PLANKTON VS. GLOBAL WARMING • SAVING DYING LANGUAGES

SCIENTIFIC AMERICAN

Asynchronous Microchips: Fast Computing without a Clock

AUGUST 2002 $4.95
WWW.SCIAM.COM

IS 95% OF THE UNIVERSE REALLY MISSING?

AN ALTERNATIVE TO

DARK MATTER

Taking the Terror Out of **Terrorism**

The Search for an **Anti-Aging Pill**

august 2002

contents

features

SCIENTIFIC AMERICAN Volume 287 Number 2

42 The puzzle of dark matter

SCIENTIFIC AMERICAN Volume 287 Number 2

departments

columns

Cover image by Cleo Vilett and concept by Ron Miller; page 3: Ron Miller

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.

A Real Pain

How much does it hurt just to *read about* the following cases?

A woman nearly faints as a physician snips tissue from the lining of her uterus—with no pain medication. A man whimpers as he endures, without drugs, a procedure to take a sample of his prostate gland through his rectum. An elderly man with early Alzheimer's has a sigmoidoscope fed into his colon and several polyps clipped off with no palliative.

Unfortunately, these are not scenes from 19th-century medicine. Nor are they departures from generally approved medical practice. Their equivalent occurs thousands of times every day in the U.S. alone.

Word seems to have gotten out that doctors should more aggressively treat severe, long-lasting pain resulting from cancer, chronic syndromes, surgery or terminal illness. After years of public information campaigns, changes in medical school curricula, and educational efforts by physicians' organizations, hospitals are finally updating their practices. Physicians are increasingly encouraged to view pain as the "fifth vital sign," along with pulse, respiration, temperature and blood pressure; in 2001 the Joint Commission on Accreditation of Healthcare Organizations issued guidelines for treating pain in patients with both terminal and nonterminal illnesses.

But the guidelines do not specifically address invasive tests or outpatient surgeries such as those cited above, and many medical practitioners still expect people to keep a stiff upper lip about the pain involved in such procedures. This despite the fact that the tests often already humiliate and frighten patients.

Dermatologists routinely deliver lidocaine when removing moles and such growths from the skin. More and more physicians are offering light anesthesia for colonoscopies. But palliative care must be made universal for people undergoing cancer-screening procedures if physicians want their patients not to avoid the tests.

Why is pain relief not routine in these situations? A major factor is a lack of knowledge. Only recently have researchers shown that lidocaine injections or nitrous oxide (laughing gas) can significantly reduce pain during a prostate biopsy. Lidocaine has not been as successful against the pain of a biopsy of the uterine lining, but a mere handful of studies have been performed, most of them outside the United States. Richard Payne, chief of pain and palliative care at Memorial Sloan-Kettering Cancer Center in New York City, attributes the paucity of U.S. research in this field in part to lack of interest among doctors.

Another reason is risk. Pain-killing drugs and sedatives can have strong effects—in rare cases, life-threatening ones. Monitoring patients to prevent bad outcomes can be costly. Administering a pain reliever or sedative during outpatient surgery could require physicians' offices to have a recovery area where patients could be monitored for side effects until they are alert and comfortable enough to leave. Such a recovery room—and the nursing staff to monitor those in it—would raise costs.

The least forgivable excuse for not alleviating pain would be for medical culture (and maybe society at large) simply to believe that pain ought to be part of medicine and must be endured. Weighing the risks and benefits of pain control should ultimately be the province of the patient. If doctors say there is no pain control for a given procedure, patients should ask why not. People undergoing invasive tests should at least be offered options for pain relief—even if they decide after all to bite the bullet.

THE EDITORS editors@sciam.com

BETTMANN CORBIS

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
Midwest
Derr Media Group
847-615-1921
fax: 847-735-1457
Southeast/Southwest
MancheeMedia
972-662-2503
fax: 972-662-2577
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

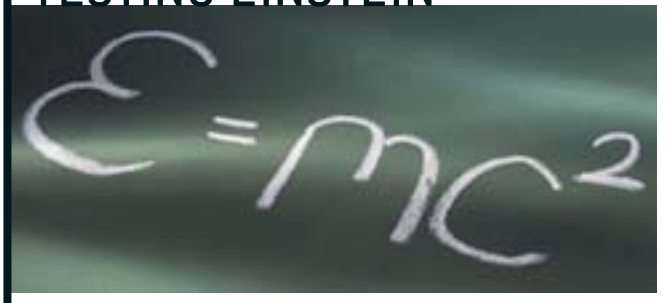
On the Web

WWW.SCIENTIFICAMERICAN.COM

FEATURED THIS MONTH

Visit www.sciam.com/explorations/ to find these recent additions to the site:

TESTING EINSTEIN



The theory of special relativity, which governs the motions of particles at speeds close to that of light, was first published in 1905. It has proved remarkably enduring. Now, nearly a century later, a new generation of high-precision experiments is under way to again put Einstein's brainchild to the test.

2002 Sci/Tech Web Awards

For the second annual ScientificAmerican.com Sci/Tech Web Awards, the editors have done the work of sifting through the virtual piles of pages on the Internet to find the top sites for your browsing pleasure. This eclectic mix of 50 sites—five sites in each of 10 subject categories—runs the gamut from the serious and information-packed to the more whimsical, and even playful, sides of science and technology.

ASK THE EXPERTS

Why is spider silk so strong?
Biologist William K. Purves of Harvey Mudd College explains.
www.sciam.com/askexpert/

SCIENTIFIC AMERICAN CAREERS

Looking to make a career change in 2002?
Visit Scientific American Careers for positions in the science and technology sectors. Find opportunities in computers, sales and marketing, research, and more.
www.scientificamerican.com/careers

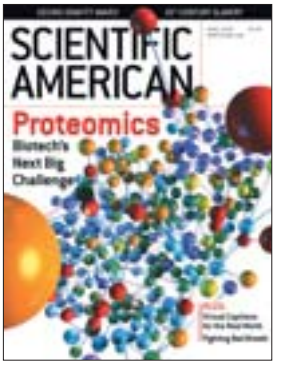
PLUS:

DAILY NEWS ■ DAILY TRIVIA ■ WEEKLY POLLS

SCIENTIFIC AMERICAN®
Established 1845

Letters

EDITORS@SCIAM.COM



"HATS OFF to your illustrator for the most prescient comment about the technology in 'Augmented Reality: A New Way of Seeing,' by Steven K. Feiner [April 2002]," writes Robert Bethune of Ann Arbor, Mich. "More than half the messages flashed to your reality-augmented pedestrian are commercials. I can see myself now, walking down the street wearing my augmented-reality gear. As I glance at the rosebushes, an advertisement for a lawn-and-garden supplier appears. As I look up at the sky, an advertisement for an airline appears. As I hear birds sing, an advertisement urges me to contribute to a wildlife conservation group. And when I shut my eyes to drown out the visual noise, I hear a jingle for a sleep aid." Reader comments on other aspects of reality (or unreality) presented in the April 2002 issue appear below.

RIPPLES VERSUS RUMBLES

After reading "Ripples in Spacetime," I bet I am not alone in coming to the following observation regarding the reported difficulties at U.S. LIGO facilities: duh.

The biggest obstacle cited to obtaining the desired results now or in the future is the low percentage of coordinated multisite online observation time, caused primarily by high ambient noise conditions—including rail and highway traffic and seismic background noise—in the environments around Hanford, Wash., and Livingston, La. This leads one to ask: What were the program managers thinking when they decided on these locations? There are vast remote and sparsely populated areas spread from western Texas to Montana or western Canada where the nearest major highway, rail line or significant commercial airport is literally hundreds of miles away and where there is little seismic activity to speak of. Certainly two such sites could have been found that would have been much "quieter" and still have had sufficient physical separation to support the global array.

Gordon Moller
Grapevine, Tex.

THE PUZZLE OF SLAVERY

The puzzle presented by modern slavery ["The Social Psychology of Modern Slavery," by Kevin Bales] is not that it is so prevalent but that it is so rare. Modern economic theory, from that as expressed in the work of Adam Smith (*The Wealth of Nations*) to the underpinnings of the free trade movement, includes as a central proposition that human labor is fungible. In modern industrial and postindustrial economies, there is little benefit in possessing any single person (with exceptions) any more than there is in possessing any single machine after the initial cost is paid off. This is why apprenticeships usually require an indenture. Likewise, modern business demands that society train its laborers (in school) so that it does not have to pay the cost of such training and the needed maintenance of the laborers during that time.

But in less developed societies, human labor is cheaper than machinery, even though it is in many cases less efficient. A way to reduce slavery is to foster industrialization and modern agriculture, systems in which untrained and unmotivated labor is generally unneeded.

Charles Kelber
Rockville, Md.

Your article omitted an obvious form of state-sponsored slavery that is thriving here in the U.S.: the grotesque growth of the prison-industrial complex that enslaves inmates. These involuntary workers, who have little choice about their situation, are made to produce products on an assembly line for a pittance. Just during the Clinton-Gore years our prison population doubled, exceeding two million incarcerated individuals. Most of this increase resulted from a greater number

of drug-war convictions, many being small-time users and marijuana farmers—hardly a violent bunch. The U.S. is number one in the world's inmate population.

Mel Hunt
McKinleyville, Calif.

I certainly can't disagree with you: slavery is bad. It definitely should be stamped out wherever it exists. Yet slavery is not a proper subject for a magazine I had assumed was devoted to popularizing difficult and leading-edge science.

I can get my politics from the *Nation* on the left and the *National Review* on the right. I can get social discourse from the *Atlantic Monthly* and *Reason*. I don't look to them for science. I do so to you. Unfortunately, I now can't seem to be able to get my science from you without having to wonder about its political slant. That is unsatisfactory. With a seriously heavy heart, I ask you to please cancel my subscription.

Terry Magrath
Marblehead, Mass.

"No social scientist," Kevin Bales writes, "has explored a master-slave relationship in depth." Perhaps not, but a celebrated novelist, who wrote under the pen name B. Traven, has. Experts in the field of slavery and slave rehabilitation may want to look at his work for some guidance. Many of the points covered in Bales's article—that slave masters view themselves as father figures, that slaves find comfort in the stability of peonage—were examined by Traven more than 70 years ago.

Known mostly for *The Treasure of the Sierra Madre*, Traven moved to Mexico in the 1920s and saw firsthand a system of debt slavery being practiced in the southern states of that country—despite the fact that a recent revolution had made debt slavery illegal. Traven lived with indigenous Mexicans for many years, learning the local languages and following the local customs, then documented what he had seen and heard in his novels. His six "jungle" books outline the psychology of slaves and slaveholders, the relation between the two groups, and the societal mechanisms that encourage such an abusive, unjust system.

Bill Bendix
Toronto

PERSPECTIVES ON PERSPECTIVES

It was with a sigh that I read the editors' introduction to the April issue [SA Perspectives], addressing the Bales article on the social psychology of modern slavery. You mention that when you run a social science article with important implications, you receive a large amount of mail complaining about "mushy" and "political" articles that are not "real science."

I have been a *Scientific American* reader since my youth. I am also a professor of sociology. I certainly do not complain when the magazine runs a physical science article with important implications. I wonder what makes your complaining readers so defensive. The social sciences and the natural sciences have their differences, but that hardly means one of them is invalid. It is true that the political implications of much social science are clear to see. But many studies of all stripes have political implications, whether acknowledged or not. Should *Scientific American* not publish articles on global warming or AIDS or nuclear technology because their political implications are widely debated?

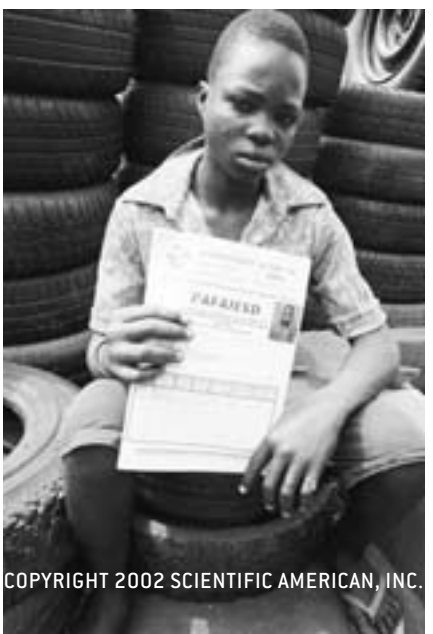
I have a background in the biological sciences, which I believe helps me in my social science research. Often when I read hard-science pieces, I feel the authors would benefit from a social science background as well.

Carrie Yang Costello
Milwaukee

In your April issue, you mention that some readers protest *Scientific American*'s becoming "more politicized." I want to assure you that many of your readers do not regard such a change as unwelcome. My favorite section of the magazine has become SA Perspectives. Almost invariably, the page discusses a problem for which the current political approach has frustrated me considerably. Therefore, the clearly *rational* approach espoused by the editors is overwhelmingly refreshing. In the best of all worlds, science and politics might progress along separate, undisturbed lines. Unfortunately, politics intimately guides future directions in science, and science obviously has much to say concerning the implications of political policy. My only hope is that ways of thinking such as those evidenced in SA Perspectives become as widespread as possible.

Jacob Wouden
Boston

CONSTANT AYITCHEOU escaped domestic servitude in Nigeria.



MICHAEL ST. MAUR SHEIL *Black Star*

ERRATA "Proteins Rule," by Carol Ezzell, failed to note that the earliest x-ray crystallography was done using x-ray beams from laboratory sources, not synchrotrons. It also erroneously stated that Syrrx uses x-ray lasers.

In a world map in "The Psychology of Modern Slavery," by Kevin Bales, the colors used for France and Italy should have been switched; slavery is worse in Italy.

Edward R. Generazio is chief of the NASA Langley Research Center's nondestructive evaluation (NDE) team ["Heads on Tails," by Phil Scott, News Scan], not Samuel S. Russell, who is a member of the NDE team at the NASA Marshall Space Flight Center.

50, 100 & 150 Years Ago

FROM *SCIENTIFIC AMERICAN*

Dead Locusts ▪ Threatened Sponges ▪ Scared Wolves

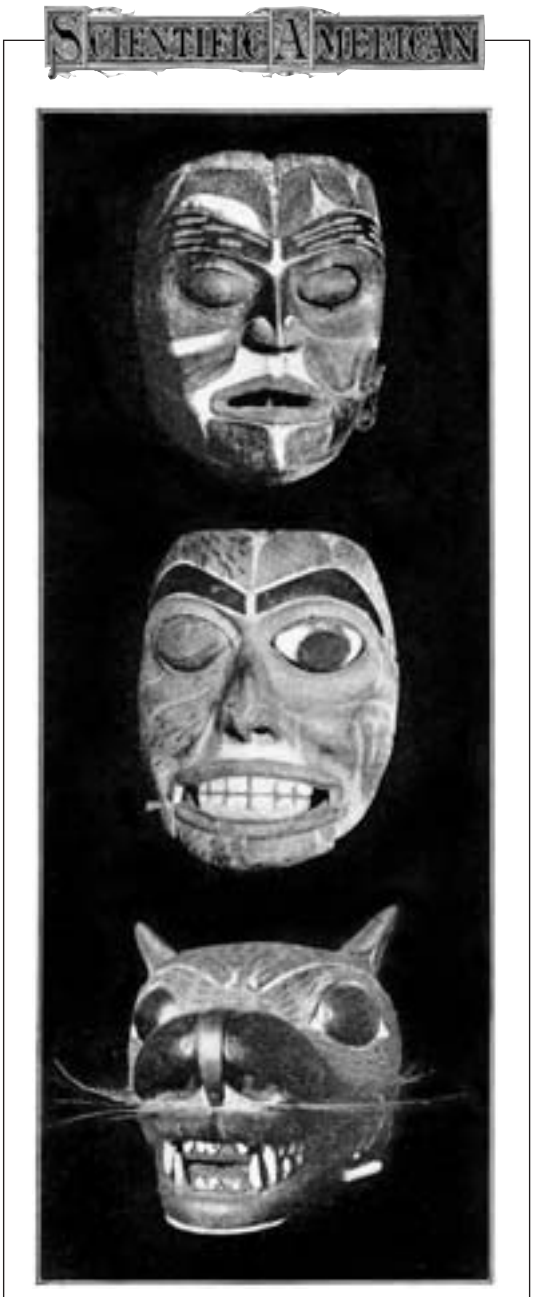
AUGUST 1952

CHEMICAL AGRICULTURE—"In March 1951, the Iranian Government asked the U.S. for immediate help in an emergency. Swarms of locusts were growing so rapidly along the Persian Gulf that they threatened to destroy the entire food crop. The U.S. responded by sending some army planes and about 10 tons of the insecticide aldrin, with which they sprayed the area. The operation almost completely wiped out the locusts overnight. For the first time in history a country-wide plague of locusts was nipped in the bud. This dramatic episode illustrates the revolution that chemistry has brought to agriculture. Chemical agriculture, still in its infancy, should eventually advance our agricultural efficiency at least as much as machines have in the past 150 years." [*Editors' note: The United Nations's Stockholm Convention of 2001 banned the production of aldrin and other persistent organic pollutants.*]

AUGUST 1902

PACIFIC COAST MASKS—"The accompanying illustrations depict curious masks worn by the Tghimpsean [Tsimshian] tribe of Indians, on the Pacific coast of British Columbia, on the Skeena River. They were secured by a Methodist missionary—Rev. Dr. Crosby—who labored among them, and these False Faces [*sic*] are now to be seen in the museum of Victoria College in Toronto."

SPONGE FISHING—"Greek and Turkish sponges have been known to the trade for hundreds of years. Syria furnishes perhaps the finest quality. During the last fifteen years, however, the output has greatly diminished, owing to the introduction by Greeks, in the seventies, of diving apparatus, which proved ruinous to fishermen and fisheries alike. The 'skafander' enables the diver to spend an hour under water. This method is a severe tax upon the sponge banks, as everything in sight—sponges large and small—is gathered, and it takes years before a new crop matures."



MASKS of Pacific Coast Indians, 1902

MODERN TIMES—"To point to the hurry and stress of modern town life as the cause of half the ills to which flesh to-day is heir has become commonplace. However, children cope more easily with the new necessities of life, and new arrangements which perplexed and worried their parents become habits easily borne. Thus we may imagine future generations perfectly calm among a hundred telephones and sleeping sweetly while airships whizz among countless electric wires over their heads and a perpetual night traffic of motor cars hurtles past their bedroom windows. As yet, it must be sorrowfully confessed, our nervous systems are not so callous."

AUGUST 1852

NAVAL WARFARE—"Recent experiments with the screw propeller, in the French navy, have settled the question of the superior economy and advantages of uniting steam with canvas [sails] in vessels of war. The trial-trip of the *Charlemagne,* which steamed to the Dardanelles and back to Toulon, surpassed all expectations. A conflict between French and American ships, the former using both steam and canvas, and our own vessel only the latter, would be a most unequal struggle. The advantage would lie altogether with the Frenchman, as he would be able to rake his adversary's decks at will and attack him on every side."

FEARFUL WOLVES—"It is said that since the completion of the railroad through Northern Indiana, the wolves which came from the North, and were so savage on flocks in the South, have not been seen south of the track. The supposition is that the wolves mistrust the road to be a trap, and they will not venture near its iron bars."

DEFENSE TECHNOLOGY

Nuclear Reactions

SHOULD NUCLEAR WARHEADS BE USED IN MISSILE DEFENSE? BY DANIEL G. DUPONT

On April 11 the *Washington Post* ran a piece asserting that U.S. Secretary of Defense Donald Rumsfeld had "opened the door" to the possible use of nuclear-tipped interceptors in a national missile defense system. The story cited comments made by William Schneider, Jr., chairman of an influential Pentagon advisory board, who told the *Post* that Rumsfeld had encouraged the panel to examine nuclear interceptors as part of a broad missile defense study.

The article kicked off the first public discussion of nuclear interceptor missiles in many years. Opponents thought they were dead and buried after a nuclear system, called Safeguard, was briefly considered in the 1970s. Military planners found them too risky because their use against even a handful of Soviet intercontinental ballistic missiles would blind American satellites and sensors, increasing the likelihood that subsequent ICBMs would hit their marks.

Today the U.S. focuses on the threat of only a few long-range missiles fired by "rogue" states or terrorists or launched accidentally from Russia or China. But the Pentagon's prototype conventional interceptors, which are scheduled to be ready no sooner than 2005, have not demonstrated much ability to discriminate between ICBMs and decoy balloons or other "countermeasures" they might release. If the administration decides that the threat is so pressing that it cannot wait for these conventional missile defenses to prove themselves, a nuclear option could gain support. "If all you're worried about is one or two North Korean warheads, and to the extent that you are concerned about near-term discrimination, then you could probably talk yourself into the possibility that a multimegaton hydrogen bomb would solve a lot of very difficult discrimination problems," says John E. Pike, a longtime missile defense critic who heads GlobalSecurity.org, an organization devoted to reducing reliance on nuclear weapons.

The Pentagon maintains that it is not, in fact, looking at nuclear interceptors. Still, Republicans in the House lauded the Pentagon's "examination of alternatives" to current missile defense plans, including nuclear interceptors, as a "prudent step, consistent with the commitment to evaluate all available technological options," as stated in a House Armed Services Committee report.

House Democrats think it's a bad idea—and indicative of a creeping, if largely unpublicized, Republican willingness to erode longstanding tenets of U.S. nuclear policy. "I think we need to be careful," says Representative Thomas H. Allen of Maine, citing Schneider's remarks about nuclear interceptors as well as other hints of policy changes. These include a House defense bill that, Democrats declared

NO NUKES—YET: Hit-to-kill is the current philosophy behind antiballistic missiles, such as this one test-launched last December in the Marshall Islands.

REUTERS/HO/USAF

in a statement, "encourages the U.S. to develop new nuclear weapons for the first time since the cold war," a reference to the nuclear "bunker busters" that officials may consider to destroy deeply buried targets.

A crude nuclear missile defense system could comprise a Minuteman ICBM equipped with new software and timed to blow up in the vicinity of an incoming missile. If done right, the explosion would destroy the missile and any "bomblets." Discrimination among missiles and decoys would not be an issue.

If the warhead atop the interceptor were too small or the missile too sophisticated, however, some bomblets—which could be packed with biological or chemical agents—could leak through. For this reason, physicists such as Kurt Gottfried, head of the Union of Concerned Scientists, and Richard L. Garwin, a senior fellow on the Council on Foreign Relations, worry that a huge warhead would be needed to make a nuclear interceptor system effective. But the bigger the warhead, the greater the potential for destruction of commercial and military satellites—and for damage on the ground. As Pike notes, explosions in space are far less constrained than they are on the earth.

Some prominent Democrats, using language harking back to the controversy over Safeguard in 1970s, oppose any discussion of deploying nuclear interceptors. They unsuccessfully pushed for legislation to make it U.S. policy not to use such systems. They have managed to call on the National Academy of Sciences (NAS) to study the possible effects of nuclear explosions in space on cities and people.

Some scientists do not believe that radioactive fallout would be a major problem—Gottfried remarks that it would be "minimal in comparison" to a ground detonation. But Allen, who backs the NAS study, isn't sure. "Raining radiation down on the American people," he says, "is a bad idea."

Daniel G. Dupont, based in Washington, D.C., edits InsideDefense.com, an online news service.

SATELLITE WIPEOUT

According to physicist K. Dennis Papadopoulos of the University of Maryland, the radiation produced by a nuclear detonation in space would have a widespread and lasting effect on satellites, especially commercial spacecraft. Citing work done by the Defense Department, among others, he says the radiation could wipe out 90 percent of the satellites in low-earth orbit in three weeks.

ATMOSPHERIC SCIENCE

A Push from Above

THE OZONE HOLE MAY BE STIRRING UP ANTARCTICA'S CLIMATE **BY SARAH SIMPSON**

About 3,250 square kilometers of Antarctica's Larsen B ice shelf shattered and tore away from the continent's western peninsula early this year, sending thousands of icebergs adrift in a dramatic testimony to the 2.5 degrees Celsius warming that the peninsula has experienced since the 1950s. Those wayward chunks of ice also highlighted a perplexing contradiction in the climate down under: much of Antarctica has cooled in recent decades.

Two atmospheric scientists have now resolved these seemingly disparate trends. David W. J. Thompson of Colorado State University and Susan Solomon of the National Oceanic and Atmospheric Administration Aeronomy Laboratory in Boulder, Colo., say that summertime changes in a mass of swirling air above Antarctica can explain 90 percent of the cooling and about half of the warming, which has typically been blamed on the global buildup of heat-trapping greenhouse gases. But this new explanation doesn't mean that people are off the hook. Thompson and Solomon also found indications that the critical atmospheric changes are driven by Antarctica's infamous ozone hole, which grows every spring because of the presence of chlorofluorocarbons (CFCs) and other human-made chemicals in the stratosphere.

Their analysis of 30 years of weather balloon measurements and additional data from stations across the continent is the first evidence strongly linking ozone depletion to climate change. It also joins the ranks of a growing body of research linking changes in the lowermost atmosphere, or troposphere, to the overlying stratosphere (between about 10 and 50 kilometers above the earth's surface).

Thompson and Solomon first noticed that

STORM CLUES FROM THE STRATOSPHERE

Keeping an eye on the stratosphere may eventually enable meteorologists to better predict weather at the earth's surface. Last year Mark P. Baldwin and Timothy J. Dunkerton of Northwest Research Associates in Bellevue, Wash., reported that large variations in stratospheric circulation over the Arctic typically precede anomalous weather regimes in the underlying troposphere. For instance, these stratospheric harbingers seem to foretell significant shifts in the probable distribution of extreme storms in the midlatitudes. If the correlations are strong enough, Baldwin says, weather prediction may be possible regardless of the driving force behind the stratospheric changes—whether it's ozone loss or something else.

an abnormally intense westerly flow of air in the troposphere seemed to be encircling the continent every summer. By bottling up cold air over the pole and restricting warm air to the outer ring, this circumpolar vortex could cool Antarctica's interior and warm its extremities. The odd thing was that the vortex—which is usually fueled by the frigid temperatures of the dark polar winter—should have dissipated in the spring, allowing the warm and cold air to mix. To discover what was prolonging the wintry behavior, the researchers looked to the overlying stratosphere.

Other scientists had already established that dramatic ozone losses—exceeding 50 percent in October throughout the 1990s—have cooled the springtime stratosphere by more than six degrees C. In examining hundreds of measurements, Thompson and Solomon saw an unmistakable correlation between the two atmospheric layers: the ozone-induced springtime cooling must propagate downward in the ensuing weeks, fueling the troposphere's wintry vortex into the summer months of December and January.

No one had made this connection before in part because conventional meteorological wisdom has long accepted that the stratosphere—which contains, on average, a mere 15 percent of the mass of the atmosphere—is too weak to push around the denser gases below it. "Most people assumed it would be like the tail wagging the dog," Thompson says. But he knew that a handful of studies from the past five years had suggested that the stratosphere also influences the Northern Hemisphere's circumpolar vortex, which modifies the climate in North America and Europe.

Some investigators think that the vortex's ultimate driving force may instead come from below. "It's quite possible that ozone loss plays a role in [Antarctica's] changes, but other candidates must be considered," says James W. Hurrell of the National Center for Atmospheric Research in Boulder. Using computer simulations, he and his colleagues discovered last year that recent natural warming of the tropical oceans can account for a vast majority of the observed changes in the Northern Hemisphere's vortex. Both Hurrell and Thompson say that such sea-surface warming could presumably work with ozone depletion to alter the Antarctic vortex as well.

If ozone loss indeed helps to drive Antarctica's summer climate, "we may see that the peninsula doesn't continue to warm in the same way," Solomon predicts. The yearly hole should begin to shrink as Antarctica's atmosphere slowly rids itself of the now restricted CFCs. Then maybe Antarctica's summertime vortex can relax a little.

SPLASHDOWN: Warmer temperatures caused the northern parts of Antarctica's Larsen B ice shelf to crumble into the ocean.

NASA/GSFC/LaRC/JPL, MISR TEAM

GENETICS

Killing the Messenger

TURNING OFF RNA COULD THWART CANCER AND AIDS

BY CAROL EZZELL

Like a jealous chef, the nucleus guards the recipes for making all the proteins a cell needs. It holds tightly to the original recipe book, written in DNA, doling out copies of the instructions in the form of messenger RNA (mRNA) to the cell's cytoplasm only as they are needed.

But in cancer cells and in cells infected by viruses, those carefully issued orders are often drowned out. Mutations can cause cancer cells to issue garbled directions that result in aberrant proteins. Similarly, viruses flood cells with their own mRNA, hijacking the cells' protein-making apparatuses to make copies of new viruses that can go on to infect other cells.

Researchers are currently attempting to harness a recently discovered natural phenomenon called RNA interference (RNAi) to intercept, or "knock down," such bad messages. They are using the technique to identify the functions of genes in fruit flies and microscopic worms—a strategy that will help them uncover what the corresponding genes

do in humans. And medical scientists are hoping to deploy RNAi as a treatment for AIDS and cancer or as a preventive against the rejection of transplants.

RNAi was first identified in 1998 in nematodes, in which it appears to serve as a way to block the proliferation of so-called jumping genes (transposons). In other organisms, cellular enzymes that are part of the RNAi machinery specifically target stray bits of double-stranded RNA, which could arise from viruses, and pull the strands apart. Then, in a process that scientists are just now beginning to understand, the freed strands go on to bind to and inactivate any mRNA that has the complementary genetic sequence.

Several groups of researchers have reported within the past few months that the same phenomenon could have therapeutic benefits in mammalian cells by shutting down viral genes or those linked to tumor formation. Two groups—one led by John J. Rossi of the Beckman Research Institute at the City of Hope Cancer Center in Duarte, Calif., and another by Phillip A. Sharp of the Massachusetts Institute of Technology—engineered human cells to make double-stranded RNAs that match sequences from HIV. They found that those RNAs completely prevented the cells from making HIV proteins after the cells were infected. Other scientists have used RNAi to knock down cancer-related proteins such as beta-catenin and p53.

Rossi says that his group plans to try the approach in people with HIV in the next couple years. One strategy he and his colleagues are considering is to remove T cells (which are normally decimated by HIV) from the individuals, use viruses to ferry into them genetic sequences for double-stranded RNAs that encode HIV proteins, and then infuse the altered cells back into the patients' bodies. "We'd like to create a pool of T cells that would resist HIV," thereby enabling HIV-infected people to stay healthy, he says.

In another application, two of Rossi's colleagues, Laurence Cooper and Michael Jensen, are using RNAi to generate a kind of cellular universal soldier for attacking cancer. Cooper, Jensen and their co-workers are invoking the technique to prevent killer T cells from making proteins that mark them as "self" or "nonself." These universal killer cells, they hope, will be infused into cancer patients without being eliminated as foreign.

Brenda L. Bass, a Howard Hughes Medical Institute investigator at the University of Utah, comments that RNAi has been shown so far in laboratory experiments to be more efficient at stifling gene expression than another technique called antisense RNA. Several antisense drugs, which also function by binding to mRNA, are now in human tests, but their mechanism of action is still poorly understood, and they are somewhat fickle. "With antisense, you don't know if it will work on your given gene," Bass says. "It is clear that RNAi is much, much more effective."

GETTING INTO THE GROOVE

Another method for shutting down genes is to bypass the messenger and go straight to the source: the DNA helix. Michael J. Hannon of the University of Warwick in England and his co-workers have devised synthetic molecular cylinders that can bind to the major groove formed by the twisting of the DNA helix. The cylinders bend the DNA so that its genes are inaccessible to the enzymes that normally convert the genetic instructions to mRNA. But so far the process is nonspecific: the scientists have yet to figure out how to target particular genes for shutdown.

PHYSICS

Coherent Computing

MAKING QUBIT SUPERPOSITIONS IN SUPERCONDUCTORS LAST LONGER **BY JR MINKEL**

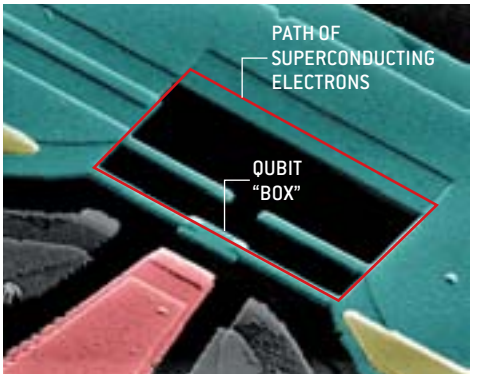
In the race for quantum computers, researchers would love to sculpt quantum bits, or qubits, with existing chip-manufacturing techniques. Qubits are able to exist as both 0 and 1 at once. But the same wires that would make qubit chips easy to manipulate and link together also connect them to the quantum-defiling messiness of the outside world. So far no one has figured out how to make the 0/1 superposition last very long in circuit form.

Three different research teams, however, have made critical breakthroughs. In May the Quantronics group at the Atomic Energy Commission (CEA) in Saclay, France, and Siyuan Han's laboratory at the University of Kansas reported qubit chip designs with coherence times at least 100 times as great as those achieved before. Investigators at the National Institute of Standards and Technology (NIST) in Boulder, Colo., have come up with a design that they think could yield similar coherence rates. "What people are beginning to understand is how to build circuits so that

these qubits aren't coupled to noise and dissipation," says NIST physicist John M. Martinis.

To make qubits, physicists have mostly relied on atoms, trapping individual ones in cavities or using nuclear magnetic resonance effects in certain liquids. In fact, they recently managed to factor the number 15 with the latter approach. These systems, however, are cumbersome compared with modern circuit designs. Unfortunately, the electrical circuits themselves are much larger than atoms, so their quantum states decohere more rapidly.

In 1999 Yasunobu Nakamura and his coworkers at NEC in Japan demonstrated coherence in a sliver of aluminum separated a short distance by an insulator from another piece of aluminum. Near absolute zero, the aluminum becomes superconducting—electrons group into twos (called Cooper pairs) and move without resistance. An applied voltage caused the Cooper pairs to tunnel back and forth across the insulator, the presence or absence of the pair defining 0 and 1. The coherent state fell apart after just nanoseconds, however, mainly because charges would periodically tunnel into the qubit from the electrode used to measure its state.



QUANTUM BIT is represented by the presence or absence of a pair of superconducting electrons in a "box" of aluminum. The gold "wings" collect stray unpaired electrons; the red port sends in microwave pulses to "write" the qubit.

Now the Saclay researchers have situated a similar Cooper pair "box" within a superconductor loop, which allowed them to measure the qubit's state using current instead of charge. By tuning the voltage and magnetic field threading through the loop, the team could adjust the energy difference between the qubit's two states such that the system resonated very well with microwave pulses of a specific frequency but poorly with others. At resonance, the two states presented no net charge or current, giving noise little room to disturb the balance between them. "We have chosen [these states] so they are as inconspicuous as possible" to the disrupting effects of the outside world, explains Saclay co-author Michel H. Devoret, now at Yale University.

The states' isolation ends—and the qubit reveals its state—at the touch of an applied current. It transforms the superposed qubit states into currents flowing in opposite directions through the loop. Noise rapidly collapses the pair of currents into just a single induced current, whose direction depends on the state of the qubit. If the applied and induced currents head the same way, they will add up to exceed the system's threshold for superconductivity, creating a measurable voltage.

Sending in microwave pulses of varying durations and time delays enabled the researchers to control the qubit and told them how it evolved. Notably, it remained coherent for around half a microsecond: still low, but good enough that the group is now working on coupling two qubits together. "That result has changed the whole landscape," states Roger H. Koch of IBM, who studies a different superconductor design.

The Kansas and the NIST qubit chips both utilized larger superconductor elements, which are currently easier to fabricate. Passing a current through such junctions isolates a pair of energy levels—0 and 1—which can be superposed. To block out the world, both groups filtered their circuits meticulously. The Kansas team inferred a coherence time of at least five microseconds, although the investigators could not yet read out the qubit's state at will. The NIST circuit lasted only around 10 nanoseconds. Martinis says that building it out of aluminum could eventually boost that to microseconds.

Researchers continue to play with various circuit designs. "At this point it's hard to say exactly which approach is better in the long run," Devoret remarks. And despite the tremendous progress, Koch adds, coupling many qubits on a chip remains an extremely challenging problem, because each individual qubit resonates at a unique frequency. "As an experimentalist, that's what scares me," he says. "There's still room for a fundamental breakthrough."

JR Minkel is based in New York City.

QUBIT BY QUBIT

Superposition is the power behind quantum computers. A single quantum bit, or qubit, can lead a double life, consisting of any arbitrary superposition of 0 and 1. But two conjoined qubits can lead a quadruple life, three can follow an eightfold path, and so on, existing as every possible combination of 0 and 1 at once. Computational power thus grows exponentially in quantum computers, not bit by bit as in classical ones.

QUANTRONICS-DSM-CEA

Crop Circle Confession

HOW TO GET THE WHEAT DOWN IN THE DEAD OF NIGHT

BY MATT RIDLEY

On August 2, Touchstone Pictures released Signs, *starring Mel Gibson as a farmer who discovers mysterious crop circles. Directed by* Sixth Sense *auteur M. Night Shyamalan, the movie injects otherworldly creepiness into crushed crops. The truth behind the circles is, alas, almost certainly more mundane: skulking humans. Herewith is the account of one such trickster.*

I made my first crop circle in 1991. My motive was to prove how easy they were to create, because I was convinced that all crop circles were man-made. It was the only explanation nobody seemed interested in testing. Late one August night, with one accomplice—my brother-in-law from Texas—I stepped into a field of nearly ripe wheat in northern England, anchored a rope into the ground with a spike and began walking in a circle with the rope held near the ground. It did not work very well: the rope rode up over the plants. But with a bit of help from our feet to hold down the rope, we soon had a respectable circle of flattened wheat.

Two days later there was an excited call to the authorities from the local farmer: I had fooled my first victim. I subsequently made two more crop circles using far superior techniques. A light garden roller, designed to be filled with water, proved helpful. Next, I hit on the "plank walking" technique that was used by the original circle makers, Doug Bower and the late Dave Chorley, who started it all in 1978. It's done by pushing down the crop with a plank suspended from two ropes. To render the depression circular is a simple matter of keeping an anchored rope taut. I soon found that I could make a sophisticated pattern with very neat edges in less than an hour.

Getting into the field without leaving traces is a lot easier than is usually claimed. In dry weather, and if you step carefully, you can leave no footprints or tracks at all. There are other, even stealthier ways of getting into the crop. One group of circle makers uses two tall bar stools, jumping from one to another.

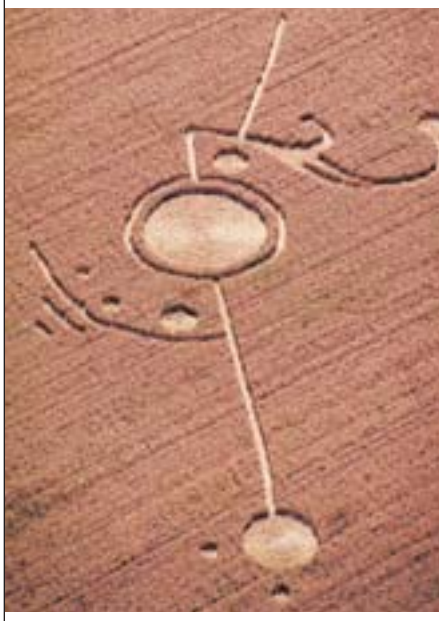
But to my astonishment, throughout the early 1990s the media continued to report that it was impossible that all crop circles could be man-made. They cited "cerealogists"—those who study crop circles—and never checked for themselves. There were said to be too many circles to be the work of a few "hoaxers" (but this assumed that each circle took many hours to make), or that circles appeared in well-watched crops (simply not true), or that circle creation was accompanied by unearthly noises (when these sounds were played back, even I recognized the nocturnal song of the grasshopper warbler).

The most ludicrous assertion was that "experts" could distinguish "genuine" circles from "hoaxed" ones. Even after one such expert, G. Terence Meaden, asserted on camera that a circle was genuine when in fact its construction had been filmed by Britain's Channel Four, the program let him off the hook by saying he might just have made a mistake this time. I soon met other crop-circle makers, such as Robin W. Allen of the University of Southampton and Jim Schnabel, author of *Round in Circles,* who also found it all too easy to fool the self-appointed experts but all too hard to dent the gullibility of reporters. When Bower and Chorley confessed, they were denounced on television as frauds. My own newspaper articles were dismissed as "government disinformation," and it was hinted that I was in the U.K. intelligence agency, MI5, which was flattering (and false).

The whole episode taught me two important lessons. First, treat all experts with skepticism and look out for their vested interests—many cerealogists made a pot of money from writing books and leading weeklong tours of crop circles, some costing more than $2,000 a person. Second, never underestimate the gullibility of the media. Even the *Wall Street Journal* published articles that failed to take the man-made explanation seriously.

As for the identity of those who created the complicated mathematical and fractal patterns that appeared in the mid-1990s, I have no idea. But Occam's razor suggests they were more likely to be undergraduates than aliens.

Matt Ridley, based in Newcastle-upon-Tyne, England, wrote Genome: The Autobiography of a Species in 23 Chapters.



DAVE OLSON *AP Photo*

SIGNS OF TERROR or just of human mischief? Wheat flattened in 1998 in Hubbard, Ore. The center circle is about 35 feet wide.

SLICED OFF BY OCCAM'S RAZOR

Scientists and skeptics like to rely on a principle first enunciated by medieval philosopher William of Occam—namely, that the simplest explanation tends to be the best. In the case of crop circles, that would be mischievous humans. Other explanations that have been proposed:

- **Spaceship landings**
- **Unusual wind patterns**
- **Tornadoes**
- **Ball lightning**
- **Strange force fields**
- **Plasma vortices (whatever they are)**
- **The Devil**
- **Rutting hedgehogs**

To learn about an organization dedicated to making crop circles, go to www.circlemakers.org

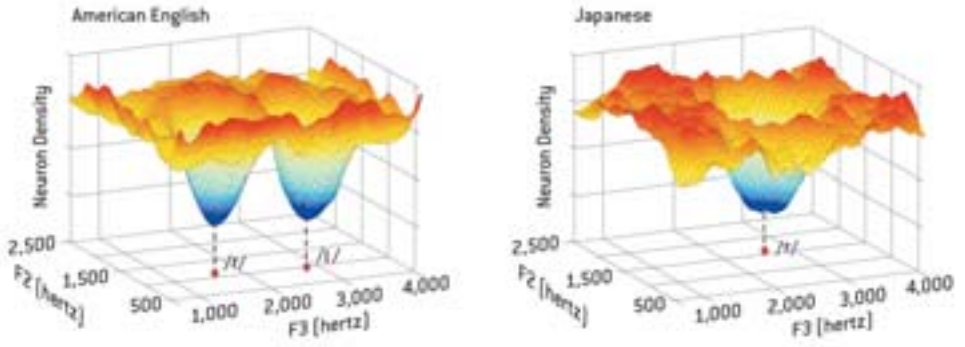
COGNITIVE SCIENCE

From Mouth to Mind

NEW INSIGHTS INTO HOW LANGUAGE WARPS THE BRAIN **BY W. WAYT GIBBS**

"Liver." The word rises from the voice box and passes the lips. It beats the air, enters an ear canal, sets nerve cells firing. Electrochemical impulses stream into the auditory cortex of a listener's brain. But then what? How does the brain's neural machinery filter that complex stream of auditory input to extract the uttered word: "liver"—or was it "river," or perhaps "lever"?

Researchers at the Acoustical Society of America meeting in June reported brain imaging studies and clinical experiments that expose new details of how the first language we learn warps everything we hear later. Some neuroscientists think they are close to explaining, at a physical level, why many native Japanese speakers hear "liver" as "river," and why it is so much easier to learn a new language as a child than as an adult.



NEURAL NETWORK, when trained on the sounds of American speech, devotes lots of cells to distinguishing /r/ from /l/. But when trained on Japanese phonemes, the neurons organize so that a few cells are sensitive to the dominant frequencies (called F2 and F3) that differ between /r/ and /l/.

At the ASA conference, Paul Iverson of University College London presented maps of what people hear when they listen to sounds that span the continuum between the American English phonemes /ra/ and /la/. Like many phonemes, /ra/ and /la/ differ mainly in the three or four frequencies that carry the most energy. Iverson had his computer synthesize sounds in which the second and third most dominant frequencies varied in regular intervals, like dots on a grid. He then asked English, German and Japanese speakers to identify each phoneme and to rate its quality.

What emerged was a map of how our experience with language warps what we think we hear. Americans labeled half the sounds /la/ and half /ra/, with little confusion. Germans, who are used to hearing a very different /r/ sound, nonetheless categorized the extreme ends of the spectrum similarly to the way English speakers did. But the map for Japanese speakers showed an entirely different perceptual landscape. "The results show that it's not that Japanese speakers can't hear the difference between /r/ and /l/," Iverson says. "They are just sensitive to differences that are irrelevant to distinguishing the two"—differences too subtle for Americans to perceive. Japanese speakers, for example, tend to pay more attention to the speed of the consonant.

Frank H. Guenther of Boston University reported building a neural network model that may explain how phonetic categories arise naturally from the organization of the auditory cortex. In his simulation, neurons are rewarded for correctly recognizing the phonemes of a certain language. In response they reorganize; most become sensitive only to ambiguous sounds that straddle categories. The simulated cortex loses its ability to distinguish slightly different, but equally clear, phonemes. Humans—as well as monkeys, chinchillas and even starlings—show just such a "perceptual magnet" effect.

When trained using Japanese speech sounds, the model neurons organized very differently from those trained on English. A pronounced dip in sensitivity appeared right at the border of /ra/ and /la/. This may reflect how "our auditory systems get tuned up to be especially sensitive to the details critical in our own native language," Iverson says. "When you try to learn a second language, those tunings may be inappropriate and interfere with your ability to learn the new categories."

Guenther scanned the brains of English-speaking volunteers as they listened to good and borderline examples of /ee/ phonemes. As predicted, malformed phoneme vowels activated more neurons than did normal ones. Collaborators in Japan are replicating the experiment with subjects who know no English. "We expect to see a very different pattern of activation," Guenther says.

COURTESY OF FRANK H. GUENTHER *Boston University*

SOUNDS OF SPEECH

Linguists surveying the world's languages have counted at least 558 consonants, 260 vowels and 51 diphthongs, according to Lori L. Holt of Carnegie Mellon University. Infants appear to be able to distinguish all 869 phonemes up to the age of about six to eight months. After that, the brain sorts all the sounds of speech into the much smaller subset of phonetic categories in its native language. American English uses just 52, Holt says. The Kalahari desert language !Xũ holds the world record at 141.

BY THE NUMBERS

Down on the Farm

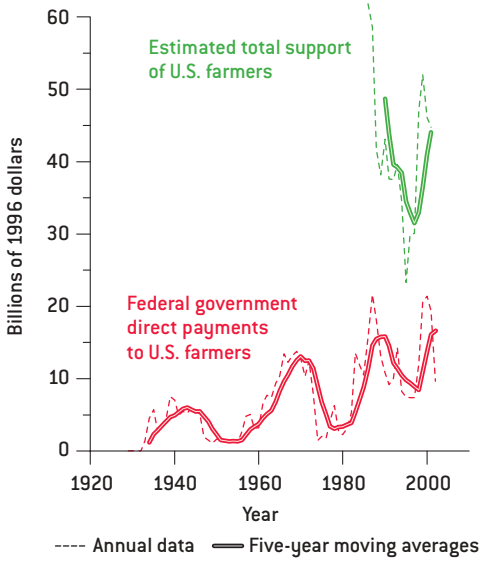
WHEN THE BIGGEST CROP IS DOLLARS **BY RODGER DOYLE**

In May, President George W. Bush signed into law the Farm Security and Rural Investment Act of 2002, the latest in a series of farmer support legislation going back to 1933. Critics say that the law, which increases spending by more than 70 percent, will further undermine the faltering free-trade movement. It gives big farmers unprecedentedly large payments that can be used to buy out small farmers and, they maintain, hurts farmers in developing countries, who cannot compete with low, subsidized American prices. On the plus side, the act gives a strong economic stimulus to agricultural states, such as Iowa.

Only farmers who produce certain crops benefit directly from subsidies: growers of corn, wheat, oilseeds, rice and cotton got more than 90 percent of payments in 1999, yet they accounted for less than 36 percent of total agricultural output. Those who produced cattle, hogs, poultry, fruit, vegetables and other products received no payments. All farmers, however, can benefit from U.S. Department of Agriculture programs, including those for conservation and subsidized crop insurance. The prices of some crops, such as sugar, are kept high by restrictive tariffs. Of the estimated 1.9 million farms in the U.S., those with $250,000 or more in sales—about 7 percent of the total number—received 45 percent of direct federal subsidies in 1999.

The chart shows the trend of all federal payments to farmers since the early 1930s, as reported by the USDA. Also shown are estimates by the Organization for Economic Cooperation and Development of total farmer support, which factors in payments by states, the market value of protective tariffs, and the value of noncash subsidies—those for energy, crop insurance, rights to graze on public lands and so forth.

The 1996 farm bill was supposed to phase out subsidies completely, but it was subverted when Congress voted to restore large-scale subsidies as grain prices plummeted in the late 1990s. How have farmers—who, together with their families, number less than the population of Manhattan, Brooklyn and Queens—been able to get more federal support than comparable groups such as union members, who outnumber them eight to one? Deft political maneuvering is one answer. Another is the cultivation of a politically potent symbol—the family farm.



SOURCES: U.S. Department of Agriculture and the Organization for Economic Cooperation and Development

The notion of the family farm appeals to the imagination of Americans, but it is now more myth than fact: as the *Des Moines Register* put it in a March 2000 article, "Family farming, as Iowa knew it in the 20th century, is irreversibly gone. A countryside dotted with 160-acre farms, each with its house and barn and vegetable garden, has all but disappeared." Agricultural economist William M. Edwards of Iowa State University says that a 160-acre farm is no longer practical for the principal commodities in Iowa: corn, soybeans and livestock. The minimum size for a viable Iowa farm growing corn and soybeans and raising livestock is, in his view, roughly 500 acres; for a farm devoted wholly to corn and soybeans, the minimum is about 1,000 acres. What has happened in Iowa is generally true of other areas: during the past 50 years, average farm size has grown nationally, primarily because of the need to buy new, more productive technology that can be more effectively employed on bigger tracts.

Rodger Doyle can be reached at rdoyle2@adelphia.net

UPS AND DOWNS OF FARMING

Net farm income (average annual, in billions of 1996 dollars)	
1930s	35
1960s	51.2
1990s	46.9
2002 (forecast)	36.7
Average farm size (acres)	
1930	157
1964	352
1997	487
Number of farms (thousands)	
1930	6,295
1964	3,157
1997	1,912

SOURCES: Bureau of Economic Analysis, the U.S. Department of Commerce and the U.S. Department of Agriculture

RODGER DOYLE

DATA POINTS: TECH SHY

You wouldn't know it from the sales of cellular phones (53.4 million in 2001, up 327 percent from 1997) and DVD players (14.1 million, up about 4,000 percent). But Americans say they feel conflicted about new technology, according to a recent poll.

Respondents who think therapeutic cloning would:
Improve the quality of life: 59%
Harm it: 36%
Percent of those aged 65 years and older who think it would improve life: 43%
Percent of those 18 to 34 who think so: 65%

Respondents who think computer chips linked to nerve cells would:
Improve the quality of life: 38%
Harm it: 52%

Percent who are very comfortable with the pace of technological change: 20%
Percent of men who say that: 25%
Percent of women: 15%

SOURCES: Consumer Electronics Association market research (cell phone and DVD player figures); Columbia University's Center for Science, Policy, and Outcomes and the Funders' Working Group on Emerging Technologies (poll data)

Stephen Jay Gould, 1941–2002

The last time I saw Stephen Jay Gould, he was checking into a San Francisco hotel for a science conference. This day happened to be some baseball teams' first day of spring training. I sidled up to Gould and gently nudged him with my elbow. "Hey. Pitchers and catchers," I said. He returned a warm smile and the ritualistically correct repetition of the phrase: "Pitchers and catchers."

Though diagnosed with mesothelioma 20 years ago, Gould died in May evidently of another, undetected cancer. He is survived by his second wife, Rhonda Shearer, and two sons from his first marriage, Jesse and Ethan. His incredibly prolific output as a writer—he authored more than 20 books, 300 columns in *Natural History* magazine and nearly 1,000 scientific papers—amazed everyone who strings words together for a living, as did his encyclopedic knowledge of, well, apparently everything.

Scientifically, Gould is probably best known for his work with Niles Eldridge on the Darwinian variation he termed punctuated equilibrium, in which new species arise swiftly after long periods of stasis. (Critics in favor of a more gradual and consistent evolutionary history called Gould's viewpoint "evolution by jerks." He fired a salvo back by referring to the critics' stance as "evolution by creeps.")

And he brought the same passion that he had for evolutionary theory to his analysis—and to his fandom—of baseball. He dissected the reasons no one has hit .400 in the major leagues since Ted Williams did so in 1941. In-

deed, in my encounters with him, we were as likely to discuss the statistical improbability of Joe DiMaggio's 56-game hitting streak as the plasticity of the genome.

At a memorial service, his son Ethan talked about how much the two of them enjoyed games together at Boston's Fenway Park. Stephen Jay Gould led a large, loud life. He had titanic battles with creationists and fantastic feuds with other scientists. He left millions of words for readers to ponder. And he was also just a guy who loved watching a ballgame with his son. —*Steve Mirsky*

DIAGNOSTICS

Eye on the Brain

The eyes are more than a window to the soul—they are also a window on health. Medical scientists in Singapore found that in a study of 8,000 patients, those with impaired mental function stemming from brain damage were roughly three times more likely to have observable lesions or other anomalies in their retinal blood vessels. The researchers, reporting in the June 2002 *Stroke,* suggest that such damage in the eye reflects similar vascular harm in the brain.

RETINA offers clues to brain health.

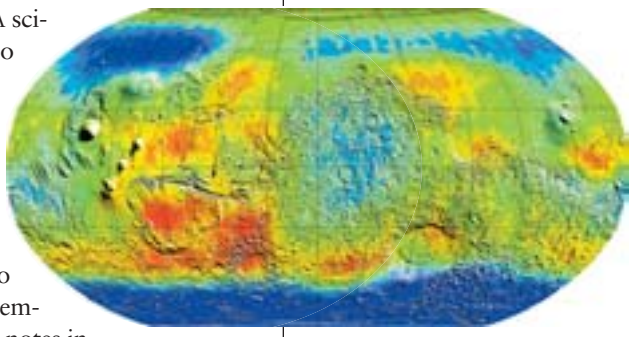
Rather than inferring brain injury, a new device could actually determine whether neurons were gasping for air. Oxygen deprivation kills brain cells but often leaves vital signs such as blood pressure unaffected; today only monitors connected to the heart can spot this condition. Sarnoff Corporation in Princeton, N.J., says that its so-called retinal eye oximeter can measure brain oxygenation noninvasively. The instrument shines in low-energy laser beams; blood appears brighter when carrying oxygen. The company hopes to develop a handheld version that can be used for human clinical trials next year. —*Charles Choi*

PAT GREENHOUSE *AP Photo (top)*; SCOTT CAMAZINE *Science News (bottom)*; ILLUSTRATION BY MATT COLLINS

ASTRONOMY

What Lies Beneath

The surface of Mars bears the scars of liquid water now gone. NASA scientists may have spotted some of that disappeared H_2O—enough to fill two Lake Michigans. A trio of studies present neutron and gamma-ray emission measurements made by the Mars Odyssey probe. The emissions, which occur after cosmic rays strike hydrogen, indicate that the element is locked away as dirty ice in at least the uppermost meter of the Red Planet's surface. The icy layer creeps up within 30 centimeters of the surface near the southern pole, which shows the widest distribution of potential ice so far. The northern latitudes also seem to contain water ice, but dry ice (frozen carbon dioxide) has temporarily obscured measurements. James F. Bell of Cornell University notes in an accompanying article that the proposed concentration of water ice implies a porous, rocky subsurface, which could make the current results the tip of a Martian iceberg. All four papers will appear in an upcoming issue of *Science*. —*JR Minkel*



BURIED WATER ICE of Mars (*dark blue*) is especially abundant near the southern pole.

NASA/JPL/UNIVERSITY OF ARIZONA/LOS ALAMOS NATIONAL LABORATORIES (*top*); U.S. DEPARTMENT OF ENERGY (*bottom*)

CHEMISTRY

Mechanical Reactions

The toxic solvents of organic chemistry could become a thing of the past. Researchers at the U.S. Department of Energy's Ames Laboratory have successfully used a mechanical process, previously employed in metallic and inorganic chemistry, to synthesize new organic compounds. The solid raw materials are placed in a vial along with a number of steel balls and are shaken vigorously. The agitation, done at room temperature 25 times a second, breaks down the materials' crystal lattices just as well as solvents do. The amorphous substances can then serve as the building blocks for polymers and pharmaceuticals. Moreover, for some processes, ball milling permits all the components to be added together in a single, efficient step rather than one by one in a series of chemical reactions. Further details are in the June *Journal of the American Chemical Society*. —*Zeeya Merali*

STEEL BALLS could replace solvents in organic chemistry.

MEDICAL REPORTING

Only the Best

Advertising campaigns routinely hype the most flattering claims to sell their products. Evidently so do papers in medical journals. Researchers at the University of California at Davis School of Medicine found too much emphasis given to favorable statistics in five of the top medical journals: the *New England Journal of Medicine*, the *Journal of the American Medical Association*, the *Lancet*, the *Annals of Internal Medicine* and the *BMJ*. The study, which looked at 359 papers on randomized trials, found that most researchers furnish a statistic only for "relative risk reduction"—the percentage difference between the effect of the treatment and a placebo. Just 18 included the more straightforward absolute risk reduction. If a treatment reduced the absolute risk from, say, 4 to 1 percent, it appears more impressive to present only the relative reduction of 75 percent. Researchers also failed to show other statistics that provide a more nuanced picture of the results of clinical trials. The article is in the June 5 *Journal of the American Medical Association*. —*Benjamin Stix*

WWW.SCIAM.COM BRIEF BITS

These items and more are at www.sciam.com/news_directory.cfm

- A methane-producing bacterium has revealed the existence of a 22nd amino acid, called pyrrolysine. There are 20 standard amino acids and a rare one, selenocysteine (discovered in 1986).
- Gravitational lensing suggests that hundreds of dark matter dwarf galaxies may girdle the Milky Way.
- Just don't do it: A study of track-and-field events concludes that wind, altitude and other random factors dictate record-breaking performances, not improved skills.
- The tip breaking the sound barrier does not cause the crack of the whip; the sonic boom comes from the loop in the whip.

Soft Manufacturing

Shaping small structures in rubber has moved beyond a Harvard lab By GARY STIX

George M. Whitesides is a towering figure in the emerging field of nanotechnology. This Harvard University chemistry professor has articulated the promise of a new discipline for building things with dimensions as small as a few atoms. Whitesides has not been content, however, to keep his work confined to an academic laboratory that today boasts about 40 graduate and postdoctoral students.

He has established a company that fabricates microsize and nano-size components in soft materials to help the pharmaceutical industry perform tests on biological samples. The technology invented by Whitesides and his collaborators, known as soft lithography, is one of a number of radically new manufacturing techniques that can make large numbers of small structures [see "Breaking the Mold," Innovations, July].

PEELING OFF a rubber sheet allows Surface Logix scientists to test how cells placed in micromolded cavities react to a drug candidate (*cells not shown*).

Soft lithography grew out of Whitesides's work in the early 1980s with Ralph G. Nuzzo of the University of Illinois. The researchers explored the making of films of highly ordered molecules called self-assembled monolayers (SAMs). These studies caught the attention of the Defense Advanced Research Projects Agency (DARPA) about a decade ago, when the agency began to investigate alternatives to conventional photolithography for patterning circuits on microchips. Using DARPA funding, Whitesides and his students bought an ordinary rubber stamp with the Harvard motto and inked it with organic molecules called thiols, printing the word "veritas" over a small area. The semiconductor community expressed deep-seated skepticism about printing with molecular ink. Whitesides remembers the prevailing opinion: "You can barely make a square-inch thing—how can you make things much smaller than that?"

These endeavors turned practical when a researcher in Whitesides's laboratory, Manoj K. Chaudhury, who had come from Dow Corning, a manufacturer of rubber materials, suggested that the team use silicone-based materials—such as poly(dimethyl siloxane), or PDMS. A group of graduate students then demonstrated that a molding technique using PDMS could produce features as small as 10 nanometers.

Soft lithography could do things that conventional lithography could not—for instance, molding or printing patterns on curved as well as flat surfaces. It may not be ideal for making microchips, though: those chips comprise stacked layers, and the soft materials used in the technique can produce misalignments between layers. But with the help of fellow professor Donald E. Ingber of Harvard Medical School, Whitesides has proved soft lithography's ability to fashion the microchannels and chambers needed to hold liquids for biological experiments. The biggest advantage is cost. Conventional lithography might require several thousands of dollars to produce a stamp or mold over a matter of weeks. "Ours could be done overnight and costs pretty much nothing," Whitesides says. Molds for these bioassays could be made quickly, tested and then remade to refine an experiment.

In late 1998 a former postdoctoral student of Whitesides's named Carmichael Roberts approached him and asked whether his onetime adviser could suggest any ideas for starting a company. "I've got a deal for you," Whitesides replied. The first $1 million for Surface Logix came from acquaintances of both Whitesides and Roberts in the Boston area—and since then, the company has amassed a healthy total of $40 million in financing.

Pharmaceutical companies have begun to inspect

EMANUELE OSTUNI *Surface Logix*

the first systems for performing bioassays produced by Surface Logix's soft lithography. Whitesides gave one example of a test that could aid in screening for new drugs. A sheet molded with hundreds of perforations, from 10 to 100 microns across, can be manufactured using soft lithography and set down on another surface. After placing mammalian cells into each of the micromolded cavities, a researcher can insert potential drug compounds into each of the openings. The rubber sheet can then be peeled off, leaving the underlying surface on which the cells rest. Now unimpeded, the cells can begin to move around. A cell can budge, however, only if it is able to synthesize a cytoskeleton, the network of fibers that gives it structural support.

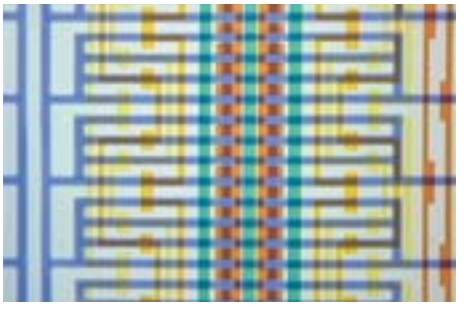
"If a compound inhibits assembly of the cytoskeleton, the cell can't walk," Whitesides says. One inhibitor of cytoskeletal growth is Taxol, an anticancer drug. The assay thus provides a method of searching for Taxol-like drugs. Soft lithography can also be deployed to stamp vast numbers of proteins and other molecules onto surfaces and then probe which compounds bind to them. The surface chemistry for these assays comes from the work of Milan Mrksich of the University of Chicago.

So far Surface Logix works with structures that vary in size from 500 nanometers to a few hundred microns. But as the technology matures, the company expects to use it to create structural elements with dimensions of less than 100 nanometers. These nanostructures might attach a small number of molecules to a surface, perhaps for use as detectors for bioweapons systems, an application that has continued to bring in DARPA money. Whitesides and Roberts have also started another company, EM Logix, that uses soft lithography to make devices for use in optics and some electronics components with larger feature sizes. The list of applications is growing. "This is a fundamentally new route to engineered, ordered materials, and the world is still understanding how to best use it," Whitesides says.

Besides serving to incubate new companies, Whitesides's work in soft lithography inspired a young professor at the California Institute of Technology. Stephen R. Quake, 33, did not set out to become an innovator in microfabrication or nanofabrication. As a biophysicist, he was interested in making images of single molecules using soft lithography. With a classmate from Stanford University in 1999, he even started a company, Mycometrix, to isolate and inspect them one by one. While building devices to perform this work, he became intrigued by the possibilities of soft lithography to make what he describes as the fluidic equivalent of the integrated circuit. Mycometrix has applied the technology to craft networks of hundreds or thousands of valves, pumps and channels. These microfluidic systems are, in essence, laboratories on a chip that can carry out steps ranging from purification to separation to detection using small quantities of liquid.

Quake's systems can sort cells, perform the polymerase chain reaction to amplify DNA, or grow protein crystals. Microfluidic components made from soft lithography are more compatible with biological systems and less expensive than the silicon and glass in other microfluidic devices. Quake has built three-dimensional networks of channels and devices. Air or water moving through an upper channel exerts downward pressure on a series of membranes. Three membranes, one after the other, compress liquid in a channel below. Quake compares this pump to someone stepping on a garden hose, causing the fluid to be propelled along.

Mycometrix, whose name later morphed into Fluidigm, has deployed the microfluidic technology in a protein-crystallization chip that could prove a boon to the field of proteomics (the cataloguing of protein interactions). Three microliters of protein are enough to conduct 144 experiments. The microfluidic system subjects nanoliter volumes of liquids taken from the initial feedstock to reactive chemicals in various mixing ratios to determine the best conditions for growing protein crystals. According to Quake, the microlaboratory allows for the crystallization of proteins that cannot be grown with conventional technologies.



GRID of microchannels and valves forms the basis of Fluidigm's laboratory-on-a-chip technology.

Run by Gajus Worthington, a former Stanford classmate of Quake's, Fluidigm has raised $50 million and has signed agreements to provide customers including GlaxoSmithKline with microfluidic systems. For his part, Quake wants to remain in the lab. "What I like to do is research with my group at Caltech," he says. "I enjoy the unconstrained environment of the university." But one thing can feed the other. Quake foresees new uses for soft lithography coming out of his laboratory, such as home testing kits and components for fuel cells—a host of applications that will move the technology beyond the realm of biology. SA

SEBASTIAN MAERKL AND TODD THORSEN *California Institute of Technology*

Deep-Sixing the Submarine Patent

Will a pending trial curb a purportedly abusive practice? By GARY STIX

A legal showdown looms over some of the world's most controversial patents. Later this year a group of manufacturing companies is scheduled to go to federal court to defang key patents held by a partnership set up before the death of Jerome Lemelson. A hero among small inventors, Lemelson garnered more than 550 patents, a number that puts him close behind Thomas Edison. Lemelson is remembered as a generous benefactor to the Massachusetts Institute of Technology, the Smithsonian Institution and others. But many corporations consider him an archvillain who misused the patent system to obtain claims on things he never invented.

In their view, Lemelson, who died at the age of 74 in 1997, was the unrivaled king of the "submarine" patent. The companies bringing the suits assert that by applying for additional or expanded claims—and by taking advantage of a procedure that extends the patent application process—he would continually delay the issuance of a patent. These ploys, they say, served to keep the details of a patent secret while allowing him to broaden its scope to encompass technologies invented and commercialized by others. Once a patent finally did surface—in one case almost 40 years after the filing date for claims purported to cover bar-code scanning—Lemelson systematically went about seeking licensing fees, and if companies refused, he sued them for infringement. By one estimate the Lemelson Medical, Education & Research Foundation, which holds these patents, has garnered $1.5 billion by pursuing this strategy.

The Lemelson legal apparatus has been good at following the money. It asked that big users of bar-code checkout scanners, such as Sears, pay licensing fees for a series of patents that it claimed covered the technology, but it left alone the scanner manufacturers, with their more modest bank accounts. Fed up at this assault on their customers—and in some cases presented with requests by retailers for reimbursement of money shelled out to Lemelson—seven bar-code manufacturers and Cognex, a machine-vision company, are now suing the Lemelson Foundation to try to put an end to what they perceive as an abuse of the patent system.

When the trial, scheduled for November, commences, the eight companies may have a potent arrow in their quiver that was unavailable for anyone who tussled with Lemelson before. The Court of Appeals for the Federal Circuit ruled in January that the companies bringing the suits were entitled to invoke a legal doctrine called prosecution laches, allowing them to argue that Lemelson's patents should not be enforced because of the protracted delaying tactics. The decision, if it is not reversed by the Supreme Court, could strengthen the companies' case during the trial. "The fat lady is warming up," says Leonard H. Goldner, executive vice president and general counsel for Symbol Technologies, a Long Island, N.Y., bar-code manufacturer that has taken a leading role in these cases.

The Lemelson Foundation disputes the idea that its founder was a submarine patenter, contending that any delays were the responsibility of the U.S. patent office. "He is the victim of those delays, and he suffered from those delays; he didn't cause them," says Gerald Hosier, the foundation's attorney, who was listed last year by Forbes as the top earner in his profession.

If the eight companies win, it may undermine lawsuits brought by the foundation against more than 400 others, including many retailers and a big computer firm. But even if they lose, the future for the type of submarine patenting that Lemelson is accused of is sinking fast. Changes in patent law now make a 40-year delay an impossibility, thus limiting the number of surprises that remain submerged at the patent office. SA

JOHN McFAUL

Why ET Hasn't Called

The lifetime of civilizations in the Drake equation for estimating extraterrestrial intelligences is greatly exaggerated By MICHAEL SHERMER

In science there is arguably no more suppositional formula than that proposed in 1961 by radio astronomer Frank Drake for estimating the number of technological civilizations that reside in our galaxy: $N = R\ f_p\ n_e\ f_l\ f_i\ f_c\ L$

In this equation, N is the number of communicative civilizations, R is the rate of formation of suitable stars, f_p is the fraction of those stars with planets, n_e is the number of Earth-like planets per solar system, f_l is the fraction of planets with life, f_i is the fraction of planets with intelligent life, f_c is the fraction of planets with communicating technology, and L is the lifetime of communicating civilizations.

Species may simply not be equipped to survive for long periods in large populations.

Although we have a fairly good idea of the rate of stellar formation, a dearth of data for the other components means that calculations are often reduced to the creative speculations of quixotic astronomers.

Most SETI (search for extraterrestrial intelligence) scientists are realistic about the limitations of their field; still, I was puzzled to encounter numerous caveats about L, such as this one from SETI Institute astronomer Seth Shostak: "The lack of precision in determining these parameters pales in comparison with our ignorance of L." Similarly, Mars Society president Robert Zubrin says that "the biggest uncertainty revolves around the value of L; we have very little data to estimate this number, and the value we pick for it strongly influences the results of the calculation." Estimates of L reflect this uncertainty, ranging from 10 years to 10 million years, with a mean of about 50,000 years.

Using a conservative Drake equation calculation, where L = 50,000 years (and R = 10, f_p = 0.5, n_e = 0.2, f_l = 0.2, f_i = 0.2, f_c = 0.2), then N = 400 civilizations, or one per 4,300 light-years. Using Zubrin's optimistic (and modified) Drake equation, where L = 50,000 years, then N = five million galactic civilizations, or one per 185 light-years. (Zubrin's calculation assumes that 10 percent of all 400 billion stars are suitable G- and K-type stars that are not part of multiples, with almost all having planets, that 10 percent of these contain an active biosphere and that 50 percent of those are as old as Earth.) Estimates of N range wildly between these figures, from Planetary Society scientist Thomas R. McDonough's 4,000 to Carl Sagan's one million.

I find this inconsistency in the estimation of L perplexing because it is the one component in the Drake equation for which we have copious empirical data from the history of civilization on Earth. To compute my own value of L, I compiled the durations of 60 civilizations (years from inception to demise or the present), including Sumeria, Mesopotamia, Babylonia, the eight dynasties of Egypt, the six civilizations of Greece, the Roman Republic and Empire, and others in the ancient world, plus various civilizations since the fall of Rome, such as the nine dynasties (and two republics) of China, four in Africa, three in India, two in Japan, six in Central and South America, and six modern states of Europe and America.

The 60 civilizations in my database endured a total of 25,234 years, so L = 420.6 years. For more modern and technological societies, L became shorter, with the 28 civilizations since the fall of Rome averaging only 304.5 years. Plugging these figures into the Drake equation goes a long way toward explaining why ET has yet to drop by or phone in. Where L = 420.6 years, N = 3.36 civilizations in our galaxy; where L = 304.5 years, N = 2.44 civilizations in our galaxy. No wonder the galactic airways have been so quiet!

I am an unalloyed enthusiast for the SETI program, but history tells us that civilizations may rise and fall in cycles too brief to allow enough to flourish at any one time to traverse (or communicate across) the vast and empty expanses between the stars. We evolved in small hunter-gatherer communities of 100 to 200 individuals; it may be that our species, and perhaps extraterrestrial species as well (assuming evolution operates in a like manner elsewhere), is simply not well equipped to survive for long periods in large populations.

Whatever the quantity of L, and whether N is less than 10 or more than 10 million, we must ensure L does not fall to zero on our planet, the only source of civilization we have known. SA

Michael Shermer is publisher of Skeptic *magazine (www.skeptic.com) and author of* In Darwin's Shadow: The Life and Science of Alfred Russel Wallace.

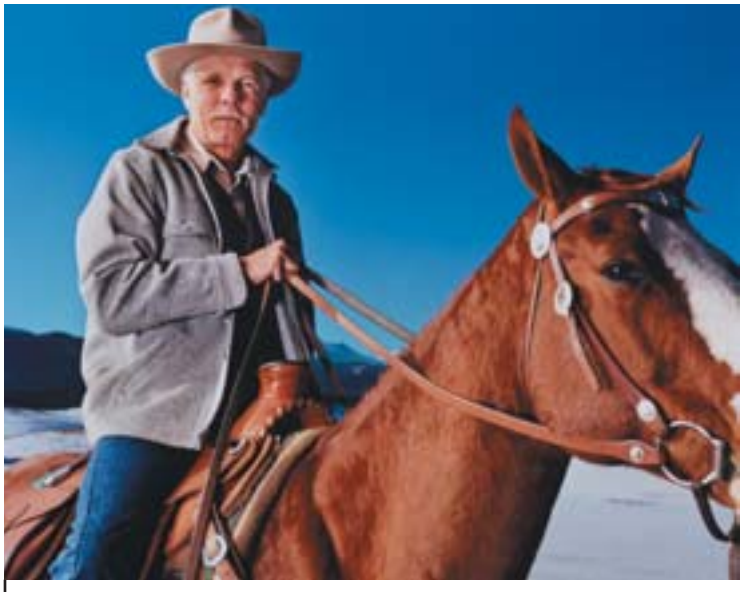
BRAD HINES

The Billionaire Conservationist

Can Ted Turner save threatened species? He is using his private lands and deep pockets to reintroduce animals driven off by development By KRISTA WEST

As the creator of CNN, the first 24-hour news network, and other cable stations, Ted Turner forever changed the landscape of American television. Now the 64-year-old "media mogul" plans to change the landscape of the American West. He is the ringleader of a giant scientific experiment to restore damaged ecosystems—specifically, to reintroduce species and to reinvigorate Western lands in an economically sustainable way. And he may just have the means and the minds to pull it off.

Turner is the largest private landowner in the nation, controlling two million acres (an area bigger than Delaware) spread across 10 states. He is using the lands as laboratories to apply existing wildlife management techniques and to develop new ones. Since 1997 his staff of traditional ranchers, former government scientists and academic researchers has produced nearly 50 scientific publications, and their impact on the science of wildlife conservation is becoming hard to ignore.



MARTIN SCHOELLER *Corbis Outline*

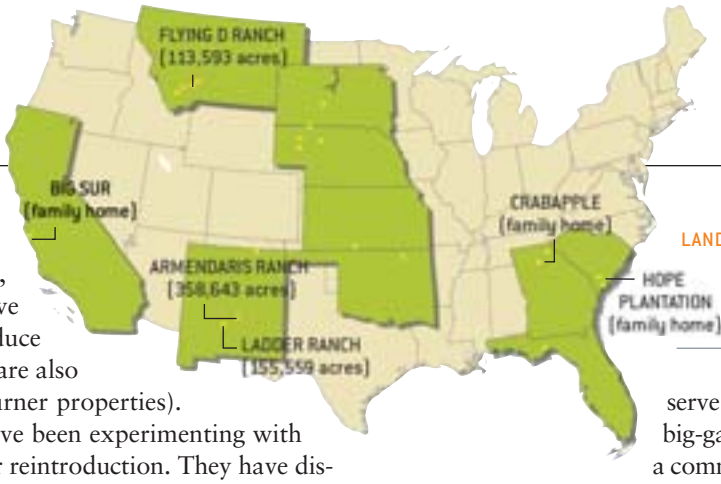
TED TURNER: RESTOCKING THE FRONTIER

- Established CNN, TNT, Turner Classic Movies, TBS and the Cartoon Network. Owns three Atlanta sports teams: the Braves, the Falcons and the Hawks.
- Vice chairman of AOL Time Warner and its largest individual stockholder.
- Net worth: $3.8 billion. Ranked 97th on *Forbes* magazine's 2001 list of the world's richest people (down from 45th in 2000).
- Loves hunting and fishing. Hates to be interviewed by journalists. (Some of his staff resent the term "media mogul.")

It all started in 1995, when the thrice-divorced father of five visited Yellowstone National Park with his then wife, actress Jane Fonda, to discover more about the federal wolf reintroduction program, an effort of the U.S. Fish and Wildlife Service to transplant wild wolves from eastern Canada to Yellowstone. They met project leader Mike Phillips and, according to Phillips, "learned that restoration could be an alternative to extinction." Two years later Turner, with his son Beau and Phillips, created the Turner Endangered Species Fund (TESF), a nonprofit organization to manage and restore wildlife on Turner's properties and adjacent public lands. Today TESF is working to reintroduce nearly two dozen species of animals, including Mexican wolves, red-cockaded woodpeckers, black-footed ferrets and Rio Grande cutthroat trout, to those lands. "You can already see the difference TESF has made on the overall health of the landscape," Turner remarks.

Of the 14 properties where restoration efforts are taking place, the Armendaris Ranch at the northern tip of the Chihuahuan Desert in south central New Mexico is among the more significant. University, state and TESF collaborations have created more projects than ranch manager Tom Waddell can keep track of. "Turner is the artist, but we're the painters," Waddell tells me in all earnestness as we rumble around the ranch in his pickup truck. Waddell, who was a biologist for the Arizona Game and Fish Department for more than 20 years, expresses unequivocal loyalty to the self-made billionaire, a trait common to most Turner employees despite the tycoon's sometimes controversial bluntness.

As we pull up to a flat plot cleared of tall grasses,



LAND ROVER: Ted Turner owns two million acres over 20 properties (*dots*); 14, including the ranches listed, host conservation efforts.

TESF biologist Joe C. Truett, an expert in desert grasslands, points out where they have been working to reintroduce prairie dogs (similar efforts are also under way on five other Turner properties). Here academic scientists have been experimenting with ways to prepare the land for reintroduction. They have discovered that setting fires and mowing work equally well in getting rid of the long grasses that make the habitat unsuitable for prairie dogs.

For the scientists, the fact that the land is privately owned allows them to experiment on a scale impossible to create in the laboratory—and to do so without the bureaucratic red tape that often accompanies work on public property. "No cookbook exists for restoring land," Truett explains. "We plan restoration based on scientific knowledge but frequently adjust our plans as experience indicates the need, and Ted's okay with that." Such adaptive management is difficult for government agencies, according to Truett and others, because legal requirements and accountability to interest groups and taxpayers make it impossible to change plans even when the need is obvious.

"Our techniques and projects are self-motivated, but we want to publish so that others can use the information," says Steve Dobrott, manager of Turner's Ladder Ranch, about 40 miles west of Armendaris. Here the staff works with government researchers to support a federal breeding facility for the Mexican wolf and a state-assisted project to reintroduce native trout. Indeed, anytime TESF wants to work with such protected species, it must seek government approval and cooperation. Technically, the animals are public resources that only happen to live on private lands, and various laws govern their management. "We won't shy away from tough species," comments Phillips, now executive director of TESF. "We judge our success by counting publications and heads"—animal heads, that is.

Still, some officials feel that Turner is trying to steamroll the process by "not waiting for the government to act," notes Wally Murphy, a federal biologist who leads a threatened and endangered species program for the U.S. Forest Service. Murphy himself is generally happy with what the billionaire is doing. "Turner's motivations may be emotional," he says, "but he's taking a very analytical approach. And he's hired good people."

For certain projects, self-interest does motivate Turner. "A large part of what we do here is study quail," Dobrott says, "mainly because Ted likes to hunt them." Turner visits the Ladder Ranch about 25 times during hunting season, from November to February, according to Dobrott. Another of his favorites is bison—a fascination Turner himself cannot explain but the main reason he began his conservation work. Today he is the biggest bison rancher in the country, with about 30,000 head, at least eight times as many as any other rancher.

More than just harking back to frontier times, the bison serve an economic role. Along with big-game hunts and timber, they are a commodity for Turner Enterprises, a for-profit group. The bison are sold to the North American Bison Cooperative in North Dakota, where Turner's two-restaurant chain, called Ted's Montana Grill, buys the meat for its bison burgers.

While Turner Enterprises pulls in the dollars, the nonprofit Turner Foundation gives money away. In 2000 it awarded 683 grants, totaling almost $50 million, to research and conservation organizations. Besides going to TESF, funds also went to the Wildlife Conservation Society and the Nature Conservancy. No other philanthropist targets conservation to this extent. Turner's one-year contribution is nearly twice what the PGA Foundations, led by Microsoft co-founder Paul G. Allen, gave for conservation from 1990 through 2000. And it is about 90 times the amount awarded in 2000 by the Bill and Melinda Gates Foundation, whose $1.2 billion yearly budget goes mainly to promote human rights, health care and education. Only the federal government spends more on wildlife-related sciences: in 2001 the National Science Foundation doled out $175 million.

Critics say Turner is simply buying the good graces of environmental groups while ignoring the concerns of neighboring property owners. Some ranchers worry that the species being reintroduced on his lands may one day bring federal land-use restrictions. Jimmy Rainey, the mayor of Truth or Consequences, N.M., a town nestled between Turner's Armendaris and Ladder ranches, doesn't necessarily agree. He says the local people aren't really concerned about what Turner is doing—they're just curious. Certainly they appreciate his financial donation to the town (in the form of a skateboard park), but, Rainey says, "It would be nice if Turner held town meetings or something to let us know what's going on."

Turner calls his land stewardship a "serious responsibility," but he isn't one to sit still for such interactions. (He barely made time to be interviewed for this story.) He relies on his staff to engage the local communities.

Turner himself thinks it's too early to judge his conservation efforts. "TESF is a new, innovative entity," he says, "and it will take time before its true impact can be clearly measured." With no other private landowner attempting to make conservation profitable on this scale, there is little by which to gauge his success. But there's no sign of Turner letting up—his staff say that they are always looking for good deals on good land. SA

Krista West, who is based in Las Cruces, N.M., likes her bison burgers well done.

JOHNNY JOHNSON

The Serious Search for an Anti-Aging Pill

In government laboratories and elsewhere, scientists are seeking a drug able to prolong life and youthful vigor. Studies of caloric restriction are showing the way

By Mark A. Lane, Donald K. Ingram and George S. Roth

As researchers on aging noted in a position statement this past May, no treatment on the market today has been proved to slow human aging—the buildup of molecular and cellular damage that increases vulnerability to infirmity as we grow older. But one intervention, consumption of a low-calorie yet nutritionally balanced diet, works incredibly well in a broad range of animals, increasing longevity and prolonging good health. Those findings suggest that caloric restriction could delay aging in humans, too.

Unfortunately, for maximum benefit, people would probably have to reduce their caloric intake by roughly 30 percent, equivalent to dropping from 2,500 calories a day to 1,750. Few mortals could stick to that harsh a regimen, especially for years on end. But what if someone could create a pill that mimicked the physiological effects of eating less without actually forcing people to go hungry? Could such a caloric-restriction mimetic, as we call it, enable people to stay healthy longer, postponing age-related disorders (such as diabetes, atherosclerosis, heart disease and cancer) until very late in life?

We first posed this question in the mid-1990s, after we came upon a chemical agent that, in rodents, seemed to reproduce many of caloric restriction's benefits. Since then, we and others have been searching for a compound that would safely achieve the same feat in people. We have not succeeded yet, but our failures have been informative and have fanned hope that caloric-restriction, or CR, mimetics can indeed be developed eventually.

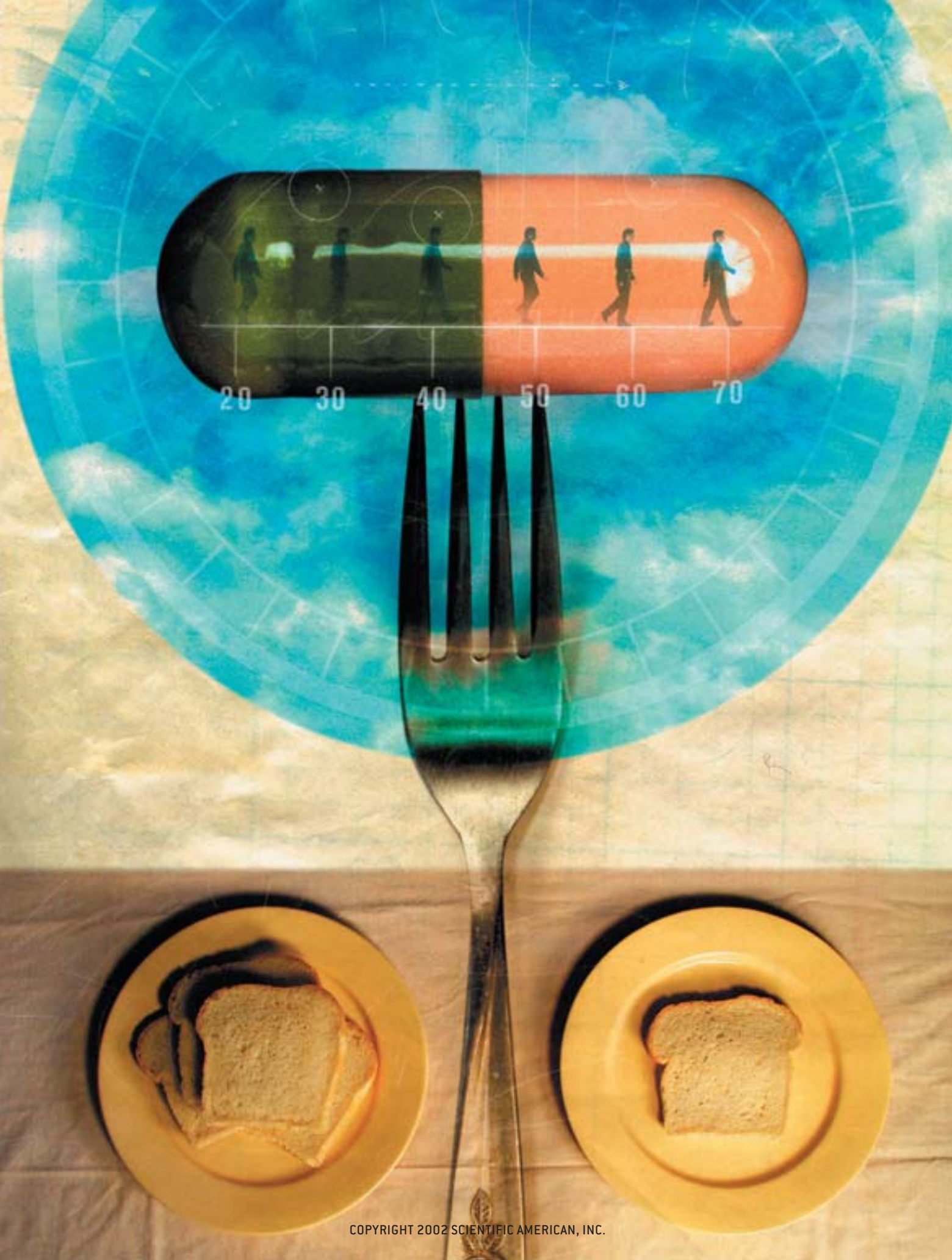
The Benefits of Caloric Restriction

OUR HUNT FOR CR MIMETICS grew out of our desire to better understand caloric restriction's many effects on the body. Scientists first recognized the value of the practice more than 60 years ago, when they found that rats fed a low-calorie diet lived longer on average than free-feeding rats and had a reduced incidence of conditions that become increasingly common in old age. What is more, some of the treated animals survived longer than the oldest-living animals in the control group, which means that the maximum life span (the oldest attainable age), not merely the average life span, increased. Various interventions, such as infection-fighting drugs, can increase a population's average survival time, but only approaches that slow the body's rate of aging will increase the maximum life span.

The rat findings have been replicated many times and extended to creatures ranging from yeast to fruit flies, worms, fish,

CALORIC-RESTRICTION MIMETIC would, if successful, enable humans to derive many of the health and life-extending benefits seen in animals on restricted diets—without requiring people to go hungry.

STUART BRADFORD



HOW A PROTOTYPE CALORIC-RESTRICTION MIMETIC WORKS

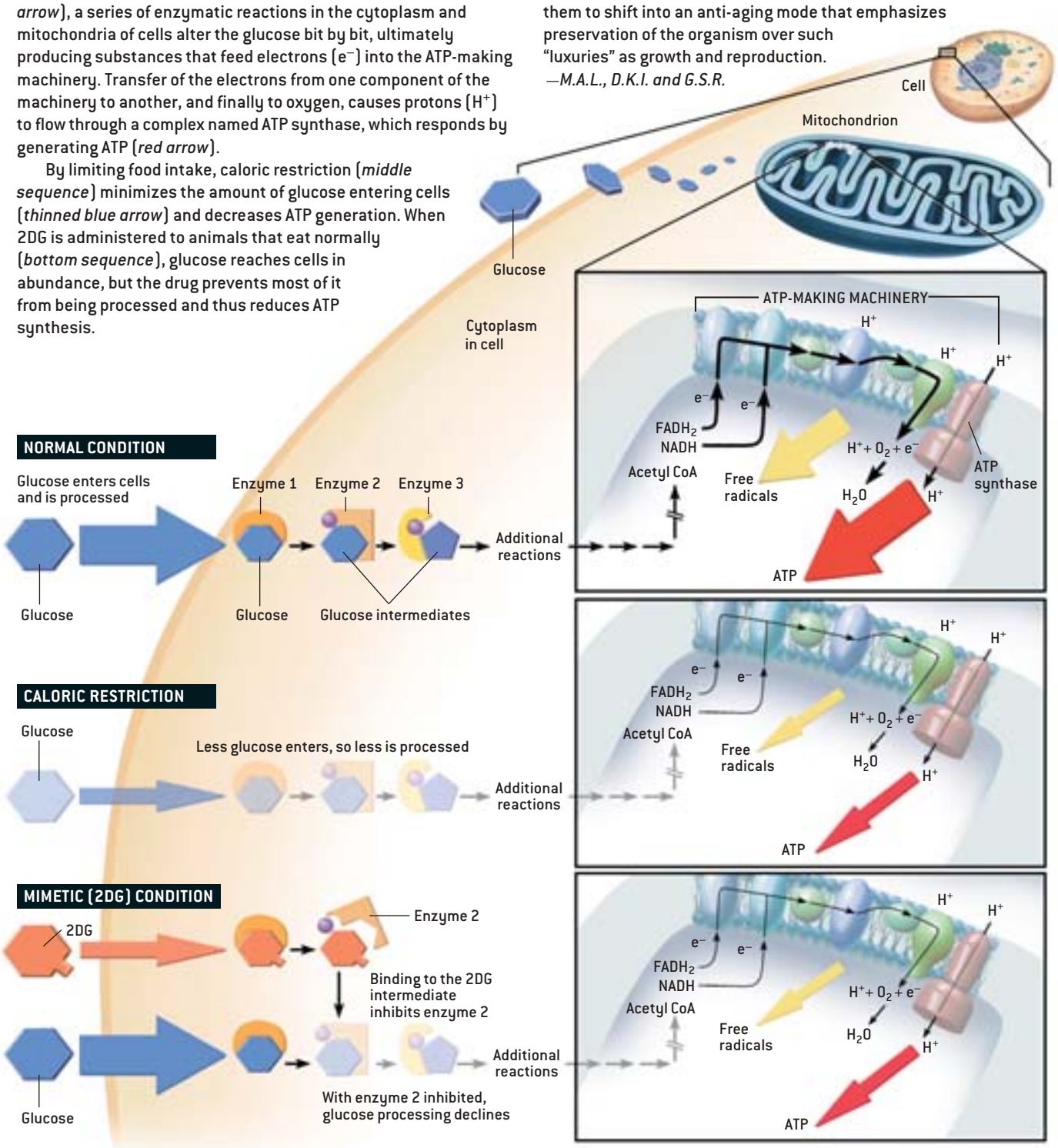
THE BEST-STUDIED CANDIDATE for a caloric-restriction mimetic, 2DG (2-deoxy-D-glucose), works by interfering with the way cells process the sugar glucose. It has proved toxic at some doses in animals and so cannot be used in humans. But it has demonstrated that chemicals can replicate the effects of caloric restriction; the trick is finding the right one.

Cells use the glucose from food to generate ATP (adenosine triphosphate), the molecule that powers many activities in the body (*top sequence*). More specifically, after glucose enters cells (*blue arrow*), a series of enzymatic reactions in the cytoplasm and mitochondria of cells alter the glucose bit by bit, ultimately producing substances that feed electrons (e^-) into the ATP-making machinery. Transfer of the electrons from one component of the machinery to another, and finally to oxygen, causes protons (H^+) to flow through a complex named ATP synthase, which responds by generating ATP (*red arrow*).

By limiting food intake, caloric restriction (*middle sequence*) minimizes the amount of glucose entering cells (*thinned blue arrow*) and decreases ATP generation. When 2DG is administered to animals that eat normally (*bottom sequence*), glucose reaches cells in abundance, but the drug prevents most of it from being processed and thus reduces ATP synthesis.

Researchers have proposed several explanations for why interruption of glucose processing and ATP production might retard aging. One possibility relates to the ATP-making machinery's emission of free radicals (*yellow arrows*), which are thought to contribute to aging and to such age-related diseases as cancer by damaging cells. Reduced operation of the machinery should limit their production and thereby constrain the damage. Another hypothesis suggests that decreased processing of glucose could indicate to cells that food is scarce (even if it isn't) and induce them to shift into an anti-aging mode that emphasizes preservation of the organism over such "luxuries" as growth and reproduction.

—M.A.L., D.K.I. and G.S.R.



TERESE WINSLOW

spiders, mice and hamsters. Until fairly recently, the studies were limited to short-lived creatures genetically distant from humans. But long-term projects under way in two species more closely related to humans—rhesus and squirrel monkeys—suggest that primates respond to caloric restriction almost identically to rodents, which makes us more optimistic than ever that CR mimetics could help people.

The monkey projects—initiated by our group at the National Institute on Aging in the late 1980s and by a separate team at the University of Wisconsin–Madison in the early 1990s—demonstrate that, compared with control animals that eat normally, caloric-restricted monkeys have lower body temperatures and levels of the pancreatic hormone insulin, and they retain more youthful levels of certain hormones (such as DHEAS, or dehydroepiandrosterone sulfate) that tend to fall with age.

The animals also look better on indicators of risk for age-related diseases. For example, they have lower blood pressure and triglyceride levels (signifying a decreased likelihood of heart disease), and they have more normal blood glucose levels (pointing to a reduced risk for diabetes, which is marked by unusually high blood glucose levels). Further, we have recently shown that rhesus monkeys kept on caloric restriction for an extended time (nearly 15 years) have less chronic disease, just as the risk data suggested. They and the other monkeys must be followed still longer, however, before we will know whether low food intake can increase both average and maximum life spans in monkeys: rhesus monkeys typically live about 24 years and sometimes up to 40; squirrel monkeys typically live about 19 years but may live for 28.

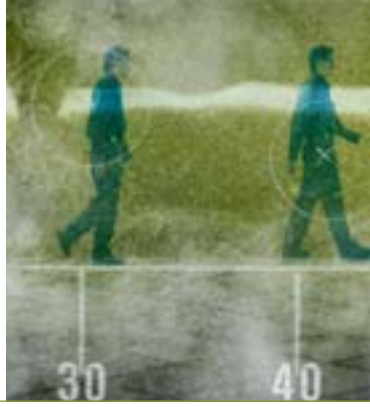
The Journey Starts

BY 1995 WE WANTED to know how the many physiological and biochemical changes induced by caloric restriction led to delaying aging in mammals. For a number of reasons, we suspected that changes in cellular metabolism would be key. By "metabolism" we mean the uptake of nutrients from the blood and their conversion to energy usable for cellular activities. We focused on metabolism in part because the benefits of caloric restriction clearly depend on reducing the overall amount of fuel coming into the body for processing. Also, caloric restriction affects the aging of a wide variety of tissues, which implies that it alters biological processes carried out by all cells. Few processes are more fundamental than metabolism.

We specifically wondered whether changes related to metabolism of the sugar glucose would account for the benefits of caloric restriction. Glucose, which forms when the body digests carbohydrates, is the primary source of energy in the body—that is, it is the main material used by cells for making ATP, or adenosine triphosphate, the molecule that directly powers most cellular activities. We also wanted to know whether alterations in the secretion and activity of insulin, which influences glucose use by cells, would be important. Insulin is secreted as glucose levels in the blood rise after a meal, and it serves as the key that opens cell "doors" to the sugar. We concentrated on glucose and insulin because reductions in their levels and increases in cellular sensitivity to insulin are among the most consistent hallmarks of caloric restriction in both rodents and primates, occurring very soon after restriction is begun.

Shortly after we decided to test the hypothesis that caloric restriction retards aging by altering metabolism, others began publishing data showing that metabolic processes involving glucose and insulin influence life span. Such findings encouraged our belief that we were on the right track. For instance, a number of investigations achieved remarkable extensions of life span in nematode worms by mutating genes similar to those involved in molecular responses to insulin in mammals. More recently researchers have found that lowered intake of glucose or disruption of glucose processing can extend life span in yeast. And in fruit flies, genes involved in metabolism, such as INDY (I'm Not Dead Yet), have been implicated in life-span control.

If 2DG could mimic caloric restriction in animals, perhaps it would do the same for people.



An "Aha!" Moment

AROUND THE TIME the nematode work came out, we began to scour the scientific literature for ways to manipulate insulin secretion and sensitivity without causing diabetes or its opposite, hypoglycemia. Our search turned up studies from the 1940s and 1950s mentioning a compound called 2-deoxy-D-glucose (2DG) that was being tested in rodents for treating cancer but that also reportedly lowered insulin levels in the blood. As we perused the literature further, we had a true "aha!" moment.

The compound apparently reproduced many classic responses to caloric restriction—among them reduced tumor growth (a response only slightly less robust than the well-known extension of life span), lowered temperature, elevated levels of glucocorticoid hormones and reduced numbers of reproductive cycles. If 2DG really could mimic many aspects of caloric restriction in animals, we

THE AUTHORS

MARK A. LANE, DONALD K. INGRAM and *GEORGE S. ROTH* researched caloric restriction together for many years at the National Institute on Aging of the National Institutes of Health. Lane, who in March became a project manager at Merck and Co., in Rahway, N.J., continues to collaborate with Ingram and Roth as a guest investigator at the NIA. Ingram is chief of the Behavioral Neuroscience Section at the institute's Laboratory of Neurosciences. Roth, who spent nearly 30 years as a full-time researcher at the NIA, is now a senior guest scientist there. He is also chief executive officer of GeroTech, a new biotechnology venture devoted to anti-aging strategies.

CALORIC RESTRICTION'S VARIED EFFECTS

CALORIE-RESTRICTED MONKEY (*left*) is shorter and leaner than its nonrestricted counterpart (*right*).

RODENTS AND MONKEYS on caloric restriction differ from their more abundantly fed counterparts in many ways, some of which are listed below (*a*–*c*). Although the influence of these shared changes on aging remains to be clarified, the close similarities in the responses of rodents and monkeys encourage hope that the health-promoting and anti-aging effects long seen in rodents (*a*–*d*) are universal among mammals, including humans. If so, caloric-restriction mimetics should help people live well longer. The effects marked by capsules (*below*) have been reproduced in rats by the compound 2DG. —*M.A.L., D.K.I. and G.S.R.*

a **EFFECTS INDICATIVE OF ALTERED GROWTH, DEVELOPMENT OR METABOLISM**

- Lower body temperatures
- Later sexual maturation
- Later skeletal maturation

b **EFFECTS INDICATIVE OF IMPROVED HEALTH**

- Lower weight
- Less abdominal fat

c **EFFECTS INDICATIVE OF REDUCED RISK FOR AGE-RELATED DISEASES (SUCH AS DIABETES AND HEART DISEASE)**

- Greater sensitivity to insulin
- Lower fasting insulin levels
- Lower fasting glucose levels
- Lower cholesterol and triglyceride levels
- Lower insulin-like growth factor 1 levels
- Higher levels of "good" (HDL) cholesterol
- Slower decline in levels of the hormone DHEAS

d **EFFECTS FOUND IN RODENTS BUT STILL UNDER INVESTIGATION IN MONKEYS**

- Later onset of age-related diseases (including cancer)
- More cell suicide (which may help limit tumor growth)
- Longer average life span
- Longer maximum life span (a strong sign of slowed aging)

thought, perhaps it would do the same for people.

While we were planning our first studies of 2DG, we scanned the literature for details of how it works at the molecular level, learning that it disrupts the functioning of a key enzyme involved in processing glucose in cells. The compound structurally resembles glucose, so it enters cells readily. It is also altered by an enzyme that usually acts on glucose itself. But the enzyme that completes the next of several steps involved in glucose processing essentially chokes on the intermediate produced from 2DG. When it tries to act on this intermediate, it fails; in addition, its ability to act on the normal glucose intermediate becomes impaired [*see illustration on page 38*].

The net result is that cells make smaller amounts of glucose's by-products, just as occurs when caloric restriction limits the amount of glucose going into cells. Certain of these products serve as the raw material for the ATP-making machinery, which is composed of a series of protein complexes located in intracellular compartments called mitochondria. Deprived of this raw material, the machinery makes less ATP. In essence, 2DG tricks the cell into a metabolic state similar to that seen during caloric restriction, even though the body is taking in normal amounts of food. As long as the amount of ATP made meets the minimum requirements of cells, this diminished operation of the ATP-making machinery is apparently beneficial.

Why might reduced functioning of the ATP-producing machinery help combat aging? We can't say with certainty, but we have some ideas. A long-standing theory of aging blames the production of molecules called free radicals. The lion's share of free radicals in the body are emitted as the ATP-making machinery operates. Over time these highly reactive molecules are thought to cause permanent damage to various parts of cells, including the protein complexes responsible for generating ATP. Perhaps by reducing the rate of ATP production, 2DG and caloric restriction slow the rate at which free radicals form and disrupt cells.

The lack of glucose's by-products might retard aging in another way as well. Certain of those substances help to induce cells in the pancreas to secrete insulin after an organism eats. Reductions in the amount of those by-products would presumably limit insulin secretion and thereby minimize insulin's unwanted actions in the body. Aside from indirectly promoting excessive operation of the ATP-making machinery and thus boosting free-radical production, insulin can contribute to heart disease and to undesirable cell proliferation.

We also suspect that cells interpret reduced levels of raw materials for the ATP-making machinery as a signal that food supplies are scarce. Cells may well respond to that message by switching to a self-protective mode, inhibiting activities not needed for cell maintenance and repair—such as reproduction—and pouring most of their energy into preserving the integrity of their parts. If that idea is correct, it could explain why caloric restriction has been shown to increase production of substances that protect

STUART BRADFORD; NATIONAL INSTITUTE ON AGING (*photographs*)

cells from excess heat and other stresses.

This adoption of a self-preservation mode would mirror changes that have been proposed to occur on an organismic level in times of food scarcity. In the generally accepted "disposable soma" theory of aging, Thomas Kirkwood of the University of Newcastle in England has proposed that organisms balance the need to procreate against the need to maintain the body, or soma. When resources are plentiful, organisms can afford both to maintain themselves and to grow and reproduce. But when food is limited, the body invokes processes that inhibit growth and reproduction and takes extra care to preserve the soma.

Testing Begins

IN OUR FIRST experiments devoted to examining 2DG's effectiveness, we delivered low doses to rats by adding it to their feed for six months. The treatment moderately reduced fasting blood glucose levels (levels measured after food was removed for 12 hours), body weight and temperature, and robustly reduced fasting insulin levels—findings consistent with the actions of caloric restriction itself. Interestingly, after an initial adjustment to the novel diet, the 2DG group did not eat significantly less food than the controls. Thus, these exciting preliminary analyses revealed that it was possible to mimic at least some sequelae of caloric restriction without reducing food intake.

Shortly after we published these results, in 1998, other groups began identifying more ways that 2DG imitates caloric restriction. For example, Mark P. Mattson, then at the University of Kentucky, and his colleagues had reported earlier that caloric restriction could attenuate damage to nerve cells and limit behavioral deficits in rodents treated with compounds toxic to brain cells. When they then treated rodents with 2DG instead of caloric restriction, they observed the same neuronal protection.

At this writing, we are in the midst of conducting long-term rodent trials of 2DG. Results from the first year of this endeavor confirm our previous findings that 2DG slightly reduces blood glucose and body temperature. We are also examining whether 2DG reduces the incidence of cancer and increases life span when fed to animals at low doses from the time they are weaned until they die.

The task becomes finding other substances that yield 2DG's benefits but are safer.

The work so far clearly provides a "proof of concept" that inhibiting glucose metabolism can re-create many effects of caloric restriction. Regrettably, however, 2DG has a fatal flaw preventing it from being the "magic pill" we were hoping for. Though safe at certain low levels, it apparently becomes toxic for some animals when the amount delivered is raised just a bit or given over long periods. The narrowness of the safety zone separating helpful and toxic doses would bar it from human use. We hope this is not a general feature of CR mimetics.

Moving On

ASSUMING OUR long-term studies confirm that inhibiting metabolism can retard aging, the task becomes finding other substances that yield 2DG's benefits but are safer over a broader range of doses and delivery schedules. Several candidates seem promising in early studies, including iodoacetate, being investigated by Mattson's group, now at the NIA's Laboratory of Neurosciences. In animals this agent appears to protect brain cells from assaults by toxic substances, just as 2DG and caloric restriction do. Treatment with antidiabetic medications that enhance cellular sensitivity to insulin might be helpful as well, as long as the amounts given do not cause blood glucose levels to fall too low.

A great deal of research implicates glucose metabolism in regulating life span, yet other aspects of metabolism can also change in reaction to caloric restriction. When the body cannot extract enough energy from glucose in food, it can switch to obtaining energy in alternative ways. For example, it may shift to breaking down protein and fat. Pharmaceuticals that targeted these processes might serve as CR mimetics, either alone or in combination with drugs that intervene in glucose metabolism. Some compounds that act in those pathways have already been identified, although researchers have not yet assessed their potential as CR mimetics. Drugs that replicate only selected effects of caloric restriction could have a role to play as well. In theory, antioxidant vitamins might fit that bill. Research conducted to date, however, indicates that this particular intervention probably will not extend longevity.

Unlike the multitude of elixirs being touted as the latest anti-aging cure, CR mimetics would alter fundamental processes that underlie aging. We aim to develop compounds that fool cells into activating maintenance and repair activities that lead to greater health and longevity of the organism. That job is difficult but no longer seems impossible. If scientists can develop agents that offer the benefits of 2DG without its drawbacks, they will finally enable people to have their cake—a longer, healthier life—and eat it, too. SA

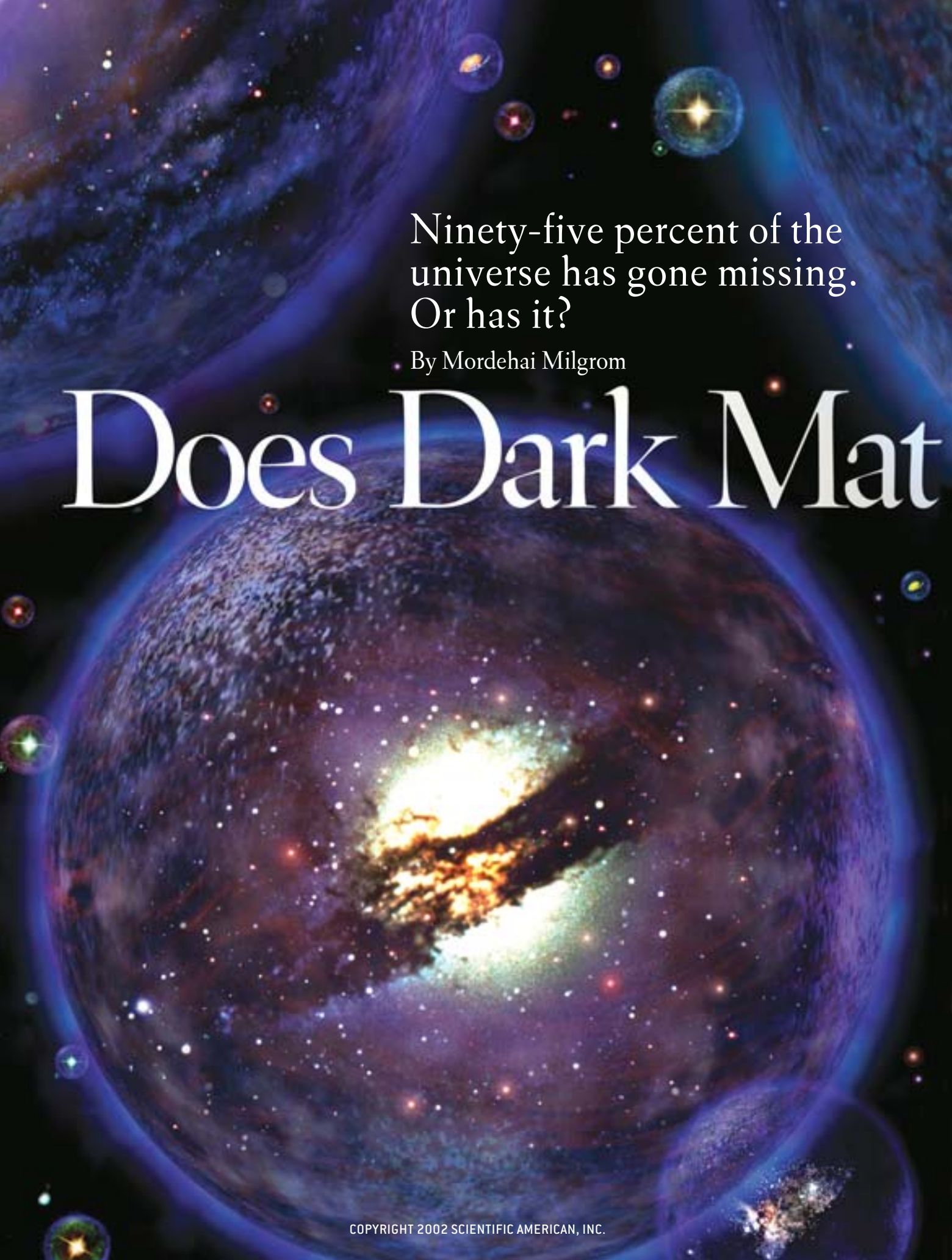
MORE TO EXPLORE

Caloric Restriction and Aging. Richard Weindruch in *Scientific American,* Vol. 274, No. 1, pages 46–52; January 1996.

2-Deoxy-D-Glucose Feeding in Rats Mimics Physiological Effects of Caloric Restriction. Mark A. Lane, George S. Roth and Donald K. Ingram in *Journal of Anti-Aging Medicine,* Vol. 1, No. 4, pages 327–337; Winter 1998.

Caloric Restriction in Primates and Relevance to Humans. George S. Roth, Donald K. Ingram and Mark A. Lane in *Annals of the New York Academy of Sciences,* Vol. 928, pages 305–315; 2001.

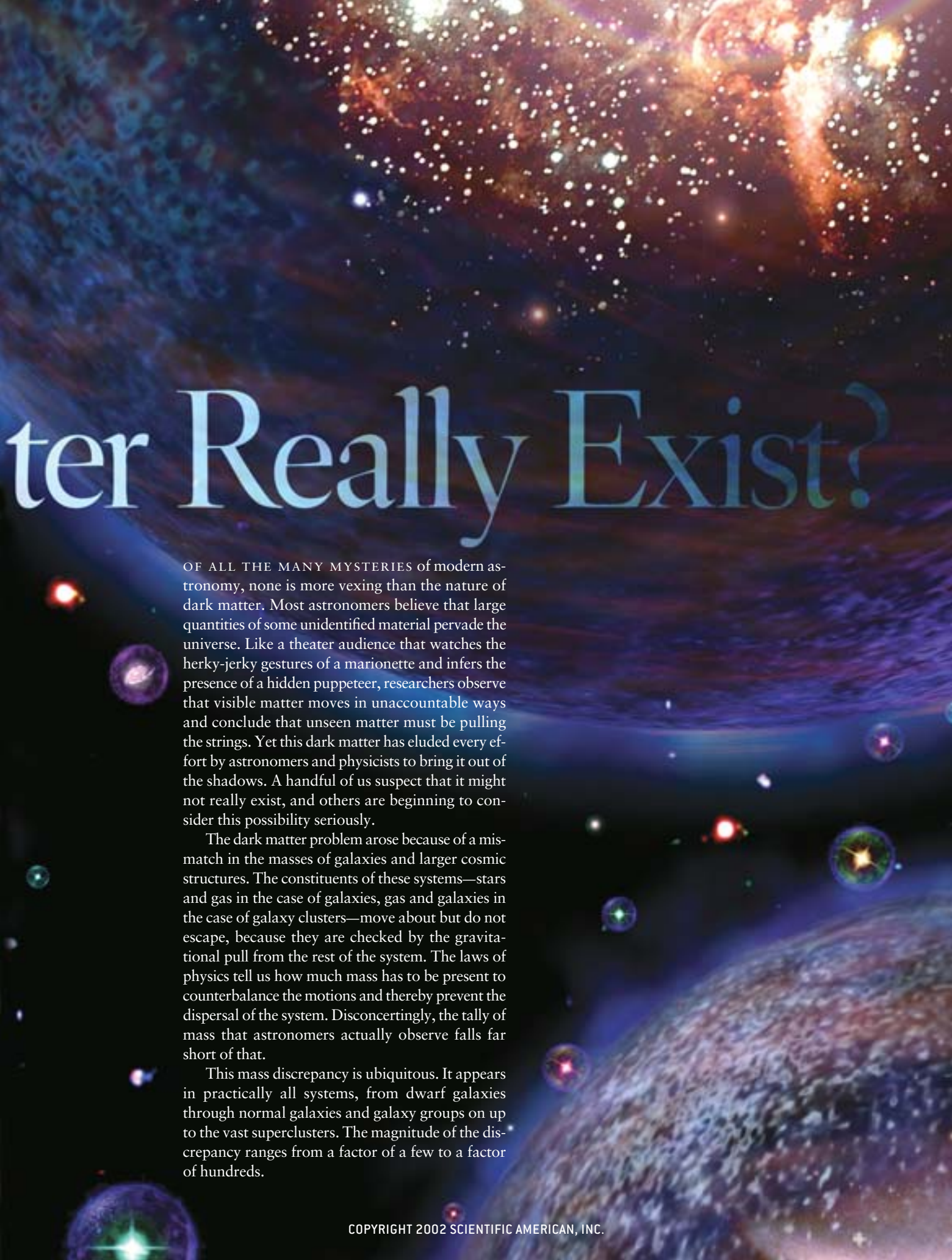
The position statement on human aging mentioned at the start of this article is available at **www.sciam.com/explorations/2002/051302aging/**



Ninety-five percent of the universe has gone missing. Or has it?

By Mordehai Milgrom

Does Dark Mat



ter Really Exist?

OF ALL THE MANY MYSTERIES of modern astronomy, none is more vexing than the nature of dark matter. Most astronomers believe that large quantities of some unidentified material pervade the universe. Like a theater audience that watches the herky-jerky gestures of a marionette and infers the presence of a hidden puppeteer, researchers observe that visible matter moves in unaccountable ways and conclude that unseen matter must be pulling the strings. Yet this dark matter has eluded every effort by astronomers and physicists to bring it out of the shadows. A handful of us suspect that it might not really exist, and others are beginning to consider this possibility seriously.

The dark matter problem arose because of a mismatch in the masses of galaxies and larger cosmic structures. The constituents of these systems—stars and gas in the case of galaxies, gas and galaxies in the case of galaxy clusters—move about but do not escape, because they are checked by the gravitational pull from the rest of the system. The laws of physics tell us how much mass has to be present to counterbalance the motions and thereby prevent the dispersal of the system. Disconcertingly, the tally of mass that astronomers actually observe falls far short of that.

This mass discrepancy is ubiquitous. It appears in practically all systems, from dwarf galaxies through normal galaxies and galaxy groups on up to the vast superclusters. The magnitude of the discrepancy ranges from a factor of a few to a factor of hundreds.

The need for extra matter arises not only in well-formed galactic systems but also through the universe at large. Long before galaxies even formed, the universe was filled with a plasma of atomic nuclei and subatomic particles. Radiation suffused the plasma and kept it extremely smooth. Fluctuations in the density of this plasma did not have a chance to grow and develop into galaxies until after the plasma had turned into a neutral gas, which does not interact with radiation as strongly. We know when this neutralization occurred and what the strength of the density fluctuations was then. The problem is that there just wasn't enough time for those fluctuations to become the galaxies we observe. Dark matter would help in that, being neutral by definition, it would not be homogenized by radiation. Therefore, it would have been contracting all along. Dark matter would have had enough time to form galaxy-mass bodies.

Common knowledge has it that part of this extra mass consists of ordinary matter that gives off too little radiation for present technology to detect: planets, dwarf stars, warm gas. Such material is more precisely called dim matter. It could represent up to 10 times as much matter as astronomers see, but even so it could account for only a small fraction of the missing mass. When researchers refer to dark matter, they usually mean an exotic breed of matter that makes up the difference. To add to the confusion, they also suspect the existence of dark energy, a distinct type of energy that would produce the observed accelerated expansion of the universe, a phenomenon that neither normal nor dark matter can explain [see "The Quintessential Universe," by Jeremiah P. Ostriker and Paul J. Steinhardt; SCIENTIFIC AMERICAN, January 2001].

In sum, astronomers widely believe the current energy content of the universe to be roughly 4 percent ordinary (or "baryonic") matter, about a tenth of which is seen as stars and gas; a third dark matter in some unknown form; and two thirds dark energy, the nature of which is even less understood.

Under Cover of Darkness

DARK MATTER is the only explanation that astronomers can conjure up for the various mass discrepancies, if we cleave to the accepted laws of physics. But if we accept a departure from these standard laws, we might do away with dark matter.

The diverse appearances of the mass discrepancy, as reflected in the motions inside galactic systems, stem from the use of a single formula in Newtonian physics. This formula combines two basic laws: Newton's law of gravity (which relates the force of gravity between bodies to the bodies' masses and separation) and Newton's second law (which relates force to acceleration). The acceleration of a body in orbit depends on the velocity and size of the orbit. Putting all this together, one derives the connection among mass, velocity, and orbital size or distance.

These laws accurately explain the flight of a ballistic missile and the motions of the planets. But their extrapolation to galaxies has never been directly tested. Might it go amiss? If the laws break down, then modifying them might obviate dark matter.

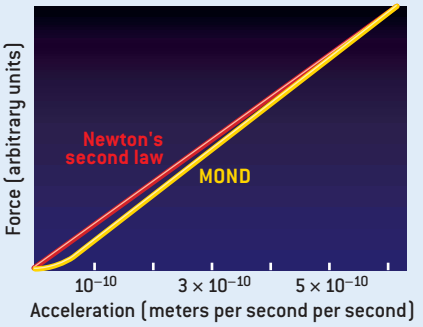
Such a modification would not be without precedent. Two drastic changes to Newtonian physics have already proved necessary. The first upgraded Newtonian dynamics to the theory of relativity—both the special theory (which changed Newton's second law) and the general theory (which altered the law of gravity). The second led to quantum theory, which accounts for the behavior of microscopic systems and even macroscopic systems under certain circumstances. The two proven extensions of Newtonian dynamics come into play under extreme conditions, such as extreme speeds (special relativity) or extremely strong gravity (general relativity). The bulk of the phenomena connected with galactic dynamics involves none of these particular conditions.

What attributes of galactic systems are so extreme that they might require yet another modification? The first possibility that comes to mind is size. Perhaps gravity departs from the Newtonian law at large distances. As early as 1923, English astronomer James H. Jeans proposed modifying the distance dependence of the law of gravity on galactic scales. But the discrepant observations he sought to explain were unrelated to dark matter and, in any event, were later refuted.

Another modified distance dependence was proposed in 1963 by Arrigo Finzi, then at the University of Rome, as a possible solution to the dark matter

Overview/*Alternative to Dark Matter*

- Astronomers have two ways to determine how much matter fills the universe. They tot up everything they see. And they measure how fast the visible objects move, apply the laws of physics and deduce how much mass is needed to generate the gravity that restrains those objects. Vexingly, the two methods give different answers. Most astronomers conclude that some invisible mass also lurks out there—the infamous dark matter.
- Perhaps the fault lies not in the matter but in the laws of physics. The author has proposed a modification to Newton's laws of motion (*see graph at right*)—or, equivalently, of gravity—that would explain away the discrepancy.
- The modification, known as MOND, does an amazingly good job of reproducing observations—better, in many ways, than dark matter does. That said, MOND has some problems, which may be unimportant or may be fatal.



NEW THEORY, known as MOND, changes Newton's second law of motion at low accelerations.

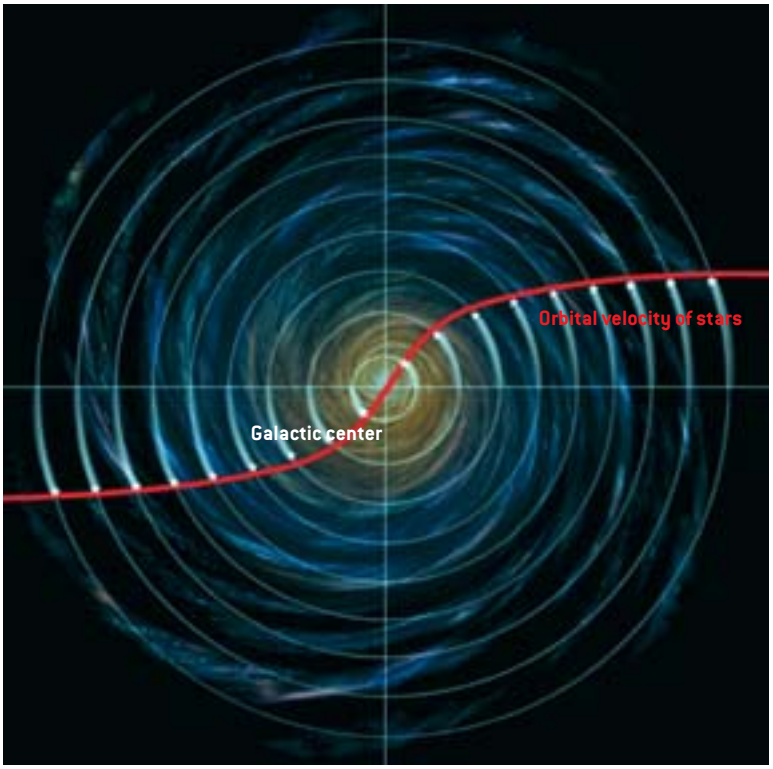
SLIM FILMS (*preceding pages*); BRYAN CHRISTIE DESIGN (*this page*)

MAKING THE ROUNDS

JUST AS THE PLANETS in the solar system orbit around the sun, stars in a galaxy orbit around the galactic center. For example, the sun completes one circuit of the Milky Way every 200 million years or so. Stars in the disks of spiral galaxies have nearly circular orbits, and their orbital speed depends on their distance from the center.

Two effects determine the speed: the distribution of mass and the weakening of the force of gravity with distance. The first effect is simply a matter of geometry—the amount of mass contained within a stellar orbit increases with distance. This effect dominates in the inner reaches of the galaxy. Farther out, the second effect becomes more important. The net result is that the orbital speed initially increases but then levels off and begins to decrease. This relation between speed and distance is known as the rotation curve.

According to Newton's laws, the rotation curve should continue to decrease forever. If MOND is correct, however, it should reach a constant value. —*M.M.*



problem in clusters. But in the early 1980s I showed that such modifications of the distance dependence of gravity fail to reproduce the observations.

What, then, can work? After systematically considering different attributes, I zeroed in on acceleration. The acceleration in galactic systems is many orders of magnitude smaller than in everyday experience. The acceleration of the solar system toward the center of our galaxy (about one angstrom, or 10^{-10} meter, per second per second) is one hundred-billionth the acceleration of the space shuttle toward the center of Earth (about 10 meters per second per second). Nearly 20 years ago I proposed a modification to Newton's second law that changed the relation between force and acceleration when the acceleration is low. This was the beginning of the idea called MOND, for Modified Newtonian Dynamics.

Building Up Speed

MOND INTRODUCES a new constant of nature with the dimensions of acceleration, called a_0. When the acceleration is much larger than a_0, Newton's second law applies as usual: force is proportional to acceleration. But when the acceleration is small compared with a_0, Newton's second law is altered: force becomes proportional to the square of the acceleration. By this scheme, the force needed to impart a given acceleration is always smaller than Newtonian dynamics requires. To account for the observed accelerations in galaxies, MOND predicts a smaller force—hence, less gravity-producing mass—than Newtonian dynamics does [*see illustration on opposite page*]. In this way, it can eliminate the need for dark matter.

In the outskirts of galaxies, the acceleration produced by gravity decreases with distance and eventually goes below a_0. Exactly where this happens depends on the value of a_0 and on the mass. The higher the mass, the farther out the effects of MOND set in. For the value of a_0 that is required by the data, and for a galaxy of typical mass, the transition takes place several tens of thousands of light-years from the center. For the mass of a typical cluster of galaxies, it happens at a few million light-years from the center.

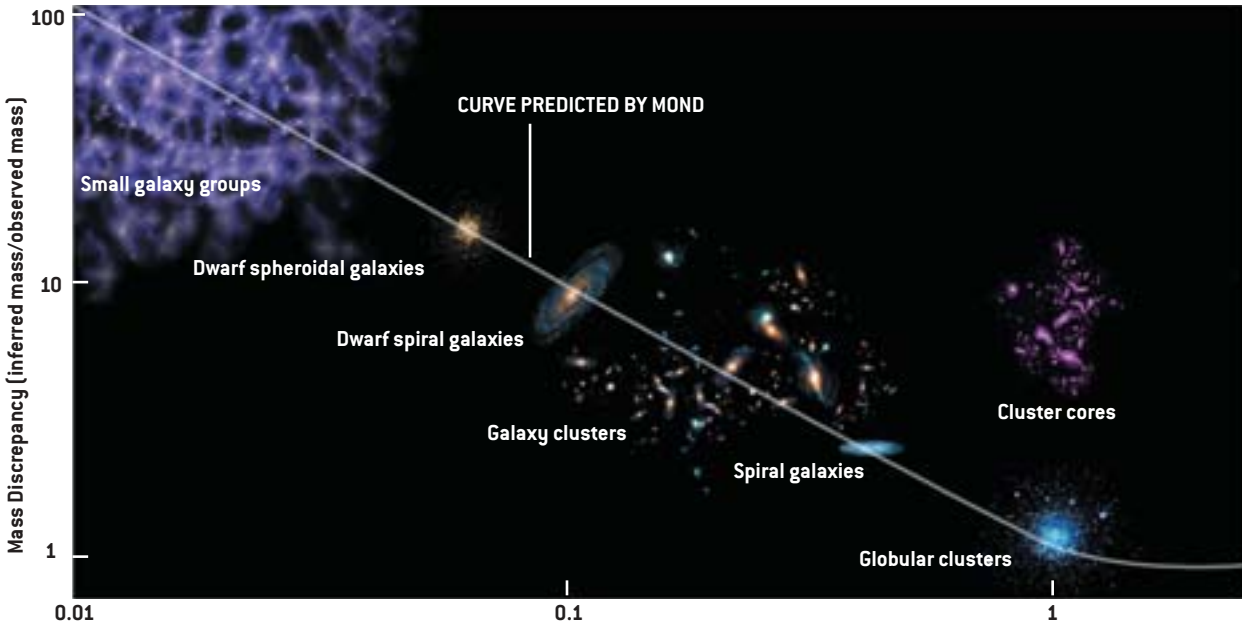
Suppose the bulk of a galaxy is contained within a certain radius. Then, by Newtonian dynamics, beyond this radius the speed of objects in circular orbits (such as gas or stars) should decrease with increasing radius. This is what happens in the solar system. The bulk of the solar system's mass is contained in the sun, and the orbital velocity of the planets decreases with distance. Mercury trundles around the sun much faster than Earth does, for example. Where MOND applies, however, the situation is radically different. At sufficiently large distances from the center of a galaxy, the orbital velocity should stop decreasing and reach a constant value. This constant velocity should be proportional to the fourth root of the galaxy's mass.

How does MOND fare when confronted with the data? Orbital velocities in spiral galaxies, instead of declining with increasing distance from the galactic center, flatten out to a constant value, as predicted by MOND. Moreover, according to an observed correlation known as the Tully-Fisher relation, this constant velocity is proportional to the fourth root of the galaxy's luminosity. This, too, emerges naturally from MOND. The key

DON DIXON

MOND IN GALACTIC SYSTEMS

ALL TYPES OF GALACTIC SYSTEMS—ranging in size from globular star clusters to dwarf galaxies to galaxy groups and clusters—suffer from a discrepancy between the observed and the inferred mass. This discrepancy scales inversely with the characteristic acceleration, just as MOND predicts. Dark-matter models have no explanation for the correlation. MOND's main failure occurs in the cores of large galaxy clusters. —*M.M.*



assumption is that the luminosity, in turn, is proportional to mass or nearly so. Recent observations vindicate the assumption: a direct velocity-mass correlation is even tighter than the velocity-luminosity correlation.

Glimpses of these regularities were already in sight when I proposed MOND; indeed, they furnished clues for its construction. What makes MOND particularly intriguing is that it predicted many effects that could not even be tested when I formulated it. One example is the nature of low-surface-brightness galaxies—stellar agglomerations so wispy that they can barely be seen at all. Whereas the acceleration in ordinary galaxies exceeds a_0 toward the center and drops below a_0 in the outskirts, the acceleration in low-surface-brightness galaxies is smaller than a_0 everywhere. According to MOND, the mass discrepancy should be seen throughout such a galaxy. At the time I propounded MOND, astronomers knew of only a few low-surface-brightness galaxies and had analyzed none in any detail. Since then, they have discovered that the mass mismatch is indeed disproportionately larger in these galaxies than in ordinary galaxies [see "The Ghostliest Galaxies," by Gregory D. Bothun; SCIENTIFIC AMERICAN, February 1997]. MOND correctly anticipated this. It also foresaw the magnitude of the discrepancies.

Another success concerns the shape of galactic rotation curves—that is, the precise variation of orbital velocity with distance. Only since the late 1980s have astronomers made observations detailed enough to compare with theoretical predictions. And the correspondence with MOND is remarkable [*see illustration on page 50*]. These comparisons involve one parameter that must be adjusted for each galaxy: the conversion factor from starlight to mass. The inferred value of this parameter agrees with theoretical expectations. For dark matter, in contrast, the comparisons involve at least two additional adjustable parameters per galaxy—namely, the extent and mass of the dark matter. Despite this flexibility, current dark matter theories do not explain the rotational data as well as MOND does.

Exception to the Rule?

IN OTHER GALACTIC systems, when one plots the mass discrepancy against the typical acceleration, the pattern almost completely agrees with MOND's predictions [*see illustration above*]. The one exception is found in rich galaxy clusters. If we consider the clusters at large, they typically show a mass discrepancy of about a factor of five to 10, which MOND can explain. If, however, we concentrate on the inner parts of these clusters, we find that a mismatch remains. MOND does not sweep away all the invisible mass. Perhaps the theory itself fails, but perhaps the observations are incomplete. Significant amounts of dim matter—ordinary matter in the form of feeble stars or lukewarm

THE AUTHOR

MORDEHAI MILGROM is professor of theoretical physics at the Weizmann Institute in Rehovot, Israel. He is the father of MOND, the most successful and persistent of the alternatives to dark matter. His other scientific interests concern cosmology and high-energy astrophysics.

DON DIXON

gas—could be lurking in these systems.

Ideally, one would like to check MOND using physics experiments as well as astronomical observations. Unfortunately, laboratory tests are infeasible. The acceleration that enters the MOND criterion is the full acceleration with respect to an absolute frame of reference. On Earth or in the near solar system, the large background acceleration—caused by Earth's gravity, its rotation, its revolution around the sun and myriad other factors—would mask the effects of MOND even if we could create a setup with small relative accelerations internally. Similarly, it would be hard to test MOND by studying the motions of the planets. The acceleration of bodies orbiting the sun does not fall below a_0 until one goes about 10,000 times as far from the sun as Earth is, far beyond the orbit of Pluto. To be sure, the structure of MOND for high accelerations—where the theory departs only minutely from Newtonian dynamics—is not yet known. It might be that the departure, though very small, is still large enough to produce observable effects. The claimed anomaly in the motions of certain spacecraft, if verified, could be naturally explained within MOND [see "Pioneering Gas Leak?" by George Musser; Science and the Citizen, SCIENTIFIC AMERICAN, December 1998].

Just as Planck's constant appears in many roles in quantum theory, so does a_0 appear in many ways in MOND's predictions for galactic systems. It is part of the success of the theory that the same value—approximately one angstrom per second per second—works for all these diverse appearances.

Successful as it may be, MOND is, at the moment, a limited phenomenological theory. By phenomenological, I mean that it has not been motivated by, and is not constructed on, fundamental principles. It was born from a direct need to describe and explain a body of observations, much as quantum mechanics (and, indeed, the concept of dark matter) developed. And MOND is limited, because it cannot yet be applied to all the relevant phenomena at hand.

The main reason is that MOND has not been incorporated into a theory that obeys the principles of relativity, either special or general. Perhaps it is impossible to do so; perhaps it is simply a matter of time. After all, it took many years for the quantum idea, as put forth by Max Planck, Einstein and Niels Bohr, to be encapsulated into the Schrödinger equation, and more time still to be made compatible with special relativity. Even today, despite long, concentrated efforts, theorists have not made quantum physics compatible with general relativity.

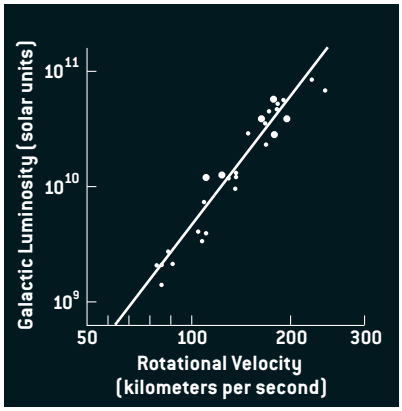
Beyond the Ken

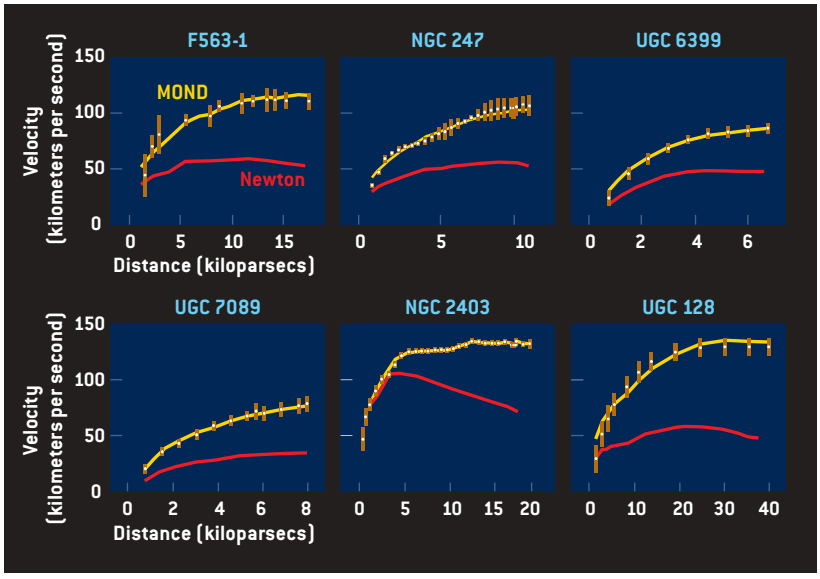
THE PHENOMENA that fall outside the present purview of MOND are those that involve, on the one hand, accelerations smaller than a_0 (so that MOND plays a role) and, on the other, extreme speeds or extremely strong gravity (so that relativity is also called for). Black holes meet the second criterion but fail the first: for the acceleration near a black hole to be smaller than a_0, the hole would have to be larger than the observable universe. Light propagating in the gravitational fields of galactic systems, however, does satisfy both criteria. MOND cannot properly treat this motion, which pertains to gravitational lensing [see "Gravity's Kaleidoscope," by Joachim Wambsganss; SCIENTIFIC AMERICAN, November 2001]. Observations that make use of gravitational lensing exhibit the same mass disparity that observations of galactic dynamics do. But we do not yet know whether MOND can explain the disparity in both cases.

THEORY VERSUS DATA

MOND reproduces key galaxy observations with remarkable precision. At the right is the so-called Tully-Fisher relation. In a sample of spiral galaxies, the orbital speed of stars in the galaxies' outskirts is correlated with the galaxies' brightness. A straight line fits the data and, within the measurement precision, matches the MOND prediction.

Below is a sample of rotation curves for various galaxies, showing how orbital speed varies with distance from the galactic center. —*M.M.*





top: DON DIXON; SOURCE: ROBERT H. SANDERS *Kapteyn Astronomical Institute* AND MARC A. W. VERHEIJEN *University of Wisconsin–Madison; bottom:* BRYAN CHRISTIE DESIGN; SOURCE: STACY S. McGAUGH *University of Maryland*

Not a Bad Idea

MOND is out of the mainstream, but it is far from wacky

By Anthony Aguirre

Although the great majority of astronomers believe that dark matter exists, an alternative hypothesis—a modification of Newtonian gravitational dynamics (MOND)—has quietly endured since its proposal in 1983. As Mordehai Milgrom discusses in the accompanying article, MOND can claim an impressive number of correct predictions regarding the dynamics of galaxies. The reactions of most astronomers fall into three categories:

1. MOND is a tautology. It explains only what it was expressly designed to explain. It has made a few fortuitous predictions, but the success of those predictions has been exaggerated by its proponents.
2. MOND describes a surprising, even mysterious, regularity in the formation and evolution of galaxies. The standard theory of gravity still applies and dark matter still exists, but somehow the dark matter emulates MOND. When applied in detail to unusual galaxies or to systems other than galaxies, MOND will eventually be shown to fail.
3. MOND replaces Newtonian dynamics under certain conditions. It is one aspect of a theory of gravitational dynamics that will supplant Einstein's general theory of relativity.

The first view, through uncharitable, was the one held by most astrophysicists for much of MOND's history. In recent years, however, outright rejection has become much less tenable. MOND's myriad predictions have been confirmed. Many of these studies have been performed by those critical of, or neutral toward, Milgrom's hypothesis. Moreover, MOND reproduces the statistics of galaxy properties at least as well as dark matter models do, even though these models describe crucial aspects of galaxy formation in an ad hoc way.

Most impressively, MOND can predict the details of galaxy rotation using only the distribution of visible matter and an assumed (fixed) ratio of mass to luminosity—a feat beyond the ability of dark matter models. These predictions and the observations they are compared with go far beyond what was available at the time of MOND's formulation. MOND is no tautology.

Meanwhile standard dark matter theory has run into difficulty when applied to galaxies. For example, it predicts that the dark matter cores of galaxies should be far denser than observations indicate. Such problems could be an artifact of computational limitations; researchers still lack computers powerful enough to simulate galaxies in full. But many theorists have taken the discrepancies seriously enough to consider modifications of the properties of dark matter.

The successes of MOND and the difficulties for dark matter have converted a number of astronomers from the first view to the second. Relatively few, though, have gone from the first or second view to the third. Why? I think there are three reasons.

First, as both its opponents and proponents point out, MOND is a modification only of Newtonian dynamics. Despite some effort, MOND's proponents have yet to formulate it in a way that can be applied to post-Newtonian phenomena such as gravitational lensing and cosmic expansion. Either no such theory exists or it is inherently difficult to develop. Whatever the reason, MOND has been unable to confront—and hence pass or fail—some key tests.

Second, it is not clear that MOND works well in systems other than galaxies. For example, its predictions about the temperature of hot gas in clusters of galaxies disagree starkly with observations, unless clusters are dominated by—what else?—undetected matter. One might hope (as do MOND's proponents) that this matter could take a recognizable but hard-to-see baryonic form such as small stars or warm gas. Those possibilities are not currently ruled out, but they are strongly constrained both observationally and theoretically. And it is rather disquieting that dark matter (even if in a prosaic form) must be postulated to save a theory devised to eliminate dark matter.

The third reason, related to the first two, is that standard dark matter theory has scored some impressive triumphs in recent years. Numerical simulations predict a spatial distribution of intergalactic gas that is in exquisite agreement with observations. Independent estimates of the mass of dark matter in clusters all agree with one another. The predicted growth of structures correctly links the galaxy distribution we see on large scales today with the tiny temperature fluctuations in the cosmic microwave background radiation from 13 billion years ago.

So what are astronomers to do? Those who are most sympathetic to Milgrom's hypothesis should continue the search for a fundamental theory of MOND, without which the idea will never draw the majority of physicists away from the standard paradigm. For others, I think that it is productive to study, test and use MOND as a convenient rule of thumb whether or not one accepts a modification of Newtonian dynamics. Perhaps we could call it Milgrom's Fitting Formula, or MIFF, to emphasize that we are using it as a practical tool while reserving judgment about whether standard physics is indeed wrong.

If general relativity is correct, and dark matter real, then as the precision of measurements increases, MIFF will ultimately fail. In the meantime, MIFF can provide a compact summary of a great deal of knowledge concerning galaxy formation and evolution.

Anthony Aguirre is a theoretical cosmologist at the Institute for Advanced Study in Princeton, N.J. He is the lead author of two critical studies of MOND, available online at ***arXiv.org/abs/astro-ph/0105184*** *and* ***arXiv.org/abs/hep-ph/0105083***

SLIM FILMS

A second system that requires MOND and relativity is the universe at large. It follows that cosmology cannot be treated in MOND. This incapacity extends to questions relating to incipient structures in the universe. MOND can be applied to well-formed systems that are detached from the global cosmological soup, but it cannot describe the early moments before galactic systems became distinct.

Researchers have made preliminary attempts to deal with what these phenomena might look like in a MOND-inspired theory. For example, Robert H. Sanders of the Kapteyn Astronomical Institute in Groningen, the Netherlands, and Adi Nusser of the Technion–Israel Institute of Technology have devised scenarios of galaxy formation by supplementing MOND with further working assumptions. But it must be said that without an underlying theory, none of these efforts can be fully trusted.

In what directions should one look for the needed underlying theory? A clue may come from the value of a_0. One angstrom per second per second would take a body from rest to near the speed of light in the lifetime of the universe. In other words, a_0 is roughly the same number as the product of two important constants: the speed of light and the Hubble constant, the present-day expansion rate of the universe. It is also near the (unrelated) acceleration produced by the dark energy. This numerical proximity, if it is not just a coincidence, may tell us either that cosmology somehow enters local laws of physics, such as the law of inertia, to produce MOND or that a common agent affects both cosmology and local physics so as to leave the same mark on both of them.

Resistance

MOND APPEARS to suggest that inertia—the responsiveness of a body to a force—is not an inherent property of bodies but is acquired by the body by dint of its interaction with the universe at large. This idea falls within the framework of an old concept, Mach's principle, which attributes inertia to such an interaction.

Physics abounds with instances in which the effective inertia of particles is not an inherent property but rather is produced by their interaction with a background medium. For example, electrons in solids sometimes behave as if their inertia has been greatly modified by the rest of the solid. Might an analogous effect be responsible for genuine inertia? If so, what could be the agent whose presence impedes acceleration and thus produces inertia?

An exciting possibility is the vacuum. The vacuum is what is left when one annihilates all matter (or, equivalently, energy) that can be annihilated. According to quantum theory, the remnant is not a complete void but rather a minimal representation of all the forms of energy. The interaction of the vacuum with particles might contribute to the inertia of objects. Intriguingly, the vacuum also enters cosmology as one explanation for dark energy. It remains unknown, however, whether the vacuum can be fully responsible for inertia and whether it can indeed account for MOND.

Many researchers maintain that although MOND neatly reproduces galactic phenomenology, it does not constitute a fundamental truth. The argument goes that MOND is perhaps an economical and useful summary of what we see in nature but that these relations and regularities will one day emerge from the complexity of the dark matter paradigm. Last year Manoj Kaplinghat and Michael S. Turner, both then at the University of Chicago, claimed that the appearance of a characteristic acceleration akin to a_0 occurs naturally in dark matter models. According to their scenario, these models predict the formation of dark matter halos of a restricted type around galaxies.

Subsequently, however, I pointed out that this scenario does not work. Kaplinghat and Turner based their work on crude approximations that disagree with observed dark matter halos and with detailed numerical simulations of dark matter behavior. Those simulations, as they now stand, do not reproduce any aspect of MOND phenomenology. To boot, their claimed result accounts for only a small fraction of the successes of MOND. But it is possible that MOND follows from the dark matter paradigm in a different way. Time will tell.

In the meantime, work is proceeding on understanding the observational consequences of MOND and improving the theory itself, with contributions from Sanders, Jacob D. Bekenstein of the Hebrew University in Jerusalem and Stacy S. McGaugh of the University of Maryland. MOND continues to be the most successful and enduring of the alternatives to dark matter. Observations, far from ruling it out, seem to prefer it over dark matter. Although people are right to be skeptical about MOND, until definitive evidence arrives for dark matter or for one of its alternatives, we should keep our minds open. SA

If we accept a departure from the standard laws of physics, we might do away with dark matter.

MORE TO EXPLORE

A Modification of the Newtonian Dynamics as a Possible Alternative to the Hidden Mass Hypothesis. Mordehai Milgrom in *Astrophysical Journal,* Vol. 270, pages 365–370; July 15, 1983. Available online from **adsabs.harvard.edu/**

How Cold Dark Matter Theory Explains Milgrom's Law. Manoj Kaplinghat and Michael S. Turner in *Astrophysical Journal,* Vol. 569, No. 1, Part 2, pages L19–L22; April 10, 2002. Available online at **arXiv.org/abs/astro-ph/0107284**

Modified Newtonian Dynamics as an Alternative to Dark Matter. Robert H. Sanders and Stacy S. McGaugh in *Annual Reviews of Astronomy and Astrophysics,* Vol. 40; 2002. Available online at **astro-ph/0204521**

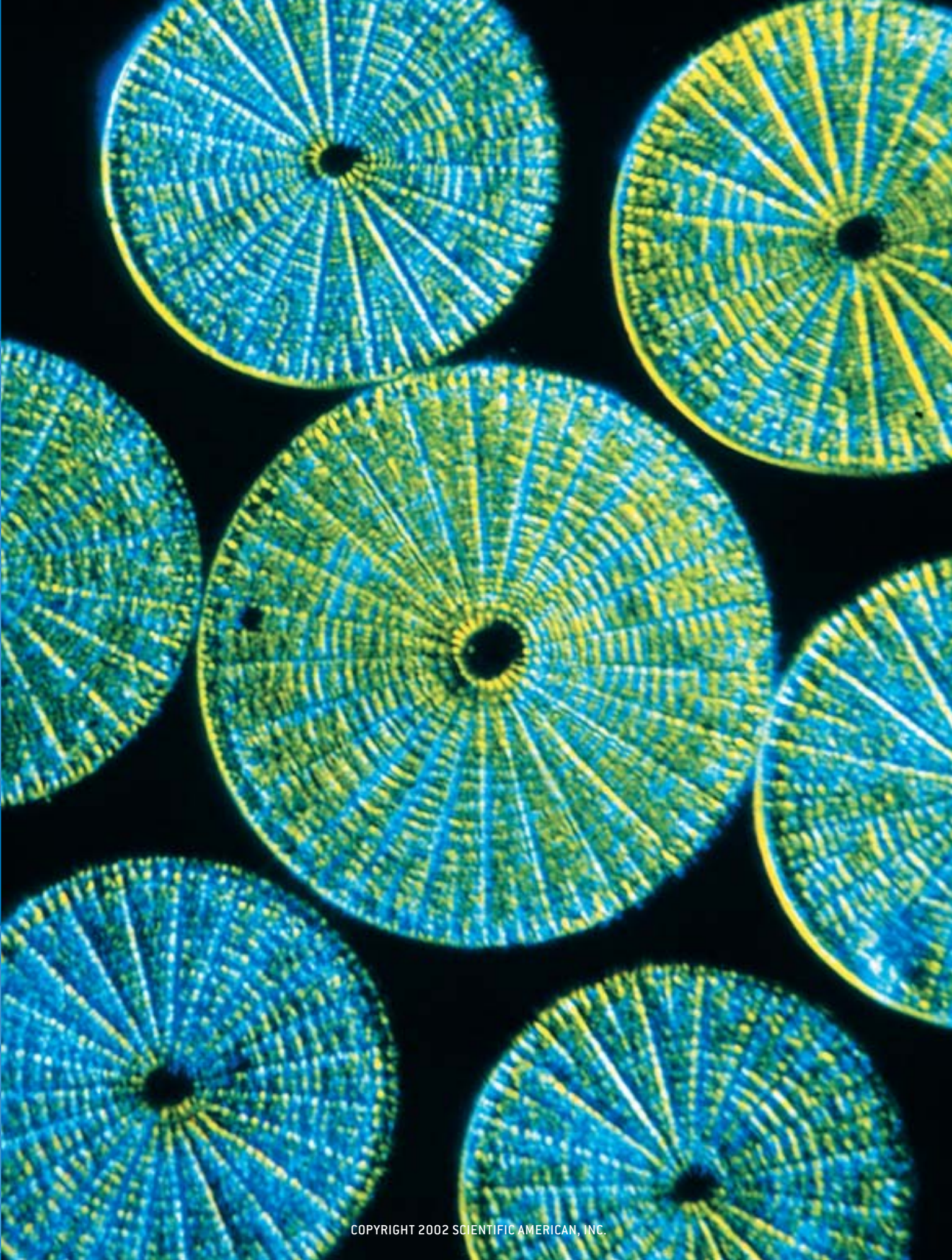
Stacy S. McGaugh's MOND Web site is at **www.astro.umd.edu/~ssm/mond**

DIATOMS ARE THE GIANTS of the phytoplankton world. The species pictured here, *Actinocyclus sp.*, can measure up to a millimeter in diameter.

The Ocean's Invisible Forest

Marine phytoplankton play a critical role in regulating the earth's climate. Could they also be used to combat global warming?

BY PAUL G. FALKOWSKI



Every drop of water
in the top 100 meters of the ocean

contains thousands of free-floating, microscopic flora called phytoplankton. These single-celled organisms—including diatoms and other algae—inhabit three quarters of the earth's surface, and yet they account for less than 1 percent of the 600 billion metric tons of carbon contained within its photosynthetic biomass. But being small doesn't stop this virtually invisible forest from making a bold mark on the planet's most critical natural cycles.

Arguably one of the most consequential activities of marine phytoplankton is their influence on climate. Until recently, however, few researchers appreciated the degree to which these diminutive ocean dwellers can draw the greenhouse gas carbon dioxide (CO_2) out of the atmosphere and store it in the deep sea. New satellite observations and extensive oceanographic research projects are finally revealing how sensitive these organisms are to changes in global temperatures, ocean circulation and nutrient availability.

With this knowledge has come a temptation among certain researchers, entrepreneurs and policymakers to manipulate phytoplankton populations—by adding nutrients to the oceans—in an effort to mitigate global warming. A two-month experiment conducted early this year in the Southern Ocean confirmed that injecting surface waters with trace amounts of iron stimulates phytoplankton growth; however, the efficacy and prudence of widespread, commercial ocean-fertilization schemes are still hotly debated. Exploring how human activities can alter phytoplankton's impact on the planet's carbon cycle is crucial for predicting the long-term ecological side effects of such actions.

Seeing Green

OVER TIME SPANS of decades to centuries, plants play a major role in pulling CO_2 out of the atmosphere. Such has been the case since about three billion years ago, when oxygenic, or oxygen-producing, photosynthesis evolved in cyanobacteria, the world's most abundant type of phytoplankton. Phytoplankton and all land-dwelling plants—which evolved from phytoplankton about 500 million years ago—use the energy in sunlight to split water molecules into atoms of hydrogen and oxygen. The oxygen is liberated as a waste product and makes possible all animal life on earth, including our own. The planet's cycle of carbon (and, to a large extent, its climate) depends on photosynthetic organisms using the hydrogen to help convert the inorganic carbon in CO_2 into organic matter—the sugars, amino acids and other biological molecules that make up their cells.

This conversion of CO_2 into organic matter, also known as primary production, has not always been easy to measure. Until about five years ago, most biologists were greatly underestimating the contribution of phytoplankton relative to that of land-dwelling plants. In the second half of the 20th century, biological oceanographers made thousands of individual measurements of phytoplankton productivity. But these data points were scattered so unevenly around the world that the coverage in any given month or year remained extremely small. Even with the help of mathematical models to fill in the gaps, estimates of total global productivity were unreliable.

That changed in 1997, when NASA launched the Sea Wide Field Sensor (SeaWiFS), the first satellite capable of observing the entire planet's phytoplankton

Overview/*Climate Regulators*

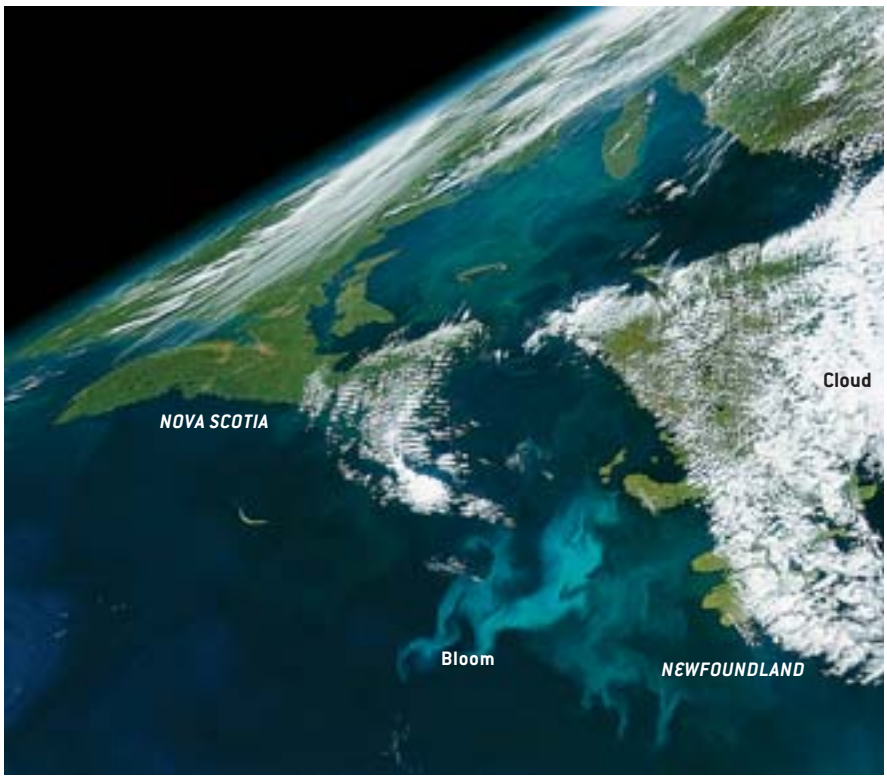
- Rapid life cycles of marine phytoplankton transfer heat-trapping carbon dioxide (CO_2) from the atmosphere and upper ocean to the deep sea, where the gas remains sequestered until currents return it to the surface hundreds of years later.
- If all of the world's marine phytoplankton were to die today, the concentration of CO_2 in the atmosphere would rise by 200 parts per million—or 35 percent—in a matter of centuries.
- Adding certain nutrients to the ocean surface can dramatically enhance the growth of phytoplankton and thus their uptake of CO_2 via photosynthesis, but whether intentional fertilization increases CO_2 storage in the deep sea is still uncertain.
- Artificially enhancing phytoplankton growth will have inevitable but unpredictable consequences on natural marine ecosystems.

PETER PARKS *imagequest3d.com (preceding page)*

populations every single week. The ability of satellites to see these organisms exploits the fact that oxygenic photosynthesis works only in the presence of chlorophyll *a*. This and other pigments absorb the blue and green wavelengths of sunlight, whereas water molecules scatter them. The more phytoplankton soaking up sunlight in a given area, the darker that part of the ocean looks to an observer in space. A simple satellite measurement of the ratio of blue-green light leaving the oceans is thus a way to quantify chlorophyll—and, by association, phytoplankton abundance.

The satellite images of chlorophyll, coupled with the thousands of productivity measurements, dramatically improved mathematical estimates of overall primary productivity in the oceans. Although various research groups differed in their analytical approaches, by 1998 they had all arrived at the same startling conclusion: every year phytoplankton incorporate approximately 45 billion to 50 billion metric tons of inorganic carbon into their cells—nearly double the amount cited in the most liberal of previous estimates.

That same year my colleagues Christopher B. Field and James T. Randerson of the Carnegie Institution of Washington and Michael J. Behrenfeld of Rutgers University and I decided to put this figure into a worldwide context by constructing the first satellite-based maps that compared primary production in the oceans with that on land. Earlier investigations had suggested that land plants assimilate as much as 100 billion metric tons of inorganic carbon a year. To the surprise of many ecologists, our satellite analysis revealed that they assimilate only about 52 billion metric tons. In other words, phytoplankton draw nearly as much CO_2 out of the atmosphere and oceans through photosynthesis as do trees, grasses and all other land plants combined.



BLOOM OF PHYTOPLANKTON called coccolithophores swirls across several hundred square kilometers of the deep-blue Atlantic just south of Newfoundland. Such natural blooms arise in late spring, when currents deliver nutrients from the deep ocean to sunlit surface waters.

Sinking out of Sight

LEARNING THAT phytoplankton were twice as productive as previously thought meant that biologists had to reconsider dead phytoplankton's ultimate fate, which strongly modifies the planet's cycle of carbon and CO_2 gas. Because phytoplankton direct virtually all the energy they harvest from the sun toward photosynthesis and reproduction, the entire marine population can replace itself every week. In contrast, land plants must invest copious energy to build wood, leaves and roots and take an average of 20 years to replace themselves. As phytoplankton cells divide—every six days on average—half the daughter cells die or are eaten by zooplankton, miniature animals that in turn provide food for shrimp, fish and larger carnivores.

The knowledge that the rapid life cycle of phytoplankton is the key to their ability to influence climate inspired an ongoing international research program called the Joint Global Ocean Flux Study (JGOFS). Beginning in 1988, JGOFS investigators began quantifying the oceanic carbon cycle, in which the organic matter in the dead phytoplankton cells and animals' fecal material sinks and is consumed by microbes that convert it back into inorganic nutrients, including CO_2. Much of this recycling happens in the sunlit layer of the ocean, where the CO_2 is instantly available to be photosynthesized or absorbed back into the atmosphere. (The entire volume of gases in the atmosphere is exchanged with those dissolved in the upper ocean every six years or so.)

Most influential to climate is the organic matter that sinks into the deep ocean before it decays. When released below about 200 meters, CO_2 stays put for much longer because the colder temperature—and higher density—of this water prevents it from mixing with the warmer waters above. Through this process, known as the biological pump, phytoplankton remove CO_2 from the surface waters and atmosphere and store it in the deep ocean. Last year Edward A. Laws of the University of Hawaii, three other JGOFS researchers and I reported that the material pumped into the deep sea amounts to between seven billion and eight billion metric tons, or 15 percent, of the carbon that phytoplankton assimilate every year.

Within a few hundred years almost all the nutrients released in the deep sea find their way via upwelling and other ocean currents back to sunlit surface waters,

COURTESY OF THE SeaWiFS PROJECT, NASA GODDARD SPACE FLIGHT CENTER AND ORBIMAGE

Phytoplankton's Influence on the Global Carbon Cycle

THE EARTH'S CARBON CYCLE can dramatically influence global climate, depending on the relative amounts of heat-trapping carbon dioxide (CO_2) that move into (*yellow arrows*) and out of (*green arrows*) the atmosphere and upper ocean, which exchange gases every six years or so. Plantlike organisms called phytoplankton play four critical roles in this cycle. These microscopic ocean dwellers annually incorporate about 50 billion metric tons of carbon into their cells during photosynthesis, which is often stimulated by iron via windblown dust (*1*). Phytoplankton also temporarily store CO_2 in the deep ocean via the biological pump: about 15 percent of the carbon they assimilate settles into the deep sea, where it is released as CO_2 as the dead cells decay (*2*). Over hundreds of years, upwelling currents transport the dissolved gas and other nutrients back to sunlit surface waters.

A tiny fraction of the dead cells avoids being recycled by becoming part of petroleum deposits or sedimentary rocks in the seafloor. Some of the rock-bound carbon escapes as CO_2 gas and reenters the atmosphere during volcanic eruptions after millions of years of subduction and metamorphism in the planet's interior (*3*).

Burning of fossil fuels, in contrast, returns CO_2 to the atmosphere about a million times faster (*4*). Marine phytoplankton and terrestrial forests cannot naturally incorporate CO_2 quickly enough to mitigate this increase; as a consequence, the global carbon cycle has fallen out of balance, warming the planet. Some people have considered correcting this disparity by fertilizing the oceans with dilute iron solutions to artificially enhance phytoplankton photosynthesis and the biological pump.

—P.G.F.

where they stimulate additional phytoplankton growth. This cycle keeps the biological pump at a natural equilibrium in which the concentration of CO_2 in the atmosphere is about 200 parts per million lower than it would be otherwise—a significant factor considering that today's CO_2 concentration is about 365 parts per million.

Over millions of years, however, the biological pump leaks slowly. About one half of 1 percent of the dead phytoplankton cells and fecal matter settles into seafloor sediments before it can be recycled in the upper ocean. Some of this carbon becomes incorporated into sedimentary rocks such as black shales, the largest reservoir of organic matter on earth. An even smaller fraction forms deposits of petroleum and natural gas. Indeed, these primary fuels of the industrial world are nothing more than the fossilized remains of phytoplankton.

The carbon in shales and other rocks returns to the atmosphere as CO_2 only after the host rocks plunge deep into the earth's interior when tectonic plates collide at subduction zones. There extreme heat and pressure melt the rocks and thus force out some of the CO_2 gas, which is eventually released by way of volcanic eruptions.

By burning fossil fuels, people are bringing buried carbon back into circulation about a million times faster than volcanoes do. Forests and phytoplankton cannot absorb CO_2 fast enough to keep pace with these increases, and atmospheric concentrations of this greenhouse gas have risen rapidly, thereby almost certainly contributing significantly to the global warming trend of the past 50 years.

As policymakers began looking in the early 1990s for ways to make up for this shortfall, they turned to the oceans, which have the potential to hold all the CO_2 emitted by the burning of fossil fuels. Several researchers and private corporations proposed that artificially accelerating the biological pump could take advantage of this extra storage capacity. Hypothetically, this enhancement could be achieved in two ways: add extra nutrients to the upper ocean or ensure that nutrients not fully consumed are used more efficiently. Either way, many speculated, more phytoplankton would grow and more dead cells would be available to carry carbon into the deep ocean.

Fixes and Limits

UNTIL MAJOR DISCOVERIES over the past 10 years clarified the natural distribution of nutrients in the oceans, scientists knew little about which ocean fertilizers would work for phytoplankton. Of the two primary nutrients that all phytoplankton need—nitrogen and phosphorus—phosphorus was long thought to be the harder to come by. Essential for synthesis of nucleic acids, phosphorus occurs exclusively in phosphate minerals within continental rocks and thus enters the oceans only via freshwater runoff such as rivers. Nitrogen (N_2) is the most abundant gas in the earth's atmosphere and dissolves freely in seawater.

By the early 1980s, however, biological oceanographers had begun to realize that they were overestimating the rate at which nitrogen becomes available for use by living organisms. The vast majority of phytoplankton can use nitrogen to build proteins only after it is fixed—in other words, combined with hydrogen or oxygen atoms to form ammonium (NH_4^+), nitrite (NO_2^-) or nitrate (NO_3^-). The vast majority of nitrogen is fixed by small subsets of bacteria and cyanobacteria that convert N_2 to ammonium, which is released into seawater as the organisms die and decay.

Within the details of that chemical transformation lie the reasons why phytoplankton growth is almost always limited by the availability of nitrogen. To catalyze the reaction, both bacteria and cyanobacteria use nitrogenase, an enzyme that relies on another element, iron, to transfer electrons. In cyanobacteria, the primary energy source for nitrogen fixation is another process that requires a large investment of iron—the production of adenosine triphosphate (ATP). For these reasons, many oceanographers think that iron controls how much nitrogen these special organisms can fix.

In the mid-1980s the late John Martin, a chemist at the Moss Landing Marine Laboratories in California, hypothesized that the availability of iron is low enough in many ocean realms that phytoplankton production is severely restricted. Using extremely sensitive methods to measure the metal, he discovered that its concentration in the equatorial Pacific, the northeastern Pacific and the Southern Ocean is so low that phosphorus and nitrogen in these surface waters are never used up.

Martin's iron hypothesis was initially controversial, in part because previous ocean measurements, which turned out to be contaminated, had suggested that iron was plentiful. But Martin and his coworkers pointed out that practically the only way iron reaches the surface waters of the open ocean is via windblown dust.

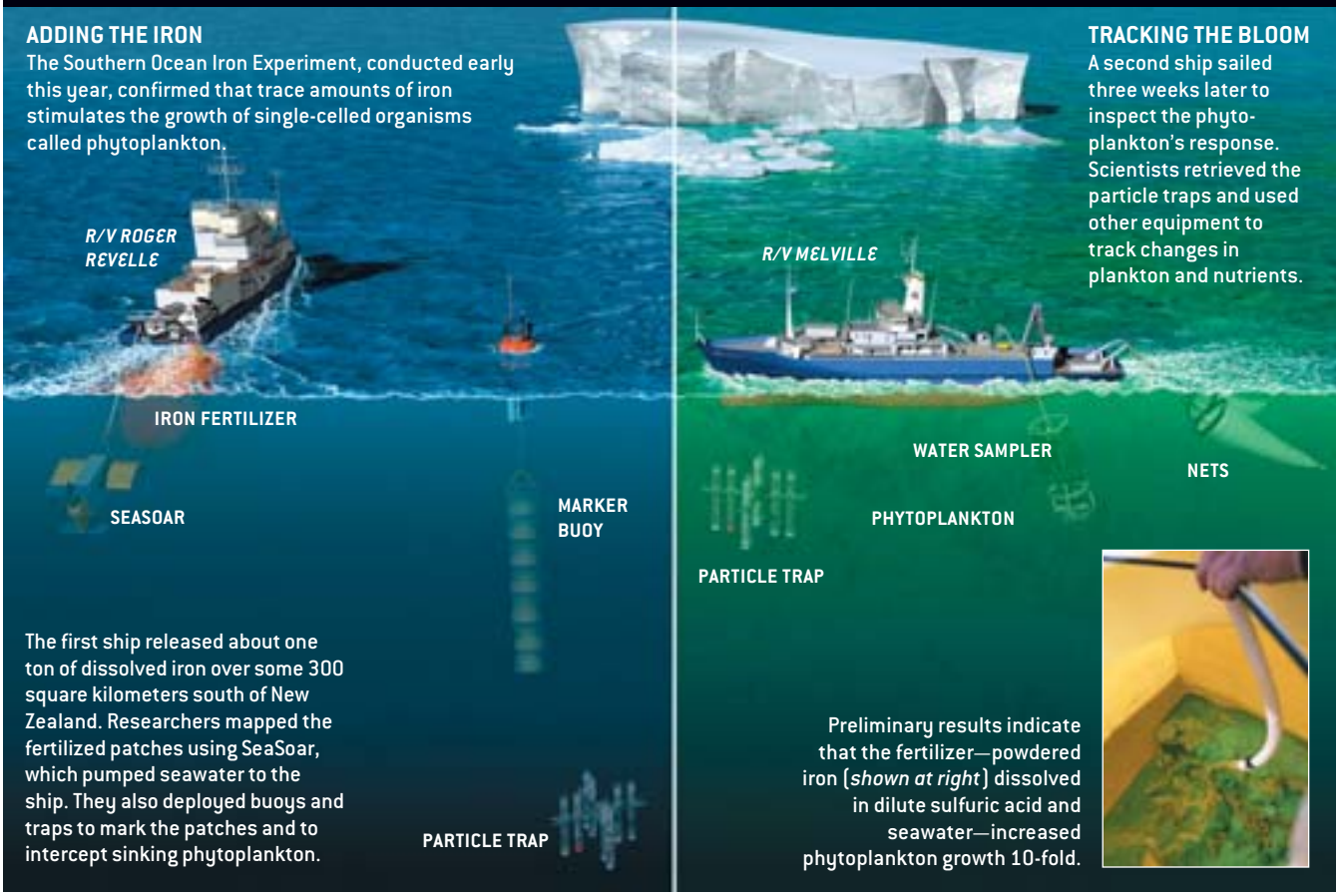
Phytoplankton draw nearly as much CO_2 out of the oceans and atmosphere as all land plants do.

DAVID FIERSTEIN (*opposite page*); PETER PARKS *imagequest3d.com* (*Triceratium*)

THE AUTHOR

PAUL G. FALKOWSKI is professor in the Institute of Marine and Coastal Sciences and the department of geology at Rutgers University. Born and raised in New York City, Falkowski earned his Ph.D. from the University of British Columbia in 1975. After completing a postdoctoral fellowship at the University of Rhode Island, he joined Brookhaven National Laboratory in 1976 as a scientist in the newly formed oceanographic sciences division. In 1997 he and Zbigniew Kolber of Rutgers University co-invented a specialized fluorometer that can measure phytoplankton productivity in real time. The next year Falkowski joined the faculty at Rutgers, where his research focuses on the coevolution of biological and physical systems.

IRON FERTILIZER FOR PHYTOPLANKTON



Consequently, in vast areas of the open ocean, far removed from land, the concentration of this critical element seldom exceeds 0.2 part per billion—a fiftieth to a hundredth the concentrations of phosphate or fixed inorganic nitrogen.

Historical evidence buried in layers of ice from Antarctica also supported Martin's hypothesis. The Vostok ice core, a record of the past 420,000 years of the earth's history, implied that during ice ages the amount of iron was much higher and the average size of the dust particles was significantly larger than during warmer times. These findings suggest that the continents were dry and wind speeds were high during glacial periods, thereby injecting more iron and dust into the atmosphere than during wetter interglacial times.

Martin and other investigators also noted that when dust was high, CO_2 was low, and vice versa. This correlation implied that increased delivery of iron to the oceans during peak glacial times stimulated both nitrogen fixation and phytoplankton's use of nutrients. The resulting rise in phytoplankton productivity could have enhanced the biological pump, thereby pulling more CO_2 out of the atmosphere.

The dramatic response of phytoplankton to changing glacial conditions took place over thousands of years, but Martin wanted to know whether smaller changes could make a difference in a matter of days. In 1993 Martin's colleagues conducted the world's first open-ocean manipulation experiment by adding iron directly to the equatorial Pacific. Their research ship carried tanks containing a few hundred kilograms of iron dissolved in dilute sulfuric acid and slowly released the solution as it traversed a 50-square-kilometer patch of ocean like a lawn mower. The outcome of this first experiment was promising but inconclusive, in part because the seafaring scientists were able to schedule only about a week to watch the phytoplankton react. When the same group repeated the experiment for four weeks in 1995, the results were clear: the additional iron dramatically increased phytoplankton photosynthesis, leading to a bloom of organisms that colored the waters green.

Since then, three independent groups, from New Zealand, Germany and the U.S., have demonstrated unequivocally that adding small amounts of iron to the Southern Ocean greatly stimulates phytoplankton productivity. The most extensive fertilization experiment to date took place during January and February of this year. The project, called the Southern Ocean Iron Experiment (SOFeX) and led by the Monterey Bay Aquarium Research Institute and the Moss Landing Marine Laboratories, involved three ships and 76 scientists, including four of my colleagues from Rutgers. Preliminary results indicate that one ton of iron solution

ILLUSTRATION BY DAVID FIERSTEIN; PHOTOGRAPH BY RON NICOLAYSEN

released over about 300 square kilometers resulted in a 10-fold increase in primary productivity in eight weeks' time.

These results have convinced most biologists that iron indeed stimulates phytoplankton growth at high latitudes, but it is important to note that no one has yet proved whether this increased productivity enhanced the biological pump or increased CO_2 storage in the deep sea. The most up-to-date mathematical predictions suggest that even if phytoplankton incorporated all the unused nitrogen and phosphorus in the surface waters of the Southern Ocean over the next 100 years, at most 15 percent of the CO_2 released during fossil-fuel combustion could be sequestered.

Fertilizing the Ocean

DESPITE THE MYRIAD uncertainties about purposefully fertilizing the oceans, some groups from both the private and public sectors have taken steps toward doing so on much larger scales. One company has proposed a scheme in which commercial ships that routinely traverse the southern Pacific would release small amounts of a fertilizer mix. Other groups have discussed the possibility of piping nutrients, including iron and ammonia, directly into coastal waters to trigger phytoplankton blooms. Three American entrepreneurs have even convinced the U.S. Patent and Trademark Office to issue seven patents for commercial ocean-fertilization technologies, and yet another is pending.

It is still unclear whether such ocean-fertilization strategies will ever be technically feasible. To be effective, fertilization would have to be conducted year in and year out for decades. Because ocean circulation will eventually expose all deep waters to the atmosphere, all the extra CO_2 stored by the enhanced biological pump would return to the atmosphere within a few hundreds years of the last fertilizer treatment. Moreover, the reach of such efforts is not easily controlled. Farmers cannot keep nutrients contained to a plot of land; fertilizing a patch of turbulent ocean water is even less manageable. For this reason, many ocean experts argue that that once initiated, large-scale fertilization could produce long-term damage that would be difficult, if not impossible, to fix.

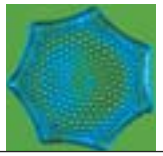
Major disruptions to the marine food web are a foremost concern. Computer simulations and studies of natural phytoplankton blooms indicate that enhancing primary productivity could lead to local problems of severe oxygen depletion. The microbes that consume dead phytoplankton cells as they sink toward the seafloor sometimes consume oxygen faster than ocean circulation can replenish it. Creatures that cannot escape to more oxygen-rich waters will suffocate.

Such conditions also encourage the growth of microbes that produce methane and nitrous oxide, two greenhouse gases with even greater heat-trapping capacity than CO_2. According to the National Oceanic and Atmospheric Administration, severe oxygen depletion and other problems triggered by nutrient runoff have already degraded more than half the coastal waters in the U.S., such as the infamous "dead zone" in the northern Gulf of Mexico. Dozens of other regions around the world are battling similar difficulties.

Even if the possible unintended consequences of fertilization were deemed tolerable, any such efforts must also compensate for the way plants and oceans would respond to a warmer world. Comparing satellite observations of phytoplankton abundance from the early 1980s with those from the 1990s suggests that the ocean is getting a little bit greener, but several investigators have noted that higher productivity does not guarantee that more carbon will be stored in the deep ocean. Indeed, the opposite may be true. Computer simulations of the oceans and the atmosphere have shown that additional warming will increase stratification of the ocean as freshwater from melting glaciers and sea ice floats above denser, salty seawater. Such stratification would actually slow the biological pump's ability to transport carbon from the sea surface to the deep ocean.

New satellite sensors are now watching phytoplankton populations on a daily basis, and future small-scale fertilization experiments will be critical to better understanding phytoplankton behavior. The idea of designing large, commercial ocean-fertilization projects to alter climate, however, is still under serious debate among the scientific community and policymakers alike. In the minds of many scientists, the potential temporary human benefit of commercial fertilization projects is not worth the inevitable but unpredictable consequences of altering natural marine ecosystems. In any case, it seems ironic that society would call on modern phytoplankton to help solve a problem created in part by the burning of their fossilized ancestors. SA

Adding iron to the Southern Ocean greatly stimulates phytoplankton productivity.



PETER PARKS *imagequest3d.com* (*Triceratium*)

MORE TO EXPLORE

Aquatic Photosynthesis. Paul G. Falkowski and John A. Raven. Blackwell Scientific, 1997.

The Global Carbon Cycle: A Test of Our Knowledge of the Earth as a System. Paul G. Falkowski et al. in *Science*, Vol. 290, pages 291–294; October 13, 2000.

The Changing Ocean Carbon Cycle: A Midterm Synthesis of the Joint Global Ocean Flux Study. Edited by Roger B. Hanson, Hugh W. Ducklow and John G. Field. Cambridge University Press, 2000.

Ocean primary productivity distribution maps and links to satellite imagery are located at **http://marine.rutgers.edu/opp/index.html**

Details about the Southern Ocean Iron Experiment are located at **www.mbari.org/education/cruises/SOFeX2002/index.htm**

COMPUTERS WITHOUT CLOCKS

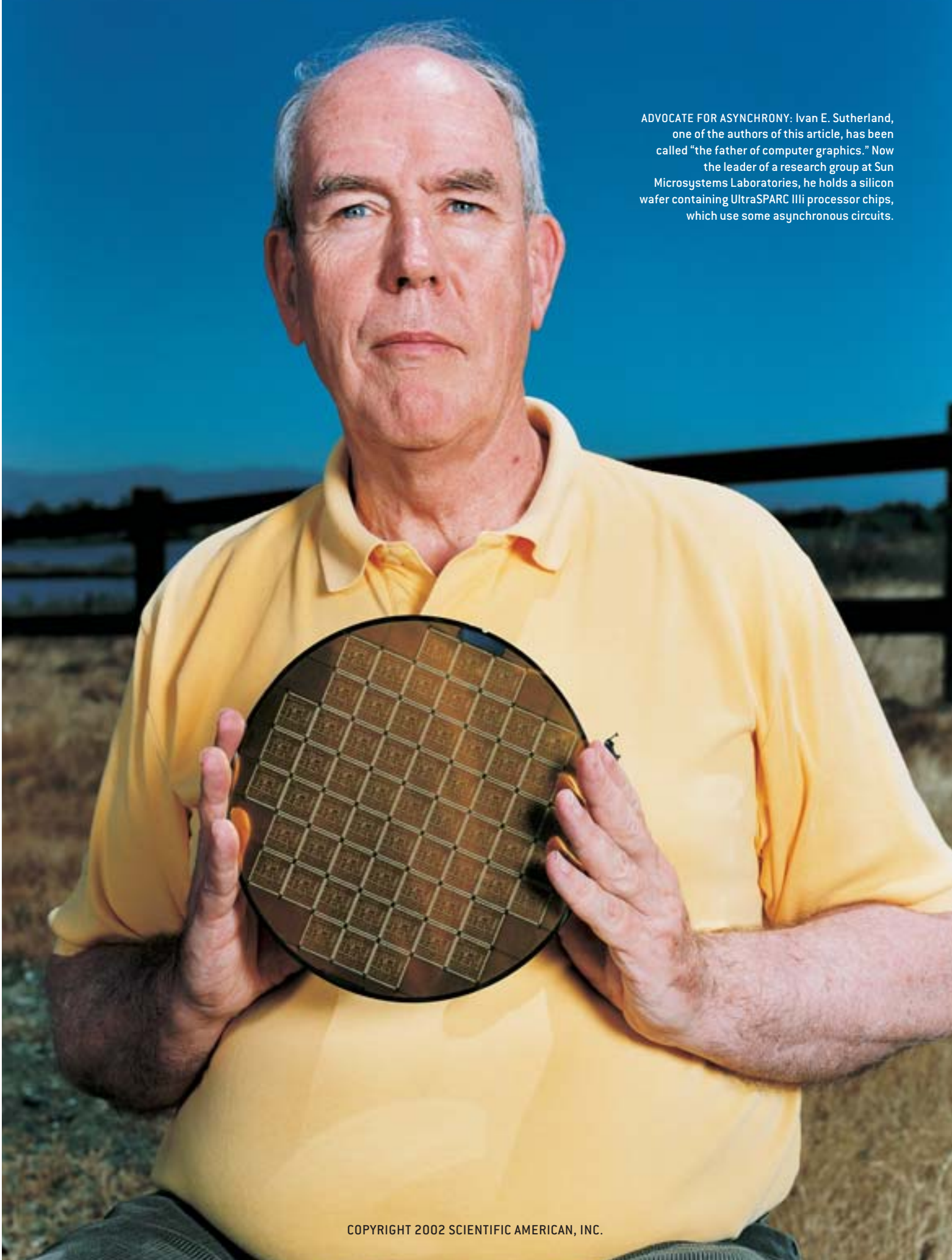
ASYNCHRONOUS CHIPS IMPROVE COMPUTER PERFORMANCE BY LETTING EACH CIRCUIT RUN AS FAST AS IT CAN

By Ivan E. Sutherland and Jo Ebergen

How fast is your personal computer?

When people ask this question, they are typically referring to the frequency of a minuscule clock inside the computer, a crystal oscillator that sets the basic rhythm used throughout the machine. In a computer with a speed of one gigahertz, for example, the crystal "ticks" a billion times a second. Every action of the computer takes place in tiny steps, each a billionth of a second long. A simple transfer of data may take only one step; complex calculations may take many steps. All operations, however, must begin and end according to the clock's timing signals.

OLIVIER LAUDE



ADVOCATE FOR ASYNCHRONY: Ivan E. Sutherland, one of the authors of this article, has been called "the father of computer graphics." Now the leader of a research group at Sun Microsystems Laboratories, he holds a silicon wafer containing UltraSPARC IIIi processor chips, which use some asynchronous circuits.

Because most modern computers use a single rhythm, we call them synchronous. Inside the computer's microprocessor chip, a clock distribution system delivers the timing signals from the crystal oscillator to the various circuits, just as sound in air delivers the beat of a drum to soldiers to set their marching pace. Because all parts of the chip share the same rhythm, the output of any circuit from one step can serve as the input to any other circuit for the next step. The synchronization provided by the clock helps chip designers plan sequences of actions for the computer.

The use of a central clock also creates problems. As speeds have increased, distributing the timing signals has become more and more difficult. Present-day transistors can process data so quickly that they can accomplish several steps in the time that it takes a wire to carry a signal from one side of the chip to the other. Keeping the rhythm identical in all parts of a large chip requires careful design and a great deal of electrical power. Wouldn't it be nice to have an alternative?

Our research group at Sun Microsystems Laboratories seeks such alternatives. Along with several other groups worldwide, we are investigating ways to design computing systems in which each part can proceed at its own pace instead of depending on the rhythm of a central clock. We call such systems asynchronous. Each part of an asynchronous system may extend or shorten the timing of its steps when necessary, much as a hiker takes long or short steps when walking across rough terrain. Some of the pioneers of the computer age, such as mathematician Alan M. Turing, tried using asynchronous designs to build machines in the early 1950s. Engineers soon abandoned this approach in favor of synchronous computers because common timing made the design process so much easier.

Now asynchronous computing is experiencing a renaissance. Researchers at the University of Manchester in England, the University of Tokyo and the California Institute of Technology have demonstrated asynchronous microprocessors. Some asynchronous chips are already in commercial mass production. In the late 1990s Sharp, the Japanese electronics company, used asynchronous design to build a data-driven media processor—a chip for editing graphics, video and audio—and Philips Electronics produced an asynchronous microcontroller for two of its pagers. Asynchronous parts of otherwise synchronous systems are also beginning to appear; the UltraSPARC IIIi processor recently introduced by Sun includes some asynchronous circuits developed by our group. We believe that asynchronous systems will become ever more popular as researchers learn how to exploit their benefits and develop methods for simplifying their design. Asynchronous chip makers have achieved a good measure of technical success, but commercial success is still to come. We remain a long way from fulfilling the full promise of asynchrony.

Overview/*Clockless Systems*

- Most modern computers are synchronous: all their operations are coordinated by the timing signals of tiny crystal oscillators within the machines. Now researchers are designing asynchronous systems that can process data without the need for a governing clock.
- Asynchronous systems rely on local coordination circuits to ensure an orderly flow of data. The two most important coordination circuits are called the Rendezvous and the Arbiter.
- The potential benefits of asynchronous systems include faster speeds, lower power consumption and less radio interference. As integrated circuits become more complex, chip designers will need to learn asynchronous techniques.

Beat the Clock

WHAT ARE THE POTENTIAL BENEFITS of asynchronous systems? First, asynchrony may speed up computers. In a synchronous chip, the clock's rhythm must be slow enough to accommodate the slowest action in the chip's circuits. If it takes a billionth of a second for one circuit to complete its operation, the chip cannot run faster than one gigahertz. Even though many other circuits on that chip may be able to complete their operations in less time, these circuits must wait until the clock ticks again before proceeding to the next logical step. In contrast, each part of an asynchronous system takes as much or as little time for each action as it needs. Complex operations can take more time than average, and simple ones can take less. Actions can start as soon as the prerequisite actions are done, without waiting for the next tick of the clock. Thus, the system's speed depends on the average action time rather than the slowest action time.

Coordinating asynchronous actions, however, also takes time and chip area. If the efforts required for local coordination are small, an asynchronous system may, on average, be faster than a clocked system. Asynchrony offers the most help to irregular chip designs in which slow actions occur infrequently.

Asynchronous design may also reduce a chip's power consumption. In the current generation of large, fast synchronous chips, the circuits that deliver the timing signals take up a good chunk of the chip's area. In addition, as much as 30 percent of the electrical power used by the chip must be devoted to the clock and its distribution system. Moreover, because the clock is always running, it generates heat whether or not the chip is doing anything useful.

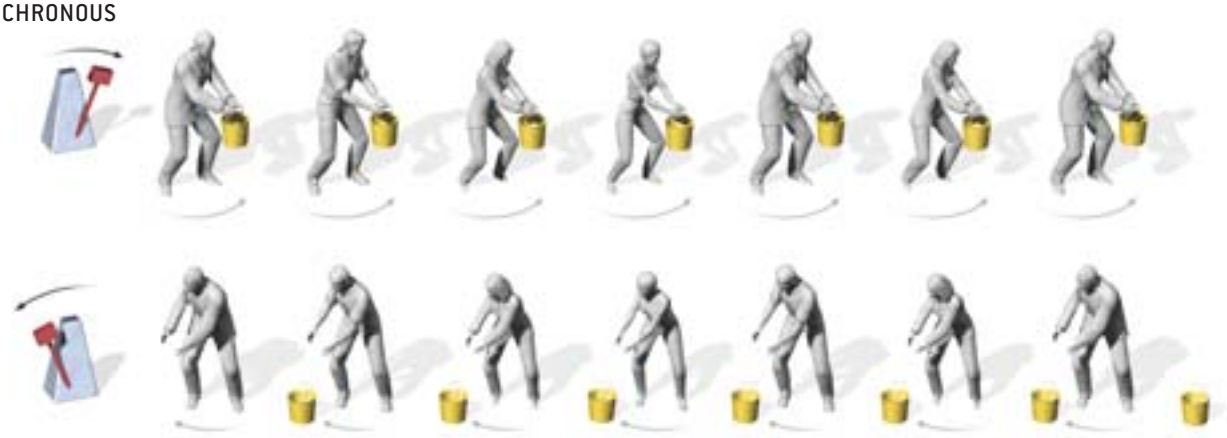
In asynchronous systems, idle parts of the chip consume negligible power. This feature is particularly valuable for battery-powered equipment, but it can also cut the cost of larger systems by reducing the need for cooling fans and air-conditioning to prevent them from overheating. The amount of power saved depends on the machine's pattern of activity. Systems with parts that act only occasionally benefit more than systems that act continuously. Most computers have components, such as the floating-point arithmetic unit, that often remain idle for long periods.

Furthermore, asynchronous systems produce less radio interference than synchronous machines do. Because a clocked system uses a fixed rhythm, it broadcasts a strong radio signal

BUCKET BRIGADES

THE METAPHOR of the bucket brigade can be used to describe the flow of data in a computer. A synchronous computer system is like a bucket brigade in which each person follows the ticktock rhythm of a clock. When the clock ticks, each person pushes a bucket forward to the next person down the line (*top*). When the clock tocks, each person grasps the bucket pushed forward by the preceding person (*middle*). An asynchronous system, in contrast, is like an ordinary bucket brigade: each person who holds a bucket can pass it down the line as soon as the next person's hands are free (*bottom*). —*I.E.S. and J.E.*

SYNCHRONOUS



ASYNCHRONOUS

at its operating frequency and at the harmonics of that frequency. Such signals can interfere with cellular phones, televisions and aircraft navigation systems that operate at the same frequencies. Asynchronous systems lack a fixed rhythm, so they spread their radiated energy broadly across the radio spectrum, emitting less at any one frequency.

Yet another benefit of asynchronous design is that it can be used to build bridges between clocked computers running at different speeds. Many computing clusters, for instance, link fast PCs with slower machines. These clusters can tackle complex problems by dividing the computational tasks among the PCs. Such a system is inherently asynchronous: different parts march to different beats. Moving data controlled by one clock to the control of another clock requires asynchronous bridges, because the data may be "out of sync" with the receiving clock.

Finally, although asynchronous design can be challenging, it can also be wonderfully flexible. Because the circuits of an asynchronous system need not share a common rhythm, designers have more freedom in choosing the system's parts and determining how they interact. Moreover, replacing any part with a faster version will improve the speed of the entire system. In contrast, increasing the speed of a clocked system usually requires upgrading every part.

Local Cooperation

TO DESCRIBE HOW ASYNCHRONOUS systems work, we often use the metaphor of the bucket brigade. A clocked system is like a bucket brigade in which each person must pass and receive buckets according to the ticktock rhythm of the clock. When the clock ticks, each person pushes a bucket forward to the next person down the line; when the clock tocks, each person grasps the bucket pushed forward by the preceding person [*see illustrations above*]. The rhythm of this brigade cannot go faster than the time it takes the slowest person to move the heaviest bucket. Even if most of the buckets are light, everyone in the line must wait for the clock to tick before passing the next bucket.

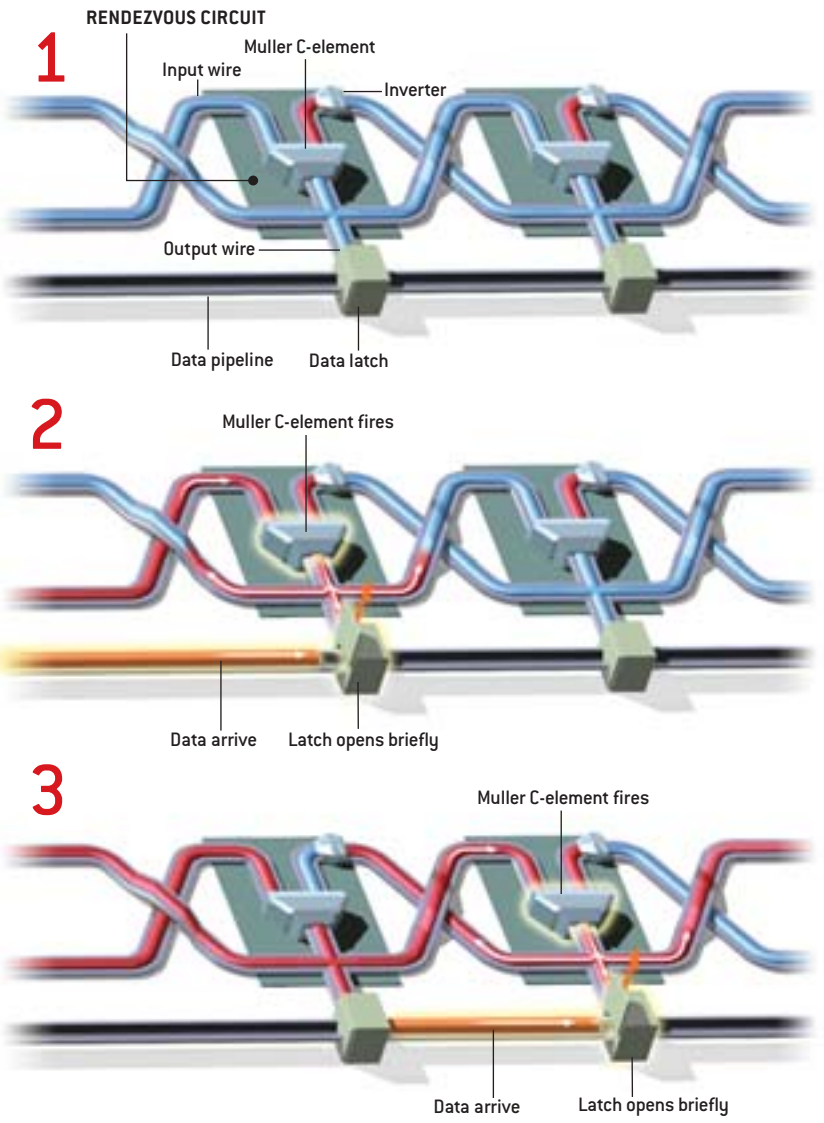
An asynchronous bucket brigade is governed by local cooperation rather than a common clock. Each person who holds a bucket can pass it to the next person down the line as soon as the next person's hands are free. Before each action, one person may have to wait until the other is ready. When most of the

BRYAN CHRISTIE DESIGN

HOW A RENDEZVOUS CIRCUIT WORKS

RENDEZVOUS CIRCUITS can coordinate the actions of an asynchronous system, allowing data to flow in an orderly fashion without the need for a central clock. Shown here is an electronic pipeline controlled by a chain of Muller C-elements, each of which allows data to pass down the line only when the preceding stage is "full"—indicating that data are ready to move—and the following stage is "empty."

Each Muller C-element has two input wires and one output wire. The output changes to FALSE when both inputs are FALSE and back to TRUE when both inputs are TRUE. (In the diagram, TRUE signals are shown in blue and FALSE signals in red.) The inverter makes the initial inputs to the Muller C-element differ, setting all stages empty at the start. Let's assume that the left input is initially TRUE and the right input FALSE (*1*). A change in signal at the left input from TRUE to FALSE (*2*) indicates that the stage to the left is full—that is, some data have arrived. Because the inputs to the Muller C-element are now the same, its output changes to FALSE. This change in signal does three things: it moves data down the pipeline by briefly making the data latch transparent, it sends a FALSE signal back to the preceding C-element to make the left stage empty, and it sends a FALSE signal ahead to the next C-element to make the right stage full (*3*). —*I.E.S. and J.E.*



buckets are light, however, they can move down the line very quickly. Moreover, when there's no water to move, everyone can rest between buckets. A slow person will still hinder the performance of the entire brigade, but replacing the slowpoke will return the system to its best speed.

Bucket brigades in computers are called pipelines. A common pipeline executes the computer's instructions. Such a pipeline has half a dozen or so stages, each of which acts like a person in a bucket brigade. Instead of handling buckets of water, however, each stage performs one action of the instruction's execution. For example, a processor executing the instruction "ADD A B C" must fetch the instruction from memory, decode the instruction, get the numbers from addresses A and B in memory, do the addition and store the sum in memory address C. A clocked pipeline executes these actions in a rhythm independent of the operations performed or the size of the numbers. Adding one and one may take just as much time as adding two 30-digit numbers. In an asynchronous pipeline, though, the duration of each action may depend on the operation performed, the size of the numbers and the locations of the data in memory (just as in a bucket brigade, the amount of water in a bucket may determine how long it takes to pass it on).

Without a clock to govern its actions, an asynchronous system must rely on local coordination circuits instead. These circuits exchange completion signals to ensure that the actions at each stage begin only when the circuits have the data they need. The two most important coordination circuits are called the Rendezvous and the Arbiter.

BRYAN CHRISTIE DESIGN

A Rendezvous element indicates when the last of two or more signals has arrived at a particular stage. Asynchronous systems use these elements to wait until all the concurrent actions finish before starting the next action. For instance, an arithmetic division circuit must have both the dividend (say, 16) and the divisor (say, 2) before it can divide one by the other (to reach the answer 8).

One form of Rendezvous circuit is called the Muller C-element, named after David Muller, now retired from a professorship at the University of Illinois. A Muller C-element is a logic circuit with two inputs and one output [*see box on opposite page*]. When both inputs of a Muller C-element are TRUE, its output becomes TRUE. When both inputs are FALSE, its output becomes FALSE. Otherwise the output remains unchanged. For the Muller C-element to act as a Rendezvous circuit, its inputs must not change again until its output responds. A chain of Muller C-elements can control the flow of data down an electronic bucket brigade.

Our research group recently introduced a new kind of Rendezvous circuit called GasP [*see box on next page*]. GasP evolved from an earlier family of circuits designed by Charles E. Molnar, our late colleague at Sun Microsystems. Molnar dubbed his creation asP*, which stands for asynchronous symmetric pulse protocol (the asterisk indicates the double "P"). We added the "G" to the name because GasP is what you are supposed to do when you see how fast our new circuits go. We have found that GasP modules are as fast and as energy-efficient as Muller C-elements, fit better with ordinary data latches and offer much greater versatility in complex designs.

Without a clock to govern its actions, an asynchronous system must rely on local coordination circuits instead.

Buridan's Ass

AN ARBITER CIRCUIT performs another task essential for asynchronous computers. An Arbiter is like a traffic officer at an intersection who decides which car may pass through next. Given only one request, an Arbiter promptly permits the corresponding action, delaying any second request until the first action is completed. When an Arbiter gets two requests at once, it must decide which request to grant first. For example, when two processors request access to a shared memory at approximately the same time, the Arbiter puts the requests into a sequence, granting access to only one processor at a time. The Arbiter guarantees that there are never two actions under way at once, just as the traffic officer prevents accidents by ensuring that there are never two cars passing through the intersection on a collision course.

Although Arbiter circuits never grant more than one request at a time, there is no way to build an Arbiter that will always reach a decision within a fixed time limit. Present-day Arbiters reach decisions very quickly on average, usually within about a few hundred picoseconds. (A picosecond is a trillionth of a second.) When faced with close calls, however, the circuits may occasionally take twice as long, and in very rare cases the time needed to make a decision may be 10 times as long as normal.

The fundamental difficulty in making these decisions is nicely illustrated by the parable of Buridan's ass. Attributed to Buridan, a 14th-century French philosopher, this parable suggests that an ass placed exactly between two equal piles of hay might starve to death because it would be unable to choose which pile to eat. Similar minor dilemmas are familiar in everyday life. For example, two people approaching a doorway at the same time may pause before deciding who will go through first. They can go through in either order, and Buridan's ass can eat from either pile of hay. In both cases, all that is needed is a way to break the tie.

An Arbiter breaks ties. Like a flip-flop circuit, an Arbiter has two stable states corresponding to the two choices. One can think of these states as the Pacific Ocean and the Gulf of Mex-

THE AUTHORS

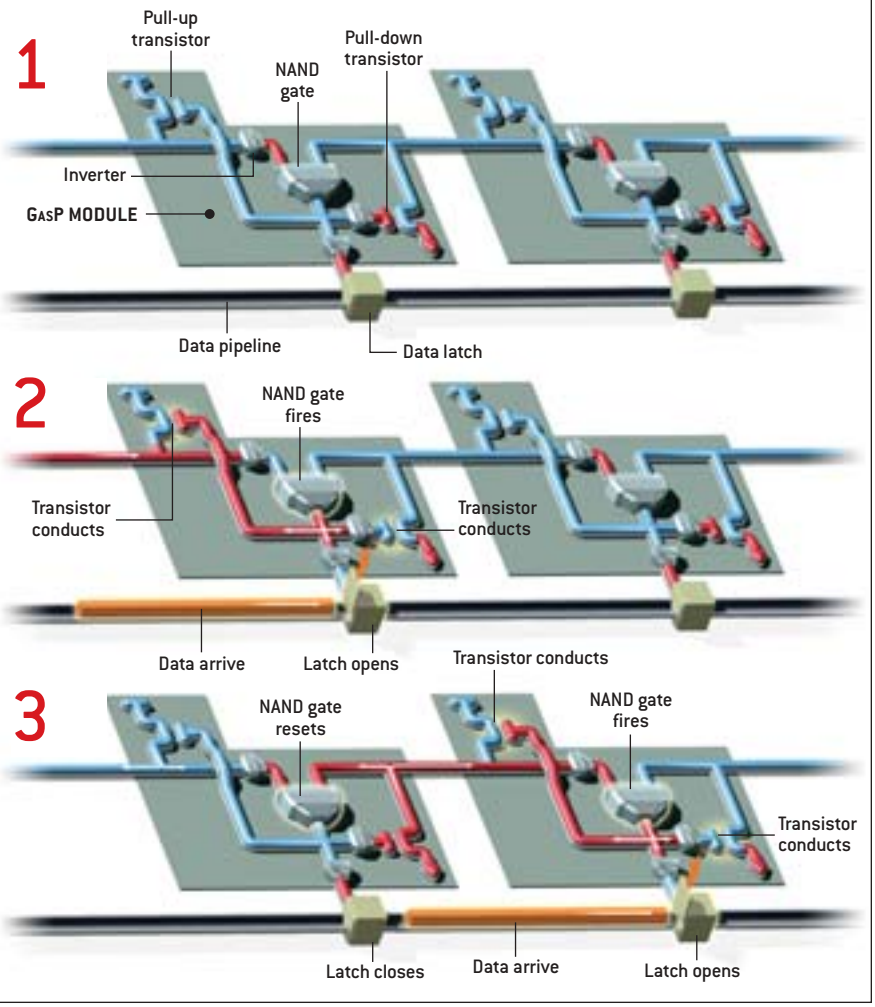
IVAN E. SUTHERLAND and *JO EBERGEN* are true believers in asynchronous computing. Although Sutherland (*middle left*) is best known as a pioneer of computer graphics—he invented the interactive graphics program Sketchpad in 1963—he became involved in asynchronous circuit design in the mid-1960s while building a graphics processor at Harvard University. He is now a vice president and fellow at Sun Microsystems, leading the Asynchronous Design Group at the company's laboratories. Ebergen (*middle right*) became fascinated by asynchronous circuit design 20 years ago during a three-month stint as a research assistant to Charles L. Seitz of Caltech. He subsequently taught at Eindhoven University of Technology in the Netherlands and the University of Waterloo in Canada before joining Sun's Asynchronous Design Group in the summer of 1996.

OLIVIER LAUDE

HOW A GASP MODULE WORKS

GASP MODULES can also act as Rendezvous elements for an asynchronous data pipeline. (The "P" in GasP is capitalized to acknowledge an earlier family of circuits.) Each GasP module has two wires connecting it to its neighbors and an output wire that drives a data latch. At the heart of the module is a NAND gate, which produces a FALSE output only when both inputs are TRUE. Otherwise the NAND produces a TRUE output.

The connection wires between modules represent the stages in the pipeline. At the start, all the signals in these wires are TRUE (*blue*), indicating that the stages are empty (*1*). The arrival of data in the pipeline (*2*) changes the incoming wire's signal to FALSE (*red*). An inverter flips the signal to TRUE and sends it to the NAND gate, which changes its output to FALSE. This makes the data latch transparent, allowing the data to move down the pipeline. The FALSE output also drives the incoming wire back to the TRUE state (which means empty) via a pull-up transistor and drives the outgoing wire to the FALSE state (which means full) via an inverter and a pull-down transistor. The NAND gate's output then returns to TRUE (*3*), making the latch opaque again. Meanwhile the FALSE signal in the outgoing wire triggers the same process in the next GasP module. —*I.E.S. and J.E.*



ico. Each request to an Arbiter pushes the circuit toward one stable state or the other, just as a hailstone that falls in the Rocky Mountains can roll downhill toward the Pacific or the Gulf. Between the two stable states, however, there must be a meta-stable line, which is equivalent to the Continental Divide. If a hailstone falls precisely on the divide, it may balance momentarily on that sharp mountain ridge before tipping toward the Pacific or the Gulf. Similarly, if two requests arrive at an Arbiter within a few picoseconds of each other, the circuit may pause in its meta-stable state before reaching one of its stable states to break the tie.

Novice Arbiter designers often seek to avoid even the occasional long delay by fashioning complicated circuits. A common error involves a circuit that notices the "hung" meta-stable state and pushes the Arbiter toward a particular decision. This is like training Buridan's ass to go left when it notices a hard choice. Such training, however, merely gives the ass three choices rather than two: go left, go right, or notice a hard choice and therefore go left. Even a trained ass will starve when unable to decide between the last two choices. Or, to use the geographic metaphor, you can move the Continental Divide with a shovel, but you cannot get rid of it. Although there is no way to eliminate meta-stability, simple, well-designed Arbiter circuits can ensure that virtually all delays are very brief. A typical contemporary Arbiter has a normal delay of 100 picoseconds and experiences a delay of 400 picoseconds less than once every 10 hours of operation. The probability of delays decreases exponentially with the length of the delay: an 800-picosecond pause occurs less than once every billion years of operation.

The Need for Speed

OUR GROUP AT SUN MICROSYSTEMS concentrates on designing fast asynchronous systems. We have found that speed often comes from simplicity. Our initial goal was to build a counterflow pipeline with two opposing data flows—like two parallel bucket brigades moving in opposite directions. We wanted the data from both flows to interact at each of the stages; the hard challenge was to ensure that every "northbound" data element would interact with every "southbound" data element. Arbitration turned out to be essential. At each

BRYAN CHRISTIE DESIGN

joint between successive stages, an Arbiter circuit permitted only one element at a time to pass. This project proved very useful as a research target; we learned a great deal about coordination and arbitration and built test chips to prove the reliability of our Arbiter circuits.

More recently, we have chosen a fresh research target, a processing structure we call FLEET. The name refers not only to speed but also to the collection of computing elements, each of which we call a ship. Each ship does its task concurrently with the others; the FLEET system controls their actions by moving data among them through an asynchronous switching network. Our work on FLEET has led to many discoveries. We wanted speed, and that led us to create the basic GasP circuits. We wanted to steer data items from one pipeline to another, like cars at highway interchanges, and that led us to design a larger family of GasP circuits that can act like a dense system of freeways. These circuits can move data about twice as fast as a clocked system could. To gauge the speed of our switching networks, we often build rings on our test chips around which the data elements rush like race cars. We measure the time it takes a data element to complete a round-trip, which is a lot easier than measuring the much shorter time it takes data to advance through one stage.

Our designs are beginning to enter Sun's computer products. The UltraSPARC IIIi processor chip contains asynchronous data queues that accept information from memory chips [*see illustration on page 63*]. This asynchronous system is the simplest and fastest way to compensate for the differences in arrival time of signals from memory chips that lie at different distances from the processor. Sun's product designers are gaining confidence that they can build asynchronous parts, that the parts work and can be tested, and that asynchrony offers important advantages over clocked design. As their confidence grows, more and more commercial products will use asynchronous parts for greater speed and flexibility, better power efficiency and reduced radio interference.

Sun is by no means the only company investigating asynchronous circuits. A group at Philips Research in the Netherlands has developed an asynchronous error corrector for a digital compact cassette player and an asynchronous version of a popular microcontroller for portable devices. The asynchronous microcontroller has been incorporated into pagers sold by Philips. The success of the Philips research group comes from three factors. First, the team learned how to create products rapidly using a programming language called Tangram that simplifies hardware design. Second, the low power consumption of their asynchronous circuits enables a pager to operate longer on a battery charge. Third, the reduced radio interference of their asynchronous circuits allowed the inclusion of both a computer and a sensitive radio receiver in a tiny device.

Furthermore, the experiments at Manchester, Caltech and Philips demonstrate that asynchronous microprocessors can be compatible with their clocked counterparts. The asynchronous processors can connect to peripheral machines without special programs or interface circuitry.

A Challenging Time

ALTHOUGH THE ARCHITECTURAL freedom of asynchronous systems is a great benefit, it also poses a difficult challenge. Because each part sets its own pace, that pace may vary from time to time in any one system and may vary from system to system. If several actions are concurrent, they may finish in a large number of possible sequences. Enumerating all the possible sequences of actions in a complex asynchronous chip is as difficult as predicting the sequences of actions in a schoolyard full of children. This dilemma is called the state explosion problem. Can chip designers create order out of the potential chaos of concurrent actions?

The technological trend is inevitable: in the coming decades, asynchronous design will become prevalent.

Fortunately, researchers are developing theories for tackling this problem. Designers need not worry about all the possible sequences of actions if they can set certain limitations on the communication behavior of each circuit. To continue the schoolyard metaphor, a teacher can promote safe play by teaching each child how to avoid danger.

Another difficulty is that we lack mature design tools, accepted testing methods and widespread education in asynchronous design. A growing research community is making good progress, but the present total investment in clock-free computing pales in comparison with the investment in clocked design. Nevertheless, we are confident that the relentless advances in the speed and complexity of integrated circuits will force designers to learn asynchronous techniques. We do not know yet whether asynchronous systems will flourish first within large computer and electronics companies or within start-up companies eager to develop new ideas. The technological trend, however, is inevitable: in the coming decades, asynchronous design will become prevalent. Eventually there will no longer be an easy answer to the question, How fast is your personal computer? SA

MORE TO EXPLORE

Micropipelines. Ivan E. Sutherland: The Turing Award Lecture. *Communications of the ACM,* Vol. 32, No. 6, pages 720–738; June 1989.

Asynchronous Circuits and Systems. Special issue of *Proceedings of the IEEE;* February 1999.

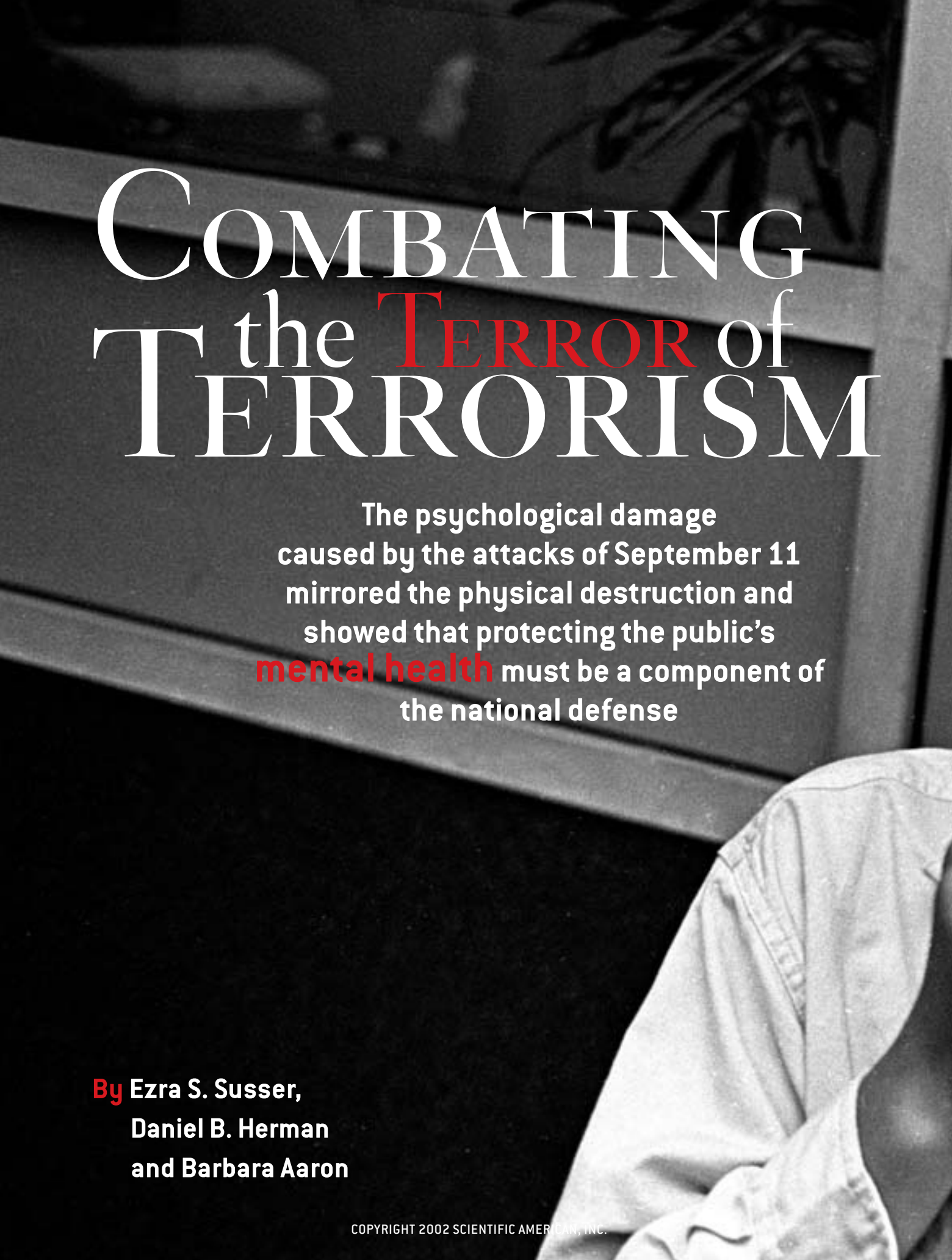
Principles of Asynchronous Circuit Design: A Systems Perspective. Edited by Jens Sparsø and Steve Furber. Kluwer Academic Publishers, 2001.

The Asynchronous Logic Home Page is at **www.cs.man.ac.uk/async/**

The Asynchronous Bibliography is at **www.win.tue.nl/~wsinap/async.html**

Combating the Terror of Terrorism

The psychological damage caused by the attacks of September 11 mirrored the physical destruction and showed that protecting the public's mental health must be a component of the national defense



By Ezra S. Susser, Daniel B. Herman and Barbara Aaron

On September 11, 2001,

the U.S. suffered the worst terrorist attacks in its history.

The destruction of the World Trade Center's twin towers was quickly followed by bioterrorism, in the form of a series of anthrax-tainted letters that killed five people and shut down the U.S. Capitol. The American response was swift, with war in Afghanistan and major appropriations for military defense. Another reaction was a welcome infusion of funding to bolster the chronically underfinanced national public health infrastructure.

Much of this spending will go toward reinvigorating infectious disease epidemiology, and this effort will improve our capacity both to detect new pathogens and to control infectious disease outbreaks. Preparedness not only defends the population against biological attacks but puts the medical community in a better position to react to natural outbreaks, such as HIV and West Nile virus. Another important response to terrorism, however, is attention to a vital medical component of national defense: the public's mental health.

"The purpose of these weapons is to wreak destruction via psychological means—by inducing fear, confusion, and uncertainty in everyday life," wrote Simon Wessely of Guy's, King's and St. Thomas's School of Medicine, London, and his colleagues in a *British Medical Journal* article on chemical and biological weapons. The same logic surely also informs physical attacks. In the terrorists' cold calculations, producing casualties is a secondary consideration to the more important goal: that the news of the horrific event gets widely disseminated and engenders a state of fear and anxiety throughout the population. An appropriate response, therefore, requires a determined effort to help the population withstand such attacks on the psyche. We must defend the intangible.

The World Trade Center, a soaring symbol of New York City and the nation, illustrates this point. The monumental structures' destruction, witnessed by millions—the burning towers were visible more than 20 miles away, with millions more watching on television (as well as the attack on another potent symbol, the Pentagon)—was clearly designed for maximum psychological effect. The devastation immediately generated a profound, widespread sense of vulnerability. Surprise was an additional element that magnified the psychological shock.

The anthrax perpetrator, still unknown as this article went to press, also carefully chose targets for maximum psychological influence: the contaminated letters went to newspapers, magazines, television stations and prominent members of the U.S. Congress. Though still unquantified, epidemic confusion and anxiety probably beset millions who wondered, "Could opening my mail kill me?" The public health system was overwhelmed with requests for antibiotics and nasal swab testing and with examinations of thousands of powder samples

PHOTOGRAPH BY ANTONIN KRATOCHVIL *VII (preceding pages)*; JOSEF CORSO *(above)*

PHOTOGRAPHS OF MISSING PEOPLE who had been in the World Trade Center were a common sight in the streets of Manhattan in the days after the attacks, constant reminders of loss and vulnerability.

across the country. And yet the vast majority of people experiencing anxiety probably had no appreciable risk of exposure to anthrax. Only 22 cases of illness were confirmed nationwide, with five unfortunate victims dying. Traffic accidents kill 115 people in the U.S. daily, but the anthrax incidents inflicted social disruption and psychological damage that traffic tragedies do not.

These examples make it clear that protecting mental health must be a central element in any terror defense. The first step in formulating a comprehensive strategy is to fully understand the problem. To that end, we considered the initial studies evaluating the psychological ramifications of the September 11 attacks, as well as other studies devoted to previous terrorism incidents and natural disasters that have traumatized large populations.

Assessing the Trauma

IN A SO-CALLED NEEDS ASSESSMENT commissioned by New York State, authored by two of us (Herman and Susser) with Chip J. Felton of the New York State Office of Mental Health, we set out to estimate the immediate psychological aftermath of the terrorist events. Our team also included colleagues from the New York City Department of Mental Health, the New York State Psychiatric Institute, the Nathan S. Kline Institute for Psychiatric Research and the New York Academy of Medicine. It is important to note that although we considered those who were directly exposed to the attacks, such as survivors of the twin towers and people who lost loved ones, the primary population we described was the general public of New York City.

For the needs assessment, we based our estimates of the psychological trauma of September 11 on three main sources. The first source was the existing literature on epidemiological disaster research, including the work of Fran H. Norris of Georgia State University. In March, Norris released a review in which she analyzed more than 200 articles, published between 1981 and 2001, concerned with the psychological consequences of 160 natural and purposeful disasters that affected 60,000 people worldwide. The second source was the research gauging the reaction of the general public in the area near the Oklahoma City bombing of 1995. The third source comprised two studies quickly conducted in New York City that examined the short-term psychological effects of the tower attacks. A study done by the New York City Department of Health in collaboration with the Centers for Disease Control and Prevention focused on some 400 residents of neighborhoods in close proximity to the World Trade Center. We also had prepublication access to a broader study that concentrated on all residents of Manhattan below 110th Street, approximately six miles north of the World Trade Center. This study, which appeared this past March in the *New England Journal of Medicine,* was led by Sandro Galea of the Center for Urban Epidemiologic Studies at the New York Academy of Medicine (who is completing his doctorate in our Columbia University department of epidemiology).

The Galea study analyzed telephone interviews with 1,008 Manhattan residents between October 16 and November 15, in which participants were questioned about their exposure to the events of September 11 and any psychological symptoms since that date. Galea and his colleagues found that "7.5 percent reported symptoms consistent with a diagnosis of current PTSD [post-traumatic stress disorder] related to the attacks, and 9.7 percent reported symptoms consistent with current depression." These percentages may be extrapolated to 67,000 Manhattanites with PTSD and 87,000 with depression. Proximity increased the reaction to the attacks: the rate of PTSD in those living close to the World Trade Center jumped to 20 percent.

Of course, some PTSD and depression existed prior to September 11. The Galea study was able to distinguish PTSD specifically related to the attacks, however, and revealed that baseline rates of PTSD roughly tripled in the entire study population during the weeks after the event.

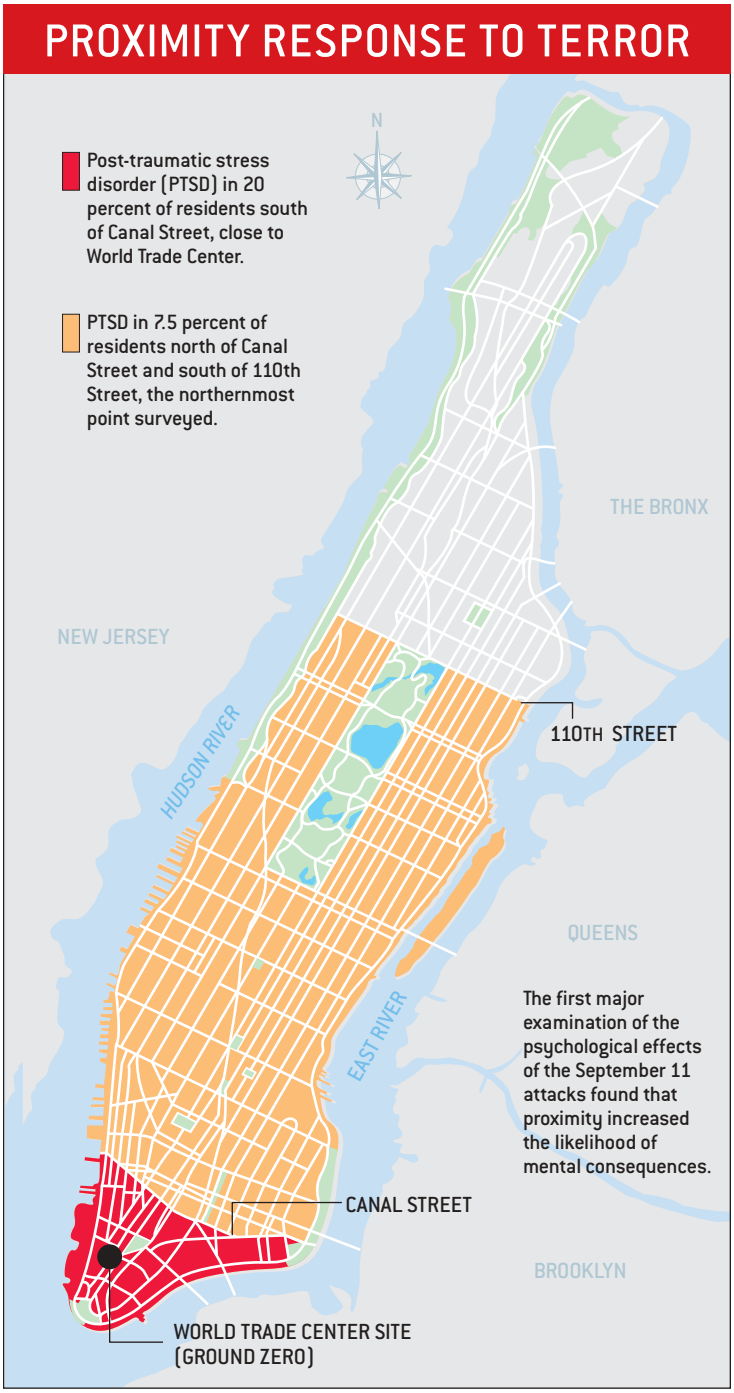
(Another sign of increased stress after September 11 was the tendency to self-medicate. A study released in June by the National Institutes of Health's National Institute on Drug Abuse showed that cigarette smoking and alcohol and marijuana use all increased in the weeks subsequent to the attacks.)

These figures are consistent with the data related to the worst previous terrorist attack in the U.S., the bombing of the Alfred P. Murrah Federal Building in Oklahoma City. Ginny Sprang of the University of Kentucky determined that 7.8 percent of 145 city residents who were not close to the building had PTSD. Carol S. North of the Washington University School of Medicine and her colleagues found PTSD in 34 percent of 182 survivors who had been in or near the building.

With the intial reports on New York City and the re-

THE AUTHORS

EZRA S. SUSSER, DANIEL B. HERMAN and *BARBARA AARON* are all at Columbia University's Mailman School of Public Health. Susser chairs the Department of Epidemiology at the Mailman School and is Head of Epidemiology of Brain Disorders at the New York State Psychiatric Institute. He has brought together disaster experts from around the country in a public-academic consortium to work on studies addressing the ongoing needs of the population most affected by the attack on the World Trade Center. Herman is the lead author of the post–September 11 New York State Mental Health Needs Assessment and is the lead investigator on a range of ongoing September 11 studies under the auspices of Susser and the Mailman School. Aaron is the administrator for World Trade Center studies at the Mailman School and has brought together a wide range of academic and governmental resources to address mental health needs in New York City.



maining literature on Oklahoma City and general disaster survival in hand, we prepared our evaluation of the psychological reaction to the attacks for the New York State needs assessment. That document contains individual estimates for the hardest-hit groups, such as surviving victims and the families of those killed, as well as rescue workers, Manhattan dwellers, other residents of the city, suburbanites and the remaining citizens of the state. The bottom line: even when making the most conservative estimates based on available data, we concluded that a minimum of approximately 422,000 New Yorkers experienced PTSD as a result of September 11.

The ongoing state assessment also drew on a separate study for the New York City Board of Education specifically examining city schoolchildren between grades four and 12. Christina Hoven of the Mailman School of Public Health of Columbia University designed the study, in which 8,266 students were surveyed about their reactions to the New York City attacks. The data, released in early May, indicated that 10.5 percent of the city's 710,000 public school students suffered from PTSD after September 11.

The Hoven study also detected substantial frequencies of other disorders, such as agoraphobia, the fear of open places. (The Galea study specifically reported on PTSD and clinical depression and not other psychological conditions, such as anxiety and subclinical depression.) Our needs assessment estimate of 422,000 cases of PTSD throughout the state is thus truly a minimum estimate of psychological trauma. Untold millions who witnessed the attacks through the media were surely shaken as well. In addition, the effects of terrorism on those already suffering from psychological conditions must be assumed to have been especially profound.

Bolstering Defenses

WITH THE SCOPE of the problem now clear, certainly the protection of the public's mental health must be a central element in any effective defense against terrorism. And yet public health leaders have, for the most part, failed to advocate strongly the integration of mental health considerations into the overall response to the terrorist threat. Medical schools, health insurance systems, disability legislation and other arenas also generally neglect the public mental health. The CDC, for example, has traditionally paid scant attention to public mental health, despite World Health Organization findings that depression is the fourth leading cause of disease and disability worldwide. And the WHO estimates that by 2020 depression will be the world's second leading cause of premature death and disability.

On the other hand, two encouraging signs of the recognition of mental illness as a public health concern came on May 21. First, the U.S. Preventive Services Task Force, sponsored by the Department of Health and Human Services's Agency for Healthcare Research and Quality, recommended that primary care physicians use a simple series of questions to screen adult patients for depression. Second, the Federal Emergency Management Agency announced a $132-million grant to continue funding Project Liberty, a program administered by the New York State Office of Mental Health that offers free counseling—but not treatment—to those most affected by September 11.

So what can be done to limit the propagation of fear, confusion and demoralization that leads to PTSD, depression and other conditions, especially in the face of a public mental health system that must still be considered

CLEO VILETT; SOURCE: SANDRO GALEA IN *NEW ENGLAND JOURNAL OF MEDICINE*, VOL. 346, NO. 13; MARCH 28, 2002

INJURIES TO LIFE AND LIMB were obvious in the aftermath of the attacks in Manhattan. Studies later revealed the extent to which New Yorkers also suffered from depression, post-traumatic stress disorder and other psychological conditions directly related to the destruction of the twin towers.

inadequate? Apart from the specific literature cited previously, relatively little research has examined the psychological effects of terrorism; even less is documented on how to protect people from these effects. A federal funding priority, therefore, should be to fully document the mental health consequences of September 11 and to devise and test strategies to minimize those consequences.

Until those studies can be done, however, the limits of the database should not limit our actions, which can be based on reasonable assumptions regarding the public interest and how best to serve it. For the needs assessment, we created guidelines aimed at protecting mental health in the first weeks after a disaster. A key aspect is the availability of trained mental health workers, to be deployed in a crisis. An example is a major initiative organized by the New York City Consortium for Effective Trauma Treatment. A panel of nationally recognized experts in current PTSD treatment techniques is training 60 clinical faculty members from numerous city mental health facilities. These faculty in turn will train local clinicians practicing privately or in community settings or employee assistance programs. The preparation emphasizes the need to reach out to school clinicians, primary health care providers and special education teachers.

Just as the military includes a large reserve force that can be called into action during crises, we also propose the creation of a mental health reserve corps made up of retired or part-time mental health professionals who would contribute their time and expertise on an emergency basis. Reserve corps members would diagnose individuals with clinical cases of mental disorders and offer them appropriate treatment, whether cognitive-behavioral therapy, pharmacotherapy, family therapy or a variety of other techniques shown to be effective. They could also take part in programs to disseminate information and to foster outreach to the public. Most people would not require professional assistance but would be helped by the understanding that their fear and sadness were normal reactions to a devastating event.

Marcelle Layton, a leader in bioterrorism preparedness with the New York City Department of Health, notes that social cohesion can be promoted by educating the community about potential threats and by informing the public as to the nature of the official response. Inspiring the populace can also be a great positive influence. Political leaders can play a significant role in caring for the public's mental health, because a sense of community and social cohesion fortifies people against terror's fundamental goal of inflicting psychological trauma. (In fact, a growing body of epidemiological literature suggests that social cohesion, or "social capital," confers *overall* protection against morbidity and mortality.)

A primary component of social cohesion and morale, probably deeply based in evolutionary psychology, is the leadership of a single, trusted authority figure. One of his-

NIKKI WEAVERS

PEOPLE CONGREGATING AT GROUND ZERO in October 2001 included small children. A study indicated that post-traumatic stress disorder afflicted more than 70,000 schoolchildren in New York City.

tory's foremost examples of the power of such social bonding dates to the 1940 Battle of Britain in World War II. Nazi bombings were designed to kill some but demoralize all. While concurrently attempting to fortify a weak air defense, Prime Minister Winston Churchill set himself the task of strengthening the resolve of his people to endure the psychological fallout of the air raids. His inspiring radio addresses, which promoted a sense of common purpose, in effect were public mental health interventions.

Likewise in New York City in the weeks after the attacks, Mayor Rudolph Giuliani well served a stricken populace, as he, too, took very visible command to personally keep the citizenry informed and to inspire with a sense of control and optimism. During the Scud missile attacks of the 1991 Gulf War, Israel also employed the paradigm of a single, familiar voice keeping the public informed with clear and consistent messages. In the midst of the attacks, most radio stations converged their broadcasts, with one senior official updating listeners.

Unless carefully delivered, however, statements by public officials and media representatives can arouse fear instead of alleviating it. Public spokespeople must be articulate and knowledgeable in their crucial roles, and the media should provide full and accurate information in a fashion that will not provoke distress and concern. For example, a comprehensive analysis of a new terrorist threat may be helpful; a 10-second promotional spot in which an anchor says, "New terrorist threat—more at 11" is not. And mixed messages in which government officials alarm us with detail-free "high alerts" but go on to advise us to travel normally and to "go shopping" are most likely counterproductive. (The U.S. Department of Justice's detentions and secretive interrogations of persons of Middle Eastern descent with no links to terrorism likewise disrupt social cohesion.)

In fact, asking the population to do something sacrificial and difficult rather than to consume conspicuously is probably a better way to increase social cohesion. Studies on fraternity hazing and military boot camps

JASON ESKENAZI

show that the shared hardships of members are responsible for much of the esprit de corps to be found in such groups [see "The Science of Persuasion," by Robert B. Cialdini; SCIENTIFIC AMERICAN, February 2001]. The scrap metal drives, war bond purchases and other contributions of noncombatants in World War II supported the war effort *and* the population's mental health. And although their principal motivation was to help others, blood donors who lined up at hospitals on the evening of September 11 were also helping themselves cope.

As this article was going to press, preliminary results were released based on additional research by Galea's group. For this study, 2,001 New Yorkers were interviewed by telephone between January 15 and February 21. Their reported symptoms indicated a discernible decrease in the number of study subjects for whom a clinical diagnosis of PTSD could be made. Many affected New Yorkers are clearly recovering naturally, a tribute to the resilience of the human psyche.

Yet also as this article was being prepared for publication, Vice President Dick Cheney and Secretary of Defense Donald Rumsfeld announced that additional terrorist attacks were virtually a certainty. It is therefore incumbent on the federal government to establish mental health teams and to call on them to devise rapid-response strategies. To reach the vast majority of the population, participants must go well beyond health institutions to schools, religious organizations, community groups, the military, and police, fire and emergency workers. We have begun to take steps to protect our lives and property. We must protect and defend our mental health as well. SA

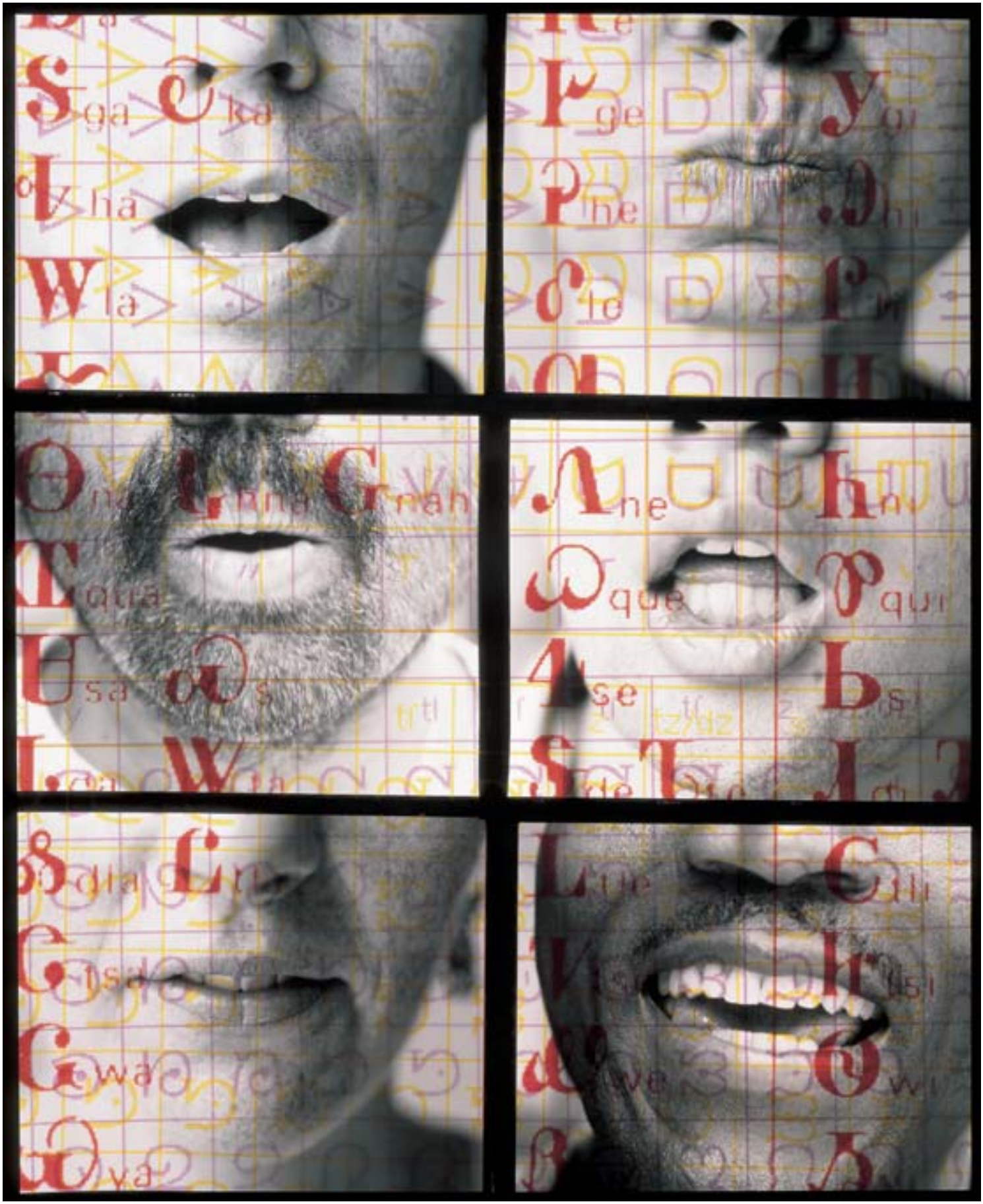
MORE TO EXPLORE

Psychological Impairment in the Wake of Disaster: The Disaster–Psychopathology Relationship. Anthony V. Rubonis and Leonard Bickman in *Psychological Bulletin,* Vol. 109, No. 3, pages 384–399; May 1991.

Psychological Sequelae of the September 11 Terrorist Attacks in New York City. Sandro Galea et al. in *New England Journal of Medicine,* Vol. 346, No. 13, pages 982–987; March 28, 2002.

The Range, Magnitude, and Duration of the Effects of Natural Disasters: A Review of the Empirical Literature. Fran H. Norris et al. A National Center for PTSD fact sheet. **www.ncptsd.org/facts/disasters/fs_range.html**

Crisis counseling is available from Project Liberty at **www.projectliberty.state.ny.us** or at 800-LIFENET (800-543-3638).



Obviously we must do some serious rethinking of our priorities, lest linguistics go down in history as the only science that presided obliviously over the disappearance of 90 percent of the very field to which it is dedicated.

—Michael Krauss, "The World's Languages in Crisis" (*Language,* 1992)

Linguists have known for years that thousands of the

SAVING

world's languages are at grave risk of extinction.

DYING

Yet only recently has the field summoned the will—

LANGUAGES

and the money—to do much about it

By W. Wayt Gibbs

Ten years ago Michael Krauss sent a shudder

through the discipline of linguistics with his prediction that half the 6,000 or so languages spoken in the world would cease to be uttered within a century. Krauss, a language professor at the University of Alaska–Fairbanks, had founded the Alaska Native Language Center to try to preserve as much as possible of the 20 tongues still known to the state's indigenous people. Only two of those languages were being taught to children. Several others existed only in the memories of a few aged speakers; the rest were rapidly falling from use. The situation in Alaska was emblematic of a global pattern, Krauss observed in the journal of the Linguistic Society of America. Unless scientists and community leaders directed a worldwide effort to stabilize the decline of local languages, he warned, nine tenths of the linguistic diversity of humankind would probably be doomed to extinction.

Krauss's prediction was little more than an educated guess, but other respected linguists had been clanging out similar alarms. Kenneth L. Hale of the Massachusetts Institute of Technology noted in the same journal issue that eight languages on which he had done fieldwork had since passed into extinction. A 1990 survey in Australia found that 70 of the 90 surviving Aboriginal languages were no longer used regularly by all age groups. The same was true for all but 20 of the 175 Native American languages spoken or remembered in the U.S., Krauss told a congressional panel in 1992.

On the face of it, the consolidation of human language might seem like a good trend, one that could ease ethnic tensions and aid global commerce. Linguists don't deny those benefits, and they acknowledge that in most cases small communities choose (often unconsciously) to switch to the majority language because they believe it will boost their social or economic status.

Many experts in the field nonetheless mourn the loss of rare languages, for several reasons. To start, there is scientific self-interest: some of the most basic questions in linguistics have to do with the limits of human speech, which are far from fully explored. Many researchers would like to know which structural elements of grammar and vocabulary—if any—are truly universal and probably therefore hardwired into the human brain. Other scientists try to reconstruct ancient migration patterns by comparing borrowed words that appear in otherwise unrelated languages. In each of these cases, the wider the portfolio of languages you study, the more likely you are to get the right answers.

"I think the value is mostly in human terms," says James A. Matisoff, a specialist in rare Asian languages at the University of California at Berkeley. "Language is the most important element in the culture of a community. When it dies, you lose the special knowledge of that culture and a unique window on the world."

In 1996 linguist Luisa Maffi helped to organize a group called Terralingua to draw attention to the apparent link be-

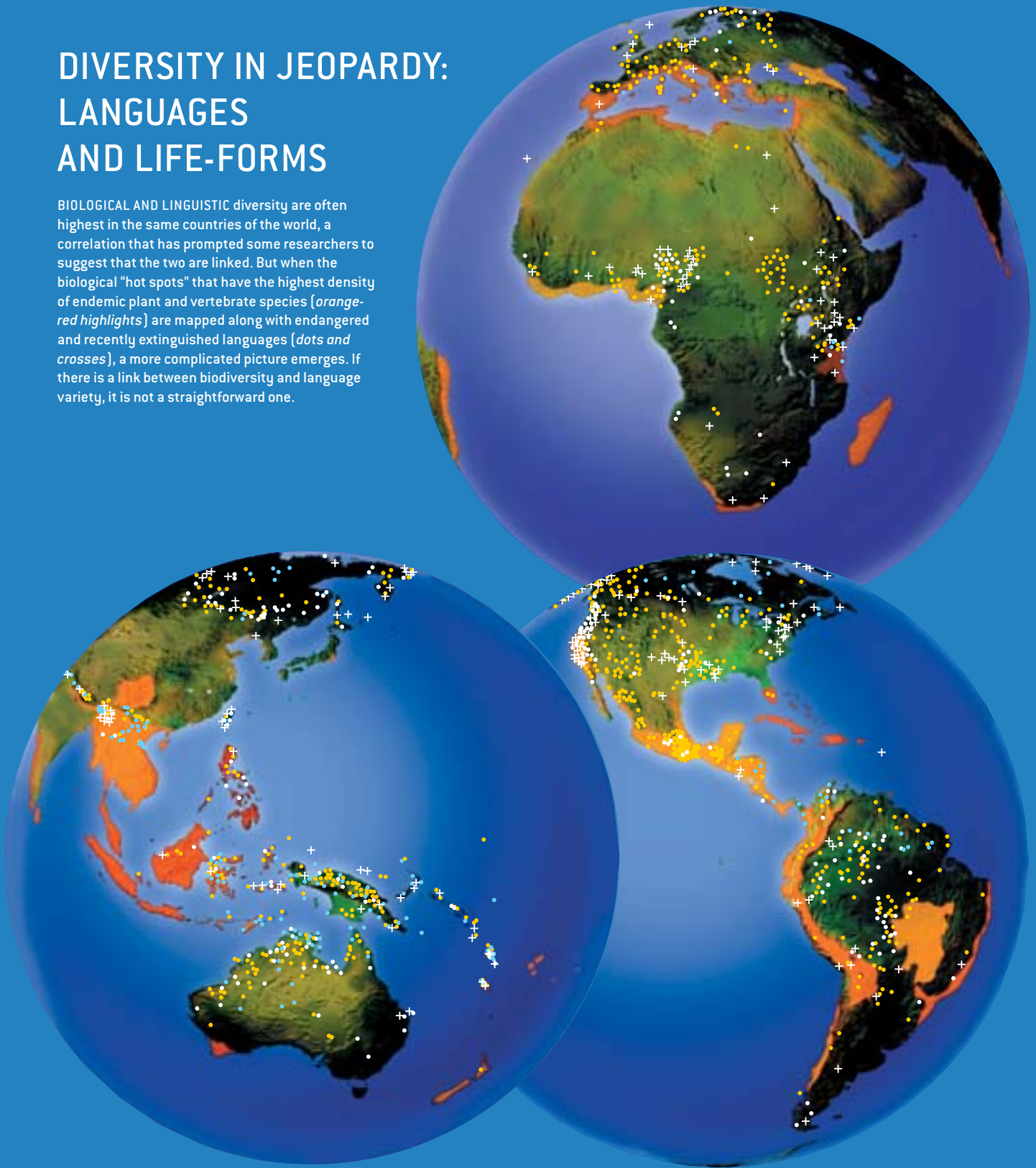
IRAIDA ICAZA (*preceding pages* and *above*)

Overview/*Endangered Languages*

- The latest edition of the *Ethnologue* lists 7,202 languages spoken worldwide, 440 of them within a generation or two of extinction. Allowing for some mislabeling of dialects, most linguists put the number of distinct languages between 5,000 and 7,000. Most also accept rough projections that without sustained conservation efforts, half or more of these will fall out of use by the end of the century.
- A small fraction of languages have been documented well enough to test theories of universal grammars, language evolution, and many other unanswered questions in linguistics and anthropology.
- Linguists have only recently begun to organize large-scale efforts to save dying languages. A new $30-million field research project set to begin early next year will increase the funding committed to such work by nearly 10-fold.

DIVERSITY IN JEOPARDY: LANGUAGES AND LIFE-FORMS

BIOLOGICAL AND LINGUISTIC diversity are often highest in the same countries of the world, a correlation that has prompted some researchers to suggest that the two are linked. But when the biological "hot spots" that have the highest density of endemic plant and vertebrate species (*orange-red highlights*) are mapped along with endangered and recently extinguished languages (*dots and crosses*), a more complicated picture emerges. If there is a link between biodiversity and language variety, it is not a straightforward one.



JOYCE PENDOLA; SOURCES: *ATLAS OF THE WORLD'S LANGUAGES IN DANGER OF DISAPPEARING*. SECOND EDITION. UNESCO PUBLISHING, 2001; "BIODIVERSITY HOTSPOTS FOR CONSERVATION PRIORITIES," BY NORMAN MYERS ET AL. IN *NATURE*, VOL. 403; PAGES 853–858; FEBRUARY 24, 2000

ETCHING A NEW ROSETTA STONE

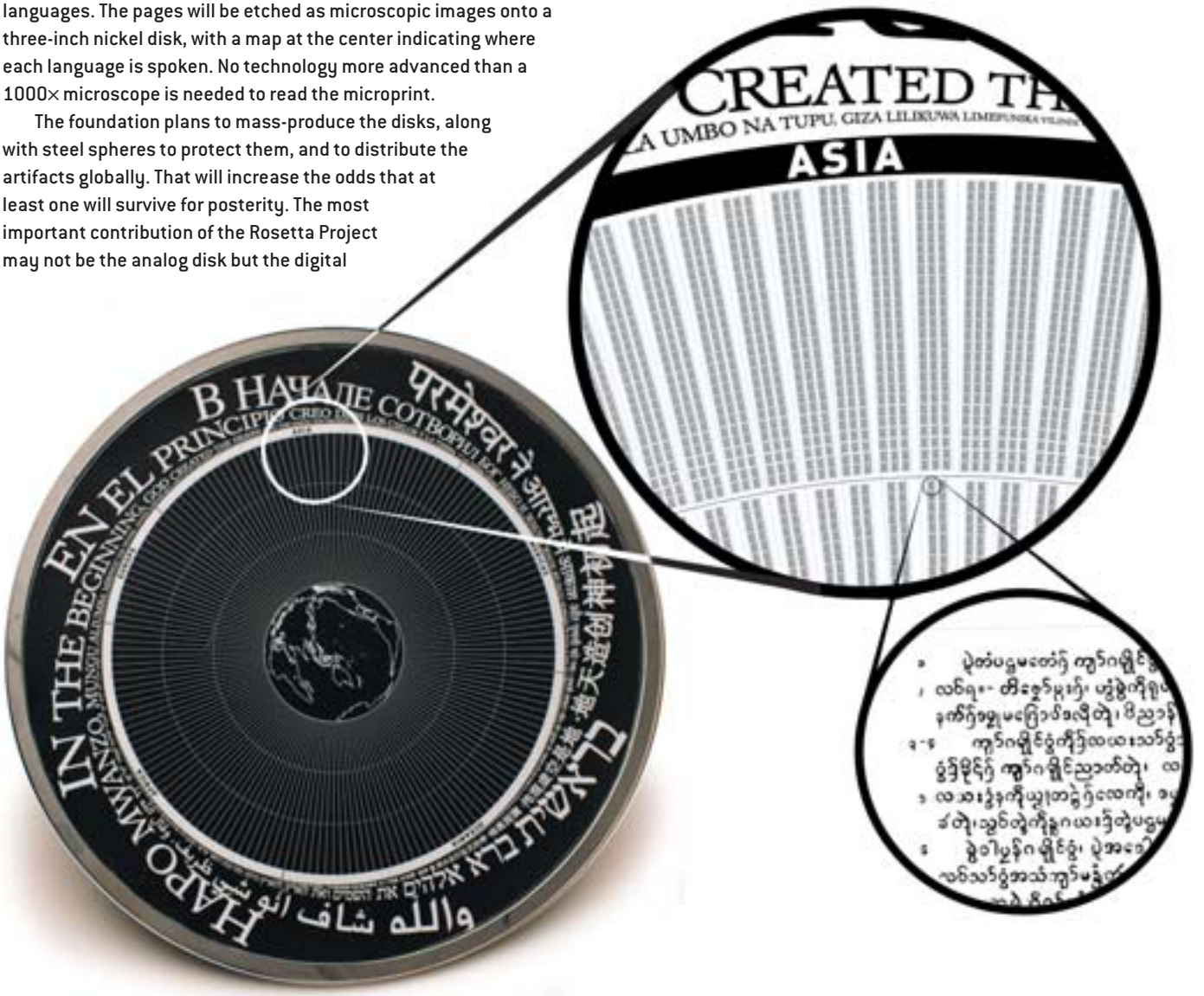
THE HIEROGLYPHIC LANGUAGE of ancient Egyptians was lost until Napoleon's troops discovered a 1,000-year-old slab of basalt in the Nile village of Rosetta. Etched into its black face were three copies of the same text: one in demotic, one in Greek and one in hieroglyphic Egyptian. With that key, scholars were able at last to unlock millennia of hidden history.

The Rosetta stone survived by chance, but it has inspired a small group of engineers and scientists to deliberately fashion a new artifact that could preserve some basic knowledge of the world's languages for anthropologists of the distant future. Jim Mason, who directs the Rosetta Project for the Long Now Foundation in San Francisco, says the group is on schedule to complete its first "stone" this autumn.

Like the original, this new Rosetta stone will carry parallel texts (the first chapter of Genesis), transliterated if the language has no native script. But its design allows it to hold much more detail—27 pages of glossed text and English description for each of 1,000 languages. The pages will be etched as microscopic images onto a three-inch nickel disk, with a map at the center indicating where each language is spoken. No technology more advanced than a 1000× microscope is needed to read the microprint.

The foundation plans to mass-produce the disks, along with steel spheres to protect them, and to distribute the artifacts globally. That will increase the odds that at least one will survive for posterity. The most important contribution of the Rosetta Project may not be the analog disk but the digital database of word lists for 4,000 to 5,000 languages that the group wants to complete next. "We already have word lists in digital form for 2,000 languages," Mason says. Scientists at the Santa Fe Institute, he adds, are keen to use the database to refine the picture of language evolution and human migration.

To fill in gaps in the database, the Rosetta team last year set up a collaborative Web site (rosettaproject.org) through which scholars and native speakers of rare languages can submit and peer-review word lists, audio recordings, grammars and other kinds of documentation. By June, 664 volunteers (25 to 30 percent of them professional linguists, Mason estimates) had contributed material. In principle, the last speakers of moribund languages could upload their knowledge for the benefit of future generations. In practice, unfortunately, last speakers are typically old, poor and computer illiterate. Few have e-mail addresses. —*W.W.G.*



ROLFE HORN (*photograph*); COURTESY OF THE LONG NOW FOUNDATION (www.longnow.org)

tween linguistic diversity and biodiversity, which seem to be highly concentrated in many of the same countries. Another international group drafted an ambitious "universal declaration of linguistic rights." The draft was submitted to UNESCO in 1996, but the organization has yet to act on the proposal.

An End to Apathy?

INDEED, DESPITE the near constant buzz in linguistics about endangered languages over the past 10 years, the field has accomplished depressingly little. "You would think that there would be some organized response to this dire situation," some attempt to determine which languages can be saved and which should be documented before they disappear, says Sarah G. Thomason, a linguist at the University of Michigan at Ann Arbor. "But there isn't any such effort organized in the profession. It is only recently that it has become fashionable enough to work on endangered languages."

Six years ago, recalls Douglas H. Whalen of Yale University, "when I asked linguists who was raising money to deal with these problems, I mostly got blank stares." So Whalen and a few other linguists founded the Endangered Languages Fund. But in the five years to 2001 they were able to collect only $80,000 for research grants. A similar foundation in England, directed by Nicholas Ostler, has raised just $8,000 since 1995. "I don't think the situation has changed in the seven years our foundation has existed," Ostler says. And no wonder. With so little research money available, says Steven Bird of the University of Pennsylvania, "anyone who wants to work on endangered languages has to forgo a more lucrative and secure career."

But there are encouraging signs that the field has turned a corner. The Volkswagen Foundation, a German charity, just issued its second round of grants totaling more than $2 million, Whalen says. It has created a multimedia archive at the Max Planck Institute for Psycholinguistics in the Netherlands that can house

A new British philanthropy has set aside $30 MILLION for a massive documentation project.

recordings, grammars, dictionaries and other data on endangered languages. To fill the archive, the foundation has dispatched field linguists to document Aweti (100 or so speakers in Brazil), Ega (about 300 speakers in Ivory Coast), Waima'a (a few hundred speakers in East Timor), and a dozen or so other languages unlikely to survive the century.

The Ford Foundation has also edged into the arena. Its contributions helped to reinvigorate a master-apprentice program created in 1992 by Leanne Hinton of Berkeley and Native Americans worried about the imminent demise of about 50 indigenous languages in California. Fluent speakers receive $3,000 to teach a younger relative (who is also paid) their native tongue through 360 hours of shared activities, spread over six months. So far about 75 teams have completed the program, Hinton says, transmitting at least some knowledge of 25 languages.

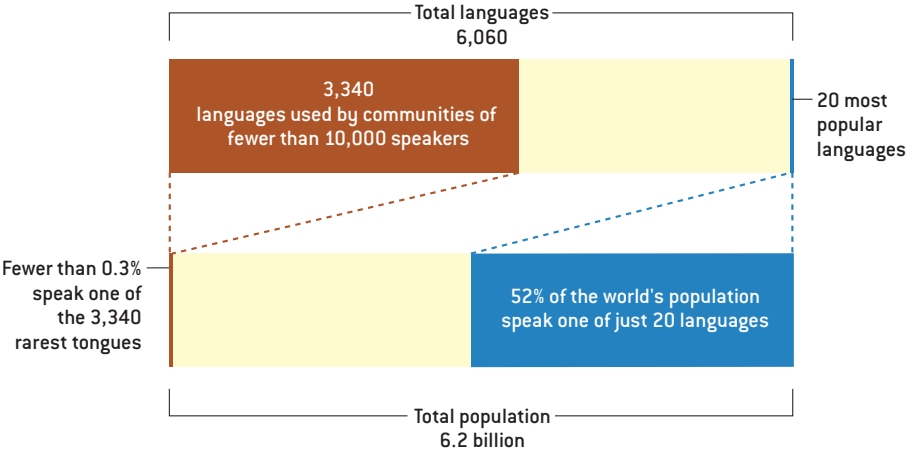
"It's too early to call this language revitalization," Hinton admits. "In California the death rate of elderly speakers will always be greater than the recruitment rate of young speakers. But at least we prolong the survival of the language." That will give linguists more time to record these tongues before they vanish.

But the master-apprentice approach hasn't caught on outside the U.S., and Hinton's effort is a drop in the sea. At least 440 languages have been reduced to a mere handful of elders, according to the *Ethnologue,* a catalogue of languages produced by the Dallas-based group SIL International that comes closest to global coverage. For the vast majority of these languages, there is little or no record of their grammar, vocabulary, pronunciation or use in daily life.

To help fill that need, the Lisbet Rausing Charitable Fund, a new British philanthropy, has set aside $30 million for a massive documentation project. Barry Supple, an adviser to the foundation, says the money will probably be doled out over the course of eight to 10 years. Part will be given to the School of Oriental and African Studies in London to train linguists specifically on field documentation of dying languages. But most of the money will go to fieldwork itself. By the time the program ends, Supple says, "we expect to document about 100 endangered languages."

A New Tower of Babel

THE RAUSING documentation project is an order of magnitude larger than any previous effort. A key test will be whether it collects the records on all these languages in a consistent way and stores them in a safe and accessible archive. "The archives we have are generally impoverished," says Bird, who is associate



FEWER THAN FOUR PERCENT of the world's people are responsible for maintaining about 96 percent of the world's languages.

JANA BRENNING; SOURCE: *LANGUAGE DEATH,* BY DAVID CRYSTAL. CAMBRIDGE UNIVERSITY PRESS, 2000

director of the Linguistic Data Consortium. "There is no archive that a university or national science foundation has committed to maintain indefinitely—say, for 25 or 50 years." He warns that languages may be recorded only to be lost again as the digital recording succumbs to obsolescence. "This is a huge issue," Whalen agrees.

Complicating matters further, dozens of institutions around the world are setting up digital libraries for data on endangered languages. This could create a tower of Babel of a new sort, because the projects plan to use inconsistent data formats, terminology and even names of languages.

Bird, Gary F. Simons of SIL International and many others have been working to bring some order to this chaos by building an "open language archives community" (OLAC) that uses metadata—a kind of digital card catalogue—to smooth out these inconsistencies. Launched in North America this past January and in Europe in May, OLAC encompasses more than 20 language repositories, including a number of those devoted to endangered languages. When the system begins operation next year, it will allow researchers to search a vast array of data to check out their theories about how languages evolved, about how the confluence of tongues reflects the migration of peoples, and about the limits of human speech.

Those are the main questions, after all, that linguists worry may become unanswerable with the loss of rare tongues. Linguistics is a young science still full of mysteries. Ostler offers one example: "Ica, spoken in northern Colombia, seems to have nothing comparable to a personal pronoun system—I, we, you, he, she, it and they. Otherwise I would have thought [that] a linguistic universal."

Bird's colleague Michael B. Maxwell is fascinated by reduplication: a feature of numerous languages in which a repetition signifies meaning, such as a plural (as if the plural of "cat" were "catcat"). Lushootseed, a nearly extinct language of the Puget Sound area, is almost unique in its use of reduplication in three different forms—as prefix, suffix and even as root—Maxwell says: "If languages like this die out, we'll never know the limits of how reduplication can work in real languages."

Or consider a different puzzle of plural variation. In several languages, such as English, most words are either singular or plural. But just a few, such as the (probably recently deceased) Australian Aboriginal language Ngan'gitjemerri, have four forms for each noun: singular, dual, trial (three of a kind) and plural. Sursurunga, Tangga and Marshallese have five forms. What's the limit? It may already be too late to know.

CULTURE OF INDIGENOUS PEOPLES, such as this G/wi tribe in Botswana, is transmitted through its native language. If the language is lost, much of the oral tradition—stories and fables, knowledge of the local environment, even a unique worldview—vanishes with it.

Better Alive Than Fossilized

EVEN IF A LANGUAGE has been fully documented, all that remains once it vanishes from active use is a fossil skeleton, a scattering of features that the scientist was lucky and astute enough to capture. Linguists may be able to sketch an outline of the forgotten language and fix its place on the evolutionary tree, but little more. "How did people start conversations and talk to babies? How did husbands and wives converse?" Hinton asks. "Those are the first things you want to learn when you want to revitalize the language."

But there is as yet no discipline of "conservation linguistics," as there is for biology. Almost every strategy tried so far has succeeded in some places but failed in others, and there seems to be no way to predict with certainty what will work where. Twenty years ago in New Zealand, Maori speakers set up "language nests," in which preschoolers were immersed in the native language. Additional Maori-only classes were added as the children progressed through elementary and secondary school. A similar approach was tried in Hawaii, with some success—the number of native speakers has stabilized at 1,000 or so, reports Joseph E. Grimes of SIL International, who is working on Oahu. Students can now get instruction in Hawaiian all the way through university. (They learn English as well.)

It is too early to tell whether this first generation of nest eggs will speak the native language to their children in the home. And immersion schools launched elsewhere have met with resistance from both within the community and without. Only one other indigenous language, Navajo, is taught this way in the U.S., according to the Center for Applied Linguistics. Leupp Public School on the Navajo reservation in Arizona started an immersion program after a survey there showed that only 7 percent of students could speak Navajo fluently. Children—initially kindergartners but now those up through fourth grade—use the language while raising sheep, tending gardens, performing traditional dances and otherwise learning about their culture. But the program has struggled to find qualified teachers, to obtain Navajo language textbooks and tests, and to garner sufficient community support.

Ofelia Zepeda of the University of Arizona, who is perhaps the most prominent Native American advocate for indigenous language revival in the U.S., describes similar troubles with her own language, Tohono O'odham. "Like every tribe in the country, our problem is that a whole generation of children are nonspeakers," she says. "The leadership supports language efforts, but the issue is funding. We've been waiting about three years to get our projects started." Even then, the small population of the tribe means that "we are essentially powerless in the grand scheme. Getting power over the schools in our own communities is a key necessity."

Just because a speech community is small does not mean that its language is doomed. At last report, notes Akira Yamamoto of the University of Kansas, there were just 185 people who spoke Karitiana. But they all lived in the same village in Brazil, which had just 191 inhabitants. So better than 96 percent of the population was still speaking the language and teaching it to their children. Because surveys of endangered languages tend to look only at the number of speakers, "there has been a history of linguists predicting the death of languages only to return 20 years later and find them still there," says Patrick McConvell of the Australian Institute of Aboriginal and Torres Strait Islander Studies in Canberra.

One factor that always seems to occur in the demise of a language, according to theorist Hans-Jürgen Sasse of the University of Cologne in Germany, is that the speakers begin to have "collective doubts about the usefulness of language loyalty." Once they start regarding their own language as inferior to the majority language, people stop using it for all situations. Kids pick up on the attitude and prefer the dominant language. "In many cases, people don't notice until they suddenly realize that their kids never speak the language, even at home," Whalen says. This is how Cornish and some dialects of Scottish Gaelic slipped into extinction. And it is why Irish Gaelic is still only rarely used for daily home life in Ireland, 80 years after the republic was founded with Irish as its first official language.

"Ultimately, the answer to the problem of language extinction is multilingualism," Matisoff argues, and many linguists agree. "Even uneducated people can learn several languages, as long as they start as children," he says. Indeed, most people in the world speak more than one tongue, and in places such as Cameroon (279 languages), Papua New Guinea (823) and India (387) it is common to speak three or four distinct languages and a dialect or two as well.

"Most Americans and Canadians, to the west of Quebec, have a gut reaction that anyone speaking another language in front of them is committing an immoral act," Grimes observes. "You get the same reaction in Australia and Russia. It is no coincidence that these are the areas where languages are disappearing the fastest." The first step in saving dying languages is to persuade the world's majorities to allow the minorities among them to speak with their own voices. SA

"Ultimately, the answer to the problem of language extinction is MULTILINGUALISM."
—James A. Matisoff, University of California, Berkeley

W. Wayt Gibbs is senior writer.

MORE TO EXPLORE

The Green Book of Language Revitalization in Practice. Edited by Leanne Hinton and Kenneth Hale. Academic Press, 2001.

On Biocultural Diversity. Edited by Luisa Maffi. Smithsonian Institution Press, 2001.

Ethnologue: **www.ethnologue.com**

Teaching Indigenous Languages: **http://jan.ucc.nau.edu/~jar/TIL.html**

PETER JOHNSON *Corbis*

SMART CARDS

Safety at a Cost

Few Americans have handled a "smart" card, but many Europeans already use them. Will the cards ever make it in the States? Experts now say yes.

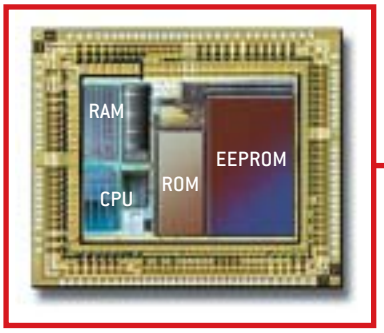
Smart cards look like credit cards but have a microprocessor and memory chip inside. In the 1990s they became popular in Europe because credit-card fraud was high and inexpensive phone service for reliable card authorization was not widespread. Smart cards can store validation and identification information that a clerk can check with a basic card reader.

In the U.S., widespread use of magnetic-stripe credit cards, bank cards and identification cards left consumers with little need for smart cards. But terrorism and identity theft in e-commerce have increased anxiety about security. A magnetic stripe can store only a simple security code, which can be read—and reprogrammed—by anyone with a $25 card reader. A smart card's microprocessor generates complex passwords that are difficult to decode and biometric markers such as fingerprints and retina patterns that are hard to forge. Also, inserting a smart card into a reader linked to a desktop computer can prevent unauthorized people from using the machine and allow the valid owner to encrypt e-mail.

Orders for smart cards in the U.S. have jumped 45 percent in the past 12 months, says Randy Vanderhoof, president of the Smart Card Alliance. Leading the way are credit cards that provide more secure transactions, fare cards for city subway and bus rides, and biometric ID cards that control access to dormitories, offices and laboratories. By the end of 2003 the Department of Defense will have issued smart-card ID badges to four million personnel. The Department of Transportation is developing standardized smart-card credentials for 15 million federal, state, airport and seaport workers. Congress is considering the Driver's License Modernization Act, which would require all states to issue smart-card licenses that improve identification and fraud detection.

The potential impediment? Institutions and retailers could track the movements and purchases of employees, students and consumers. Is tighter security worth the loss of privacy? The debate may grow as fast as smart-card usage. —*Mark Fischetti*

SMART CARD
contains an integrated circuit with a processor (CPU), dynamic memory (RAM), permanent memory (ROM) and programmable memory (EEPROM). The IC is embedded with epoxy. Wires connect it to electrical contacts that brush against contacts in a card reader or to an antenna that communicates with the reader via radio waves, or both. The reader powers the card.



ILLUSTRATIONS BY ALAN DANIELS; SOURCE: GEMPLUS (*IC and cross section*)

DID YOU KNOW . . .

➤ KILLER COUPON: Americans haven't warmed to smart cards in part because there hasn't been a "killer application." Marketers at the Target chain are issuing 2.5 million smart Visa cards to customers and offering them free card readers. Their hope? That you will connect the reader to your computer, access Target's Web site and download discounts into your card to redeem at your local store. Target executives claim that no one clips coupons anymore. Furthermore, they say, consumers come out ahead because the company can tailor promotions to individual buying preferences.

➤ COURT CARD: High school student Omar Lovell was arrested in 1998 for allegedly fare-jumping a New York City subway turnstile. In court he presented his student MetroCard record, which showed that he had paid. In 2000 prosecutors used MetroCard records to dispute the alibi of Marco Valencia in the December 1999 beating and robbery of a Manhattan store manager. Valencia later pleaded guilty.

➤ POCKET ROCKET: Gemplus and SchlumbergerSema are the world's leading smart-card providers. SSP Litronic in Reston, Va., makes smart cards in conjunction with the U.S. National Security Agency. Its latest, dubbed the Forté, features a 32-bit microprocessor instead of the ubiquitous eight-bit model; the Forté can perform high-level encryption and fast, massive data transfers to card readers. Litronic predicts that the supercard will trickle down into consumer applications.

EMBOSSED CARD
holds only 16 digits of information. In one scheme, digit 1 identifies the card issuer, 2 through 7 is your bank ID number, 8 through 15 is your account, and 16 is a verification number.

MAGNETIC-STRIPE CARD
has three parallel storage tracks. In one scheme, they hold roughly 80 six-bit characters, 40 four-bit characters, and 100 four-bit characters, respectively.

What's in Your Wallet

Card Type	*Defining Feature*	*Example*
Embossed	Raised characters only, on front	Old credit card
Magnetic stripe	Magnetic bar on back, characters on front	Bank card
Memory	Electronic memory inside	Prepaid phone card
Smart	Electronic memory and processor inside	Biometric ID card
Contact	Electrical contacts exposed on surface	
Contactless	Radio antenna embedded inside	

This month's topic was suggested by reader Jason Jansky. Have an idea for a future column? Send it to workingknowledge@sciam.com

Machine Chic

THE POMA WEARABLE COMPUTER IS FLASHY BUT NOT VERY FUNCTIONAL **BY MARK ALPERT**

It isn't easy to stand out in a place like New York City, where outrageously dressed people are as common as pigeons. Even if you wear a nose ring and dye your hair purple, most of the locals won't give you a second glance. But I recently devised a plan to rise above my anonymous status and become the most sensational person in Manhattan, attracting stares of wonder and bewilderment every time I walked down the street. The secret to my new celebrity would be a device called the Poma, a computer that looks like a futuristic fashion accessory.

Unlike the now pedestrian PDA (for personal digital assistant)—the Palm, the Visor and so on—the Poma is a portable device meant to be worn, not held. A shiny silver band clamped to the forehead suspends an inch-wide computer screen in front of one of the user's eyes, like a high-tech monocle. A black wire connects the screen to the computer itself, an 11-ounce, six-inch-high unit that can be carried in a coat pocket or clipped to a belt.

Another wire links the computer to an optical mouse that fits in the user's hand; just rub your thumb across the beam of light, and the cursor moves across the screen. In theory, at least, you could write stories, play games, watch videos or surf the Internet while casually strolling through the park. You'd see the real world with one eye and the virtual world with the other.

The concept of a wearable computer is not exactly new. Researchers at the Massachusetts Institute of Technology, Carnegie Mellon and other universities have been developing prototypes of such devices for years [see "Wearable Intelligence," by Alex P. Pentland; SCIENTIFIC AMERICAN PRESENTS: EXPLORING INTELLIGENCE; Winter 1998]. In the late 1990s Xybernaut, a company based in Fairfax, Va., developed a line of wearable computers called Mobile Assistant that was targeted at business users—for example, maintenance crews and utility workers who need easy access to information while working with their hands. The Poma is Xybernaut's first wearable device aimed at the consumer market. As soon as I saw the pictures on the company's Web site—showing a beautiful young woman and a clean-cut executive posing with the computer—I knew I had to try it.

FIRST WEARABLE COMPUTER for the masses, the Poma packs a lot of power into a small package. The six-inch-high computer (*center*) contains a 128-megahertz RISC processor, 32 megabytes of random-access memory and an equal amount of read-only memory. The head-mounted screen (*top*) weighs less than three ounces. The optical mouse (*lower left*) allows you to move the cursor with your thumb.

Strapping on the Poma for the first time is definitely fun. Feeling like a cyborg getting dressed for work, I tightened the silver band around my head and spent a few minutes adjusting the tiny computer screen until it was properly positioned in front of my right eye. (If you happen to favor your left eye, you can place the screen there.) The device contains a mirror that reflects a liquid-crystal display into the user's line of sight; because the image is projected so close to the eye, the inch-wide screen appears to be the same size as an ordinary desktop screen viewed from about two feet away. The Poma's operating system is Windows CE, the modified version of Windows used in many PDAs, so the screen icons and menus look comfortingly famil-

BETH PHILLIPS

iar. The system allows you to write text by clicking the mouse on the letters of a "software keyboard" that pops up on the screen.

My disappointment began when I tried to do something useful with the device. I was able to create text files using Microsoft Pocket Word, but I found that writing with the software keyboard is maddeningly laborious. I couldn't deftly maneuver the mouse with my thumb, and I had to put the letters in a gigantic font to make them legible. Even so, I still had hopes for the Poma; although I couldn't use it to write a novel while walking to work, perhaps I could dash off brief e-mail messages or check out my favorite Web sites. The Poma can access the Internet wirelessly via the popular Wi-Fi standard (also known as 802.11b). If you slip a wireless LAN card into the device, you can connect with the Wi-Fi networks now available in many airports, hotels and coffee shops. But this feature is of dubious value, because in most of these places you're more likely to be sitting with the device than walking. And if you're sitting, it would be a lot easier to surf the Internet with a PDA or a laptop than with the Poma.

I was amazed to see that the Poma contained no games—not even solitaire, which would be a nice distraction for a bored commuter wearing the device. The people at Xybernaut said I could download games to the Poma using a serial input/output card hooked up to a desktop PC. But this was easier said than done. When I tried to connect the Poma to a PC, I was stymied by a series of error messages. After several hours of effort and numerous calls to the technical folks at Xybernaut, I ran out of patience. This failure also made it impossible for me to evaluate the Poma's presentation of video and audio clips, because I couldn't download any multimedia files.

Nevertheless, I was determined to walk through the streets of Manhattan with my Poma, even if I couldn't do much

SCIENTIFIC AMERICAN Help Desk

For your convenience, we have summarized the answers to often-asked customer service questions and how to get in touch with us.

Scientific American Frontiers
The magazine's PBS television series, Scientific American Frontiers, is hosted by actor and life long science buff, Alan Alda. There are now 10 one-hour episodes each season , with Frontiers becoming an integral part of a new PBS science programming initiative. Visit the Scientific American Frontiers Web site at www.pbs.org/saf

www.sciam.com
Our award-winning Internet resource, updated weekly and thoroughly linked to related sites. Here you will find timely and interesting current news, recent illustrated scientific articles, current issue highlights, ask the editors, our on-line store, subscriptions, e-mail, other features not found in Scientific American, and much more.

Scientific American subscription rates
In the U.S., Scientific American is $34.97. In Canada, $49. Elsewhere, $55. Your 12-issue subscription includes the annual September single-topic issue.

Subscription inquiries
Call us: In the U.S. and Canada: (800) 333-1199. In other countries: (515) 247-7631. E-mail: subscriptions@sciam.com. Or write: Subscription Manager; Scientific American; PO Box 3187; Harlan, IA 51537. The date of the last issue of your subscription appears on the mailing label. For change of address, please notify us at least eight weeks in advance, and include both old and new addresses.

Back issues
In U.S. $9.95 each. Elsewhere $12.95 each. Many issues are available. Fax your order with your Visa, MasterCard or AMEX information to: (212) 355-0408.

Reprints
$4 each, minimum order 10 copies, prepaid. Limited availability.
Write to: Reprint Department; Scientific American; 415 Madison Avenue; New York, NY 10017-1111; fax: (212) 355-0408 or tel. (212) 451-8877.

Photocopying rights
Granted by Scientific American, Inc., to libraries and others registered with the Copyright Clearance Center (CCC) to photocopy articles in this issue for the fee of $3.50 per copy of each article plus $0.50 per page. Such clearance does not extend to photocopying articles for promotion or other commercial purpose.

Subscriber alert!
Scientific American has been made aware that some subscribers have received notifications/subscription offers from companies such as Publishers Services Exchange, United Publishers Network and Publishers Access Service. These are not authorized representatives/ agents of Scientific American and are selling subscriptions at a much higher rate than the regular subscription or renewal price. Please forward any correspondence you may receive from these companies to: L. Terlecki, Scientific American, 415 Madison Ave., NY, NY 10017.

more than tap out a few words on the screen. When the device is turned on, the computer worn on the belt and the handheld mouse emit a neon-blue light, which nicely complements the wearer's otherworldly appearance. For one brief, glorious moment, I would be the strangest-looking person in New York!

But even this goal was frustrated. As I walked down Madison Avenue, trying very hard to keep a straight face, many of the passersby did double takes and gaped at me. But many others didn't even notice the thing, and quite a few jaded individuals took one look and turned away, unimpressed. What's more, this street test revealed that the Poma has several practical limitations. As every New Yorker knows, weaving through the crowds on the sidewalk requires alacrity as well as agility. Although the Poma didn't completely block the view from my right eye—the computer screen is semitransparent, allowing a dim vista of the outside world to filter through—I quickly discovered that I risked serious injury if I focused too much attention on the screen. Only at the street corners, while waiting for the light to change, was I able to input a few pithy observations, such as "now crossing the street."

TOO WEIRD to wear? The author discovered that the Poma didn't faze most New Yorkers.

I also learned that the Poma works best on overcast days. Whenever the sun broke through the clouds, the images on the computer screen paled and the text became unreadable. I felt as if I were waking from a dream as my virtual world thinned into nothingness. For the rest of the walk, I tried to stay in the shadows. But after I'd traveled about 10 blocks, I noticed a more intractable problem. A sharp pain was spreading across my forehead. My ocular muscles were feeling the strain of keeping my right eye focused on the screen. The pain became so intense that I finally had to rip the Poma off my head. For the next half a minute, I lurched dizzily down the sidewalk as my eyes came back into alignment.

Perhaps I was wearing the device incorrectly. Or perhaps the trouble was my nearsighted, astigmatic vision. In any event, I lost enthusiasm for the Poma—it had literally nauseated me. I was also put off by its eye-popping price tag of $1,499. When I asked Xybernaut how many Pomas had been sold so far, I expected a modest number, maybe several hundred. But the answer was depressingly low: just over 20.

Still, I haven't given up on wearable computers. Voice-recognition software may be the key to improving the devices; it would be a lot easier to compose letters and sonnets on the fly if one could dictate to the machine. Until manufacturers work out the kinks, however, I'm not going to make any more high-tech additions to my wardrobe. SA

ASTRID BROWN

Riffs on General Relativity

BLACK HOLES, TIME TRAVEL, RIPPLES IN SPACETIME, THE BIG BANG **BY CHET RAYMO**

THE FUTURE OF SPACETIME
by Stephen W. Hawking, Kip S. Thorne, Igor Novikov, Timothy Ferris and Alan Lightman. Introduction by Richard Price
W. W. Norton, 2002
($25.95)

Put Stephen Hawking, Kip Thorne, Igor Novikov, Timothy Ferris and Alan Lightman in a room together, and I would imagine that the intellectual sparks would fly lively and thick. The five essays collected in this book are adapted from those sparks, talks given at the California Institute of Technology in June 2000 to honor the 60th birthday of physicist Kip Thorne.

If there is a unifying theme to the essays, it is the possibility of time travel, one of Thorne's obsessions as a theoretician of general relativity and, of course, a topic of perennial popular interest. None of the authors was paid for his contribution, and royalties will go to a Caltech scholarship fund in Thorne's name.

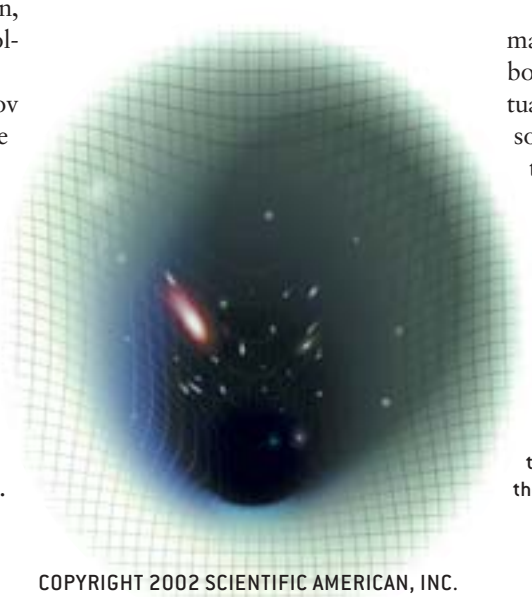
Theoretical physicist Igor Novikov starts by asking, "Can we change the past?" He shows how curious foldings and warpings of spacetime apparently allow the possibility of traveling back in time and considers the so-called grandfather paradox: What if I travel back in time and kill my grandfather? Then, logically, I would never have been born to make my journey into the past. Novikov argues that the laws of nature would prevent such logical paradoxes from happening. Stephen Hawking is perhaps the world's most famous theorist of spacetime. He is less sanguine than Novikov that time travel is possible, except on the scale of individual atomic particles, which is not of much use for science-fiction fantasies. If Hawking's take on the physics is correct, grandfather is doubly safe. Thorne uses his commanding presence at the heart of the book to address the question implicit in the title: How will our understanding of spacetime evolve in the near future, theoretically and experimentally? The final two essays, by writers Timothy Ferris and Alan Lightman, though excellent in themselves, have nothing directly to do with the topic at hand. Ferris considers how science is communicated to the general public, and Lightman muses on relations between science and art.

It all adds up to less than the sum of its parts. The word "hodgepodge" comes to mind, and the fact that the editors decided the book needed a long preparatory introduction (longer than all but one of the five contributions) and a puffed-up glossary suggests that the problems were apparent from the beginning. Anyone who wants the skinny on time travel and the future of spacetime would do well to go directly to Thorne's excellent popular book *Black Holes and Time Warps: Einstein's Outrageous Legacy* (W. W. Norton, 1994).

Still, there is a terrific story lurking among the disparate parts of the present volume, but readers will have to dig it out for themselves. I would suggest skipping the introduction and going straight to Lightman's piece on science and art. He was trained as a physicist and has transformed himself into a successful novelist, so he knows both sides of which he speaks. He takes us to the heart of the creative process and shows us what physicists and novelists have in common. For one thing, they both make up stories, and they both want their stories to be true.

Ferris tells us why scientific story making is essential to a healthy and free body politic. "Technologically, intellectually, and even politically, science resides somewhere near the center of our culture, by which I mean the society of all those persons who value their freedom, honor their responsibilities, appreciate their ignorance, and are willing to keep learning," he writes.

Now go to the essays by Novikov and Hawking and watch two outra-



GRAVITY of a huge black hole warps the fabric of spacetime. From the cover of the July 1996 issue of *Scientific American.*

ILLUSTRATION BY SLIM FILMS

geously clever minds at play in the fields of knowledge and ignorance. They take Einstein's supreme story—his theory of gravity and spacetime, called general relativity—and make delightful riffs on the theme. What if? they ask. They agree on this: even if it turns out that time travel is impossible, it is important that we understand *why* it is impossible. Finally, turn to Thorne's central essay, where it all comes together. We have in Einstein's legacy a fabulously inventive story: black holes, time travel, ripples in spacetime, the big bang—stuff any novelist would have been proud to invent. But the story must be put to the experimental test, and so far general relativity has passed muster. Soon new tests of a most exquisite sensitivity will come on line, and these are the focus of Thorne's crystal-ball gazing.

The Laser Interferometer Gravitational Wave Observatory (LIGO)—three huge instruments at Hanford, Wash., and Livingston, La.—and similar devices in Italy, Germany and Japan promise the possibility of detecting gravitational waves rippling through spacetime from colossal events (imploding stars, colliding black holes, even the big bang itself) unfolding across the universe. Then, sometime around 2010 if physicists get their way, the Laser Interferometer Space Antenna (LISA) will be launched into space. Three intercommunicating spacecraft arrayed across millions of kilometers of the solar system, bobbing like corks in water as gravitational waves roll by, will map in fine detail the bending of space, the warping of time, and the whirl of spacetime around distant black holes. LISA will detect ripples in spacetime as small as one hundredth the diameter of an atom.

What a story! What a test! This is story making that lifts the human spirit out of our sometimes petty terrestrial concerns and places us among the stars. SA

Chet Raymo is emeritus professor at Stonehill College in Massachusetts and a science columnist for the Boston Globe.

THE EDITORS **RECOMMEND**

SAKHAROV: A BIOGRAPHY

by Richard Lourie. Brandeis University Press, published by University Press of New England, Hanover, N.H., 2002 ($30)

A pillar of Soviet science—father of that nation's hydrogen bomb—becomes an outspoken dissident of the regime and wins the Nobel Peace Prize for his efforts. That is the unusual story of physicist Andrei Sakharov (1921–1989). Lourie, a biographer and translator of contemporary Russian and Polish authors, tells that story well and enriches the picture with accounts of Sakharov's family life and friendships. "It is not false modesty," Sakharov said of himself, "but the desire to be precise that prompts me to say that my fate proved greater than my personality. I only tried to keep up with it."

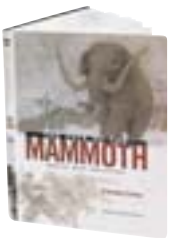
THE GHOST WITH TREMBLING WINGS: SCIENCE, WISHFUL THINKING, AND THE SEARCH FOR LOST SPECIES

by Scott Weidensaul. North Point Press, New York, 2002 ($26)

Weidensaul writes superbly about animals and his role as an amateur naturalist in studying them, as he showed in *Living on the Wind* (1999) on migratory birds. Here his subject is the search for lost species, often as a participant. Among the species are Semper's warbler, the golden toad and the ivory-billed woodpecker. Most searches fail, but Weidensaul describes some successes: the rediscovery of lost species, including Gilbert's potoroo, the large Jamaican iguana and the Congo bay owl. "Imagining [a rediscovery] leads to a germ of hope," he writes, "and hope sometimes leads to belief, to obsession and piles of old maps, to fruitless expeditions and squandered life savings. All of which would seem a sad and farcical pathology, except that just often enough, some lucky searcher hits pay dirt, and the world stands surprised and delighted with the discovery."

THE FATE OF THE MAMMOTH: FOSSILS, MYTH, AND HISTORY

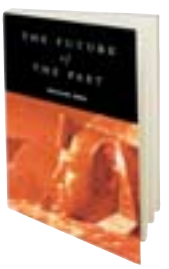
by Claudine Cohen. University of Chicago Press, 2002 ($30)



The mammoth, Cohen writes, "shares traits both of the elephant and the teddy bear." She tells the story of this extinct creature as a means of telling the story of paleontology. Both stories read well. Cohen, who teaches the history of science at the School of Higher Studies in Social Sciences in Paris, examines the possible reasons for the extinction of the mammoths and considers the possibilities for reviving the species through cloning. "By giving life to extinct species," she writes, "paleontology would surpass itself and move from being a science of death to truly being a science of life."

THE FUTURE OF THE PAST

by Alexander Stille. Farrar, Straus and Giroux, New York, 2002 ($25)



It is a paradox, historian Stille says, that human society forgets its past while steadily gaining technology that would help to study and preserve it. His absorbing book seeks "to show the double-edged nature of technological change in a series of different contexts and from a number of odd angles." His explorations include the Sphinx, the looting of Roman artifacts, the Vatican library and "the museum of obsolete technology"—the U.S. National Archive, where technicians "try to tease information out of modern media that have long vanished from circulation."

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

Repellanoids

BY DENNIS E. SHASHA

A virus called a repellanoid has been implicated in a host of diseases. Biochemists have determined that it forms a cylinder, but they don't know the circumference of the cylinder. Here is what they know: There are five different strand sizes, which the scientists refer to by colors: aqua strands of length 4; blue, 5; crimson, 6; dark yellow, 7; and emerald, 8. Each ring of the cylinder consists of strands laid out end to end.

The biochemists have identified the following constraints: Taking the vertical direction to be along the length of the cylinder, if X is a ring and Y is the ring immediately above X, two above X, or three above X, then any vertical line between X and Y must touch a different color on X than on Y. That is, like repels like up to a distance of three. Of course, all the rings must have the same circumference. What is the smallest circumference of the cylinder?

A warm-up problem is illustrated below: Suppose that like repels like at a distance of two, and the only strand types are aqua, blue, crimson and dark yellow. What would be the smallest circumference?

A further question: Another group has discovered a second virus, called a mini repellanoid. Here is what the researchers know: Every strand has a length of 2 or greater, there are five different strand sizes, the circumference is 7, and like repels like at a distance of three. How could this be? SA

Dennis E. Shasha's third book of puzzles, Dr. Ecco's Cyberpuzzles, *has just been published by W. W. Norton.*

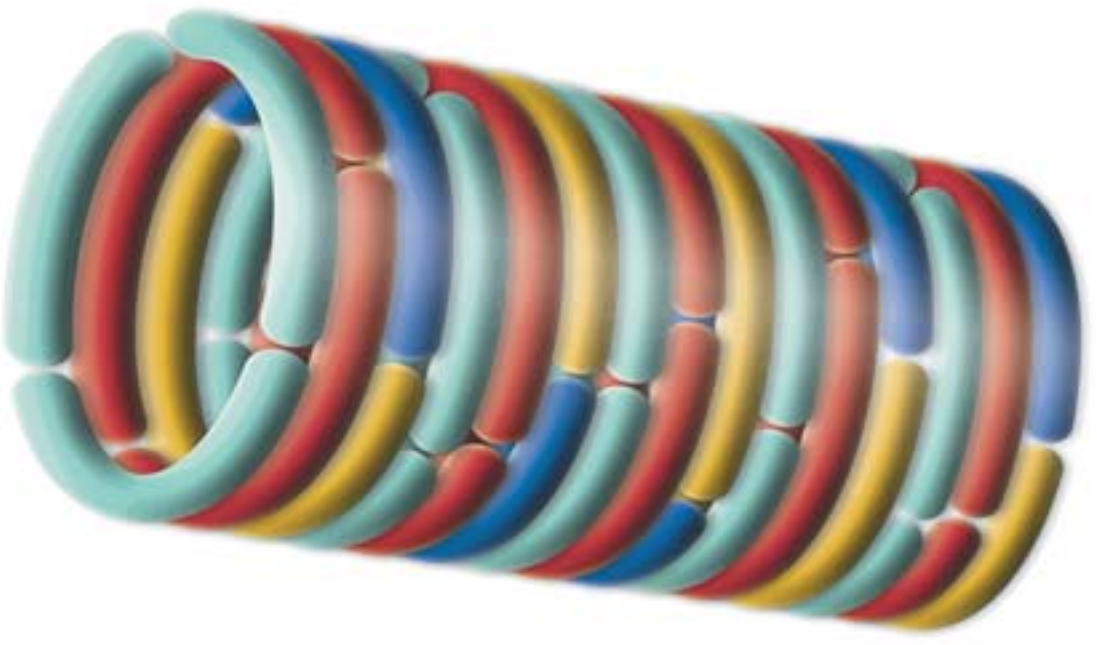
Answer to Last Month's Puzzle

The plaintiff will maximize his monetary award by asking for $10 million, the high end of the judge's range, even if he suspects that the defendant can sneak a peek at his request. For a full explanation, visit www.sciam.com

Web Solution

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

THE WARM-UP REPELLANOID



SARA CHEN

Solution to warm-up: 12. One circle could be aqua-aqua-aqua-aqua; its neighbor, crimson-crimson-crimson; and *its* neighbor blue–dark yellow. This pattern could be repeated.

Poultry and Poetry

O CHICKEN NEW WORLD, THAT HAS SUCH WEIRD STUFF IN IT BY STEVE MIRSKY

Woody Allen called one of his books *Without Feathers,* an angst-ridden response to Emily Dickinson's buoyant comment that "hope is the thing with feathers." In late May a creature seemingly brought forth from Woody's tormented subconscious was revealed to the public: the featherless chicken. Bred by a geneticist in Israel, the bird would allegedly withstand heat better—while alive, anyway. The featherless chicken thus joins the boneless chicken ("How did it walk?" Oscar Madison famously asked) and the rubber chicken as a comedy staple.

As the perpetrator of this column, I was immediately seized with gratitude, recognizing the exposed poultry for the surefire wacky-science page-filler it is. But I felt guilty taking advantage of this chicken served up on a silver platter—where's the challenge? So, having reported the existence of the undressed bird, I'm going to pass the buck and change the subject.

One of the more unpleasant tasks at *Scientific American* is to inform some people that we will not be publishing their theory/discovery/life's work. Therefore, in an attempt to flick chickens from my thoughts, I began to experiment with the pastoral Japanese haiku form as a gentler vehicle than the standard rejection letter. Here are the preliminary results of that experiment:

Your science studies
Fall from the tree like ripe fruit.
There shall we leave them.

Our greatest respect
For your research endeavors
Must remain unknown.

Feynman on bongos
Made a better argument
Than your article.

Photosynthesis
Makes oxygen from forests.
So don't waste paper.

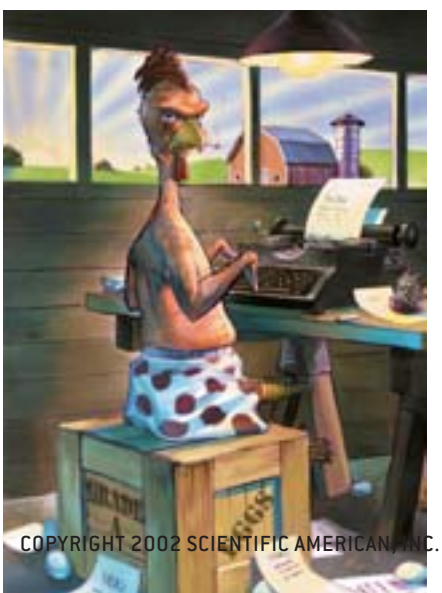
Those examples are general, good for any occasion. Next we have specific replies addressing some common themes from prospective authors:

Proving Einstein wrong
Isn't all that easy, pal.
You took him lightly.

So you cloned yourself.
Please ask your new twin brother
To write the report.

Planet 10, you say?
Boy, do we have news for you.
Pluto got the ax.

We wouldn't have guessed
That your cold fusion device
Would work this badly.



Messages from Mars?
Please check your tinfoil-lined hat.
We think it got wet.

You found no new gene.
A-U-G started this mess.
Toe TAG your sequence.

Though the specs look nice,
A stop sign halts your machine:
Thermo's second law.

Nice three-color map,
With adjacent states both blue.
We knew you'd need four.

Heavy hydrogen?
Congrats on your spanking-new
FBI folder.

What a waste of time
Making transgenic white mice
That contract pinkeye.

Your quantum paper
Was filled with uncertainty:
Neither here nor there.

Stop reinventing
Electromagnetism.
All's well with Maxwell.

We applaud your zeal.
But they never walked abreast,
Dinosaurs and man.

No global warming?
I guess you might be correct.
But where's all the snow?

And finally:

Featherless chicken.
How does one get down off you?
Just think about it. SA

ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

ASK THE EXPERTS

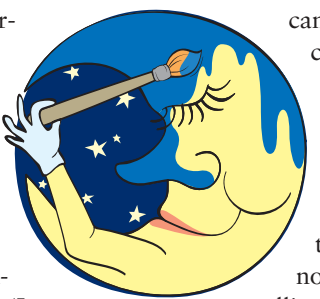
How can an artificial sweetener contain no calories?

—A. Rivard, Argyle, Minn.

Arno F. Spatola is professor of chemistry and director of the Institute for Molecular Diversity and Drug Design at the University of Louisville. His current research focuses on peptides, including artificial sweeteners. He offers this answer:

Sweetness is a taste sensation that requires interaction with receptors on the tongue. Many sugar substitutes, such as saccharin and acesulfame K, also known as Sunette, do not provide any calories. This means that they are not metabolized as part of the normal biochemical process that yields energy in the form of adenosine triphosphate, or ATP. In some cases, small quantities of additives such as lactose are present to improve the flow characteristics or to give bulk to a product. But the amounts are so small that they do not represent a significant source of energy.

The low-calorie approach of the sugar substitute aspartame, also called NutraSweet, is more interesting. This synthetic compound is a dipeptide, composed of the two amino acids phenylalanine and aspartic acid. As with most proteins, which are chains of amino acids, it can be metabolized and used as an energy source. In general, we obtain energy in the amount of four calories (more correctly termed kilocalories) per gram of protein. This is the same value as the number of calories acquired from sugars or starches. (In contrast, each gram of fat consumed provides more than twice that amount, or about nine calories a gram.)



So if aspartame has the same number of calories per gram as common table sugar (sucrose), how is it a low-calorie sweetener? The answer is that aspartame is 160 times as sweet as sugar. That is, a single teaspoon of aspartame (four calories) will yield the same sweetening effect as 160 teaspoons of sugar (640 calories). If 3,500 extra calories is equivalent to a gain of one pound in weight, it is easy to see why so many people turn to artificially sweetened beverages in an effort to maintain some control over their amount of body fat.

But does that actually lead to weight loss? Perhaps not. Either by a physical effect, or perhaps a psychological one, many of us seem to make up the loss of sugar calories by eating or drinking other foods. For this reason, artificially sweetened diet drinks alone are hardly likely to have much of an effect on the problem of obesity in the U.S.

What is a blue moon?

—B. Purvis, Carlisle, Pa.

George F. Spagna, Jr., chair of the physics department at Randolph-Macon College, supplies an explanation:

The definition has varied over the years. A blue moon once meant something virtually impossible, as in the expression "When pigs fly!" This was apparently the usage as early as the 16th century. Then, in 1883, the explosion of Krakatau in Indonesia released enough dust to turn sunsets green worldwide and the moon blue. Forest fires, severe drought and volcanic eruptions can still do this. So a blue moon became synonymous with something rare—hence the phrase "once in a blue moon."

The more recent connection of a blue moon with the calendar apparently comes from the 1937 *Maine Farmer's Almanac*. The almanac relies on the tropical year, which runs from winter solstice to winter solstice. In it, the seasons are not identical in length, because the earth's orbit is elliptical. Further, the synodic, or lunar, month is about 29.5 days, which doesn't fit evenly into a 365.24-day tropical year or into seasons roughly three months in length.

Most tropical years have 12 full moons, but occasionally there are 13, so one of the seasons will have four. The almanac called that fourth full moon in a season a blue moon. (The full moons closest to the equinoxes and solstices already have traditional names.) J. Hugh Pruett, writing in 1946 in *Sky and Telescope*, misinterpreted the almanac to mean the second full moon in a given month. That version was repeated in a 1980 broadcast of National Public Radio's *Star Date*, and the definition stuck. So when someone today talks about a blue moon, he or she is referring to the second full moon in a month. SA

For a complete text of these and other answers from scientists in diverse fields, visit www.sciam.com/askexpert

SARA CHEN