends an electrical pulse to an electrostatic actuator that rotates wheel 1, then wheel 2 and so on. If the wheels' notches align properly, a lever arm opens. Sandia says hackers have a one-in-a-million chance of opening the lock, and one wrong number freezes it. The lab may use the lock on weapons-control systems.

➤ WHICH TO PICK? To pick key locks, locksmiths and crooks insert a thin lever to apply a slight lateral tension to the cylinder. Then they insert a curved pick to push each pin to the shear line. Under the right tension, the pins will stay put. When all the pins reach the shear line, the cylinder turns completely, opening the lock. There's no easy way to reach the disks of a combination lock to pick it.

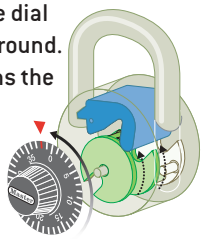
STEP 1

"Turn dial right two or more turns and stop at 20." This allows the dial nub to engage the other two nubs so the disks spin in unison. Stopping at 20 puts the first disk notch in the proper release position.



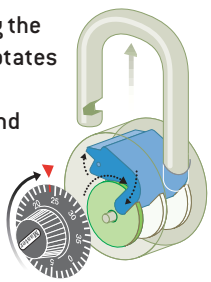
STEP 2

"Turn dial left one whole turn past previous number and stop at 38." This removes nub pressure from the first disk but pushes the second disk nub to the left once the dial nub comes back around. Turning to 38 aligns the second disk notch with the first disk notch.



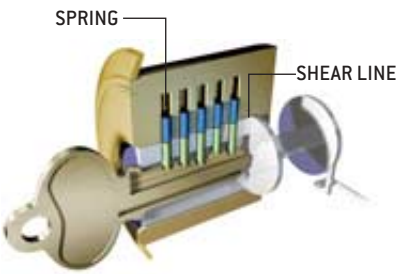
STEP 3

"Turn dial right and stop at 24, then pull shackle open." This removes nub pressure again. Stopping at 24 aligns the third notch. Pulling on the shackle rotates a locking lever, which allows the shackle to be released. Pushing the shackle back in rotates the lever into the shackle's clasp and pushes an upset arm that spins the first disk, scrambling the disk alignment.



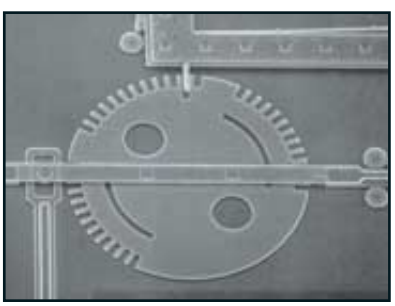
COMMON KEY LOCK

opens when its central cylinder can be turned. Springs in small shafts push pairs of pins down into the cylinder, holding it in place. But a key with the correct profile of ridges will push the pins upward just enough so that the top set of pins clears the shear line, allowing the cylinder to rotate and pull open a door latch, unhinge a padlock shackle or engage a car ignition.



MICROELECTROMECHANICAL LOCK

made by Sandia National Laboratories is less than 300 microns in diameter.



Have a topic for a future column? Send it to workingknowledge@sciam.com

Ancient Rituals on the Atlantic Coast

FULL MOON IN MAY BRINGS HORSESHOE CRABS ASHORE TO MATE AND MIGRATING BIRDS IN TO FEAST **BY MARGUERITE HOLLOWAY**

LAUGHING GULLS are among the many scores of birds that feed on the abundant horseshoe crab egg clusters found every spring on the New Jersey shore.

FRED BAVENDAM *Minden Pictures*

It is a gentle clicking, tapping, scraping concert on the upper reaches of a beach and in the long grass of a marsh. Only a soft sound, like muffled castanets, but it speaks of thousands—thousands of years, thousands of horseshoe crabs and thousands of birds.

As has occurred for aeons, every spring on the New Jersey shore and along the Delaware Bay, hundreds of thousands of horseshoe crabs emerge from the depths of the ocean to mate. The large females, sometimes attended by as many as four smaller males, walk and are washed onto the beach with the full-moon and other spring high tides. They dig holes at the high-water mark and lay their clusters of tiny olive-colored eggs. The males then fertilize them. As the tide ebbs, the crabs are swept back to sea.

This annual coming of the carapaces can be seen and heard at night, by the light of the moon. Or during the day, when shorebirds raucously feed on the wealth of green protein. More than a million birds traveling from South and Central America on their way to the Arctic Circle stop to gorge themselves on horseshoe crab eggs. Sanderlings, ruddy turnstones, semipalmated sandpipers, dunlins and more than half the world's population of red knots are among the species that double their weight eating the nutritious eggs. According to the New Jersey Division of Fish and Wildlife, the birds—some of which will travel 11,000 miles in total—eat more than 300 tons of eggs during

their stay. It is all over in June. The full moon comes again, and the tides wash out the horseshoe crab larvae that have newly hatched after incubating in the sun for a month or so. The birds continue their journey north.

There are many places along the Delaware Bay and the New Jersey shore to see the rising tide of horseshoe crabs and their avian attendants. Several of the side roads off the main routes at the southern end of New Jersey lead directly to good viewing spots: Reed's Beach, Thompson's Beach, Hand's Landing and Sunray Beach among them. Visitors have only to be careful of the piping plover nesting grounds. But there are walkways at some of the sites, and naturalists are stationed at certain points to answer questions. For those who want a more formal visit, several conservation groups organize tours.

Last spring there was much to see, even a day or so after the highest tide. In some of the inlets of the bay, dozens of terrapins poked their heads out of the water as if watching the horseshoe crabs shuffle on the sand, and the herring gulls and black-backed gulls chorused as they stalked the beach, looking for the golf ball–size clusters of crab eggs. In a nearby marsh, hundreds of crabs could be seen stranded in the eelgrass beds. Some were not lucky enough to make the ebb tide; others were trying to turn themselves over with their long spiked tail, or telson, which exists for that purpose.

The horseshoe crabs, *Limulus polyphemus,* can be found at various points from Mexico to Maine, but the largest population resides around the New Jersey shore. In 2001 the National Oceanic and Atmospheric Administration banned fishing for the crabs in a part of Delaware Bay because of concern that their numbers were dwindling dramatically. Commercial fishermen have been harvesting the crabs—which can bring as much as $1 apiece—to sell as bait for lucrative eel and conch fisheries in other countries. Although there are no good data on the extent of the population—or how it has changed over the past few decades—anecdotal reports suggest that there are fewer crabs coming in to nest. That perception, coupled with a report from some wildlife groups that the population of red knots was down 33 percent, spurred protection. Other states have reduced catches as well.

Horseshoe crabs are still taken for medical research, however. Their blood, which turns blue when exposed to the

FRED BAVENDAM *Minden Pictures (opposite page)*

RELICS of the Paleozoic age, horseshoe crabs are the closest living relatives of trilobites. The females, which are much larger than the males, can produce up to 88,000 eggs a year.

air, coagulates around certain forms of bacteria, those that are gram-negative, and as a result, the blood is important for medical diagnoses. Horseshoe crabs have been around for about 350 million years—they were the contemporaries of dinosaurs—and for that reason they offer insight into early bacteria and immune responses. They are not, in fact, crabs at all but are more closely related to scorpions and spiders. A wonderful book from 1987, *The Year of the Crab: Marine Animals in Modern Medicine,* by William Sargent, describes the horseshoe crab's importance to various fields of research.

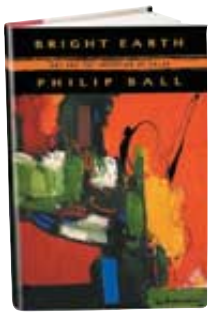
To see this amazing spectacle of ancient creatures hauling themselves up on beaches, take the Garden State Parkway toward Cape May, keeping an eye out for Route 9. Take it south to Route 83; follow 83 east until it hits Route 47 South, which runs along the Delaware Bay side of the peninsula. Look on your road map for the beaches mentioned above or for Norbury's Landing Road and Kimble's Beach Road. Make sure you go on the full or new moon tide.

For information about guided tours and other beaches where you can view the horseshoe crabs and the birds, contact these organizations: the American Littoral Society, 732-291-0055; the Nature Conservancy, 609-861-0600; the Cape May Bird Observatory, 609-861-0700, which is run by the New Jersey Audubon Society; and the New Jersey Division of Fish and Wildlife, 609-292-9400 and www.njfishandwildlife.com. If you are interested in learning more about the region, its history and its many creatures, read Bruce Stutz's *Natural Lives, Modern Times: People and Places of the Delaware River* (1992). SA

Marguerite Holloway is a contributing editor for Scientific American.

A Good Blue Is Hard to Find

CHEMISTRY AND ART HAVE ADVANCED HAND IN HAND THROUGH THE AGES **BY PETER G. BROWN**



BRIGHT EARTH: ART AND THE INVENTION OF COLOR
by Philip Ball
Farrar, Straus and Giroux, New York, 2002 ($30)

Shortly after my debut as the editor of *The Sciences,* I complained to our printer about the colors of a painting we were planning to reproduce. The printer's proof was rendering the rich, deep blues of the original in drab, lifeless tones. As a demanding neophyte, I insisted he do better. Two more proofs—and several hundred dollars—later I was forced to acknowledge what the printer, ever so patiently, had been explaining all along: the color I wanted was simply outside the "color space" accessible to the four process colors of offset printing. I could buy a closer approximation (never an exact match) for several thousand dollars more. But I could have saved my breath (and my publisher a lot of money) if only I'd known what Philip Ball, in his wonderful new book, *Bright Earth,* devotes an entire chapter to explaining: "A good blue [is] hard to find."

Ball's canvas is grand yet effectively framed: how pigment materials have enhanced, and constrained, the "coloristic possibilities" available to artists through the ages. Anyone familiar with painting or graphic arts today—not to mention the millions more who must accept the compromises of Kodachrome or the color phosphors of their television or computer screens—knows that when you try correcting "too pink," you often get "too green." That is no failing of Kodak or the makers of video terminals. All colors—the light-sensitive dyes in film, the phosphors on a video screen, the pigments made from beetle exoskeletons—impose their own discipline on the artist. And every artist must resolve the tension between the mind's eye and the material framework of color. Ball's *Earth* is a fertile field to plow.

Among the many charms of the book are its etymological surprises. Who would have thought that the word "miniature" traces its origins not to any synonym for "small" but to the Latin word *minium,* a lead-based red pigment often used in the scenes (usually small, to be sure) depicted in illuminated medieval manuscripts? Or that "crimson" has such a tangled past? The word comes from the name of an insect of the genus *Kermes,* from which a red compound was extracted. But because, when seen from a distance, the kermes insects look like kernels of grain, Pliny the Elder called the pigment *granum,* Latin for "grain." By the time of Chaucer, "dyed in grain" had come to mean dyed crimson.

Ball is a writer educated in physics and chemistry at the University of Oxford and at the University of Bristol in England, and so in his hands the history of pigment becomes a stained-glass window into the history of chemistry and materials science. By today's standards of scientific specialization, that connection might seem a stretch, but in earlier manifestations those disciplines were motivated by a genuine interest in artists' colors. A love of color, Ball reports, "ushered in the Bronze Age." In Napoleonic France the interior ministry took artists' need for a stable and inexpensive blue seriously enough to commission a leading chemist to search for such a pigment.

In return, of course, artists' palettes owe as much to chemistry as chemistry does to art. Artists' colors have long been piggybacking on the demand for

ALINARI Art Resource, New York

GIOTTO: *The Betrayal of Christ,* early 14th century.

paints and dyes in the building, clothing and furnishing trades. Inevitably, then, the artist is as much a creature of the industrial economy as anyone else and benefits from the same advances and excesses. Zinc white (zinc oxide) and "permanent white" (barium sulfate) were both developed to spare factory workers—and artists—the toxic effects of lead-based white pigment. I find no entry for "poison" or "poison gas" in the extensive and useful index, but it is worth recalling that one of the darker chapters in human cruelty can be linked to the manufacture of coloring. During World War I, both sides rained poison gas on their enemies, but the Germans enjoyed a substantial advantage. The peacetime German dye industry had extensive experience with chlorine and phosgene, both chemical precursors of the battlefield gases; the German military simply diverted those capacities to the production of gas.

The painter is both limited and liberated by his materials. One sad consequence of J.M.W. Turner's headlong rush to seize the newest and most brilliant colors of his day has been the serious discoloration in parts of his paintings. Frank Stella, in contrast, has been quite content to work with "primaries and secondaries straight from the can." What matters, somehow, is the way the artist reshapes materials according to an imaginative vision. After I'd been checking color proofs for a few years, I worried less about "matching" an original than about being "true" to it. In that, I was delighted to find support in Ball's book by no less a coloristic eminence than Georgia O'Keeffe. As she wrote to the printers preparing some of her paintings for reproduction: "It doesn't really matter if the color isn't absolutely right, if the picture feels right when you finish the print." SA

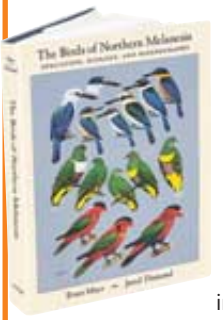
Peter G. Brown is a writer based in New York City. He is the former editor in chief of The Sciences.

THE EDITORS **RECOMMEND**

THE BIRDS OF NORTHERN MELANESIA: SPECIATION, ECOLOGY, AND BIOGEOGRAPHY

by Ernst Mayr and Jared Diamond. Oxford University Press, New York, 2001 ($55)

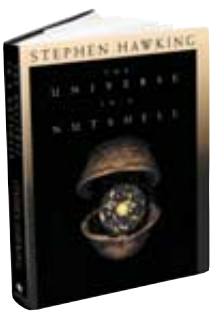
"In this book we present a comprehensive, detailed study of speciation for all resident land and freshwater bird species of Northern Melanesia." Broad as it is, that statement does not fully convey the solidity of the book or the amount of work and time that went into it. The encyclopedic result covers 195 bird species on 76 islands of the Bismarck and Solomon archipelagos, which lie east of New Guinea and constitute Northern Melanesia. Mayr, the eminent evolutionary biologist who is professor emeritus of zoology at Harvard University, has studied Melanesian birds since 1929. Diamond, professor of physiology at the medical school of the University of California at Los Angeles (and author of the Pulitzer Prize–winning *Guns, Germs, and Steel,* among many other popular books), has made 19 expeditions to the South Pacific to study birds in pursuit of his interest in ecology and biogeography.



THE UNIVERSE IN A NUTSHELL

by Stephen Hawking. Bantam Books, New York, 2001 ($35)

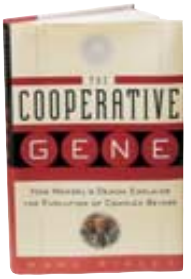
Hawking's *A Brief History of Time,* published in 1988, sold millions of copies even though it was a tough read for nonscientists. As Hawking says now, many readers "got stuck in the early chapters and never reached the more exciting material later on." And so he set out to write "a different kind of book that might be easier to understand." It is. Hawking, the renowned theoretical physicist who holds Isaac Newton's former chair as the Lucasian Professor of Mathematics at the University of Cambridge, has organized the book so that the two opening chapters provide a foundation and that each remaining chapter stands largely on its own as an essay on a cutting-edge topic in cosmology. The illustrations, copious and imaginative, add illumination to the illuminating text.



THE COOPERATIVE GENE: HOW MENDEL'S DEMON EXPLAINS THE EVOLUTION OF COMPLEX BEINGS

by Mark Ridley. Free Press, New York, 2001 ($26)

Ridley asks how living things vastly more complex than the simple life-forms that first populated Earth can arise. His answer is that they require mechanisms to deal with mistakes in the copying of DNA and a means of taming "selfish genes that harm the body by uncooperative and subversive acts." The main solution to the copying problem is sexual reproduction; the means of taming selfish genes is Mendelian inheritance, which ensures "the probabilistic rather than certain inheritance of genes." Ridley (a biologist in the zoology department of the University of Oxford) invokes a hypothetical "Mendel's demon" as the enforcer of chance inheritance; the demon "stands over each gene in a parent and decides whether it will be inherited in the next generation, and which other genes it will be passed on with." And, having laid out his argument, he intriguingly explores the possibility that "superhuman life forms" may someday arise.



All the books reviewed are available for purchase through www.sciam.com

JOHNNY JOHNSON

Card Counters

BY DENNIS E. SHASHA

Alice and Bob play a game using 14 cards—the ace through seven of hearts and the ace through seven of spades. Bob picks a number between one and seven, then tells Alice the number. Alice deals that many cards, finishing the first round. Bob turns over the last card dealt; the number on that card determines how many cards Alice deals in the second round (an ace is considered one). Again, Bob turns over the last card dealt, and Alice then deals that many more cards, and so on. Alice wins if the last card dealt in some round is the very last of the 14 cards and if it is the ace of spades. Bob wins otherwise.

First, consider a warm-up problem: How can Alice arrange the cards in advance so that no matter which number Bob chooses at the beginning, the game will end with the ace of spades as the last card turned up? One possible answer is shown in illustrations A through E below.

Here's a distinctly harder problem. Suppose that Bob arranges all the cards that bear numbers greater than four (the five through seven of hearts and spades). Then Alice, without looking at what Bob has done, arranges the remaining cards and puts her cards after Bob's in the deck. Can she still force a win?

I've saved the hardest problem for last. Say that Bob takes any seven of the cards except the ace of spades. Alice may look at what Bob has done and insert the remaining cards among Bob's cards, but she cannot change their order. For example, if Bob orders the cards as shown in illustration F, Alice can win the game by making the right insertions (I'll leave it to you to figure out how). But can Alice win in general, no matter how Bob arranges the cards? My conjecture is yes, but I know of no proof. Perhaps some clever reader can enlighten me. SA

Dennis E. Shasha is professor of computer science at the Courant Institute of New York University.

Answer to Last Month's Puzzle

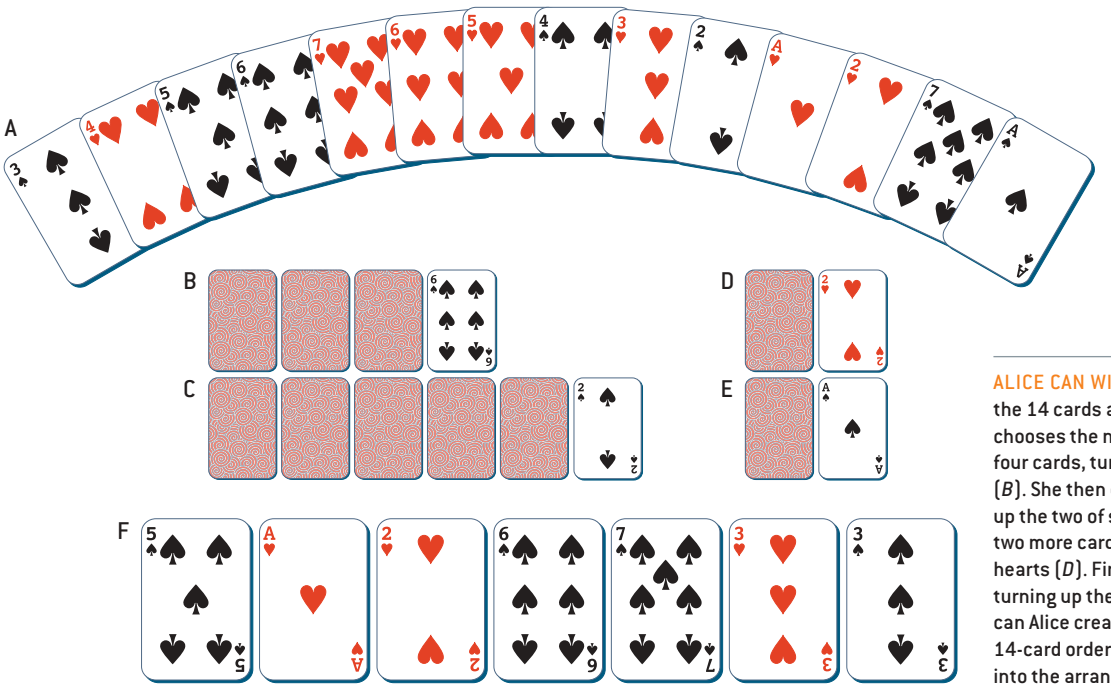
Of the nine lies told by the shifty witnesses, at least five must falsely assert innocence ("the suspect has no drugs"). The only possible scenario that will yield this result is if suspects 1, 4, 5, 7, 8 and 10 have drugs and the rest do not. For a more detailed explanation, go to www.sciam.com

Correction

In the solution to the "Truck Stop" puzzle [November 2001], the number of traversals required across both BC and CD is 12, not 11.

Web Solution

For a peek at the answer to this month's problem, visit www.sciam.com



ALICE CAN WIN the game by ordering the 14 cards as shown (*A*). Suppose Bob chooses the number four. Alice deals four cards, turning up the six of spades (*B*). She then deals six cards, turning up the two of spades (*C*). She deals two more cards, turning up the two of hearts (*D*). Finally, she deals two cards, turning up the ace of spades (*E*). But can Alice create another winning 14-card order by making insertions into the arrangement shown (*F*)?

SARA CHEN

Divining Comedy

CAN RESEARCHERS DISSECT HUMOR WITHOUT KILLING THE PATIENCE? BY STEVE MIRSKY

Last September the British Association for the Advancement of Science (BAAS) announced a plan to discover the world's funniest joke. This quest for the joke of all jokes, the wisest of cracks, the topper de tutti toppi, was to be conducted using a double-pronged approach in which visitors to a Web site could submit jokes and vote on those already available for judgment. According to the journal *Nature,* visitors to the site (www.laughlab.co.uk) were asked to fill out a brief questionnaire "about their age, gender and nationality, as well as a brief cognition quiz." The associations between jokes and survey responses would theoretically make possible "the largest-ever look at the psychology of humor."

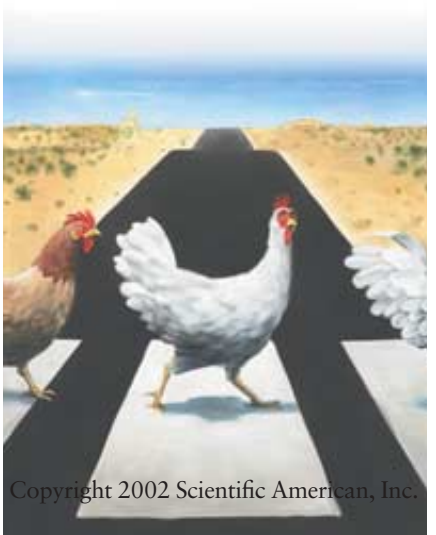
As the writer of what passes for a humor column (which at *Scientific American* is like making the best sloppy joes at the culinary institute), I naturally took an interest. My initial reaction was summed up by taking advantage of certain principles of fluid dynamics to produce what is known as a Bronx cheer, or, in England, a raspberry. I felt this way because jokes tend to be on the low end of the funnymeter. (The accent is on the second, not first, syllable.) Confirming my prejudice, last December the BAAS sheepishly announced the Laugh Lab's winning entry, the funniest joke in the world, according to science:

> Sherlock Holmes and Dr. Watson are going camping. They pitch their tent under the stars and go to sleep. Sometime in the middle of the night Holmes wakes Watson up.

I have to interrupt the joke already to talk about the Reuters news service coverage of this research. *Their* version of the joke opens with, "Famed fictional detective Sherlock Holmes and his gruff assistant Dr. Watson …" (Neil Simon just called, weeping.) I mean, if you're telling the world's funniest joke to someone who doesn't know who Holmes and Watson are, you're like the guy at the aforementioned culinary institute who was just wasting thyme.

Anyway, let us return to the yuckfest. Holmes and Watson are camping. It is dark. We begin:

> HOLMES: Watson, look up at the stars, and tell me what you deduce.
> WATSON: I see millions of stars, and if there are millions of stars, and if even a few of those have planets, it is quite likely that there are some planets like Earth, and if there are a few planets like Earth out there, there might also be life.
> HOLMES: Watson, you idiot, somebody stole our tent!



It's really not an awful joke, especially when you tell it right, which means that you probably don't mention the tent until the punch line, and Holmes just says to look up, not to look up at the stars. Then the listener gets to make all the connections, for an even more delectable aesthetic guffawing experience. (How do I know so much about comedy, you wonder? Easy. I watch a lot of television. Especially the Sunday-morning political roundtables.)

Laugh Lab researchers, whom I rib in the spirit of good humor, might benefit from a reading of Robert R. Provine's book *Laughter: A Scientific Investigation*. Or better yet, they can read Frans B. M. De Waal's review of it in the December 2000 issue of *Scientific American,* as I did. (The mark of the true faux intellectual is to read reviews rather than the actual books, allowing one to comment as if one really knew anything. The faux intellectual also uses words like "faux.")

De Waal notes, "One of the revelations of this book is that the stand-up comedy model of laughter as a response to jokes is mistaken. The large majority of laughs measured by Provine and his students … occurred after statements that were far from humorous," which reveals that social relationships are probably the biggest key to humor. You can tell a joke, but you can't really "tell" Kramer exploding into Jerry's apartment, or Norton addressing the golf ball, or the funeral of Chuckles the Clown. As Heisenberg did not say to the electron, "You just had to be there." SA

ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

How do neon lights work?

—S. Depue, Buffalo, N.Y.

Eric A. Schiff, chair of the department of physics at Syracuse University, provides this explanation:

By definition, the atoms of inert gases such as neon, helium and argon never (well, almost never) form stable molecules by chemically bonding with other atoms. But it is pretty easy to build a gas discharge tube—for instance, a neon light—that reveals that inertness is a relative matter. One need apply only a modest electric voltage to electrodes at the ends of a glass tube containing the inert gas, and a light begins to glow.

At the atomic level, the voltage across a discharge tube accelerates a free electron up to some maximum kinetic energy. The energy provided by the voltage must be larger than that required to ionize the atom. An ionized atom has had an electron plucked out of an orbital to make it a free particle, and the atom it leaves behind has become a positively charged ion. The resulting plasma of positively charged ions and negatively charged electrons carries the electric current between the tube's electrodes.

So why do these gas discharges emit light? As an alternative to being removed by an energetic collision, an electron in an atom can instead become excited. In atomic physics, one speaks of the electron as having been promoted to an orbital of higher energy. When the electron eases back down to its original orbital, a particle of light (a photon) carries away the energy of excitation—and the discharge tube glows. A photon's energy (its wavelength, or color) depends on the energy difference between orbitals. A given atom can emit photons at many energies, corresponding to its orbitals' energy levels. This series of photon energies—the emission lines, to a spectroscopist—is unique to a particular atom. Mercury discharge tubes, for example, have a very different hue (blue) than do neon discharge tubes (orange). The inert gas helium was actually discovered this way. While taking spectral measurements of sunlight during an eclipse in 1868, Norman Lockyer, an English astronomer, saw an emission line that had never before been seen from sources on the earth. He therefore named the new element *helios*, after the Greek word for "sun." SA

For the complete text of this and many other answers from scientists in diverse fields, visit Ask the Experts (www.sciam.com/askexpert).

> *"The most incomprehensible thing about our universe is that it can be comprehended."*
>
> —Albert Einstein

SIDNEY HARRIS

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. 242764. Canadian BN No. 127387652RT; QST No. Q1015332537. Subscription rates: one year $34.97, Canada $49, International $55. 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 or send e-mail to sacust@sciam.com **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.** Printed in U.S.A.