CANCER-KILLING VIRUSES • DANGEROUS MELTING IN THE ARCTIC

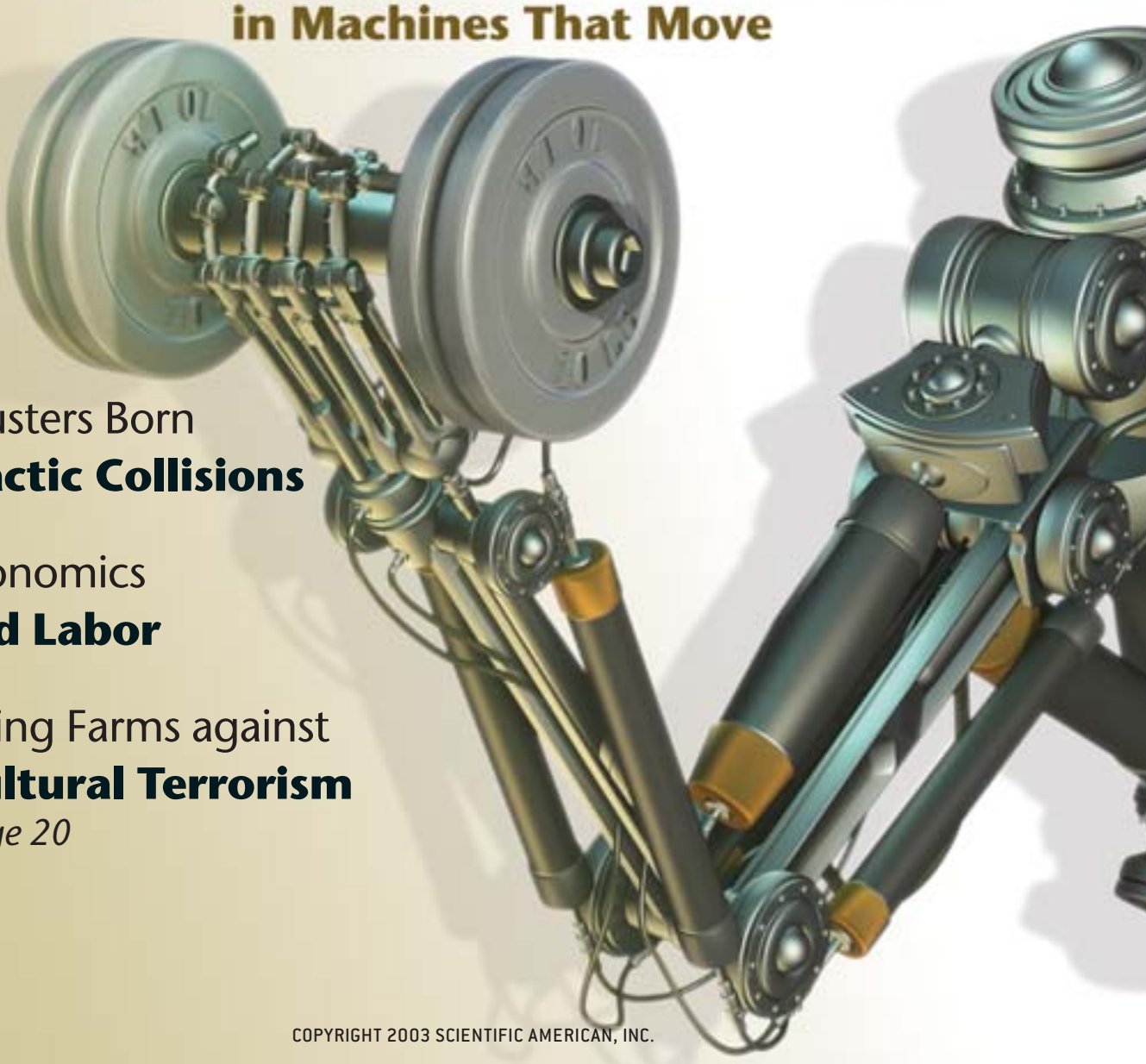
SCIENTIFIC AMERICAN

China:
The Next
Space
Superpower

OCTOBER 2003 $4.95
WWW.SCIAM.COM

ARTIFICIAL MUSCLES

Shape-Shifting Plastics Replace Motors in Machines That Move



Star Clusters Born
of **Galactic Collisions**

The Economics
of **Child Labor**

Protecting Farms against
Agricultural Terrorism
—see page 20

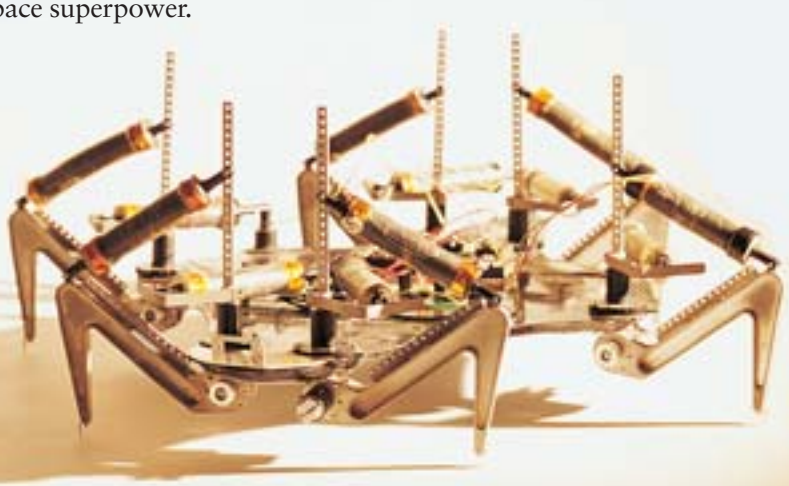
october 2003

contents

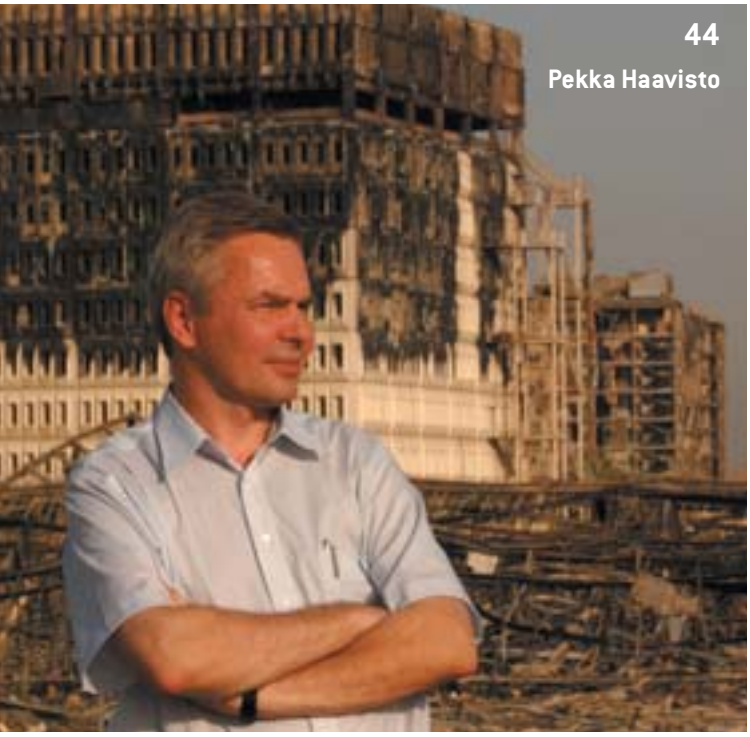
SCIENTIFIC AMERICAN Volume 289 Number 4

features

52 Robot walks using artificial muscles



departments



columns

Cover image by Kenn Brown; SRI International (*preceding page*); Kate Brooks (*left*)

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 USD, International $55 USD. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. **Reprints available:** write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; (212) 451-8877; fax: (212) 355-0408 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.

Biotech's Clean Slate

When President George W. Bush scolded Europe last June for resisting genetically modified crops, he was acting out his part in what has become one of the most boring controversies in biotechnology. The U.S. always plays rah-rah cheerleader to Europe's pouty bench-warmer. Much of the public on both sides of the Atlantic continues to reject transgenic food despite science and industry arguments about its benefits. Findings that point to risks from transgenic organisms are waved away as manageable by biotech's advocates; findings that reinforce biotech's safety never surmount the precautionary principle (do nothing new until its safety is perfectly assured). The story hasn't progressed in years.

VATS OF CELLS grown for industrial biotech suffer little critical scrutiny.

What a relief, then, to consider an often overlooked segment of biotech that has so far escaped the fracas. Industrial biotechnology applies the life sciences to manufacturing—for instance, by using cells to synthesize materials or by substituting enzymes for caustic reagents.

Last April in Lyon, France, the World Life Sciences Forum (BioVision 2003) dedicated much of its program to the subject, which some European authorities call "white biotechnology" to distinguish it from the agricultural ("green") and medical ("red") varieties. Similarly, at the Biotechnology Industry Organization in Washington, D.C., in June, it was sometimes called the "third wave of biotechnology."

The sobriquets may be new, but the essence of industrial biotechnology is as old as bread and beer. Yeasts, molds and other microorganisms have been used to produce goods throughout history. Today's manufacturers can tweak cells in unprecedented ways, however. The technology's proponents argue that life-science solutions can lower production costs, create jobs, conserve resources and reduce pollution, to boot.

McKinsey and Company has projected that by 2010, between 10 and 20 percent of all chemical production might involve biotechnology (up from roughly 5 percent now), reflecting about $280 billion in sales value. Such growth would require favorable circumstances—not the least of which, of course, is the continued avoidance of a backlash like that tormenting genetic crop developers. How likely is that?

Industrial biotech's best ace in the hole is the public's casual confidence in it. No one seems to protest when industrial chemicals or other nonfood products come from genetically modified sources, probably because the organisms are cloistered inside facilities.

Still, growth in industrial biotech may require more exposure of altered organisms to the outside world. As the complexity of the modified organisms rises, so, too, might ethical objections. In 2002, when Nexia Biotechnologies created goats that had spider silk proteins in their milk, the ethics of passing genes between such dissimilar species bothered many animal welfarists. At a more mundane level, just as claims that agricultural biotech can feed the starving in poor countries incite debate, the proposition that industrial biotech will help in the developing world needs to engage criticisms that not all biotechnologies are easily exportable.

It seems inevitable that eventually a conflict will oblige the public to see similarities between white and green biotech. When that happens, perhaps this new wave of biotech will suffer. Or perhaps—here's a wild thought—good experiences with industrial biotech will reassure the public over its fears of genetically modified crops. Maybe the biotech impasse will budge yet.

THE EDITORS editors@sciam.com

JAMES HOLMES/*SPL/Photo Researchers, Inc.*

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*:
www.sciam.com/permissions
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
Derr Media Group
847-615-1921
fax: 847-735-1457
U.K.
The Powers Turner Group
+44-207-592-8331
fax: +44-207-630-9922
France and Switzerland
PEM-PEMA
+33-1-46-37-2117
fax: +33-1-47-38-6329
Germany
Publicitas Germany GmbH
+49-211-862-092-0
fax: +49-211-862-092-21
Sweden
Publicitas Nordic AB
+46-8-442-7050
fax: +46-8-442-7059
Belgium
Publicitas Media S.A.
+32-(0)2-639-8420
fax: +32-(0)2-639-8430
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/ontheweb** to find these recent additions to the site:

Mare Gives Birth to Own Clone

Scientists have for the first time successfully cloned a horse. What is more, the mare from which the original cell material was taken—not a surrogate—gave birth to it. This challenges the idea that the early success of a pregnancy depends on the mother's immune system responding to a developing fetus and placenta as something different from itself.

GIOVANNA LAZZARI

E-mail Study Corroborates Six Degrees of Separation

Chances are, you don't personally know any Australian policemen, Estonian archival inspectors or Norwegian army veterinarians. But you could probably get in touch with one of these distant individuals through a friend, or a friend of a friend, or a friend of your friend's friend. Experimental data from the Internet confirm that every person on the planet is separated from everyone else by a chain of about six people.

Ask the Experts

What causes feedback in a guitar or a microphone?
Robert L. Clark of Duke University's Pratt School of Engineering explains.

New to the Site: Astronomy Channel

Travel to the far reaches of space straight from ScientificAmerican.com. This section houses recent articles on topics ranging from comet watching and planet hunting to stellar black holes and cosmic baby booms.

New and Improved Scientific American DIGITAL

- Faster and easier access to Digital
- All current issues before they reach the newsstands
- More than 140 issues of *Scientific American* from 1993 to the present
- Full-text search for all stories and captions—no more flipping through the issues to find an article

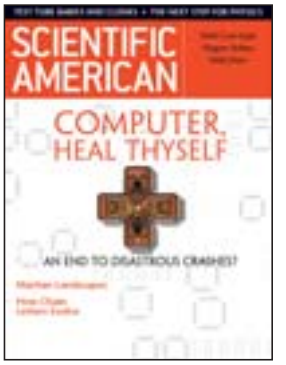
TRY IT TODAY AT WWW.SCIAMDIGITAL.COM

SCIENTIFIC AMERICAN®
Established 1845

Letters

EDITORS@SCIAM.COM

"TRUTH," WROTE AMERICAN POET Ralph Waldo Emerson, "is our element." With a flow of queries and comments, readers essayed to distill this element from nearly every topic in the June issue of *Scientific American*. Some, after reading "The Dawn of Physics beyond the Standard Model," by Gordon Kane, questioned just how close to true a unified theory based on particles could get. In response to the profile of a U.N. weapons inspector ["One Last Look," by Gary Stix, Insights], many insisted that judgments about Iraqi weapons programs be based on tangible, verifiable evidence. Several joined resident Skeptic Michael Shermer in challenging what does count as actual evidence, vis-à-vis Michael Drosnin's book *The Bible Code* ["Codified Claptrap"]. Reactions to these and other elements in our periodical table of contents fill this column.



PARTICULAR INTERESTS

Gordon Kane's excellent piece "The Dawn of Physics beyond the Standard Model" ends with the suggestion that "particle physics might increase our understanding of nature to the point where the theory can be formulated with no inputs." He uses the word "inputs" to include not only properties of fundamental particles but also the existence of spacetime and the rules of quantum theory. Is this really probable?

In his book *Unended Quest,* the late philosopher of science Karl Popper says that "the evolution of physics is likely to be an endless process of correction and better approximation. And even if one day we should reach a stage where our theories were no longer open to correction, since they were simply true, they would still not be complete—and we would know it. For [mathematician Kurt] Gödel's famous incompleteness theorem would come into play: in view of the mathematical background of physics, at best an infinite sequence of such true theories would be needed in order to answer the problems which in any given (formalized) theory would be undecidable."

Crispin Rope
Woodbridge, England

BIBLE CODE CODA

As author of *The Bible Code* and *Bible Code II: The Countdown,* I am replying to Michael Shermer ["Codified Claptrap," Skeptic]. He states that "to be tested scientifically, Bible codes would need to predict events *before* they happen. They won't, because they can't.…"

It appears that Shermer never read my books, or he would know that in several cases the Bible code did predict events before they happened. The most dramatic was my prediction that Israeli prime minister Yitzhak Rabin would be assassinated. I personally warned the prime minister a year before he was killed.

Shermer then goes on to cite a misquotation of me that appeared in *Newsweek:* "When my critics find a message about the assassination of a prime minister encrypted in *Moby Dick,* I'll believe them." *Newsweek* left out two key words, "in advance." In correspondence with me, the magazine described my prediction as accurate and also stated that "we believe it is clear enough in our story that the timing of your achievement is what distinguishes it from other claims."

I could list many examples of predictions found in advance—the test Shermer states. The exact dates of the Gulf War, before the war started; the exact date of impact of a comet, with its exact name, and the planet Jupiter, found months before the impact; the outcomes of elections in the U.S. and Israel that every poll got wrong. But of course Shermer cannot admit the reality of things that he cannot explain.

Michael Drosnin
New York City

THE EDITORS REPLY: Newsweek *has never published a clarification of Drosnin's quote from 1997 or acknowledged having misquoted him.*

Shermer overgeneralized in saying that all Drosnin's predictions were really after-the-fact "postdictions"; Drosnin did send Rabin a warning before he was assassinated. Shermer and other critics nonetheless maintain that Drosnin's predictions are less than they seem. For example, according to The Bible Code, *Drosnin predicted the July 1994 comet strike on Jupiter two months before the impact, but astronomers had announced the calculated date of that collision by mid-1993.*

Scientific evaluations of Drosnin's prophetic abilities must be based on all his predictions—not just the ones that came true but also the ones that were wrong, and discounting those that are post hoc, imprecise or explainable as lucky guesses.

Thanks to Michael Shermer for another wonderful Skeptic column. His work is comforting to many of us exposed to the everyday mentality.

Gary Smith
via e-mail

SIGNING OFF

As the parent of two girls with cochlear implants, I was encouraged to see this wonderful technology getting needed exposure in your magazine ["To Hear Again," by Mark Fischetti, Working Knowledge]. But I was discouraged to see you print the statement of the National Association of the Deaf. The association discourages such implants in children born deaf "because, even with the technology, it is very hard for them to develop the cognition for spoken language," and the children often aren't taught sign language early enough as a result, causing "developmental delays that can be extremely difficult to reverse."

Both my daughters were born deaf and received cochlear implants at very young ages. Neither has ever learned, or needed, sign language. Both girls hear so well that they can converse on the phone with ease. In fact, I would counter the association's statement by saying that learning how to sign would have hindered the development of my daughters' spoken language.

Melissa K. Chaikof
Atlanta, Ga.

STILL LOOKING

In Insights, "One Last Look," by Gary Stix, U.N. weapons inspector Rocco Casagrande remarks that Iraq "wasn't behaving like a country that doesn't have biological weapons." How would Iraqis behave if they did not possess chemical or biological weapons, had significant information or technology but no weapons, or technology whose true purpose they did not know? What if orders were simply to reveal nothing and to be as obstructive as possible? Beyond this, the language and culture of Iraq would be unfamiliar to many inspectors. People living in the surreal atmosphere of a ruthless, secretive, repressive, deeply suspicious and no doubt extremely resentful regime seem likely to behave in ways that others might regard as unusual.

Rather than speculating about the state of weapons based on behavior, inspectors' conclusions—not to mention decisions about war—should be based on real information.

Howard Eaton
New Westminster, B.C.

JUST OIL? U.N. weapons inspector Rocco Casagrande examines dilapidated oil barrels in Juwesma, Iraq, in January.

HUSSEIN MALLA *AP Photo*

RULING THE ROOT

In "Chain Letters and Evolutionary Histories," by Charles H. Bennett, Ming Li and Bin Ma, a diagram suggests that man descended from the chicken and is accompanied by the disclaimer that "obviously the mammals did not evolve from the chicken." Then why does the chicken occupy the root position of the tree? I think the tree is a setback for evolutionary theory and will provide creationists with yet another object of ridicule.

Thomas J. Kelanic
Turtle Creek, Pa.

BENNETT, LI AND MA REPLY: A standard practice in constructing a phylogeny for a group of present-day organisms, such as mammals, is to use a present-day unrelated organism as a stand-in for the group's extinct common ancestor, which is generally not available for study. Thus, we used the chicken as a stand-in for the unavailable common ancestor of all mammals and birds. Confusingly, phylogenies are often drawn with the outgroup as the root, even though everyone knows it is not an ancestor of the other organisms. We regret that we followed this custom in our diagram.

ERRATA In "The Unearthly Landscapes of Mars," by Arden L. Albee, the scale bar on page 45 should have read "80 meters," not "10 meters." The caption should have said that crossing the area shown in the image would take about "half an hour," not "five minutes."

"Chain Letters and Evolutionary Histories," by Charles H. Bennett, Ming Li and Bin Ma, omitted mentioning Gregory Chaitin's contributions to the foundations of algorithmic information theory.

The opening artwork for "Self-Repairing Computers," by Armando Fox and David Patterson, should have been credited to Frank Ippolito, not Slim Films.

50, 100 & 150 Years Ago

FROM *SCIENTIFIC AMERICAN*

Edible Algae ▪ Safer Phosphorus ▪ Cheap Anthracite

OCTOBER 1953

GREENS FOR DINNER—"Many scientists all over the world are interested in the food possibilities of the water plants called algae. On the basis of laboratory experiments it is estimated that each acre given to cultivation of Chlorella could produce an annual yield of 20 tons of protein and three tons of fat per acre—astronomical figures compared with present rates of production in conventional agriculture. Whether algae can be an important contribution to the world food supply will depend on the cost and the yield of large-scale culture. The production of each ton of algal protein requires about 1.1 tons of potassium nitrate and .75 ton of ammonium sulfate."

OCTOBER 1903

LANGLEY'S FAILURE—"Those who have the interests of aerial navigation at heart will regret the failure of Prof. Samuel Pierpont Langley's last experiment, not so much because the aerodrome refused to fly, but because of the adverse newspaper comment which the trial has prompted. This aerodrome of his is the result of years of arduous study and ceaseless experimentation. That it should have failed is to be regarded simply as one step in the solution of the problem of aerial navigation, and not altogether as an abject failure. On the report of Prof. C. M. Manly, it appears the clutch which held the aerodrome on the launching ways [*see illustration*] and which should have released at the instant of the fall, was found to be injured." [*Editors' note: The failure of this test and the one on December 8, 1903, led to such scathing public criticism that Langley gave up aviation research.*]

GOT A LIGHT?—"By a law of May 10, 1903, Germany forbade the use of white phosphorus in the making of matches. A new material, made of non-poisonous red phosphorus and potassium chlorate, has been bought by the government and is to be substituted for the deleterious and dangerous white phosphorus. In spite of its high igniting point, the new material may be lighted by scratching on almost any material—sandpaper, bricks, soles of shoes, rough clothing, etc. It is a great gain that it does not ignite easily, important when one is reminded of fires caused by the ignition of white phosphorus matches by the sun's rays."

LANGLEY'S AIRPLANE on the catapult, 1903

SANITAS AMERICANA—"Since the American occupation of Cuba, yellow fever is gradually being eradicated. This remarkable sanitary change is due partly to the explosion of the old superstitious beliefs by the army surgeons and partly to a systematic extermination of the mosquito. Dissipating the common notion that yellow fever is a deadly filth disease, highly contagious, our army experts showed that yellow fever was actually spread by the mosquito. Attempts at the extermination of the mosquito in Cuba have borne such fruitful results that in time the Cuban cities will be as free from yellow fever as our Southern ports."

ANTS—"An unlooked-for sequence in the drainage of New Orleans is the appearance of hordes of ants, which, according to the Iron Age, have become as threatening as the plagues of Egypt. They attack the woodwork of houses and speedily destroy it, and make their way into warehouses where costly goods are stored. When the soil was saturated the ants could not breed in it; now that it is no longer wet they defy suppression."

OCTOBER 1853

COAL FOR TRAINS—"With very few exceptions, wood is the only fuel used for locomotive engines. It is becoming so scarce and dear that some substitute must be sought. Anthracite coal suggests itself first, because it is the cheapest and most free from smoke and waste. An impression, however, is that this fuel destroys the steam firebox so quickly that it cannot be used with economy. Other objections grow out of the intensity of the heat. But all of these objections have been removed by the Millholland engine. There are now in daily use on the Reading Railway, Pa., twenty-eight first class locomotives on the Millholland plan; these use anthracite coal exclusively. No engineer will run a wood burning locomotive if he can get a coal burning one, as they cause far less work and also make better time."

FASHIONS OF THE DAY—"The importation of monkey skins is an important business in Salem. The 'Gazette' says: 'Monkey skins have formed an article of commerce for several years, and we dare say that many a fair lady has strutted her brief hour in all the glory of a monkey skin muff and rat skin gloves, without suspecting the quality of her finery.'"

DEFENSE

Food Fears

THE THREAT OF AGRICULTURAL TERRORISM SPURS CALLS FOR MORE VIGILANCE **BY DANIEL G. DUPONT**

CATTLE CONCERN: A terrorist attack with foot-and-mouth disease would devastate the U.S. beef industry.

Earlier this year the discovery of a single cow with bovine spongiform encephalopathy, better known as mad cow disease, crippled the Canadian cattle market. In 2002 the mere rumor of foot-and-mouth disease in Kansas sent shock waves through the American cattle industry. And the discovery of exotic Newcastle disease in southern California led to the destruction of millions of chickens and prompted many countries to ban poultry coming from the area—and, in some cases, from the entire U.S.

Terrorists probably had nothing to do with the incidents, but agriculture and Homeland Security officials cite these and similar events in describing the possible effects of a bioterror attack on domestic agriculture. Officials take such a threat seriously—the terrorist group Al Qaeda long ago put the U.S. food supply on its list of potential targets. The federal government is working to bolster the nation's readiness for an agroterror attack—and some of their assessments suggest significant vulnerabilities that critics say are not getting enough attention.

From farm crops and animals through the processing system to the grocery store, the food supply chain provides numerous opportunities for attack. Moreover, the system would ensure rapid disease progression: animals are moved often and quickly, and anticrop agents can be spread by the wind. Since the September 11 attacks, the U.S. Department of Agriculture has hired new inspectors and strengthened its diagnostic capabilities around the country. The Food and Drug Administration has bolstered food safety rules and made it easier for investigators to trace the origins of an outbreak. The Department of Homeland Security has assumed responsibility for the inspection of agricultural

products entering the country. And states are working on their own efforts to educate and ensure proper coordination in the event of an outbreak.

In part to determine the effectiveness of such preparations, the Pentagon organized two classified exercises called Silent Prairie, part of an ongoing series of simulations run by the National Defense University (NDU). In the February exercise, members of Congress, state officials and government representatives dealt with foot-and-mouth disease. The USDA had already calculated that the highly virulent sickness could spread to as many as 25 states in as little as five days. The Silent Prairie simulation produced equally horrifying results: more than one third of the nation's cattle herds wound up infected, according to Zdenka Willis, a navy captain at the NDU. Representative Devin Nunes of California, who hails from a district heavily dependent on farming, remarks that such an outbreak "would be devastating to our food supply and our economy."

Thomas McGinn, head of the North Carolina emergency programs office and a Silent Prairie consultant, says that too many people see agricultural terror incidents as local events, akin to "lobbing a grenade over enemy lines." But the rapid spread of foot-and-mouth makes it "a homeland security issue, immediately," he insists. Indeed, according to an NDU paper outlining the Silent Prairie results, "response to an agricultural bioterrorism attack could require significantly more resources than the attack on the World Trade Center."

The expense of coping with agroterror is why Peter Chalk, an analyst with the RAND Corporation, a Santa Monica, Calif., think tank, remains concerned about what he deems insufficient federal focus on the threat of "economically catastrophic" attacks. Chalk notes that some improvements have been made in the past year, notably in such areas as security at food-processing facilities. But "I haven't really seen much in terms of concrete policy taking place," he says. Chalk wants the U.S. to undertake a "comprehensive threat analysis" as well as an assessment of how much money is needed and where it should be spent.

McGinn believes that the federal government should do for agriculture what it has done for human health since 9/11: dramatically increase state and local capabilities to detect diseases and educate medical personnel as well as the public. "The ability to feed ourselves has become part of the critical infrastructure of our country," McGinn states. "We've got to increase the security in our food system all the way from the farm to the fork."

Daniel G. Dupont, who edits InsideDefense.com, an online news service in Washington, D.C., wrote about nonlethal knockout gases in the January issue.

CIVILIANS ON THE FRONT LINES

No amount of government action can protect the nation's entire agricultural infrastructure, experts warn. That puts the farmers, the veterinarians, the feedlots and the processing facilities on the front lines. Several members of Congress have pushed for more federal funds and emphasis on heightened awareness at those levels.

A problem, notes Robert E. Brackett, director of food safety for the FDA's Center for Food Safety and Applied Nutrition, is that many studies of the vulnerabilities in the U.S. food supply have been classified—which has made it difficult for the agency to work with industry to shore up weak spots. The FDA is working to ensure that at least key members of the food industry have access to the essential information, even if some of it is classified.

POLICY

Return of the Fleece

SCIENCE FEELS THE HEAT FROM THE POLITICS OF MORALITY

BY SALLY LEHRMAN

In March 1975 psychologist Elaine Hatfield and her $84,000 study on romantic attraction became the centerpiece of U.S. Senator William Proxmire's inaugural "Golden Fleece Award," designed to skewer government waste. Proxmire complained that scientists would never find the answer to the mystery of love. Even if they did, the senator from Wisconsin insisted, no one else would want to know it. Bags of mocking letters deluged Hatfield's University of Wisconsin office, and Proxmire asked to see her expense records plus the confidential names and addresses of the several thousand students she had interviewed.

Undeterred, the scientist went on to develop a popular instrument to measure the intensity of obsessive romantic love—and never again applied for government funding. The senator issued another 150 awards for squandering taxpayer money over the years, before retiring in 1989 at age 73.

Now scientists are preparing for what they say could be a far more dangerous sort of attack—not based on perceived frivolity but on

politicians' moral aversion to research topics that touch on sexuality and sex roles. Unlike Proxmire, today's congressional representatives are "not laughing about it," says John Bancroft, director of the Kinsey Institute, which is affiliated with Indiana University. "They're expressing outrage and disgust."

AX SWINGING: Representative Patrick Toomey wanted to cut funds to five NIH-approved studies.

PUT ON THE CHOPPING BLOCK

The House of Representatives almost voted to withdraw 2002 National Institutes of Health funding for five studies:

- **Mechanisms Influencing Sexual Risk-Taking (Kinsey Institute, Indiana University): $237,038**
- **Longitudinal Trends in the Sexual Behavior of Older Men (New England Research Institutes): $69,000**
- **HIV Risk Reduction among Asian Women Working in Massage Parlors (University of California, San Francisco): $641,000**
- **Health Survey of Two-Spirited Native Americans (University of Washington): $521,022**
- **Spatial and Temporal Interrelationships between Human Population and the Environment (Michigan State University): $157,500**

The most recent round came in July during House debate over the 2004 Labor, Health and Human Services and Education Appropriations bill. Republican representatives Patrick Toomey of Pennsylvania and Chris Chocola of Indiana proposed that Congress excise five grants totaling about $1.6 million from National Institutes of Health funding. "Who thinks this stuff up?" complained Toomey, emphasizing four inquiries that delved into sexuality, arousal and sexual orientation. He condemned the research as studies into forays of drug use and sex among marginal groups such as prostitutes and transgendered Native Americans. "If they want to do this sort of research, we need to fund this privately and not with taxpayer dollars," he said.

In fact, two of the studies investigate HIV-related sexual risk-taking and prevention strategies in communities where the epidemic is advancing most rapidly. Another, at the Kinsey Institute, focuses on people in which negative mood improves arousal and could shed light on compulsive sexual behavior and assault. A fourth analyzes sexual activity in aging men to see if dysfunction could be an early indicator of conditions such as diabetes and heart disease.

Toomey's measure was defeated by only two votes, 212 to 210, with one House member later apologizing to constituents that he had voted "no" in error. "It was alarming to me that this did not go down in a tumultuous negative reaction," says Harold E. Varmus, president of Memorial Sloan-Kettering Cancer Center in New York City and former NIH director. Scientists clearly need to educate Congress about peer review, the scientific process and the damage that political interference can do, he warns.

Before they can win approval, NIH grant proposals are ranked by study sections of 15 to 20 outside scientists and then sent on to be reviewed by advisory councils made up of scientists and members of the public. Duane Alexander, director of the NIH's National Institute of Child Health and Human Development, says that all three NICHD grants earmarked by Toomey had received "outstanding" and "excellent" ratings. They also answered urgent calls—including one by then Surgeon General David Satcher in 2001—to study sexual behavior and address problems such as HIV/AIDS, unwanted pregnancies, sexual abuse and rape. In 2002 the House Appropriations Committee commended the NICHD for its studies on healthy adult sexuality, Alexander points out, adding, "That's a pretty clear mandate to do this very kind of research. Then they go back and try to stop it."

Alexander worries that efforts such as Toomey's could chill free inquiry and turn the public against science. Researchers have already begun to carefully select the terms they use in grant titles and correspondences and are gearing up for another battle over appropriations in the Senate. They point to a series of incidents that they say indicate an acceleration of pressure on studies involving HIV prevention, reproductive health, and sexuality.

"They think the peer-review process has been hijacked by some liberal political agenda," says Craig Hogan, vice provost for research at the University of Washington, home to the study on transgendered Native Americans. But "it's just about public health. That's why American science is so good—there's a lot of integrity in the system."

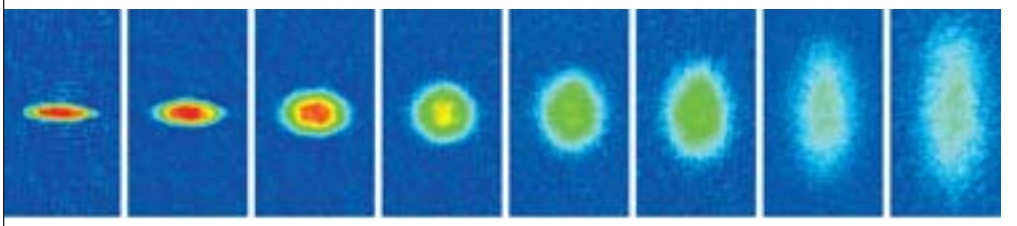
Sally Lehrman is based in San Francisco.

DOUGLAS GRAHAM *Roll Call*/NewsCom

PHYSICS

The Next Big Chill

PHYSICISTS CLOSE IN ON A NEW STATE OF MATTER **BY GRAHAM P. COLLINS**



SUPERFLUID? An ultracold gas of lithium 6, initially compressed in a thin cylinder, expands radially when released—a result that is suggestive of superfluidity but is not conclusive. The sequence runs from 0.1 to 2.0 milliseconds after the release.

It occurs in objects as diverse as superconductors, atomic nuclei and neutron stars. Several research groups are in a race to recreate it in the laboratory in microscopic specks of ultracold gas. If they succeed, it will enable experimental studies of processes that have heretofore been the domain of theorists. "It" is a superfluid state of matter predicted to occur when quantum particles that normally shun one another pair up and behave en masse as a single body of fluid.

This superfluid state involves a broad class of quantum particles called fermions. According to quantum mechanics, all particles in nature are either bosons or fermions. The distinct characters of these two classes become most accentuated at very low temperatures: Bosons sociably gather all in a single quantum state, forming a Bose-Einstein condensate. Fermions, in contrast, act as individualists, no two occupying the same quantum state. As things cool, fermions increasingly occupy the lowest energy states, but they stack up one to a state, like people crowded onto a narrow flight of stairs. This state, in which most of the lowest energy states are occupied by one fermion each, is called a degenerate fermi gas.

In 1999 Deborah S. Jin and Brian DeMarco of JILA in Boulder, Colo., produced the first degenerate fermi gas of atoms in a tiny cloud of potassium atoms in a magnetic trap. But such a degenerate gas is only half the story. In similar degenerate systems that occur in liquid helium 3 and among electrons in superconductors, something new happens—some of the fermions form up in pairs called Cooper pairs. These pairs, which are bosonic, then form a superfluid state very similar to a Bose-Einstein condensate: in helium 3 it is responsible for the liquid's superfluid properties; in a superconductor it allows the resistanceless flow of electricity.

Can such a superfluid state be made in the gaseous fermion systems? Theory predicts that atomic Cooper pairs usually will form only at a temperature much colder than that required for degeneracy, a temperature that seems beyond the reach of experiment at the moment. Recently, however, an alternative method was suggested, based on the fact that Cooper pairing depends on not just the temperature but also the interaction between the atoms. So instead of making the gas colder, why not increase the interaction? Nature has fortuitously provided a convenient way to adjust the interaction—by applying a magnetic field of just the right strength to create what is called a Feshbach resonance, which generates a powerful attractive or repulsive interaction between the atoms. (An attractive one is needed for Cooper pairs to form.)

In late 2002 a group led by John E. Thomas of Duke University used these techniques with lithium 6 atoms to produce results highly suggestive of superfluidity. The trapped gas formed a thin cylinder, and when the trapping laser beams were turned off, the gas expanded radially to form a disk shape—very little expansion took place along the axis of the cylinder. Such anisotropic expansion had previously been predicted to be a hallmark of the superfluid state.

As the Duke group pointed out, however, other effects can also generate such anisotropic expansion. Indeed, experiments conducted earlier this year by Jin's group and by Christophe Salomon and his co-workers at the École Normale Supérieure in Paris have exhibited similar anisotropic expansions in situations where a superfluid cannot be present.

A technique for directly detecting the Cooper pairs or the superfluid is needed. Jin, as well as Wolfgang Ketterle's group at the Massachusetts Institute of Technology, recently reported using radio waves to study the precise states of the atoms in the trapped

JOHN E. THOMAS *Duke University*

A BUNCH OF DEGENERATES

A degenerate gas of fermions occurs in diverse situations, as described below:

- **Superconductors:** The electrons are degenerate and form loosely correlated Cooper pairs, which produce the superconductivity. Something similar must happen in high-temperature superconductors, but that process remains a mystery.
- **Neutron stars:** The refusal of neutrons (which are fermions) to occupy identical quantum states generates a repulsion that prevents the star from collapsing under its own immense weight. A similar repulsion stabilizes the laboratory-made degenerate fermi gases against collapse.
- **Quark-gluon plasma:** As created at the Relativistic Heavy Ion Collider at Brookhaven National Laboratory, the exploding cloud of free quarks (which are fermions) and gluons has properties similar to a gas of fermionic atoms released from the confines of a trap.

degenerate gas; if Cooper pairs were present, the binding energy of the pairs should show up clearly. Neither group saw such signs of Cooper pairs, but both uncovered useful new details of how fermionic atoms interact near a Feshbach resonance.

Several teams have recently studied the formation of loosely bound two-atom molecules in their gases. "We hope that we can turn [the molecules] into Cooper pairs," Ketterle says. And in August theorist Yvan Castin and his co-workers at the École Normale Supérieure suggested just how that might be done: first let the molecules Bose-condense, then adjust the Feshbach resonance. If that is true, experimenters are just two steps from their goal.

HEALTH

Musical Medicine

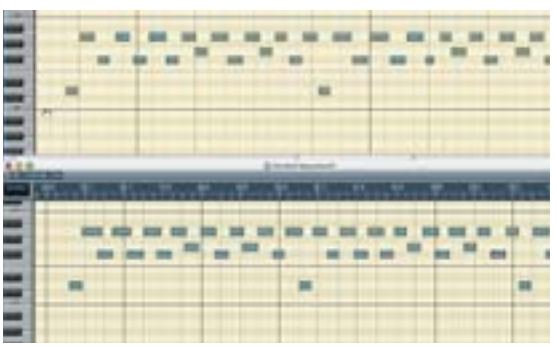
A HIGH-TECH PIANO TREATS A REPETITIVE STRESS DISORDER **BY W. WAYT GIBBS**

IT GETS STUCK IN YOUR HEAD

The stress injury called dystonia appears to originate in the brain, not the muscles. "If you do an MRI on someone with focal dystonia," says Edgar E. Coons of the Center for Neural Science at New York University, "you see a change in the parts of the brain that receive touch and motor feedback for each finger. Those zones are normally physically separated in the brain. But in people with dystonia, the regions merge." Curiously, the cramping and rigidity of focal dystonia often disappear during everyday activities but resurface as soon as the musician starts to play.

My left forearm twinges as I sit down at Kathleen M. Riley's piano. An hour of scribbling notes and two days of working on a laptop computer have inflamed my repetitive stress injury, an ailment common among journalists—and musicians. In fact, hardworking musicians can develop a much more severe condition, known as focal dystonia, which cramps the hands so badly that it often ends a promising career. Injections of botulinum toxin can relieve dystonia for some, but the effect lasts only a couple of months.

I never had a career at the piano. But I have played Haydn's Sonata No. 50 more than 100 times over the past 20 years and at one point had even committed much of it to memory. What is unnerving me, in part, is the computer attached to the Yamaha Disklavier piano that will record just how I touch each key. Also unsettling are Riley's refereelike gaze and the video camera trained on my left hand. But mainly my trepidation is fed by a gloomy certainty that that sore hand will lag through the opening bars. As indeed it does: what should be a quiet, perfectly even motif of sixteenth notes comes out as a skewed, off-tempo jangle.



PIANO ROLL of author's performance shows notes that are inconsistent in spacing and duration (*top*) and subsequent improvement (*bottom*).

Riley, a music technologist and dystonia therapist at New York University, can help. By linking the instrumented instrument with software and a precisely synchronized video recording, she has turned the piano into a medical machine. The system captures the time and velocity at which each note is struck and released. Even more important, it captures the position of the performer's hands, arms and body. Bad habits—slouching, angled wrists, rigid forearms, raised elbows—can over years of playing contribute to focal dystonia.

"Athletes are coached about how to hold and move their bodies," Riley says. "But musicians rarely get that kind of instruction from their teachers. And unlike athletes, musicians tend not to warm up before practicing, take breaks to rest, or stretch out afterward." Riley, who so far has helped five musicians ease their dystonia, uses the computer's "piano roll" display of a performance to detect which fingers are cramping in certain passages. The synchronized video reveals unhealthy postures and overly tense muscles. Riley then coaches the musician to play in ways that allay the cramps.

After a minute of the Haydn sonata, for example, she stops me and rewinds the video. As the Disklavier replays my performance—the keys moving themselves, ghostly—she points to the monitor. "See how your left wrist drops?" It is half an inch lower than my right wrist, forcing the left hand to cock upward and its fingers to flatten. "Also, you are sitting much too close," she says. "Your left elbow and wrist are locked, so your forearm is full of tension."

Riley has me move the bench six inches, arch my back to shift forward my center of gravity, and straighten and raise my wrists so that the piano keys, rather than my sore joints, bear the weight of my arms. I play the passage again, and she brings up both record-

KATHLEEN M. RILEY *New York University*

ings on the laptop screen for comparison.

Blocks represent each note: the size of each block shows the note's duration; color indicates the velocity with which it was struck. After coaching, the notes are much more even and overlap slightly with one another, producing the desired legato sound. Note velocities, which initially ranged wildly from 32 to 62 (in arbitrary MIDI units), now all cluster close to 40.

Of course, it can take months or years to break bad habits developed over decades. But at least now there is a good tool for those who must try, in order to save their career. And the therapy may ease other kinds of repetitive stress injuries, Riley says. "One girl I worked with took what she learned at the piano and applied it to her typing and mousing. It relieved her carpal tunnel syndrome."

MEDICINE

Hormone Hysteria?

HORMONE REPLACEMENT THERAPY MAY NOT BE SO BAD

BY DENNIS WATKINS

Postmenopausal women have for decades relied on estrogen supplements to control the hot flashes, memory loss, osteoporosis and other ailments that can occur when their bodies no longer produce the compound. But hormone replacement therapy (HRT) is no longer considered the best way to treat menopause, ever since a report last year found that women receiving a certain type of HRT were at increased risk for dangerous side effects, such as breast cancer. Many health professionals have concluded that altering a woman's physiology will always increase risks over time. But a handful of respected scientists are calling for another look at HRT, arguing that not all therapies are created equal.

The largest blow to HRT appeared in the July 17, 2002, *Journal of the American Medical Association*. It presented important results of the Women's Health Initiative's long-term study of more than 16,000 women taking estrogen and a progesterone derivative. The study was halted prematurely, the authors reported, because too many women were encountering serious medical problems. "I believe that the drug we studied has more harms than benefits when used for the prevention of chronic diseases such as osteoporosis in generally healthy women," notes Jacques Rossouw, project officer of the initiative. In the past year a steady cascade of articles has enumerated all the higher risks that patients in the study experienced: an 81 percent increase in heart disease in the first year of therapy, a 24 percent increase in invasive breast cancer and a 31 percent increase in stroke. The therapy also doubled the risk of dementia. (A study of more than 800,000 women published in *Lancet* on August 9 also found an increased risk of breast cancer in postmenopausal women receiving a wide variety of HRT but noted that the risk of mortality from breast cancer related to HRT could not be determined.)

SAFETY OF DAILY DOSES of hormones for menopausal women may depend on the particular type of hormone combination used.

SATURN STILLS SPL/Photo Researchers, Inc.

The essential ingredient of hormone replacement therapy is estrogen. Taken alone and without interruption, however, estrogen causes cell division in the uterus, which in many women leads to uterine cancer. Women who have had hysterectomies can take estrogen by itself without fear of harmful side effects. (In fact, an estrogen-only arm of the Women's Health Initiative has continued because few participants have developed breast cancer.) For other women, though, the solution is to include a progestin, which blocks estrogen action in the uterus. Prempro, the Wyeth-manufactured drug used in the study, combines a cocktail of conjugated horse estrogens called Premarin with a synthetic derivative of progesterone called Provera, or medroxyprogesterone acetate. This pill, taken daily, was the most widely prescribed hormone replacement therapy drug in the U.S. when the initiative started during the 1990s.

For many scientists, a critical question yet remains: To what extent do the results of the initiative study apply to other forms of hormone replacement? "We cannot be sure whether other hormone combinations will have the same effects," Rossouw cautions, "but in my opinion we should assume they do until proven otherwise." But neuroendocrinologist Bruce S. McEwen of the Rockefeller University is unequivocally critical of the study: "I think that it borders on a tragedy that Premarin and Provera were chosen as the only HRT treatments."

A growing number of researchers believe that Provera is a poor substitute for progesterone. For example, medroxyprogesterone will bind in the breasts to progesterone receptors, which causes breast cells to divide after puberty and during the menstrual cycle, and also to glucocorticoid receptors, which causes cell division during pregnancy. This double-barreled assault on breast cells, explains C. Dominique Toran-Allerand, a developmental neurobiologist at Columbia University, probably led to the high rates of breast cancer in the study. "With Provera you are activating two receptors involved with cell division in the breast," she says, "and that's the culprit, not estrogen."

In addition, recent research shows

that Provera interferes with estrogen's ability to prevent memory loss and dementia. "Estrogen is able to protect neurons against toxic assaults that are associated with Alzheimer's disease," notes Roberta Diaz Brinton, a neuroscientist at the University of Southern California. Using in vitro studies of several types of progestin, she found that Provera—and no other progestin—blocks the mechanisms that allow estrogen to fight the brain's immune response to Alzheimer's. This immune response wears away at brain cells and causes them to leak neurotransmitters such as glutamate, which overloads and kills neurons. "It's basically as if someone were to open your mouth and shove down gallons" of soft drink, Brinton explains. "It's caustic, and you can't metabolize it enough."

Several researchers believe in the need for a study similar in scale to the Women's Health Initiative that tests hormones that more closely represent natural human hormones. Others suggest looking for better, more selective isotopes of the hormones. Until more research is completed, they agree, HRT deserves careful consideration.

GEOPHYSICS

Weight Watching

SATELLITE MAPS REVEAL THE VARIATIONS IN EARTH'S GRAVITY **BY DANIEL CHO**

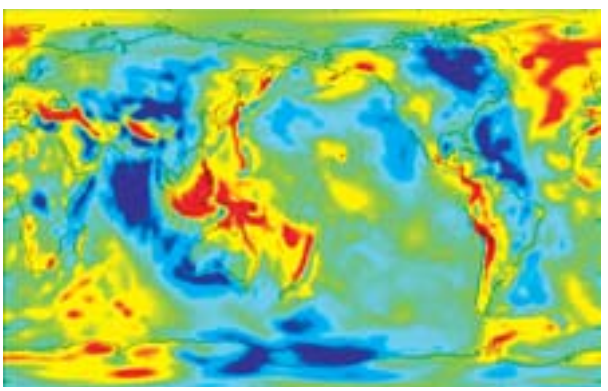
To lose weight fast, sail off the coast of Sri Lanka. That's one area of the earth where gravity is weakest, thanks to the deep Mid-Indian Basin. This geologic feature, among others—mountains, valleys, ridges, trenches and such—distributes mass unevenly about the planet's surface, thereby making the pull of gravity vary slightly. Recently a pair of satellites launched in March 2002 have yielded the most detailed map of the planet's gravitational field yet. These satellites make up the Gravity Recovery and Climate Experiment (GRACE), a joint effort by NASA and the German Aerospace Center.

The seas in particular draw interest because water sloshes around. "We can use these gravity measurements as a new way to track down water," says GRACE project scientist Michael Watkins. "It's exciting to have a new data type like that." At any given location, the gravitational field determines an ideal height where the water surface would rest if it weren't for winds, tides and other influences. By knowing this hypothetical rest position for the ocean, called the geoid, scientists can better understand how ocean currents behave, a key to predicting weather and climate. Gravity can also provide clues about underground water sources and help track the disappearance of the polar ice caps.

Previous maps of the earth's gravity relied on the tracking of dozens of individual satellites, launched into orbit for other purposes, and on ground-based sensors at a handful of locations. GRACE's maps are far more accurate than the patchwork models produced before—in some places nearly 100 times as accurate. The project will continue to release more refined maps throughout its five-year lifetime.

The secret to GRACE's accuracy lies in its twin satellites, which are spaced about 220 kilometers apart and maintain constant contact with each other through a microwave beam. The distance between them is measured to within one micron. As their orbit carries them through a nonuniform gravity field, the satellites either speed up or slow down. The difference in position and speed between the two indicates the strength of the gravity disturbances they encounter—a variation of about one part in 10,000 at most. So you would lose only a quarter of an ounce or so in the Indian Ocean. You may want to forget the sailing trip and buy a rowing machine instead.

GRAVITY ANOMALIES vary by about one part in 10,000 at most. In many spots, the disturbances are much smaller.



NASA/JPL/UNIVERSITY OF TEXAS CENTER FOR SPACE RESEARCH/GFZ, POTSDAM

The Progress of Love

AMERICANS ARE DISCARDING TABOOS AGAINST MIXED UNIONS **BY RODGER DOYLE**

Legal marriages between black and white people, which were rare before the Civil War, rose after emancipation, peaked about 1900, and declined until 1940. Beginning sometime after World War II, black-white marriages rose once again, but so slowly that by 2002, they accounted for only 0.7 percent of all marriages.

The taboo against black-white marriage built up over three centuries, first by the distinction between slave and free and later by segregation in the Jim Crow era. Laws against such intermarriage, which date to a 1691 Virginia statute, were declared illegal by the Supreme Court only in 1967. The historic legacy of stigma, together with scant opportunity to meet on an equal footing in offices, schools and neighborhoods, has been the main obstacle to black-white unions. Another has been the repudiation of intermarriage by black people who view it as an expression of racial disloyalty.

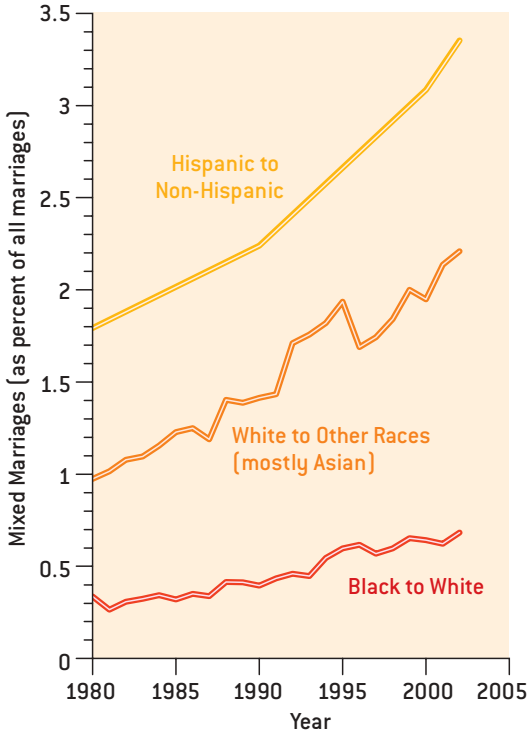
Asians were also stigmatized—for example, by the Chinese Exclusion Act of 1882—but they were subject to fewer antimiscegenation laws than were African-Americans. In 1963, for example, five states barred Asians from marrying whites, compared with 16 states barring blacks. As Asians assimilated and became better educated, intermarriage became unremarkable. Marriage between Japanese-Americans and whites has become so common that some observers believe that Japanese-Americans will eventually lose their distinctive ethnic identity. Native Americans, who have rarely been legally barred from intermarriage, have long had very high rates of marriage with whites. Hispanics have never been barred from intermarriage by law, a fact reflected in their high and increasing rates of marriage to whites.

Marriages between black males and white females are far more common than marriages between white males and black females. The opposite is true for Asian-white unions. In Hispanic-Anglo marriages, the wife is somewhat more commonly the Hispanic, whereas in marriages of Native Americans and whites, the husband is more likely to be white.

Official statistics on race are becoming increasingly meaningless. According to one estimate, up to 70 percent of Americans classified as black have a white ancestor; another estimate finds that as many as 21 percent of whites have African blood. When the husband is white and the wife Japanese, three quarters of the children are labeled white. If, by some miracle of genetic testing, the U.S. Census Bureau could establish the ancestry of every American, it would be apparent that the U.S. is much further down the road to a mixed-race society than most would imagine.

Rodger Doyle can be reached at rdoyle2@adelphia.net

Mixed Marriages



SOURCE: U.S. Census Bureau, Current Population Survey. Year-to-year variations may reflect sampling error.

RODGER DOYLE

news SCAN

AMERICAN WEDDINGS

MARRIED COUPLES, 2002: 57,919,000

Number of white-nonwhite marriages per 100 white-white marriages: 7.6

BLACK-WHITE MARRIAGES
Per 100 white-white marriages: 0.7
Per 100 black-black marriages: 8.1

HISPANIC-WHITE MARRIAGES
Per 100 white-white marriages: 4.0
Per 100 Hispanic-Hispanic marriages: 32.6

NATIVE AMERICAN–WHITE MARRIAGES
Per 100 white-white marriages: 1.2
Per 100 Native American–Native American marriages: 195.4

ASIAN-WHITE MARRIAGES
Per 100 white-white marriages: 1.6
Per 100 Asian-Asian marriages: 31.0

Note: Asian includes Pacific Islander.

FURTHER READING

Interracialism: Black-White Intermarriage in American History, Literature, and Law. Edited by Werner Sollors. Oxford University Press, 2000.

Interracial Intimacy: The Regulation of Race and Romance. Rachel F. Moran. University of Chicago Press, 2001.

Love's Revolution: Interracial Marriage. Maria P. P. Root. Temple University Press, 2001.

Interracial Intimacies: Sex, Marriage, Identity, and Adoption. Randall Kennedy. Pantheon Books, 2003.

DATA POINTS: RIP 'N' ROAR

On November 14, 2001, the Kunlunshan earthquake shook the Tibetan plateau in northwestern China. Centered about 440 miles from Lhasa, the powerful quake produced the longest surface rupture ever. Michel Bouchon and Martin Vallée of Joseph Fourier University in Grenoble, France, have analyzed the seismic data and conclude that the earth split with a velocity much faster than previously thought possible. The rupture started normally but then sped up after 100 kilometers.

Magnitude of Kunlunshan earthquake: 8.1

Length of surface rupture, kilometers: 400

Expected rupture speed, kilometers per second: 3 to 3.2

Calculated average rupture speed, kilometers per second: 3.7 to 3.9

Calculated maximum, kilometers per second: 5

Calculated maximum, miles per hour: 11,250

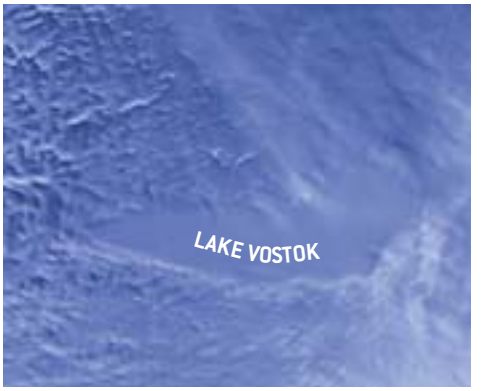
Cruising jetliner, miles per hour: 500

SOURCES: Science, *August 8, 2003; seismowatch.com*

BIOCHEMISTRY

Leaning Left

Amino acids twist to the left and right, but life has chosen only the lefties from which to build proteins. Experiments by Purdue University researchers suggest how the amino acid serine may have steered things eons ago. Unlike other amino acids, serine clumps strongly in groups of eight; moreover, it bonds only with those serine molecules exhibiting the same handedness (called chirality). The lefty serine clusters began bonding with other southpaw amino acids in the primordial soup, shutting out the righties. The left-handed groups also preferentially grabbed right-handed sugar molecules. Just why the left-handed serine won out over its righty counterpart—the chemical properties of amino acids remain unaffected by chirality—is still a mystery. The study is described in the August 4 international edition of *Angewandte Chemie*. —*Philip Yam*



SODA UNDER ICE: The subglacial Lake Vostok, imaged here by radar, might fizz.

GEOPHYSICS

Vostok Pop Top

Drilling into Lake Vostok, the alien environment nestled 2.5 miles below the Antarctic ice, could make it blow like a punctured soda can. Vostok is prized for its uniquely cold and isolated state (and the possibility that it harbors exotic microbes), which could serve as preparation for exploring the Jovian moon Europa. Last year Russian scientists announced plans to drill into the lake. Now, based on the gas content of the surrounding ice, a NASA team reports that every liter of Vostok water contains 2.5 liters of compressed gas, about the same pressure as an unopened soda. A borehole would have to be pressurized or allowed to refreeze, the group says; otherwise a geyserlike explosion would drain the lake and permit contaminants to enter. The July *Geophysical Research Letters* has more. —*JR Minkel*

PHYSICS

Missing: One-Quarter Hydrogen

When is water not H_2O? Why, when it's $H_{1.5}O$. In 1995 German and British physicists began bombarding water molecules with energetic neutrons to analyze proton behavior and found 25 percent fewer scattered neutrons than expected, suggesting that a quarter of the hydrogen nuclei (protons) became invisible. Although the result held up in benzene (C_6H_6) and hydrogenated metals, the group wanted to verify it by an independent technique. Now the scientists have struck a solid polymer called formvar ($C_8H_{14}O_2$) with electrons instead of neutrons, and the hydrogen gap remains. Theorists attribute the partial transparency to short-lived quantum entanglement between protons. In the less than 10^{-15} second required for scattering, a proton exists in a delicate, interconnected quantum state. The entanglement causes the proton to intefere destructively with itself, effectively wiping out a portion of itself. The scattering particles feel only the remaining portion. Details are in the August 1 *Physical Review Letters*. —*JR Minkel*

NOT SO for H_2O?

CANADIAN SPACE AGENCY (*top*); OSCAR BURRIEL *SPL/Photo Researchers, Inc.* (*bottom*); ILLUSTRATION BY MATT COLLINS

NANOTECH

Barrier-Free Nanotubes

A big challenge for any transistor is pumping electrons into it from a metal wire. Engineers overcome this so-called Schottky barrier in silicon semiconductors by replacing the metal wire with a strand of silicon doped with other elements. Now researchers at Stanford and Purdue universities have found a way around the Schottky barrier in semiconducting carbon nanotubes, which are difficult to dope in the required way. The scientists connected wide tubes (three nanometers in diameter) to palladium wires, which conduct readily and stick to nanotubes mysteriously well. The nanotubes could then carry about five times as much electricity as was previously possible—close to their theoretical ballistic limit (at which electrons travel without ricocheting off other particles). High currents are key to manufacturing high-powered computer chips. The work appears in the August 7 *Nature*. —*JR Minkel*

BRIEF POINTS

■ **A study of cadmium-resistant worms suggests that creatures that evolve a resistance to pollution lose that ability once the environment is cleaned up. The loss of resistance may therefore serve as an indicator of ecological recovery.**

Proceedings of the National Academy of Sciences USA, August 19, 2003

■ **A new vaccine protects macaque monkeys from the Ebola virus in four weeks with just a single shot—fast enough to contain outbreaks should the vaccine work in humans. Previous vaccines required multiple inoculations over several months.**

Nature, August 7, 2003

■ **The bacterium that causes gastric ulcers, *Helicobacter pylori,* remains active for years because the hole-puncturing toxin it secretes also blocks the proliferation of immune cells, thereby preventing the body from clearing the infection.**

Science, August 22, 2003

■ **Hanging with Mr. Cooper Pairs: Want to see stem cell debates live or hear scientists discuss climate change? Using C-SPAN as the model, researchers hope to develop a 24-hour cable network for science.**

www.csntv.org

MOSQUITOES spread the West Nile virus, but they may also harbor the vaccine.

VIROLOGY

A Shot against West Nile

In 1796 Edward Jenner discovered that infection with the relatively benign cowpox virus granted immunity against its fatal cousin smallpox. A similar strategy might work against the deadly, mosquito-borne West Nile virus. After scientists unraveled West Nile's genetic code, they learned that its sequence strongly resembles that of the Australian Kunjin virus, which is nonlethal and less debilitating, causing mostly fever and aches. Microbiologist Roy Hall of the University of Queensland in Brisbane and his colleagues injected mice with varying levels of Kunjin DNA that had weakened virulence. Investigators found that even 0.1 microgram of Kunjin DNA triggered antibodies against both Kunjin and West Nile and protected mice injected with lethal doses of West Nile. The results appeared online in the *Proceedings of the National Academy of Sciences USA* the week of August 11. —*Charles Choi*

ORIGINS

Rethinking Siberian Americans

Archaeologists have long held that the first Americans descended from prehistoric big-game hunters who tracked mammoths from Siberia across the Bering Land Bridge over millennia to colonize the New World. The latest carbon dating suggests that this idea could be wrong. The earliest people in the Americas are thought to be the Clovis (named after the location near Clovis, N.M., where fluted spear points characteristic of the population were first discovered) that date back 13,600 years. Evidence that the Clovis settlers came from Asia was unearthed at the Lake Ushki site in Russia, where charcoal in a grave there was previously dated back 16,800 years.

After reanalysis, researchers in the U.S. and Russia now find the Ushki charcoal is only about 13,000 years old, indicating that the two groups lived concurrently. If the Clovis descended from the Ushki ancestors, then they would have had to migrate down very rapidly, in less than four centuries. The observations are in the July 25 *Science*. —*Charles Choi*

CLOVIS ARTIFACTS from the first Americans.

TONY BRAIN *SPL/Photo Researchers, Inc. (top);* PETER BOSTROM *(bottom)*

Alchemy of a Supermetal

Serendipity delivers a process that may cut the cost of a high-tech material

By STEVEN ASHLEY

Titanium often ranks as the engineer's first choice as a structural material for jet aircraft, racecars, oil-drilling equipment or prosthetic body implants. And it's little wonder: titanium alloys are light and strong, as well as heat- and corrosion-resistant. The silvery-gray metal is pricey, however, compared with stainless steel and aluminum, a fact that limits its use. Scarcity is not the issue—titanium is the ninth most common element on earth—but the high cost of wresting the pure metal from the ore translates into expensive products.

TITANIUM SHEATHES the Guggenheim Museum Bilbao in Spain. If the metal were cheaper, designers and engineers would use it more frequently.

This past March the U.S. Defense Advanced Research Projects Agency (DARPA) tapped three materials research groups to address this persistent problem. Agency managers awarded separate contracts totaling $5 million to Titanium Metals Corporation (TIMET) and two others to fund parallel efforts to develop potentially low-cost production routes for titanium and its alloys.

Chemists have two ways to pry metal from an oxide ore. One, electrolysis, decomposes the ore into its elementary constituents with electricity. Aluminum manufacturing employs this method. The alternative, called chemical reduction, involves reacting the ore with a substance that has a greater affinity for oxygen than the metal to be extracted. This procedure is used to refine iron.

In current industry practice, titanium ore undergoes chemical reduction. But unlike iron ore, from which the oxygen is removed cheaply by reaction with carbon coke, extracting titanium requires a laborious, two-stage procedure. Plant technicians heat the ore in the presence of carbon and chlorine to create titanium tetrachloride, from which titanium is extracted by reacting the tetrachloride with magnesium. The result is titanium sponge, a porous form of the metal with salt compounds entrapped in the spaces. This process, invented by William J. Kroll in the late 1930s, has remained the chief titanium-refining route since industrial production began after World War II.

The Kroll process has drawbacks, however. The reducing agents for titanium are more expensive than coke. Kroll production is a batch process that requires the reaction vessels to be repeatedly emptied, refilled and sealed, rather than a continuous operation. And titanium tetrachloride is a volatile, corrosive liquid that requires special handling. In fact, soon after the Kroll process was introduced in the 1950s, its inventor reportedly predicted that an electrolytic process would replace it within 15 years. This shift never occurred, despite many attempts and millions in investment.

Thanks to a bit of serendipity, the most prominent of the electrolytic extraction techniques could eventually lead to cheaper titanium. In 1993 University of Cambridge metallurgists Derek J. Fray, Tom W. Farthing and George Zheng Chen were experimenting with electrolysis in an attempt to eliminate the oxide film that forms when titanium is exposed to air. The trio hoped

RALPH RICHER/ESTO/ARCHITEKTURPHOTO

that electrical flow through the titanium would pull the oxygen ions to the surface, where they could be removed. Instead the team observed an unexpected side effect: the process converted titanium oxides directly into the pure metal—an astonishing result.

In standard electrolysis, chemists dissolve the compound that is to be broken down in a conducting fluid called an electrolyte, which conveys charged ions from one electrode to another. For the electrolysis of titanium dioxide, metallurgists prefer an electrolyte of molten calcium chloride. Previous unsuccessful experiments along these lines relied on dissolving the tetrachloride (or the dioxide form) into the molten salts.

The Cambridge group's calculations showed, however, that it should be possible to reduce titanium dioxide electrolytically without having to dissolve it. The team used a cathode made of titanium dioxide. Other materials scientists had neglected to test this cell design because they believed that solid titanium dioxide—an insulator—could not be electrolyzed, Fray says. But the team's observations suggested that this electrolysis could in fact occur because titanium dioxide conducts electricity once some oxygen is taken out of the compound. When they tried it out, it worked. "It was shocking to see the little pellet of white titanium dioxide, which looks like an aspirin pill, being transformed into a piece of titanium," Fray recalls. "We sat around asking, 'Why hasn't this been done before?'"

The FFC Cambridge process converted titanium oxides directly into the pure metal—an astonishing result.

That the electrolysis converted oxide straight to metal would have gratified even a medieval alchemist. And if the process could be scaled up to industrial levels, the kind of riches for which alchemists always strove might be attainable. In addition to producing titanium more cheaply, the method might also work for other premium-priced metals, such as chromium and zirconium. Further, by forming the cathode from mixed metal oxide precursors, it might be possible to create titanium alloys in a single run, rather than via the conventional method of melting together alloy ingredients.

The U.K. defense ministry soon took notice of the Fray/Farthing/Chen process, which by then had come to be known as the FFC Cambridge method (after the inventors' initials and their employer). The U.K.'s military research agency licensed the technology and supported the team's investigations until 1998, when a company—British Titanium—was established to sublicense and commercialize the technology. This step led to a pilot plant that produces kilogram-size quantities of titanium.

The high costs of fully developing the commercial process hindered further progress for several years, however, despite considerable attention from industry. After gauging DARPA's interest in providing funds for the development of cheaper refining routes, TIMET sub-sublicensed the technology and proposed leading a U.S. government-funded R&D project. By March, DARPA had opted to invest in the FFC Cambridge approach.

The TIMET research syndicate will include scientists from defense contractors General Electric Aircraft Engines, United Defense Ltd. Partners, and Pratt & Whitney, as well as experts from Cambridge and the University of California at Berkeley. "By the third quarter of 2004, we're to have demonstrated a process capable of producing 50 pounds of metal a day," says Stephen Fox, U.S. director of research at TIMET. Success in that effort could lead to further DARPA money to subsidize a scale-up to 500 pounds a day and eventually to commercial levels measured in tons a day.

Fox points to the possibility of great payoffs if FFC Cambridge–based manufacture of titanium and its alloys can be achieved. "The process offers control over the resulting product form—powdered metal or pieces of sponge in a tailored size range," he says. "These could feed right into existing part-manufacturing processes, or, potentially, traditional remelting and fabrication steps might be avoided by directly consolidating the metal into a near-finished form."

Many hurdles still exist. Fox lists as key the detailed engineering of the cell for operations on a mass scale, the development of process controls, and the manner by which the reactive precursors and final materials are transported in and out of the cell. "Much remains to be done to get costs down," he notes.

Not everyone is optimistic. "I take a skeptical view of these efforts because these new process technologies are extremely high risk and very costly to develop," comments Firoze E. Katrak, metallurgist and metals market analyst at Charles River Associates. Near-term price reductions of titanium sponge, he believes, may come more readily from converting the Kroll process from a batch process into a semicontinuous one. So the debate about the best way to achieve low-cost titanium persists. But the big payoffs—such as making a next-generation airliner or a less weighty SUV—ensure that the quest will continue. SA

Kick Me, Myself and I

An inveterate tinkerer creates a technology for self-flagellators By GARY STIX

The independent inventor is a symbol of American ingenuity who can justly claim credit for creations such as the photocopier and the implantable cardiac pacemaker. But this archetypal figure, whose ranks receive nearly one in five patents issued by the U.S. Patent and Trademark Office, may sometimes be more Rube Goldberg than Thomas Edison.

Take Joe Armstrong. The 70-year-old shares traits of both utilitarian and prankster. While residing in Phoenix, Armstrong made a living for almost 25 years off a patented invention that he devised for mounting aluminum-coated Mylar screens in car and truck windows to deflect the brutal Southwestern sun. Elvis Presley's Cadillac and seven of Imelda Marcos's Mercedes were adorned with the screens, according to Armstrong.

When he retired to Tennessee in 1989, he set up a shop in his garage in Lenoir City, about 30 miles southwest of Knoxville. Armstrong has never been able to suppress an impulse to tinker. "If I see something mechanical, I always study how I could make it work better," he says. A fan of the sports teams of the University of Tennessee, a school where he spent two and a half years in the 1950s, he would often hear the expression "to kick butt." He would also hear athletes and even ordinary mortals mutter to themselves, "I'm so sorry I did that, I could just kick my own butt."

Armstrong marveled at the anatomical impossibility of this saying. Inspired, he set about to try to rectify technologically this evolutionary shortfall, even applying for a patent on what he ultimately invented. In 2001 the patent office issued patent number 6,293,874 for a "User-Operated Amusement Apparatus for Kicking the User's Buttocks." The self-flagellator consists of a crank positioned at a level that the user must bend over to reach. The patent notes that the crank connects by a drive train to a post behind the user that has four rotating arms "with a central axis of the rotating arms positioned at a height generally level with the user's buttocks.... As the user rotates the crank, the user's buttocks are paddled by flexible shoes located on each outboard end of the elongated arms to provide amusement to the users and viewers of the paddling." Armstrong has clad the four arms with footwear ranging from cowboy boots to clown shoes, depending on the occasion. The adjustment of the machine ensures that the impact does not actually hurt the user.

Smokey the hound dog, the mascot of the University of Tennessee, has deployed the butt kicker to taunt fans of rival Vanderbilt at a basketball game, beckoning them to descend from the stands for posterior stimulation. "It was lucky we won that game; otherwise we really would have been embarrassed," Armstrong says. He has sold several machines for $600 to $800, including one to an amusement park in Blackpool, England, and another to a Christian fun park in North Carolina. The latter requested that labels on the machine that used the word "butt" be changed to "rear."

The inventor once traveled to Portland, Ore., to appear on a local television show in which the host enticed passersby on the street to give themselves the boot on camera. Recently he crafted a five-foot-tall Uncle Sam who, when placed on the machine, appears to turn the crank (operated remotely by Armstrong), which results in an effigy of Osama bin Laden getting kicked in the bum.

Armstrong's was not the first invention to target the seat of the pants. A patent search turned up a device used for initiation rites dating back as far as 1900 that shocked and spanked its user. Nevertheless, the patent office apparently acknowledged Armstrong as helping to advance the state of the art for self-administering a good swift kick to the behind. SA

JENNIFER KANE

Remember the Six Billion

For millennia we have raged against the dying of the light. Can science save us from that good night? By MICHAEL SHERMER

Between now and the year 2123 a tragedy of Brobdingnagian proportions will befall humanity, causing the death of more than six billion people. I'm serious.

According to Carl Haub, a demographer at the Population Reference Bureau in Washington, D.C., between 50,000 B.C. and A.D. 2002, about 106 billion people were born. Earth's population is currently around 6.3 billion. Of the approximately 100 billion people born before us, every one has died. To the extent that the past is the key to the future, that means that within the next 120 years (today's maximum life span), more than six billion humans will suffer the same fate. And there is not a damn thing we can do about it. Or is there?

In the next 120 years, six billion people will die.

For most of our history, humans could turn only to prayer and poetry to help cope with this reality. Today we are offered scientistic alternatives—if not for immortality itself, then at least for longevity of biblical proportions. All have some basis in science, but none has achieved anything like scientific confirmation. Here is a short sampling, from the almost sublime to the near ridiculous:

Virtual immortality. According to Tulane University physicist Frank J. Tipler, in the far future we will all be resurrected in a virtual reality whose memory capacity is 10 to the 10^{123} bytes. If the virtual reality were good enough, it would be indistinguishable from our everyday experience. Boot me up, Scotty. One problem, among many, is that Tipler's resurrection machine requires so much energy that the universe must one day collapse, which present data show is not going to happen.

Genetic immortality. Oh, those pesky telomeres at the ends of chromosomes that prevent cells from replicating indefinitely. If only we could genetically reprogram normal cells to be like cancer cells. Alas, this is no solution, because biological systems are so complex that fixing any one component does not address all the others that play a role in aging.

Cryonics immortality. Freeze. Wait. Reanimate. It sounds good in theory, but you're still a corpsicle. And when your tissue is thawed, your cells will be mush. Don't forget to pay the electric bill in the meantime.

Replacement immortality. First we replace our organs (which today are often rejected), then our cells and molecules nano-a-nano (not yet technologically feasible), eventually exchanging flesh for something more durable, such as silicon. You can't tell the difference, can you?

Lifestyle longevity. Because this is a goal we can try to implement today, the hucksters are out in force offering all manner of elixirs to extend life. To cut to the chase, S. Jay Olshansky, Leonard Hayflick and Bruce A. Carnes, three leading experts on aging research, have stated unequivocally in the pages of this magazine that "no currently marketed intervention—none—has yet been proved to slow, stop or reverse human aging, and some can be downright dangerous" ["No Truth to the Fountain of Youth," SCIENTIFIC AMERICAN; June 2002].

It has never been satisfactorily demonstrated, for example, that antioxidants—taken as supplements to counter the deleterious effects of free radicals on cells—attenuate aging. In fact, free radicals are necessary for cellular physiology. Hormone replacement therapy, another popular antiaging nostrum, helps to counter short-term problems such as loss of muscle mass and strength in older men and postmenopausal women. But the therapy's influence on the aging process is unproved, and the long-term negative side effects are unknown.

As a lifelong cyclist, I am pleased to report that proper diet and sufficient exercise are tried-and-true methods of increasing the length of your life. These, along with modern medical technologies and sanitation practices, have nearly doubled the average lifetime over the past century. Unfortunately, this just means that more of us will get closer to the outer wall of 120 years before inexorably succumbing to the way of all flesh.

As 20th-century English poet Dylan Thomas classically admonished, "Do not go gentle into that good night . . . / Rage, rage against the dying of the light." Rage all you like, but remember the six billion—and the 100 billion before. Until science finds a solution to prolonging the duration of healthy life, we should instead rave about the time we have, however fleeting. SA

Michael Shermer is publisher of Skeptic *(www.skeptic.com) and author of* How We Believe.

BRAD HINES

Cleaning Up after War

Bombs and bullets can kill years after the battles have ended, by leaving behind toxins and contaminants. It's up to Pekka Haavisto to figure out how to handle the mess By MARC AIRHART

During its springtime assault against Saddam Hussein, the Pentagon played videos showing the deadly precision of U.S. weaponry. Guided by satellites and lasers, missiles found their targets without hitting nearby buildings. Yet even if civilians were spared, they could face dangers from spent munitions. For many weapons, U.S. forces have for the past two decades relied on depleted uranium, which, being nearly twice as dense as lead, can penetrate materials more effectively than conventional alloys can.

PEKKA HAAVISTO: POSTCONFLICT FIXING

- Toured Europe at age 15 via a 25-nation train pass. "Traveling taught me to understand a country's culture and history. When offering solutions to the environmental problems, different traditions have to be understood."
- On his job: "One third is lobbying, one third is fund-raising, and one third is the real environmental work."
- Depleted uranium used in battles against Iraq since 1991: 400 to 450 metric tons. (Estimate by Dan Fahey, an independent policy analyst in Berkeley, Calif.)

The metal, a by-product of uranium enrichment for nuclear power plants and warheads, is toxic when ingested and slightly radioactive, and that worries Pekka Haavisto. "Do you think that people in the postconflict situation are somehow harder people and they can take more burden?" Haavisto asks. "Or do you think that they are human beings like us, and whatever you can avoid, you should avoid?"

It's clear what his answer would be. The 45-year-old Finn chairs the Geneva-based Post-Conflict Assessment Unit (PCAU), a division within the United Nations Environment Program. His team goes to places where conflicts have just ceased, looks for environmental trouble spots and sets priorities for cleanup and reconstruction.

The PCAU began in 1999 following the war in the Balkans (it was known then as the Balkan Task Force). Some of the NATO bombings resulted in the release of toxic chemicals. The executive director of the U.N. Environment Program, Klaus Toepfer, needed someone to determine the severity of the war pollution. He remembered that, while serving as a German environmental official, he had met a young environment minister from Finland who was enthusiastic and well respected. "So I came to the conclusion that this would be a great chance to bring Pekka Haavisto on board," Toepfer recalls.

Haavisto had recently finished his term in office and was considering returning to environmental journalism when Toepfer called. "And of course that was an opportunity to which you could not say no," Haavisto says. "And I arrived to an empty room with nobody to help me that first day."

Haavisto, who cofounded the Green Party in Finland, pulled together 60 experts from around the world. Through that summer and fall, the team searched for toxic or radioactive pollution in river sediments, groundwater, soil and air. In the end, they concluded that the war had not resulted in an environmental catastrophe.

PHOTOGRAPH BY KATE BROOKS, BAGHDAD, AUGUST 11, 2003

But they found four "hot spots"—industrial sites where pollution posed a threat to human health. Since then, most of the necessary cleanup has been completed. "After Kosovo came the Serbia work and then the Bosnia work," Haavisto says. "Then we were asked to do similar work in the occupied Palestinian territories and Afghanistan and now just lately in Iraq. I don't know when I'm returning home to Helsinki."

At first, U.N. member nations were skeptical about the need for assessing a postconflict environment. "People were always saying, 'Well, why are you coming with the environmental portfolio? We have a humanitarian crisis, we have the refugees, and we have social issues and the schools,' and so on," recalls Haavisto, who talks virtually nonstop at times. But if you don't take care of the environment immediately, before reconstruction, Haavisto points out, it will be much costlier later. Plus, contaminants may prolong the suffering of people. "And I'm quite convinced that this is the approach that the international community should have in each and every region and after each and every conflict," he insists.

Larger, more chronic issues persist in places such as Afghanistan, where more than 20 years of fighting has taken its toll. Land mines continue to kill people and animals. Clean drinking water is in short supply because of drought, contamination from poorly located dump sites, past bombings and even simple neglect. Biodiversity loss and deforestation add to the environmental woes.

SMOKE from oil fires around Baghdad and other wartime pollution could create long-term health hazards.

Haavisto's latest project is an assessment of Iraq. In a perfectly safe region, Haavisto and his PCAU team would need three months to complete the fieldwork and another two months to analyze the samples. Haavisto had hoped to be in Iraq by June, but frequent attacks on U.S. troops have delayed his efforts until August. He says that assessing Iraq will cost about $850,000, much of it from the Humanitarian Flash Appeal, a relief fund to which U.N. countries are asked to contribute.

Of major concern is the depleted uranium of some ammunition. When such a projectile makes impact, a bit of the uranium gets pulverized, turning into airborne radioactive dust that could be dangerous to breathe. Fragments of depleted-uranium weapons sitting on the ground can corrode and leach into the soil and groundwater. But the public health dangers of depleted uranium in the environment are not fully known. Some argue that it causes birth defects, cancers and Gulf War Syndrome. Military experts counter that no conclusive evidence links it to disease. But that may have more to do with the relatively recent use of the material and the lack of actual studies.

In any case, the PCAU team has begun mapping the areas exposed to the metal. Haavisto explains that the British government was providing information on where depleted-uranium ammunition had been used in southern Iraq. But the U.S. military was so far not helping in this regard. Distinguishing which depleted-uranium contamination resulted from this year's bombings and which from the 1991 Gulf War may also be hard.

Uranium is just one of several hazards in postwar Iraq. Haavisto's team will undoubtedly find that some industrial and military targets released toxic chemicals into the air, soil and water. The black smoke from burning oil trenches around Baghdad, meant to shroud targets, contained many toxic substances that might affect the soil and drinking water.

In addition, Haavisto expects to find a disaster in the Mesopotamian marshes: the nourishing water that once made this area the Fertile Crescent has been dammed up and siphoned off by the ousted regime. "It has not only influenced or affected the biodiversity but also the livelihoods and the situation of the marsh Arabs," he says.

Ironically, one of the biggest environmental problems in Iraq may stem not from direct military conflict but from a decade of U.N.-imposed sanctions. Haavisto explains that as replacement parts became harder to acquire, proper maintenance of oil drilling and production facilities became more difficult. When pipelines developed leaks, they were simply ignored, paving the way for widespread contamination of soil and groundwater.

Besides pointing out the problems, each assessment recommends specific solutions. In certain cases, it might mean just removing contaminants from soil in a certain place. In others, it might mean creating an entirely new administrative infrastructure for monitoring wildlife or habitats.

Other nations have begun seeing the value of environmental assessments. Tanzania wants an evaluation of the impact of refugees on the country. After years of civil strife, Somalia, Ivory Coast and Congo badly need this kind of appraisal. There is no shortage of work, yet "I still have a one-month contract," Haavisto remarks. "People are always asking, 'When are you finished?' And I say that I'm finishing every month on the 11th." For nearly five years, that contract has been renewed, fortunately—or perhaps, unfortunately. Remarks Klaus Toepfer: "We were still optimistic enough to believe that postconflict assessment would not be something like a growing market." SA

Marc Airhart is a producer for the Earth and Sky *radio series in Austin, Tex. Daniel Cho contributed to the reporting.*

COURTESY OF SPACE IMAGING

The Unexpected Youth of Globular Clusters

Conventional wisdom says that globular star clusters are the stodgy old codgers of the universe, but it turns out that many of these clusters are young

By Stephen E. Zepf and Keith M. Ashman

THROUGHOUT MOST OF A GALAXY, STARS ARE SPRINKLED LIKE FARMHOUSES on an open plain. Separated by vast distances, they lead their lives almost independently of one another. Some areas of a galaxy, though, look more like cities than countryside. These are globular clusters—groups of a million or so stars crammed into a volume that elsewhere would accommodate barely a single star. Not only are they congested, they are aged. The roughly 200 globular clusters in our Milky Way galaxy contain some of the oldest known stars in the universe; young inhabitants are nowhere to be seen. Accordingly, astronomers have generally seen globulars as ancient cities, like the historic districts of Rome or Istanbul, formed long ago and little changed—cramped, worn-out municipalities that tell us much about bygone times but little about the way modern galaxies organize themselves.

RON MILLER

COLLISION OF TWO GALAXIES is one of the most momentous events in the universe. This artist's conception shows a collision from the vantage point of an old globular cluster (*foreground*) on the outskirts of one of the doomed galaxies. Astronomers used to think that all globular clusters were old, having formed early in cosmic history, but recent findings indicate that clusters continue to form even today—in events such as this collision.

NGC 1316: Postcollision galaxy

It is as though a MARTIAN CAME TO VISIT EARTH and landed

At least that is what astronomers used to think. Lately they have been busy rewriting the galactic charts (and textbooks). With the penetrating vision of the Hubble Space Telescope, they have seen the lights of new burgs—newfound cities full of bustling activity. Globulars can and do continue to form, apparently when galaxies smash into one another. These discoveries give researchers a handle on key questions in astronomy, such as how and when massive galaxies originate and evolve.

For most of the past century, astronomers have known that globular clusters are strikingly ubiquitous, found within nearly all galaxies. With few exceptions, all the stars in an individual cluster appear to have a very similar age and composition. Globulars thus appear to result from a stellar baby boom—the formation of a large number of stars at the same time within a small region. Within the Milky Way, that process coincided with the formation of the galaxy itself, making these clusters ideal for probing the age of the universe and the galactic formation process [see "Rip Van Twinkle," by Brian C. Chaboyer; SCIENTIFIC AMERICAN, May 2001; and "How the Milky Way Formed," by Sidney van den Bergh and James E. Hesser; January 1993]. Interestingly, globular clusters in the Milky Way are scattered throughout a spherical volume that extends far beyond the flattened disk where most stars lie. This sphere presumably traces the shape our galaxy used to have. Therefore, the galaxy must have collapsed significantly during its creation and early evolution.

The old age typical of the globular clusters in our galaxy has had a profound impact on the study of globulars in general. It is as though a Martian came to visit Earth and landed in an old-age home where everyone was an octogenarian. Not knowing any other people, the Martian might conclude that human birth had ceased 80 years ago. In the early 1990s this was essentially the situation facing astronomers. Most models that tried to explain the origins of globular clusters focused on conditions in the early universe and assumed that the process was a thing of the past.

Overview/*Great Balls of Stars*

- Globular clusters are extremely dense groupings of stars that contain a few hundred stars per cubic light-year, compared with about 0.01 per cubic light-year in the sun's neighborhood. Astronomers once believed that they came in only one variety—ancient. Stars in the clusters are so old that researchers use them to place limits on the age of the universe.
- But recent results show that some clusters are young or middle-aged. Rather than emerging in a single, primordial growth spurt, globulars have come into being over nearly the entire history of the universe. They form whenever conditions within galaxies are sufficiently violent. They are thus providing insights into what happens when lumbering galaxies collide and merge.

Merge a Galaxy, Make a Cluster

WE BEGAN OUR WORK in this field a decade ago by suggesting a somewhat different model. Specifically, we, along with François Schweizer of the Carnegie Observatories in Pasadena, Calif., and others, independently proposed that globular cluster formation need not, in fact, have ended. Perhaps the lack of younger clusters was simply an artifact of limited observations. The globulars in our galaxy might all be old, but what about those in other galaxies? Maybe globulars go on forming whenever conditions permit.

That might happen during the collision of two spiral galaxies, which have lots of spare gas available for making new stars. Although galaxies are separated by large distances, they have been observed to collide. Needless to say, dramatic changes can result. Mergers of two colliding galaxies often cause starbursts, a huge but temporary increase in the rate of star formation. Collisions also can transform the shape of galaxies involved; specifically, the merger of two spiral galaxies may produce a single elliptical galaxy.

Astronomers used to doubt that galactic collisions could turn spirals into ellipticals. Their skepticism was fed by the observation that ellipticals have more globular clusters than spirals of the same mass; the simple addition of two spirals should keep the number of globular clusters per galaxy mass the same. But this excess of globulars is explained if the merger itself created globular clusters.

To verify our model, observers had to look at gas-rich galaxies to see whether they indeed contained recently formed globular clusters. Those observations were made possible by the Hubble Space Telescope. Unhampered by Earth's atmosphere, Hubble could resolve distant galaxies—previously seen as little more than a blur—into individual star clusters. When Hubble examined galaxies undergoing starbursts, it revealed massive, dense, young star clusters. These clusters had sizes and inferred masses similar to those of globular clusters in the Milky Way. They formed only in regions of intense starburst activity. Therefore, galactic collisions may lead to the birth not just of new stars but also of new globular clusters.

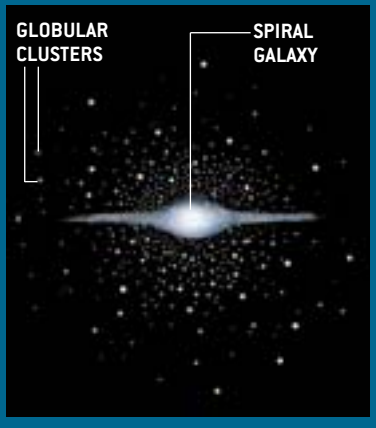
As compelling as these images were, they did not close the case. How could we be sure that these globularlike clusters really were younger versions of the globulars in our own galaxy? Further evidence linking these two sets of systems was required. To return to our earlier analogy, if the Martian became accustomed to a world exclusively populated by octogenarians and then was suddenly presented with a group of babies, he might

C. GRILLMAIR (IPAC/CALTECH) ET AL., WFPC2, HST, NASA

in an old-age home where EVERYONE WAS AN OCTOGENARIAN.

LIKE DIAMONDS IN THE SKY

GLOBULAR CLUSTERS ARE A FAVORITE observing target of both amateur and professional astronomers. Two years ago the Hubble Space Telescope took this picture of NGC 1850, a young globular cluster found in the Large Magellanic Cloud, a sidekick galaxy of our Milky Way (*right*). In general, globular clusters are scattered around a spiral galaxy in a vast sphere, which astronomers call a halo. Most other stars reside in a flattened disk (*below*), which also contains the pinwheel-like arms (*not visible in this edge-on view*).



MARTINO ROMANIELLO *European Southern Observatory,* EUROPEAN SPACE AGENCY, NASA (*Hubble image*); LUCY READING (*diagram*)

ask for proof that these babies and the senior citizens were even the same type of creature.

Young and Old Are the Same

ONE WAY TO DEMONSTRATE a link would be to identify two populations of globular clusters in old elliptical galaxies. If such galaxies formed by the merger of spirals, they should contain old globular clusters (those originally associated with the spirals) as well as younger ones (created during the merger). In the specific model we developed, elliptical galaxies have roughly equal numbers of each type.

We also suggested that the two types could be distinguished by their colors. The stars in newer globular clusters should have a higher fraction of heavy elements than stars in older globulars, because they formed out of gas that had become "polluted" with heavy elements, which are made as stars die in supernova explosions and other events. Indeed, the old globular clusters in spiral galaxies such as our own Milky Way are known to be mostly poor in heavy elements. The fraction of heavy elements determines the color of a star. For two stars of the same age, the one with the higher fraction of heavy elements will be redder because of the way nuclear fusion operates within stars. Fusion at the center of the star generates radiation, which is absorbed by gas inside the star, producing an outward pressure. This pressure is what prevents the star from collapsing under its own weight. When more heavy elements are present, the gas absorbs radiation more efficiently, allowing the star to maintain its balance at a lower temperature. The lower temperature in turn causes a redder color.

Since we made this prediction, Hubble and ground-based telescopes have analyzed the colors of the globular clusters in elliptical galaxies. Definitive results by Brad Whitmore of the Space Telescope Science Institute, Arunav Kundu—now a postdoctoral fellow working with one of us (Zepf) at Michigan State University—and Søren Larsen and Jean Brodie of the Uni-

THE AUTHORS

STEPHEN E. ZEPF and *KEITH M. ASHMAN* began working together on globular cluster systems when they were both in Baltimore, Zepf as a graduate student at Johns Hopkins University and Ashman as a postdoctoral researcher at the Space Telescope Science Institute. Zepf is now a professor of physics and astronomy at Michigan State University; he worked elsewhere in the U.S. and in Britain before returning to the Midwest where he grew up. Ashman, a native of London, recently joined the physics faculty of the University of Missouri–Kansas City. He lives in Lawrence, Kan., with his cat, Zoot, and spends his free time gardening and playing bass guitar.

A BIRTHING GROUND OF GLOBULAR CLUSTERS

SPIRAL GALAXIES that get a little too close to each other collide and merge, as this simulation depicts. Our own Milky Way is destined to collide with the nearby Andromeda galaxy in several billion years' time. The two galaxies bring along their ancient globular clusters—

HUBBLE HERITAGE TEAM (AURA/STScI/NASA) (*top left*); JOHN DUBINSKI *University of Toronto* (*collision simulation*)

versity of California at Santa Cruz have verified the bimodal picture. Most of these systems show clear evidence for two distinct populations of globular clusters, one blue and one red. This finding corroborates the link between the merger process and the formation of globulars.

Another way to link the young and old globulars would be to find clusters of intermediate age—the "thirtysomethings." The observational challenge is that these intermediate-age systems are not so obviously different from older systems. When globular clusters are young, they are bright, reflecting the properties of the massive stars they contain. These massive stars evolve quickly and die quickly, and stellar evolution proceeds more slowly after that. As a result, the differences between a globular cluster of a thirtysomething age and one of "senior citizen" age are subtle and hard to distinguish from the heavy-element effect described above.

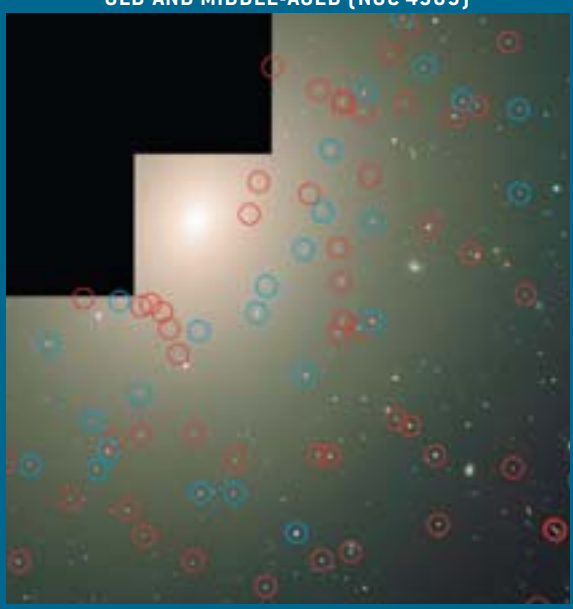
Recently astronomers have discovered these intermediate-age globular clusters in several elliptical galaxies. A group including Zepf and astronomers Markus Kissler-Patig of the European Southern Observatory in Germany and Thomas Puzia of the Munich University Observatory uncovered a large, intermediate-age globular cluster population in a fairly ordinary elliptical galaxy. Meanwhile studies by Schweizer, Whitmore, Paul Goudfrooij of the Space Telescope Science Institute, and others focused on ellipticals that were suspected of being thirtysomething galaxies because of their slightly perturbed appearance—which indicates that they had not reached full maturity—and the presence of some younger stars. Detailed studies confirm that they also host intermediate-age globular clusters.

THREE GENERATIONS OF GLOBULARS

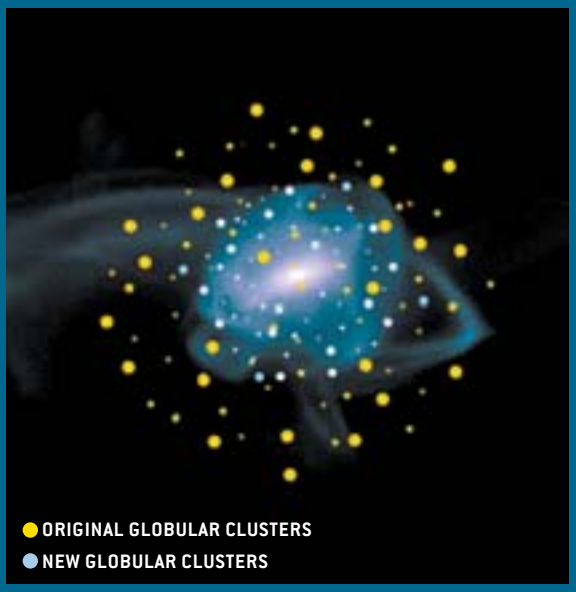
THE GLOBULAR CLUSTERS in the Milky Way galaxy, such as M80 (*right*), are ancient—the stars in the cluster are nearly as old as the universe itself. But the globulars in the elliptical galaxy NGC 4365 (*center*) come in two varieties: ancient (*red circles*) and intermediate age (*blue circles*). Those in the Antennae, a pair of galaxies—NGC 4038 and 4039 (*far right*)—currently undergoing a collision, are even younger (*inset*).

OLD (M80)

OLD AND MIDDLE-AGED (NGC 4365)



those that formed around the same time as the galaxies themselves (*yellow circles*). The galaxies swoop around each other until settling down into a single galaxy with an elliptical, rather than spiral, shape. The collision raises the gas pressure within the galaxy, triggering the formation of new globular clusters (*blue circles*). Astronomers have indeed observed elliptical galaxies with two distinct populations of globulars.



Dense Surprises

THUS, GLOBULAR CLUSTERS of all ages have now been observed, ranging from the youngest systems in ongoing mergers a few to hundreds of millions of years old, to intermediate-age clusters a few billions of years old, to the old clusters about 12 billion years old, which were once the only kind known.

The discovery of young and intermediate-age globular clusters opens up a whole new way to study how globulars form. Instead of being objects that arose only in the murky past, they can now be studied up close and in detail. This work has led to much progress and a few surprises.

Gas in starbursts is observed to have high pressure, about 100 to 1,000 times higher than typical pressures in spiral galaxies. The high pressure compresses the gas to the density necessary to form stars. That explains why star formation in the disk of our galaxy no longer produces globulars: the pressure has dropped too low. Starbursts are not the only way to generate high pressure. An alternative mechanism is that star formation early in the history of the universe heated gas clouds in dwarf and spiral galaxies, thereby boosting pressures. That could be why dwarf galaxies and the outer regions of spirals contain globular clusters even though they have not undergone massive mergers. (If they had undergone such mergers, they would no longer be dwarfs or spirals.)

One of the important surprises is that the size of a recently formed globular cluster appears to be unrelated to its mass. That is, the more massive clusters are not any larger; they are just denser. This fact is unexpected. For gravitationally bound objects such as ordinary stars and rocky planets, more massive objects are larger. The best current explanation is that lightweight clusters start off smaller than heavier ones but then lose some of their stars. The mass loss weakens the gravitational glue that holds the clusters together, so the clusters expand in size.

But perhaps the most interesting implication is that globulars can serve as a tool to trace what has transpired in the cosmos since the big-bang expansion began. The oldest globulars are fossils of the universe at the earliest epochs of star and galaxy formation, and the younger clusters reflect the evolution of galaxies right to the present. Globulars transport us mentally through time, opening our eyes to the rich layering of history in the heavens. SA

HUBBLE HERITAGE TEAM (*M80*); EUROPEAN SOUTHERN OBSERVATORY (*NGC 4365*); BRAD WHITMORE *Space Telescope Science Institute* (*NGC 4038/4039*)

MORE TO EXPLORE

Globular Cluster Systems. Keith M. Ashman and Stephen E. Zepf. Cambridge University Press, 1998.

Star Clusters. Bruce W. Carney and William E. Harris. Springer-Verlag, 2001.

The Young, the Old, and the Globular: Special Section on Globular Clusters. *Science,* Vol. 299, pages 59–75; January 3, 2003.

Formation of Globular Clusters in Merging Galaxies. François Schweizer in *New Horizons in Globular Cluster Astronomy.* Edited by Giampaolo Piotto, Georges Meylan, George Djorgovski and Marco Riello. Astronomical Society of the Pacific conference series, Vol. 296; 2003. Available online at **arXiv.org/abs/astro-ph/0212243**

ROBOTIC ARM driven by electroactive polymers may eventually be pitted against a human's in an arm-wrestling match.

Artificial Muscles

By Steven Ashley

Novel motion-producing devices—actuators, motors, generators—based on polymers that change shape when stimulated electrically are nearing commercialization

It's only a $100 toy—an aquarium of swimming robotic fish developed by the Eamex Corporation in Osaka, Japan. What makes it remarkable is that the brightly colored plastic fish propelling themselves through the water in a fair imitation of life do not contain mechanical parts: no motors, no drive shafts, no gears, not even a battery. Instead the fish swim because their plastic innards flex back and forth, seemingly of their own volition. They are the first commercial products based on a new generation of improved electroactive polymers (EAPs), plastics that move in response to electricity.

For decades, engineers who build actuators, or motion-generating devices, have sought an artificial equivalent of muscle. Simply by changing their length in response to nerve stimulation, muscles can exert controlled amounts of force sufficient to blink an eyelid or hoist a barbell. Muscles also exhibit the property of scale invariance: their mechanism works equally efficiently at all sizes, which is why fundamentally the same muscle tissue powers both insects and elephants. Something like muscle might therefore be useful in driving devices for which building tiny electric motors is not easily accomplished.

EAPs hold promise for becoming the artificial muscles of the future. Investigators are already ambitiously working on EAP-based alternatives to many of today's technologies. And they aren't afraid to pit their creations against nature's. A few years ago several individuals, including Yoseph Bar-Cohen, a senior research scientist at the Jet Propulsion Laboratory (JPL) in Pasadena, Calif., posted a challenge to the electroactive polymer research community to drum up interest in the field: a race to build the first EAP-driven robotic arm that could beat a human arm-wrestler one on one. Later, they began searching for sponsors to subsidize a cash prize for the winner.

Perhaps the most promising of the current EAP efforts is being conducted by SRI International, a nonprofit contract-research laboratory based in Menlo Park, Calif. Within a few months, SRI management hopes to complete arranging the initial $4 million to $6 million in investment funding needed to launch a spin-off company—tentatively called Artificial Muscle Incorporated—to commercialize the EAP technology it has patented. Even now, SRI is working on half a dozen R&D contracts with the U.S. government and companies

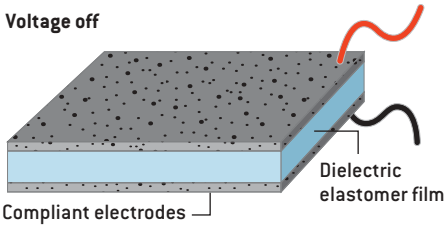
HOW ELECTRICITY MAKES A PLASTIC EXPAND

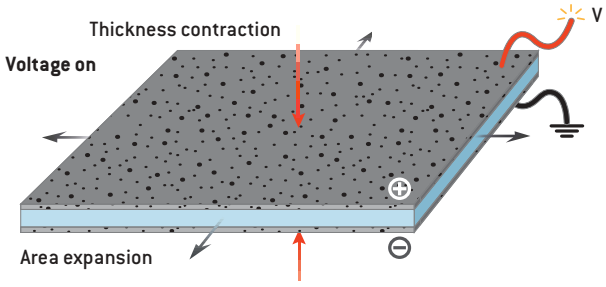
THE FUNDAMENTAL MECHANISM underlying new artificial muscle products is relatively simple. When exposed to high-voltage electric fields, dielectric elastomers—such as silicones and acrylics—contract in the direction of the electric field lines and expand perpendicularly to them, a phenomenon physicists term Maxwell stress. The new devices are basically rubbery capacitors—two charged parallel plates sandwiching a dielectric material. When the power is on, plus and minus charges accumulate on the opposite electrodes. They attract each other and squeeze down on the polymer insulator, which responds by expanding in area.

Engineers laminate thin films of dielectric elastomers (typically 30 to 60 microns thick) on the front and back with conductive carbon particles suspended in a soft polymer matrix. When connected by wires to a power source, the carbon layers serve as flexible electrodes that expand in area along with the material sandwiched in the middle. This layered plastic sheet serves as the basis for a wide range of novel actuation, sensory and energy-generating devices.

Dielectric elastomers, which can grow by as much as 400 percent of their nonactivated size, are by no means the only types of electroactive materials or devices, although they represent some of the more effective examples.

The graph at the right compares the performance of various classes of actuation materials and devices. These include well-established motion-generating products driven by electric current as well as applied electrostatic and electromagnetic fields. Strain refers to the amount of displacement or travel per unit length the devices can





in the toy, automotive, electronics, medical product and footwear industries to bring artificial muscles to market.

The start-up firm's goal? Only to replace a substantial number of the myriad electric motors we use regularly, not to mention many other common motion-generating mechanisms, with smaller, lighter, cheaper products using SRI's novel actuators. "We believe this technology has a good chance to revolutionize the field of mechanical actuation," states Phillip von Guggenberg, the lab's director of business development. "We'd like to make the technology ubiquitous, the kind of thing you could pick up in hardware stores."

Materials That Move

BAR-COHEN HAS SERVED as the unofficial coordinator for the highly diverse community of international EAP researchers since the mid-1990s. Back during the field's infancy, "the electroactive polymer materials I read about in scientific papers didn't work as advertised," he recalls, chuckling slyly. "And as I already had obtained NASA funding to study the technology, I was forced to look around to find who was working in this area to find something that did." Within a few years Bar-Cohen had learned enough to help establish the first scientific conference on the topic, start publishing an EAP newsletter, post an EAP Web site and edit two books on the nascent technology.

Sitting among arrays of lab tables strewn with prototype actuation devices and test apparatus in a low-slung research building on the JPL campus, Bar-Cohen reviews the history of the field he has come to know so well. "For a long time," he begins, "people have been working on ways to move objects without electric motors, which can be too heavy and bulky for many applications. Until the development of EAPs, the standard replacement technology for motors were piezoelectric ceramics, which have been around for some time."

In piezoelectric materials, mechanical stress causes crystals to electrically polarize and vice versa. Hit them with electric current and they deform; deform them and they generate electricity.

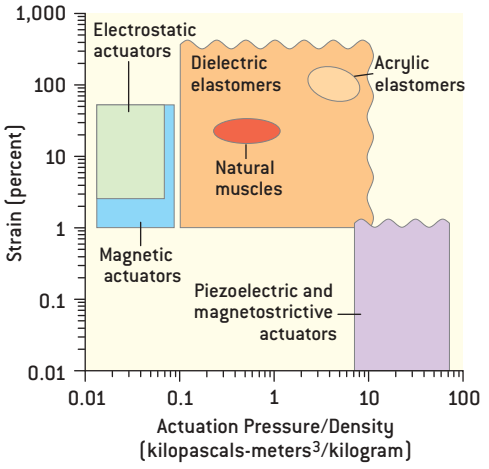
Bar-Cohen lifts a small grayish disk off one of the lab benches, saying, "This one's made of PZT—lead zirconate titanate." He explains that electric current makes the

Overview/*Electroactive Polymers*

- Physicists and chemists have long sought to develop lightweight materials that grow or shrink significantly in length or volume when subjected to electric stimulation. Such substances could serve as the drivers for novel motion-generating devices (generally called actuators)—possible replacements for the ubiquitous electric motor, which is often too bulky and heavy for smaller-scale applications.
- A new generation of electroactive polymer materials displays sufficient physical response to electrical excitation to power new classes of actuators as well as innovative sensors and energy generators. Products based on this "artificial muscle" technology are just starting to hit the market.

JAMIE KRIPKE (*preceding pages*); SARA CHEN (*above and opposite page*)

create. Actuation pressure/density is a measure of the force they produce. Dielectric elastomers can generate more strain and force than many of the competing technologies. Their properties in this regard are similar to those of natural animal muscle—hence the moniker "artificial muscles."



piezoelectric PZT shrink and expand by a fraction of a percent of its total length. Not much motion but useful nonetheless.

In an adjoining room, Bar-Cohen shows off foot-long impact drills driven by PZT disks that he is building with his JPL colleagues and Cybersonics. "Inside this cylinder is a stack of piezoelectric disks," he states. "When activated with alternating current, the stack beats ultrasonically on a mass that hops up and down at a high rate, driving a bit into solid rock." To one side sit piles of stone blocks into which drill bits have cut deep holes.

As a demonstration of how effectively piezoceramics can perform as actuators, it is impressive. But many applications would demand electroactive materials that grow by more than just a fraction of a percent.

Plastics That React

POLYMERS THAT change shape in response to electricity, according to Bar-Cohen, can be sorted into two groups: ionic and electronic types, each with complementary advantages and disadvantages.

Ionic EAPs (which include ionic polymer gels, ionomeric polymer-metal composites, conductive polymers and carbon nanotubes) work on the basis of electrochemistry—the mobility or diffusion of charged ions. They can run directly off batteries because even single-digit voltages will make them bend significantly. The catch is that they generally need to be wet and so must be sealed within flexible coatings. The other major shortcoming of many ionic EAPs (especially the ionomeric polymer-metal composites) is that "as long as the electricity is on, the material will keep moving," Bar-Cohen notes, adding: "If the voltage is above a certain level, electrolysis takes place, which causes irreversible damage to the material."

In contrast, electronic EAPs (such as ferroelectric polymers, electrets, dielectric elastomers and electrostrictive graft elastomers) are driven by electric fields. They require relatively high voltages, which can cause uncomfortable electric shocks. But in return, electronic EAPs can react quickly and deliver strong mechanical forces. They do not need a protective coating and require almost no current to hold a position.

SRI's artificial muscle material falls into the electronic EAP classification. The long, bumpy and sometimes serendipitous road to its successful development is a classic example of the vagaries of technological innovation.

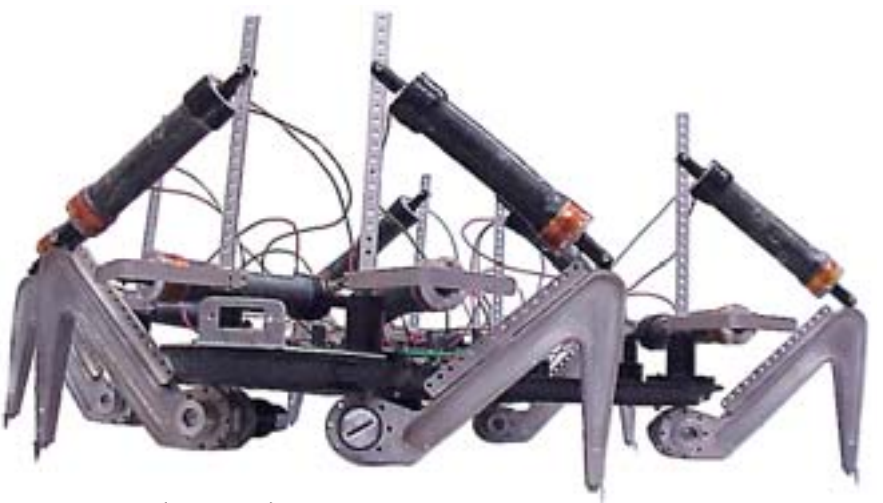
Electrifying Rubber

"SRI INTERNATIONAL began work on artificial muscles in 1992 under contract to the Japanese micro-machine program," says Ron Pelrine, the physicist-turned-mechanical engineer who leads the SRI team. Japanese officials were looking for a new kind of micro-actuator technology. A few SRI researchers started searching for a motion-generating material that resembled natural muscle in terms of force, stroke (linear displacement) and strain (displacement per unit length or area).

"We looked at a whole bunch of possible actuation technologies," Pelrine recalls. Eventually, however, they considered electrostrictive polymers, a class of materials then being investigated by Jerry Scheinbeim of Rutgers University. The hydrocarbon molecules in those polymers are arranged in semicrystalline arrays featuring piezoelectriclike properties.

When exposed to an electric field, all insulating plastics, such as polyurethane, contract in the direction of the field lines and expand perpendicularly to them. This phenomenon, which differs from electrostriction, is called Maxwell stress. "It had been known for a long time but was regarded generally as a nuisance effect," Pelrine says.

He recognized that polymers softer than polyurethane would squash more under electrostatic attraction and thus would provide greater mechanical strains. Working with soft silicones, the SRI scientists soon demonstrated quite acceptable strains of 10 to 15 percent. With further research those numbers rose to 20 to 30 percent. To distinguish the new actuator materials, silicones and other softer plastics were christened dielectric elas-



INSECTLIKE ROBOT (named Flex) walks on legs powered by extension roll-type artificial muscles.

SRI INTERNATIONAL

tomers (they are also called electric-field-actuated polymers).

Having identified several promising polymer materials, the group focused for much of the remainder of the 1990s on the nuts and bolts of building devices for specific applications. Much of the SRI team's new external funding support and research direction came at the time from the Defense Advanced Research Projects Agency (DARPA) and the Office of Naval Research, whose directors were primarily interested in using the technology for military purposes, including small reconnaissance robots and lightweight power generators.

As the elastomers began to exhibit much larger strains, the engineers realized that the electrodes would have to become expandable as well. Ordinary metal electrodes cannot stretch without breaking. "Previously, people didn't have to worry about this issue, because they were working with materials that provided strains of 1 percent or so," Pelrine notes. Eventually, the team developed compliant electrodes based on carbon particles in an elastomeric matrix. "Because the electrodes expand along with the plastic," he points out, "they can maintain the electric field between them across the entire active region." SRI International patented this concept, one of the keys to subsequent artificial-muscle technology.

Eager to demonstrate, Pelrine holds out what looks like a six-inch-square picture frame with plastic sandwich wrap stretched tautly across it. "See, this polymer material is very stretchy," he says, pushing a finger into the transparent film. "It's actually a double-sided adhesive tape that's sold at low cost in large rolls." On both sides of the middle of the sheet are the black, nickel-size compliant electrodes, trailing wires.

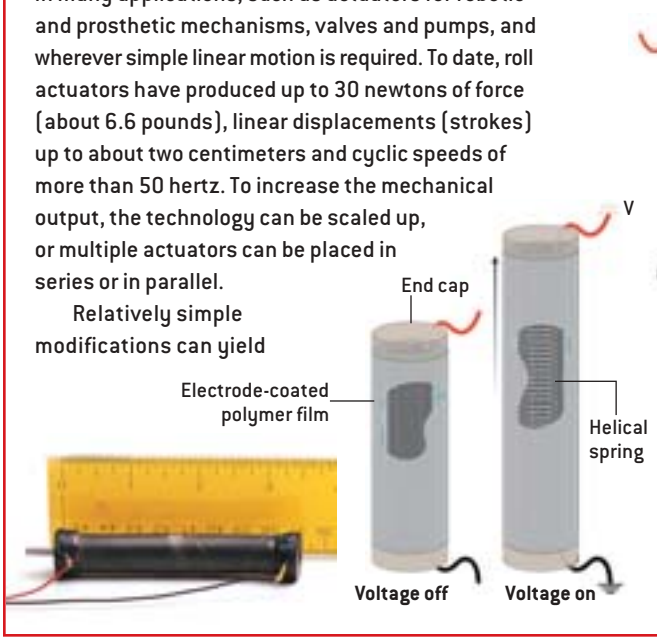
Pelrine turns a control knob on the electric power supply. Instantly, the dark circle of the paired electrodes grows to the diameter of a quarter. When he returns the knob to its original position, the disk shrinks back immediately. He flashes a grin and repeats the sequence a few times, explaining: "Fundamentally, our devices are capacitors—two charged parallel plates sandwiching a dielectric material. When the power is on, plus and minus charges accumulate on the opposite electrodes. They attract each other and squeeze down on the polymer insulator, which responds by expanding in area."

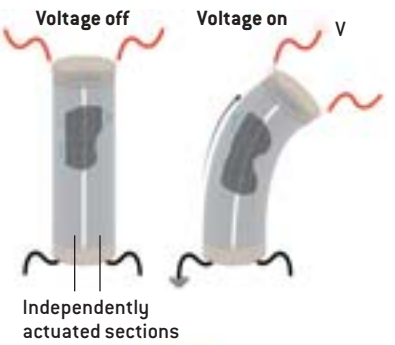
Although several promising materials had been identified, achieving acceptable performance in practical devices proved to be a challenge. A couple of breakthroughs in 1999 drew considerable interest from government and industry, however. One arose from the observation that stretching the polymers before electrically activating them somehow vastly improved their performance. "We started to notice that there seemed to be a sweet spot at which you get optimum performance," remembers engineer Roy Kornbluh, an-

SPRING ROLLS, SNAKES AND ROBOT ARMS

POLYMERS THAT EXPAND in response to electricity make possible roll actuators that extend—or bend—on command. Engineers first roll up two layers of dielectric elastomer sheet (laminated on both sides with flexible electrodes) into compact cylinders. Often the film materials are wrapped around a compressed helical spring that holds a high circumferential prestrain on the films, thereby enhancing device performance. So-called spring rolls can serve in many applications, such as actuators for robotic and prosthetic mechanisms, valves and pumps, and wherever simple linear motion is required. To date, roll actuators have produced up to 30 newtons of force (about 6.6 pounds), linear displacements (strokes) up to about two centimeters and cyclic speeds of more than 50 hertz. To increase the mechanical output, the technology can be scaled up, or multiple actuators can be placed in series or in parallel.

Relatively simple modifications can yield devices that bend on command. Researchers spray specially patterned electrodes onto the dielectric elastomer film in such a fashion that the roll incorporates two independently energized actuators on either side (split lengthwise). If only the left half receives voltage, the right one inhibits the resulting motion and causes the device to bend toward the right (*below*). If only the right half is activated, the roll bends left. If both halves are energized, the roll extends. More complicated arrays of independent electrodes can create more complex motion. Applications for bending rolls include snakelike robots and manipulators, steerable catheters and endoscopes, legged robots, and pointing mechanisms for antennas.





SARA CHEN (*illustrations*); JAMIE KRIPKE (*photographs*)

other team member. "No one was sure exactly why, but prestretching the polymers increased breakdown strengths [resistance to the passage of current between electrodes] by as much as 100 times." Actuation strains improved to a similar degree. Although the reason is still unclear, SRI chemist Qibing Pei believes that "prestretching orients the molecular chains along the plane of expansion and also makes it stiffer in that direction." To achieve the prestraining effect, SRI's actuator devices incorporate an external bracing structure.

The second key discovery came about primarily because the researchers "were testing every stretchy material we could find—what we call an Edisonian approach," Pelrine says with amusement. (Thomas Edison systematically tried all kinds of materials for suitability as lightbulb filaments.) "At my home, we had placed a polymeric door lock on the refrigerator to keep my toddler from getting in. As he got older, we didn't need the lock anymore, so I removed it. But since it was made of a stretchy material, I decided to test its strain properties. Surprisingly, it gave very good performance." Tracking down the material and determining its composition took no small effort, but in the end the mystery polymer "turned out to be an acrylic elastomer that could provide tremendous strains and energy output—as much as 380 percent linear strain. These two developments allowed us to start applying the dielectric elastomers to real-world actuator devices," the researcher says.

Making It Real

THE SRI TEAM'S general approach is flexible, encompassing many designs and even different polymers. As Pei says, "This is a device, not a material." According to Pelrine, the team can produce the actuation effect using various polymers, including acrylics and silicones. Even natural rubber works to some extent. In the extreme temperatures of outer space, for example, artificial muscles might best be made of silicone plastics, which have been demonstrated in a vacuum at –100 degrees Celsius. Uses that require larger output forces might involve more polymer or ganging up several devices in series or in parallel.

"Because the dielectric elastomers can be purchased off the shelf and we'd use at most only a few square feet of material in each device, the actuators would be very low cost, particularly in volume production," SRI's von Guggenberg estimates.

The voltages required to activate dielectric elastomer actuators are relatively high—typically one to five kilovolts—so the devices can operate at a very low current (generally, high voltage means low current). They also use thinner, less expensive wiring and keep fairly cool. "Up to the point at which the electric field breaks down and current flows across the gap [between the electrodes], more voltage gives you greater expansion and greater force," Pelrine says.

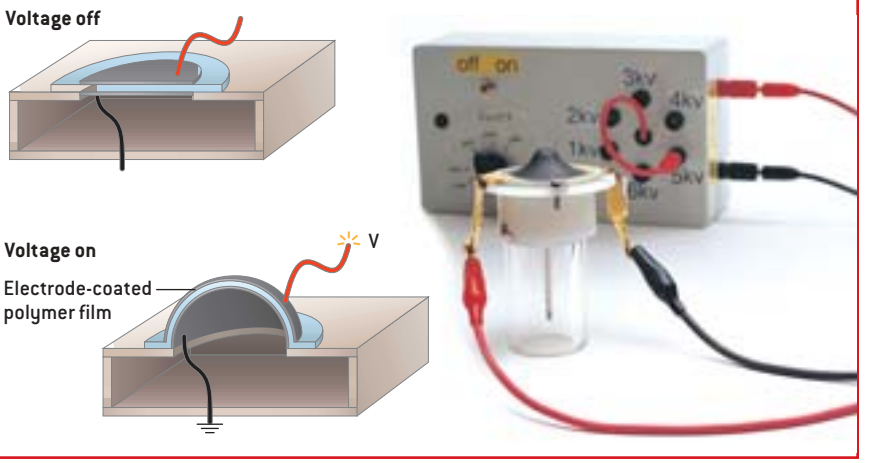
"High voltage can be a concern," Kornbluh comments, "but it's not necessarily dangerous. After all, fluorescent lights and cathode-ray tubes are high-voltage devices, but nobody worries about them. It's more of an issue for mobile devices because batteries are usually low-voltage, and thus additional electric conversion circuits would be needed." Moreover, at Pennsylvania State University, Qiming Zhang and his research group have managed to lower the activation voltages of certain electrostrictive polymers by combining them with other substances to create composites.

When asked about the durability of SRI's dielectric elastomer actuators, von Guggenberg acknowledges a need for more study but attests to a "reasonable indication" that they continue to work sufficiently long for commercial use: "For example, we ran a device for one client that produces moderate, 5 to 10 percent strains for 10 million cycles." Another generated 50 percent area strains for a million cycles.

Although artificial-muscle technology can weigh significantly less than comparable electric motors—the polymers themselves have the density of water—efforts are ongoing at SRI to cut their mass by reducing the need for the external structure that prestrain the polymers. Pei, for instance, is experimenting with chemical

PUMP UP THE MEMBRANE

DIAPHRAGM ACTUATORS are made by stretching dielectric elastomer films over an opening in a rigid frame. Typically the diaphragm is biased, or pushed up or down, by a spring, light air pressure, foam or other means. Biasing makes the diaphragm actuate in one direction (up or down) rather than simply wrinkling randomly when voltage is applied. Diaphragm actuators can displace volume, making them suitable for pumps or loudspeakers. Alternative drive technologies such as piezoelectric materials have long been in use, but dielectric elastomer diaphragms offer larger displacements. Some designs can deflect from an initial flat position to a convex shape (*below*).



SARA CHEN (*illustrations*); JAMIE KRIPKE (*photograph*)

processing to eliminate the need for the relatively heavy frame.

Building Products

HAVING DEVELOPED a basic mechanism, the SRI team soon began work on a flood of application concepts:

Linear actuators. To make what they call spring rolls, the engineers wrap several layers of prestrained laminated dielectric elastomer sheet around a helical spring. The tension spring supports the circumferential prestrain, whereas the lengthwise prestrain of the film holds the spring compressed [*see illustration on page 56*]. Voltage makes the film squeeze in thickness and relax lengthwise so that the device extends. The spring rolls can therefore generate high force and stroke in a compact package. Kornbluh reports that automakers are interested in these mechanisms as replacements for the many small electric motors found in cars, such as in motorized seat-position controls and in the valve controls of high-efficiency camless engines.

Bending rolls. Taking the same basic spring roll, engineers can connect electrodes to create two or more distinct, individually addressed sections around the circumference. Electrically activating that section makes its side of the roll extend, so the entire roll bends away from that side [*see illustration on page 56*]. Mechanisms based on this design could engage in complicated motions that would be difficult to accomplish using conventional motors, gears and linkages. Possible uses would be in steerable medical catheters and in so-called snake robots.

Push-pull actuators. Pairs of dielectric elastomer films or of spring rolls can be arranged in a "push-pull" configuration so that they work against each other and thus respond in a more linear ("one input yields one output") fashion. Shuttling voltage from one device to the other can shift the position of the whole assembly back and forth; activating both devices makes the assembly rigid at a neutral point. In this way, the actuators act like the opposing bicep and tricep muscles that control movements of the human arm.

Loudspeakers. Stretch a dielectric elastomer film over a frame that has an aperture in it. Expanding and contracting rapidly according to the applied voltage signal, the diaphragm will then emit sound. This configuration can yield a lightweight, inexpensive flat-panel speaker, whose vibrating medium is both the driver and sound-generating panel. Current designs offer good performance in the mid- and high-frequency ranges. The speaker configuration has not yet been optimized as a woofer, although no obstacle prevents it from operating well at low frequencies [*see illustration on preceding page*].

Pumps. The design of a dielectric elastomer diaphragm pump is analogous to that of a low-frequency loudspeaker to which engineers have added a fluid chamber and two one-way check valves to control the flow of liquid. Artificial muscles are well suited to powering microfluidic pumps, for example, on the lab-on-a-chip devices prized by medicine and industry.

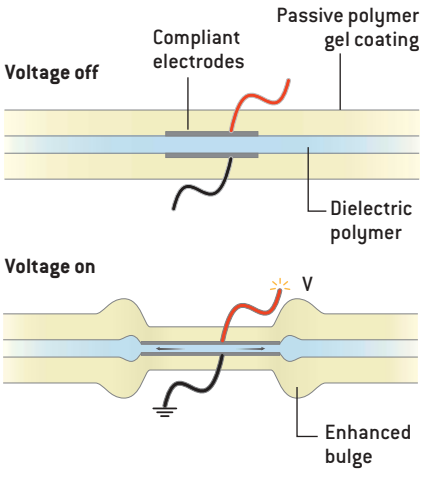
Sensors. Because of their nature, all SRI's dielectric elastomer devices exhibit a change in capacitance when they are bent or stretched. Thus, it is possible to make a sensor that is compliant and operates at low voltage. According to Kornbluh, the team came close to getting an automaker to adopt the technology as a sensor for measuring the tension of a seat belt. Such sensors could similarly be incorporated in fabrics and other materials as fibers, strips or coatings, he says.

Surface texturing and smart surfaces. If the polymers are imprinted with patterns of electrodes, arrays of dots or shapes can be raised on a surface on demand. This technology might find use as an active camouflage fabric that can change its reflectance as desired or as a mechanism for making "riblets" that improve the aerodynamic drag characteris-

CONTROLLABLE SURFACE TEXTURES

CHANGING the texture of a surface can be desirable in a variety of applications, such as "active" military camouflage materials that can alter their reflectance. Surface texturing can also help control air or water flow over the surfaces of airplanes or ships. Touch-based, or haptic, displays could be based on changes in texture.

Most dielectric elastomer actuators take advantage of large-scale deformations in the plane of the film. Alterations in the thickness of the film, on the other hand, are barely perceptible. By coating the thin-film sheets and patterned compliant electrodes with a much thicker and softer layer of polymeric gel, however, thickness changes can be greatly amplified so that they are readily apparent. As the film grows in the plane, the gel spreads out along with the expanding film and bunches up at the points at which the film compresses.



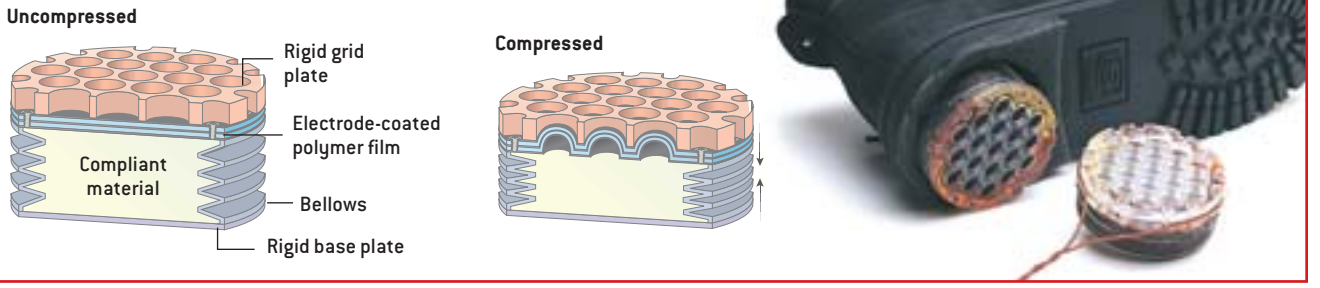
SARA CHEN (*illustrations*); JAMIE KRIPKE (*photograph*)

ELECTRIC BOUNCE IN EVERY STEP

DIELECTRIC ELASTOMERS can produce electric power. In generator mode, a voltage is applied across the dielectric elastomer, which is deformed by external force. As the shape of the elastomer changes, the effective capacitance of the device also alters and, with the appropriate electronics, electrical energy is generated. The energy density of these materials when used as a generator is high, which means that they can be made lighter than other technologies.

Dielectric elastomers are well suited to applications in which electrical power comes from relatively large motions, such as those produced by wind energy, waves and human activity. Capturing the compression energy of a shoe heel when it strikes the ground during walking or running is a good way to generate portable electrical power. This energy is free in the sense that it does not place an additional burden on the wearer. The heel-strike generator effectively couples the compression of the heel to the deformation of an array of multilayer diaphragms.

SRI engineers expect that, with further development, a device will be able to generate about a watt during normal walking. A unit in each shoe should provide enough electricity to power a cellular phone, for example. Such a device is being developed for the U.S. military to supply power to soldiers in the field, but the technology has civilian uses as well.



tics of airplane wings [*see illustration on preceding page*].

Power generators. Again, because these materials act as soft capacitors, variable-capacitance power generators and energy harvesters can be built from them. DARPA and the U.S. Army funded development of a heel-strike generator, a portable energy source that soldiers and others in the field could use to power electronic devices in place of batteries. An average-size person taking a step each second can produce about a watt of power using a device now under development [*see illustration above*]. Von Guggenberg says this concept has caught the interest of footwear companies. The devices could similarly be attached to backpack straps or car-suspension components. In principle, this approach could also be applied to wave generators or wind-power devices.

SRI researchers recently tested a more radical concept—"polymer engines." Propane fuel was burned inside a chamber, and the pressure from the resulting combustion products distorted a dielectric elastomer diaphragm, generating electricity. Such designs might eventually lead to efficient, extremely small generators in the centimeter-or-less size range.

But truly marketable products are still to come. "At this point we're building turnkey devices that we can place in the hands of engineers so they can play with them and get comfortable with the technology," von Guggenberg notes. "We hope it's just a matter of time before every engineer will consider this technology as they design new products."

Bar-Cohen says that he is impressed by the progress the SRI group has made on its actuator technology. But success has also created problems in one regard: the arm-wrestling challenge. "We expected it would take 20 years or so for anyone to develop a mechanical arm that would be strong enough to compete against a human," he laughs. "Now SRI says they're ready to build one, and we haven't raised the prize money yet!" SA

Steven Ashley is a staff writer and editor.

MORE TO EXPLORE

High-Speed Electrically Actuated Elastomers with Over 100% Strain. Ron Pelrine, Roy Kornbluh, Qibing Pei and Jose Joseph in *Science,* Vol. 287, pages 836–839; February 4, 2000.

Electroactive Polymer (EAP) Actuators as Artificial Muscles: Reality, Potential, and Challenges. Edited by Yoseph Bar-Cohen. SPIE Press Monograph, Vol. PM98; 2001.

Polymer Actuators. Peter Sommer-Larsen and Roy Kornbluh in *Proceedings of Actuator 2002.* 8th International Conference on New Actuators, Bremen, Germany; June 2002.

An All-Organic Composite Actuator Material with High Dielectric Constant. Q. M. Zhang et al. in *Nature,* Vol. 419, pages 284–287; September 19, 2002.

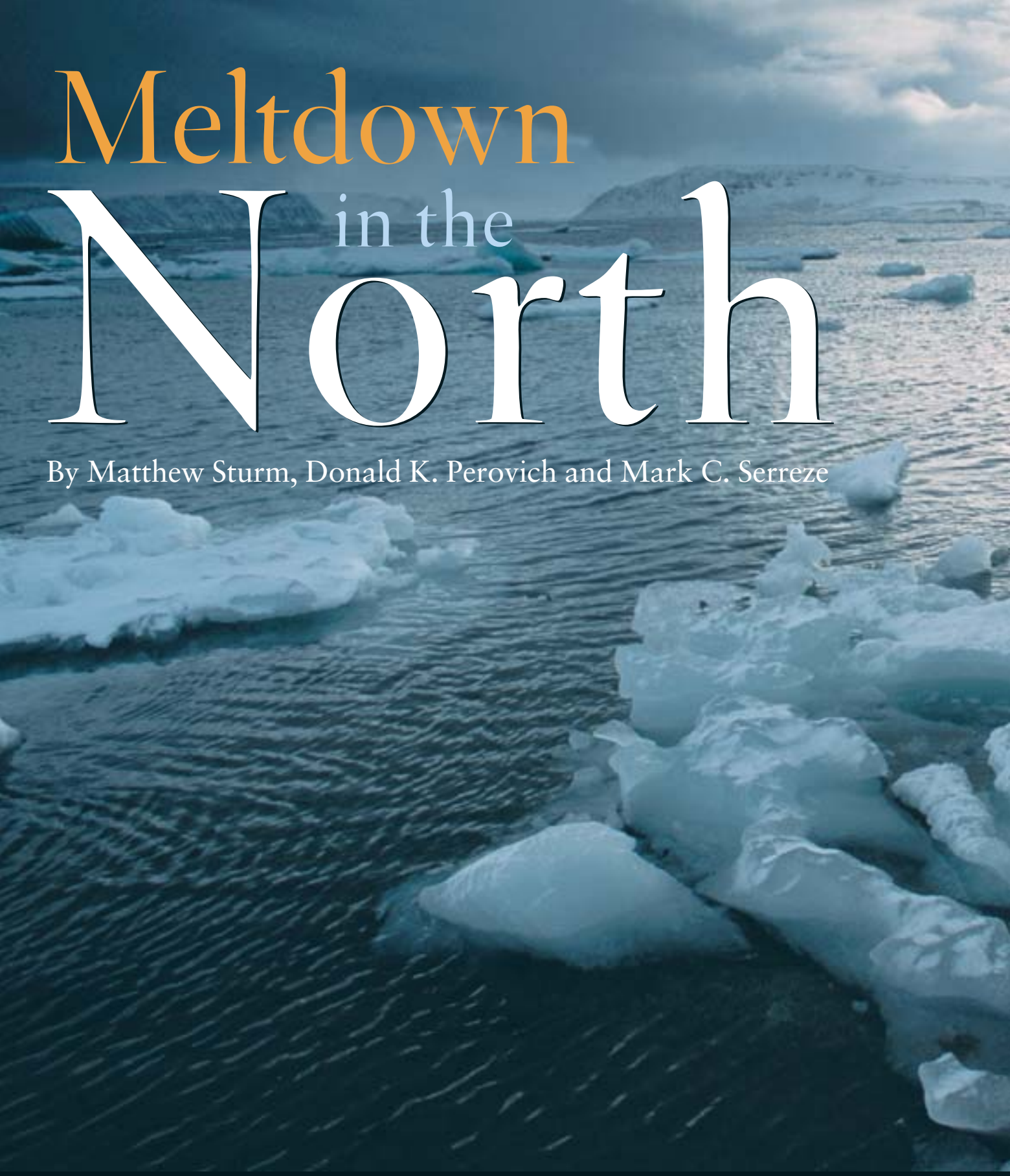
Engineering a Muscle: An Approach to Artificial Muscle Based on Field-Activated Electroactive Polymers. Roy Kornbluh et al. in *Neurotechnology for Biomimetic Robots.* Edited by Joseph Ayers, Joel Davis and Alan Rudolph. MIT Press, 2002.

Smart Structures and Materials 2003: Electroactive Polymer Actuators and Devices (EAPAD). Edited by Yoseph Bar-Cohen. *Proceedings of the SPIE,* Vol. 5051; 2003.

Simple muscle demos are available online at **www.erg.sri.com/projects/muscle/recipe** and **ndeaa.jpl.nasa.gov/nasa-nde/lommas/eap/EAP-recipe.htm**

The WorldWide ElectroActive Polymer Newsletter is at **ndeaa.jpl.nasa.gov/nasa-nde/lommas/eap/WW-EAP-Newsletter.html**

SARA CHEN (*illustrations*); JAMIE KRIPKE (*photograph*)



Meltdown in the North

By Matthew Sturm, Donald K. Perovich and Mark C. Serreze

Sea ice and glaciers are melting, permafrost is thawing, tundra scientists are struggling to understand how these

REDUCTION IN SEA ICE is one of the most striking measures of change in the Arctic. The total area covered by sea ice has shrunk by 3 percent during each of the past three decades. Thickness has decreased even more over this same period—as much as 40 percent in some places. The image shows the Arctic Ocean near Russia.

is yielding to shrubs—and
changes will affect **not just the Arctic but the entire planet**

Snow crystals sting my face and coat my beard and the ruff of my parka. As the wind rises, it becomes difficult to see my five companions through the blowing snow. We are 500 miles into a 750-mile snowmobile trip across Arctic Alaska. We have come, in the late winter of 2002, to measure the thickness of the snow cover and estimate its insulating capacity, an important factor in maintaining the thermal balance of the permafrost. I have called a momentary halt to decide what to do. The rising wind, combined with –30 degree Fahrenheit temperatures, makes it clear we need to find shelter, and fast. I put my face against the hood of my nearest companion and shout: "Make sure everyone stays close together. We have to get off this exposed ridge."

At the time, the irony that we might freeze to death while looking for evidence of global warming was lost on me, but later, snug in our tents, I began to laugh at how incongruous that would have been. —*Matthew Sturm*

The list is impressively long: The warmest air temperatures in four centuries, a shrinking sea-ice cover, a record amount of melting on the Greenland Ice Sheet, Alaskan glaciers retreating at unprecedented rates. Add to this the increasing discharge from Russian rivers, an Arctic growing season that has lengthened by several days per decade, and permafrost that has started to thaw. Taken together, these observations announce in a way no single measurement could that the Arctic is undergoing a profound transformation. Its full extent has come to light only in the past decade, after scientists in different disciplines began comparing their findings. Now many of those scientists are collaborating, trying to understand the ramifications of the changes and to predict what lies ahead for the Arctic and the rest of the globe.

What they learn will have planetwide importance because the Arctic exerts an outsize degree of control on the climate. Much as a spillway in a dam controls the level of a reservoir, the polar regions control the earth's heat balance. Because more solar energy is absorbed in the tropics than at the poles, winds and ocean currents constantly transport heat poleward, where the extensive snow and ice cover influences its fate. As long as this highly reflective cover is intact and extensive, sunlight coming directly into the Arctic is mostly reflected back into space, keeping the Arctic cool and a good repository for the heat brought in from lower latitudes. But if the cover begins to melt and shrink, it will reflect less sunlight, and the Arctic will become a poorer repository, eventually warming the climate of the entire planet.

Figuring out just what will happen, however, is fraught with complications. The greatest of these stems from the intricate feedback systems that govern the climate in the Arctic. Some of these processes are positive, amplifying change and turning a nudge into a shove, and some are negative, behaving as a brake on the system and mitigating change.

Chief among these processes is the ice-albedo feedback [*see box on page 66*], in which rising temperatures produce shorter winters and less extensive snow and ice cover, with ripple effects all the way back through the midlatitudes. Another feedback is associated with the large stores of carbon frozen into the Arctic in the form of peat. As the climate warms and this peat thaws, it could release carbon dioxide into the atmosphere and enhance warming over not just the Arctic but the whole globe—a phenomenon commonly referred to as greenhouse warming.

The key problem is that we don't fully understand how some of these feedback processes work in isolation, let alone how they interact. What we do know is that the Arctic is a complex system: change one thing, and everything else responds, sometimes in a counterintuitive way.

Overview/*Arctic Warming*

- Signs of warming are everywhere in the Arctic, and these changes are bound to affect conditions on the rest of the planet because the polar regions exert an outsize control on the earth's heat balance.
- A complex web of climatic feedback systems makes it extremely difficult to know whether greenhouse warming is the primary cause of the transformation in the Arctic.
- But whatever is causing the changes, scientists face the urgent task of predicting what lies ahead rather than waiting to react to the consequences as they unfold.

Heating Up

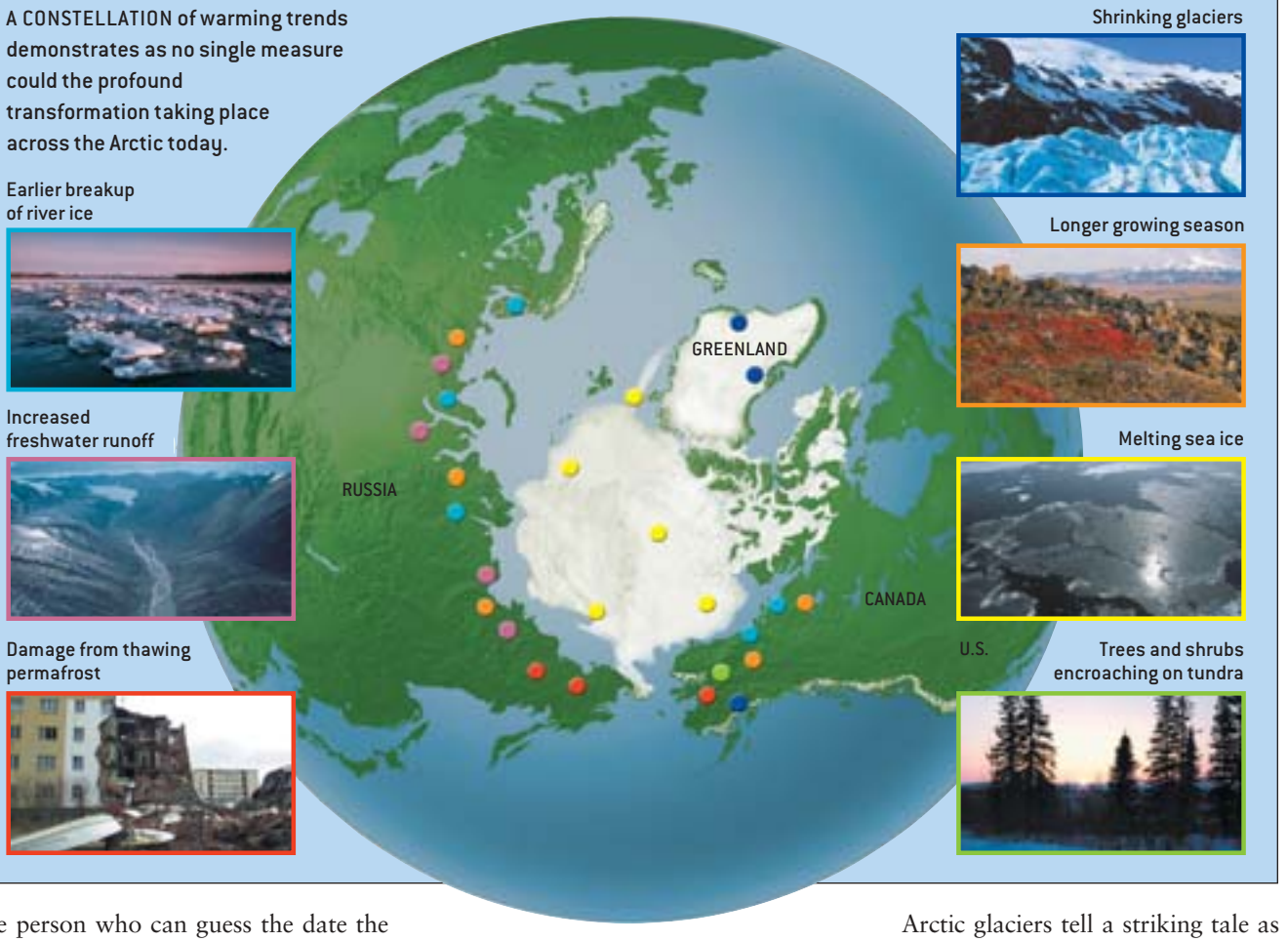
THE MORE WE LOOK, the more change we see. Arctic air temperatures have increased by 0.5 degree Celsius each decade over the past 30 years, with most of the warming coming in winter and spring. Proxy records (ice and peat cores, lake sediments), which tell us mostly about summer temperatures, put this recent warming in perspective. They indicate that late 20th- and early 21st-century temperatures are at their highest level in 400 years. The same records tell us that these high levels are the result of steady warming for 100 years as the Arctic emerged from the Little Ice Age, a frigid period that ended around 1850, topped off by a dramatic acceleration of the warming in the past half a century.

The recent temperature trends are mirrored in many other time series. One example is that Arctic and Northern Hemisphere river and lake ice has been forming later and melting earlier since the Little Ice Age. The total ice-cover season is 16 days shorter than it was in 1850. Near one of our homes (Sturm's) in Alaska, a jackpot of about $300,000 awaits

GALEN ROWELL *Corbis (preceding pages)*

A CATALOGUE OF CHANGES

A CONSTELLATION of warming trends demonstrates as no single measure could the profound transformation taking place across the Arctic today.



the person who can guess the date the Tanana River will break up every spring. The average winning date has gotten earlier by about six days since the betting pool was instituted in 1917. Higher-tech data—satellite images—show that the snow-free season in the Arctic has lengthened by several days each decade since the early 1970s. Similarly, the growing season has increased by as much as four days.

Shrinking Glaciers, Thawing Permafrost

THERE WAS NOTHING *complex about my first research in Arctic climate change: march around a small glacier on Ellesmere Island, drill holes in the ice, insert long metal poles in the holes, measure them, come back a year later and see if more pole was showing.*

We put in most of the pole network in the warm summer of 1982 and returned in 1983 to a very different world—week after week of cold, snow and fog. It looked like the start of a new ice age. Our plan had been to go back annually, but as so often happens, funding dried up, and my Arctic experiences became fond memories.

But memories sometimes get refreshed. In 2002 I got a call from an excited graduate student. He had revisited the glacier. It was rapidly wasting away. 1983 had been an anomaly. My stakes were there, except they were all lying on the surface of the ice. How deeply had I installed them? Did I still have my field notes? He need not have worried. There was my field book, dusty but safe in my bookcase. Now I'm going back to Ellesmere Island, to see what's left of the glacier that in 1983 seemed like such a permanent feature of the landscape but that I now realize may well die before I do.

—Mark C. Serreze

Arctic glaciers tell a striking tale as well. In Alaska, they have been shrinking for five decades, and more startlingly, the rate of shrinkage has increased threefold in the past 10 years. The melting glaciers translate into a rise in sea level of about two millimeters a decade, or 10 percent of the total annual rise of 20 millimeters. Determining the state of the much larger and more slowly changing Greenland Ice Sheet has been something of a Holy Grail for Arctic researchers. Older field and satellite results suggested that the ice sheet was exhibiting asymmetrical behavior—the west side thinning in a modest way and the east side remaining in balance. Recent satellite images indicate that the melt rate over the entire ice sheet has been increasing with time. The total area melting in a given summer has increased by 7 percent each decade since 1978, with last summer setting an all-time record. Winter snowfall appears insufficient to offset

LUCY READING (*illustration*); CLOCKWISE FROM TOP RIGHT: CORBIS; SCOTT T. SMITH *Corbis*; JACQUES LANGEVIN *Corbis Sygma*; GLEN LISTON; VLADIMIR ROMANOVSKY; JIM BRANDENBURG *Minden Pictures*; GALEN ROWELL *Corbis*

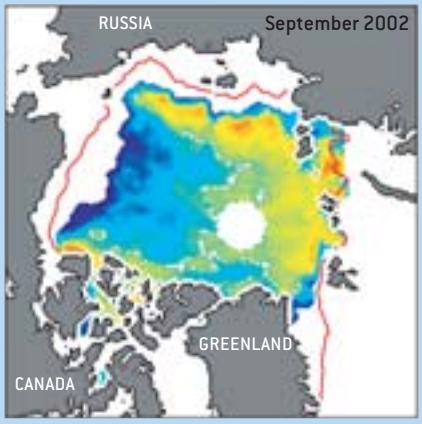
this heavy summer melt, so the sheet is shrinking.

The permafrost—the permanently frozen layer below the surface—is thawing, too. In a study published in 1986, researchers from the U.S. Geological Survey carefully logged temperature profiles in deep oil-exploration boreholes drilled through the permafrost of northern Alaska. When they extrapolated the profiles to the surface, they found an anomalous curvature that was best explained by a warming at ground level of two to four degrees C during the preceding few decades. More recently, preliminary results suggest an additional increase of two to three degrees C has occurred since 1986. Because the Arctic winter lasts nine months of the year, snow cover controls the thermal state of the ground as much as air temperature does, so these borehole records almost certainly reflect a change in the amount and timing of winter precipitation as well as an increase in temperature. More snow means thicker insulation and therefore better protection for the ground from frigid winter temperatures. Ground that is not chilled as much in the winter is primed for more warming in the summer.

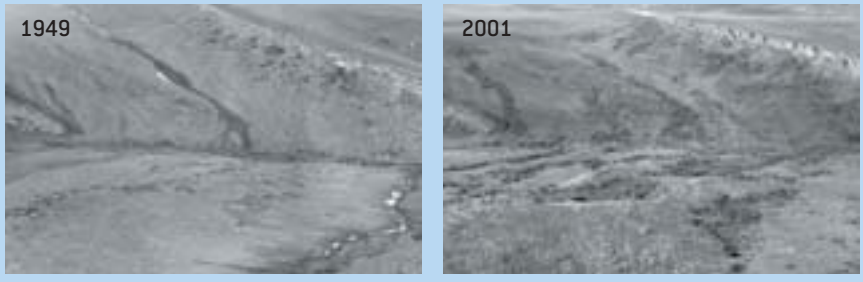
Regardless of why it is occurring, one thing is certain. Thawing permafrost is trouble. It can produce catastrophic failure of roads, houses and other infrastructure. It is also implicated in another recently detected change: over the past 60 years, the discharge of freshwater from Russian rivers into the Arctic Basin has increased by 7 percent—an amount equivalent to roughly one quarter the volume of Lake Erie or three months of the outflow of the Mississippi River. Scientists attribute the change partly to greater winter precipitation and partly to a warming of the permafrost and active layer, which they believe is now transporting more groundwater. This influx of freshwater could have important implications for global climate: the paleo-record suggests that when the outflow of water from the Arctic Basin hits a critical level of freshness, the global ocean circulation changes dramatically. When ocean circulation changes, climate does as well, because the circulation system—essentially a set of moving rivers of water in the ocean, such as the Gulf Stream—is one of the prime conveyors of heat northward toward the pole.

THE HARD, NOT SO COLD FACTS

SCIENTISTS PUZZLING over alterations in the Arctic now have the benefit of many years of data from various sources to help them. The patterns they have extracted from these records reveal that the warming trend is far greater than would be expected if the climate were following a natural progression from the Little Ice Age to a less frigid temperature regime—which indicates that greenhouse warming cannot be ruled out as a cause.



RECORD for minimum sea ice was measured in September 2002; the red line shows average sea ice cover for this time of year. Colors indicate the deviation in ice concentration from the historic norm. For example, the dark blue areas to the north of Alaska had between 30 and 40 percent lower concentration of ice than the average value for that time of year.



INVASION OF TUNDRA by shrubs (*dark bands*) over the past half a century is evident in these photographs taken of the same area near the Chandlar River in Alaska in 1949 and in 2001.

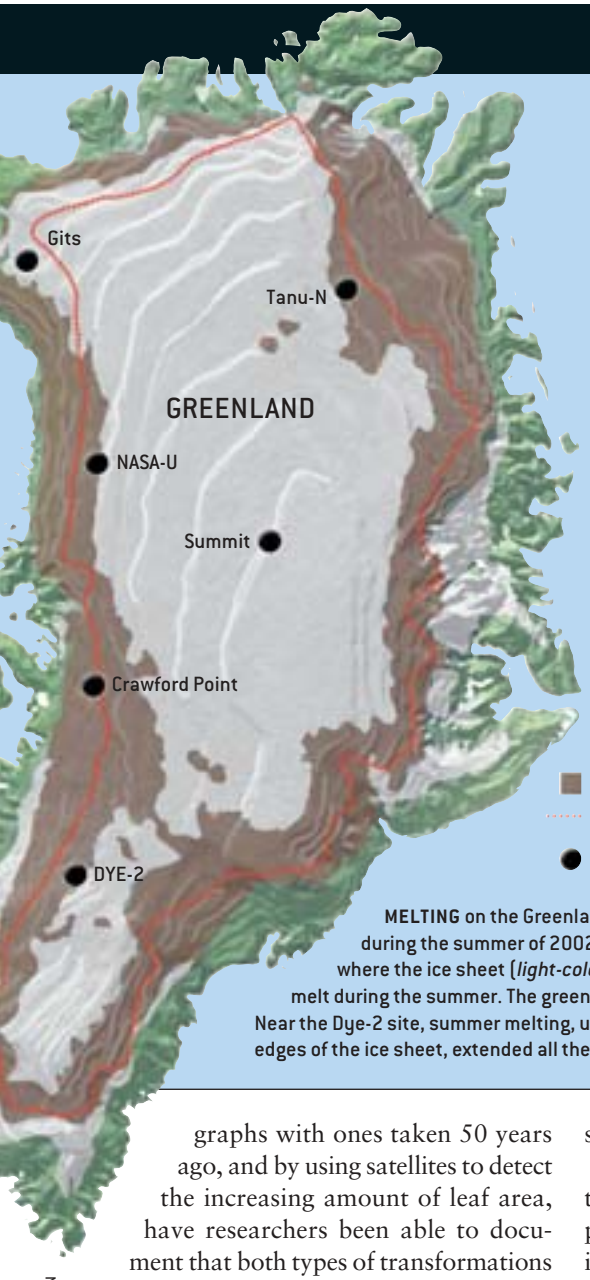
MARK C. SERREZE ET AL., 2003 (*map*); U.S. NAVY (*1949*); KEN TAPE (*2001*)

Greening of the Arctic

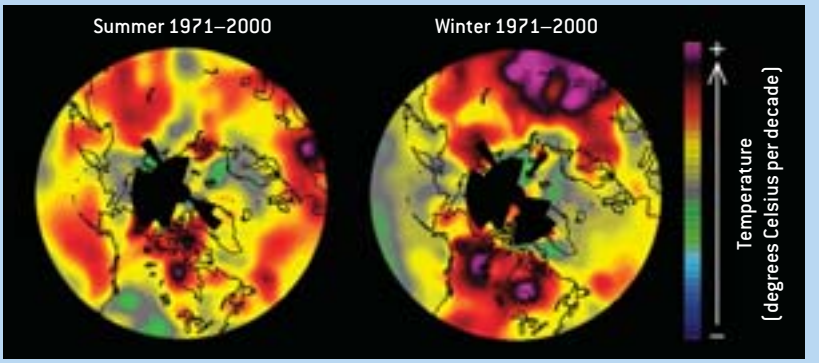
THE ARCTIC LAND COVER is also shifting. Based on warming experiments using greenhouses, biologists have known for some time that shrubs will grow at the expense of the other tundra plants when the climate warms. Under the same favorable growing conditions, the tree line will migrate north. Researchers have been looking for these modifications in the real world, but ecosystem responses can be slow. Only in the past few years, by comparing modern photo-

THE AUTHORS

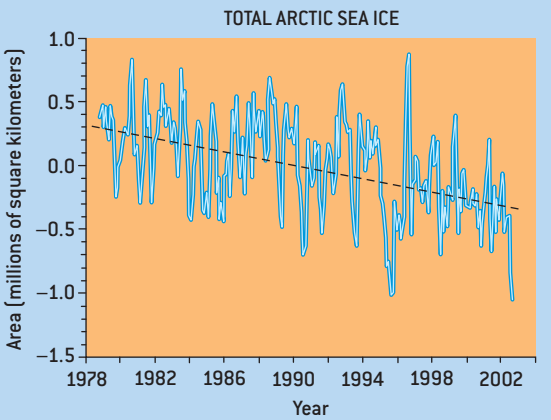
MATTHEW STURM, DONALD K. PEROVICH and *MARK C. SERREZE* have spent most of their research careers trying to understand the snow, ice and climate of the Arctic. In 16 years at the U.S. Army Cold Regions Research and Engineering Laboratory–Alaska, Sturm has led more than a dozen winter expeditions in Arctic Alaska, including most recently a 750-mile snowmobile traverse across the region. Perovich is with the New Hampshire office of the U.S. Army Cold Regions Research and Engineering Laboratory. His work has focused on sea ice and the ice-albedo feedback. Perovich was chief scientist on Ice Station SHEBA, a yearlong drift of an icebreaker frozen into the Arctic pack ice. Since 1986 Serreze has been with the National Snow and Ice Data Center at the University of Colorado at Boulder. His studies have emphasized Arctic climate change and interactions between sea ice and the atmosphere.



MELTING on the Greenland ice sheet set a record during the summer of 2002. The brown color shows where the ice sheet (*light-colored area*) underwent melt during the summer. The green indicates ice-free areas. Near the Dye-2 site, summer melting, usually confined to the edges of the ice sheet, extended all the way to the summit.



RATE AT WHICH Arctic air temperatures have been warming is shown for the past three decades. Warming has been greater in winter. Greens and blues indicate cooling; yellows and reds, warming. The scale bar runs from –1 to +1 degree Celsius per decade.



STRIKING REDUCTIONS in the extent of Arctic sea ice have been recorded since 1978.

graphs with ones taken 50 years ago, and by using satellites to detect the increasing amount of leaf area, have researchers been able to document that both types of transformations are under way. As the vegetation alters, so does the role of the Arctic in the global carbon cycle. Vast stores of carbon in the form of peat underlie much of the tundra in Alaska and Russia, evidence that for long periods Arctic tundra has been a net carbon sink; about 600 cubic miles of peat are currently in cold storage. In recent years, warming has produced a shift: the Arctic now appears to be a net source of carbon dioxide. The change is subtle but troubling because carbon dioxide and methane constitute the primary greenhouse gases in the atmosphere, returning heat to the earth instead of allowing it to escape into space.

Warmer winters have driven some of the shift. When the air is warmer, more precipitation falls from the sky, some of it coming as snow. The thicker snow holds more warmth in the earth, resulting in a longer period during which the tundra is releasing carbon dioxide. But as the tundra becomes shrubbier and as the soil becomes drier in the summer as a result of higher temperatures, the balance could swing back the other way, because plants, particularly woody ones, will fix more carbon and lock it back into the Arctic ecosystem. The most recent studies suggest, in fact, that the magnitude and direction of the Arctic carbon balance depend on the time span that we are examining, with the response varying as the plants adapt to the new conditions.

Melting Sea Ice

"THIS SEA ICE *is ridiculously thin," I thought as I broke through the ice for the second time that morning in August 1998. There was no real danger, now that personal flotation devices had become the de rigueur fashion accessory, but the thin ice was troubling for other reasons.*

My journey to this place, 600 miles from the North Pole, had begun 10 months earlier on board the icebreaker Des Groseilliers, *which we had intentionally frozen into the pack to begin a yearlong drift. Our mission was to study ice-albedo and cloud-radiation feedbacks. When we started the journey, I was surprised at how thin the ice was. Now, after a much longer than expected summer melt season, it was thinner still, even though we had been drifting steadi-*

LUCY READING, *SOURCE:* K. STEFFAN AND R. HUFF *(map)*; CHAPMAN AND WALSH, 1993, UPDATED TO THE PRESENT BY CHAPMAN *(temperature maps)*; LUCY READING, *SOURCE:* NATIONAL SNOW AND ICE DATA CENTER, BOULDER, COLO. *(graph)*

ly north. I was uncertain which would come first: the end of the summer or the end of the ice. Little did I know that this summer the record for minimum ice cover was being set throughout the entire western Arctic Ocean. Unfortunately for the long-term survival of the ice pack, it was a record that was easily broken in 2002. —*Donald K. Perovich*

Of all the changes we have catalogued, the most alarming by far has been the reduction in the Arctic sea-ice cover. Researchers tracking this alteration have discovered that the area covered by the ice has been decreasing by about 3 percent each decade since the advent of satellite records in 1972. This rate might be low for a financial investment, but where time is measured in centuries or millennia, it is high. With the sea ice covering an area approximately the size of the U.S., the reduction per decade is equivalent to an area the size of Colorado and New Hampshire combined, the home states of two of us (Perovich and Serreze). The change in the thickness of the ice (determined from submarines) is even more striking: as much as 40 percent lost in the past few decades. Some climate models suggest that by 2080 the Arctic Ocean will be ice-free in summer.

The melting sea ice does not raise sea level as melting glaciers do, because the ice is already floating, but it is alarming for two other reasons. Locally, the demise of the sea ice leads to the loss of a unique marine ecosystem replete with polar bears, seals and whales. Globally, an ice-free Arctic Ocean would be the extreme end point of the ice-albedo feedback—far more solar radiation would be absorbed, warming not just the Arctic but eventually every part of the earth.

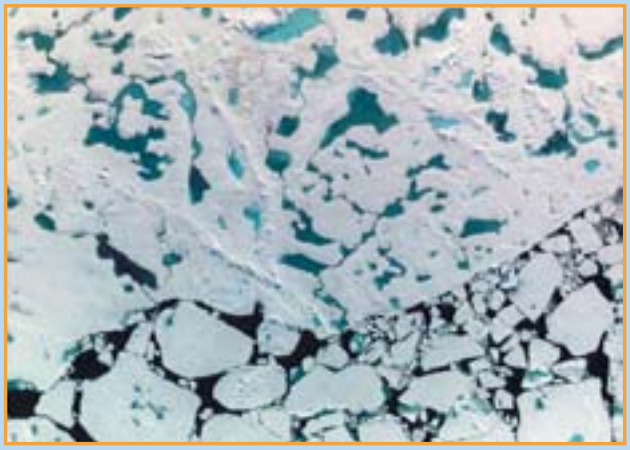
The shrinking sea-ice cover has not escaped the attention of businesspeople, tourists and politicians. Serious discussions have been under way about the feasibility of transporting cargo via Arctic waters—including through the fabled Northwest Passage, now perhaps close to being a practical shipping route because of climate change. Roald Amundsen, the redoubtable Norwegian polar explorer, took more than three years to complete the first transit of the passage in 1906, when the Arctic was still under the influ-

A COMPLEX WEB

THE MANY FEEDBACK SYSTEMS operating in the Arctic make predicting the future state of the region a challenge. The ice-albedo feedback is the granddaddy of all these systems. It works this way: land, ocean and ice reflect a fraction of the incoming sunlight, which consequently escapes into space and does not contribute to heating the climate. This fraction is called the albedo. A surface with an albedo of 1 reflects all light, and a surface with an albedo of 0 reflects none. Strikingly, the Arctic Ocean spans nearly this entire range. Where it is frozen and snow-covered, it has the highest albedo of any naturally occurring material, about 0.85, but where it is ice-free, it has the lowest, around 0.07.

In late spring the ice pack is snow-covered—bright and white. The surface reflects most, but not quite all, of the incident sunlight. Some of the ice melts, causing the ice edge to retreat and replacing the bright, highly reflecting snow-covered ice with the dark, absorbing ocean water. Moreover, away from the ocean's edge, melting snow produces ponds of water that also have a low albedo. Melting in both these areas decreases the albedo, which leads to even greater melting, and so on and on.

If the ice-albedo feedback operated in isolation, predicting its ramifications on global climate might be possible even now. But it does not. Instead multiple feedbacks, some positive and some negative, work in concert, and their net effect is difficult to assess. For example, if the albedo is reduced, the effect is to warm the climate, but then the atmosphere can hold more water vapor, and cloud cover will increase. Clouds act as an umbrella that reduces the amount of sunlight reaching the surface (resulting in cooling), but they also trap long-wave radiation from the surface like a blanket (resulting in warming). In the winter the effect is clear—no sunlight, no umbrella, only the blanket. The cloud feedback is positive.



BIRD'S-EYE VIEW of melting sea ice: turquoise water is melt pond, freshwater from melting snow on the surface of the ice; the nearly black water is ocean. Both types are much more heat-absorbing than the bright ice.

But what about in summer, when sunlight is plentiful? Field studies have shown that the feedback depends on the nature of the clouds. For high, thin clouds composed primarily of ice, the umbrella effect dominated and the cloud-radiation feedback was negative. But for the low, liquid-water clouds that are prevalent in the summer, the blanket dominated and the feedback was positive. Indeed, when these low clouds were present, more ice melted than on sunny days.

Scientists are now trying to sort out which of the feedbacks in the complex web that constitutes the Arctic are the ones we need to worry about the most. These are the ones—such as the ice-albedo feedback—that can amplify changes already under way, speeding them up and magnifying them. They are the ones that can push the system over the edge. —*M.S., D.K.P. and M.C.S.*

DONALD K. PEROVICH

ence of the Little Ice Age. Many explorers before him had died trying to make the journey. In the past few years, however, dozens of ships have completed the route, including Russian icebreakers refurbished for the tourist trade. These events would have been unimaginable, even with icebreakers, in the more intense ice conditions of 100 years ago.

Is Greenhouse Warming the Culprit?

THIS INVENTORY of startling transformation in the Arctic inevitably raises the question of whether we are still emerging from the Little Ice Age or whether something quite different is now taking place. Specifically, should we interpret these changes as being caused by the increased concentration of atmospheric greenhouse gases overriding a natural temperature cycle? Or are they part of a longer-than-expected natural cycle?

The intricate web of feedback interactions renders this question exceedingly complicated—and we don't know enough yet to answer it unequivocally. But we know enough to be very worried.

Whatever is causing the melting and thawing now wracking the Arctic, these modifications have initiated a cascade of planetwide responses that will continue even if the climate were suddenly and unexpectedly to stop warming. Imagine the climate as a big, round rock perched on uneven terrain. The inventory tells us that the rock has been pushed a little—either by a natural climate cycle or by human activity—and has started to roll. Even if the pushing stops, the rock is going to keep rolling. When it finally does stop, it will be in a completely different place than before.

To cope with the constellation of changes in the Arctic in a concerted fashion and to develop an ability to predict what will happen next rather than just react to it, several federal agencies have begun to coordinate their Arctic research in a program called SEARCH (Study of Environmental Arctic Change). Early results give some promise for success in teasing out the linkages among the tightly coupled systems that shape the climate of the Arctic and thus the earth. A recent discovery about the patterns of wind circulation, for example, helps to explain previously puzzling spatial patterns of increasing temperature [*see box above*]. Equally important, high-quality records of climate change now extend back 30 to 50 years.

Soon these records and other findings should allow us to determine whether the Arctic transformation is a natural trend linked to emergence from the Little Ice Age or something more ominous. Our most difficult challenge in getting to that point is to come to grips with how the various feedbacks in the Arctic system interact—and to do so quickly. SA

WINDS OF CHANGE

EVEN IN THE DAYS of the Vikings, people knew that when winters in northern Europe were mild, they tended to be severe in southern Greenland, and vice versa. Today we know that this seesaw in temperature affects more than just Greenland and Europe. It is related to an atmospheric circulation pattern known as the North Atlantic Oscillation (NAO, which may be part of an even larger pattern, the Arctic Oscillation, or AO). The NAO describes the linked variation of a major low pressure area centered near Iceland with a major high pressure area centered near the Azores. When both pressure features are strong, the NAO is in its positive mode. When both are weak, it is in its negative mode.

A key feature of the NAO is that winds blow counterclockwise around the Icelandic Low, while they blow clockwise around the Azores High. In positive mode, the winds around the Icelandic Low are stronger than normal and warm air from the south streams over northern Europe and northern Eurasia. At the same time, the circulation pattern sweeps cold air down from the high Arctic over parts of Greenland, the North Atlantic and northeastern North America. When the NAO is negative, the wind pattern weakens (and at times can even reverse), which leads to roughly the opposite temperature pattern. Since about 1970, the winter NAO has been largely "stuck" in its positive mode, which helps to explain why we have observed widespread warming over Alaska, western Canada and Eurasia but regional cooling in eastern Canada and southern Greenland. The pattern has also caused increased precipitation in northern Eurasia and contributed to the reduction in sea ice. *—M.S., D.K.P. and M.C.S.*

NORTH ATLANTIC OSCILLATION is depicted in its positive mode.

A segment based on this article will air September 25 on *National Geographic Today*, a program on the National Geographic Channel. Please check your local listings.

NATIONAL GEOGRAPHIC CHANNEL

MORE TO EXPLORE

Year on Ice Gives Climate Insights. D. K. Perovich et al. in *EOS, Transactions of the American Geophysical Union,* Vol. 80, No. 481, pages 485–486; 1999.

Global Warming and Terrestrial Ecosystems: A Conceptual Framework for Analysis. G. R. Shaver et al. in *BioScience,* Vol. 50, No. 10; 2000.

Observational Evidence of Recent Change in the Northern High-Latitude Environment. Mark Serreze et al. in *Climatic Change,* Vol. 46, pages 159–207; 2000.

The Surface Heat Budget of the Arctic Ocean (SHEBA). Special section in *Journal of Geophysical Research,* Vol. 107, No. 15; October 2002.

SEARCH Web site: **http://psc.apl.washington.edu/search/**

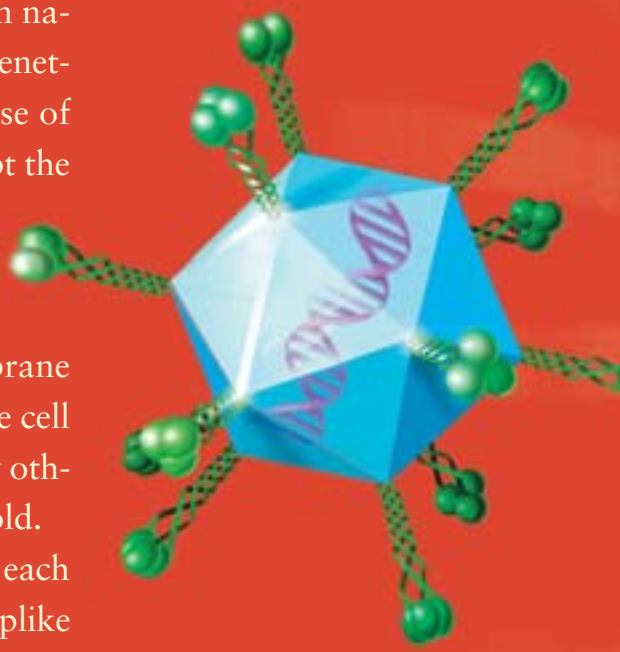
NOAA Arctic Climate Change Web site: **www.ngdc.noaa.gov/paleo/sciencepub/front.htm**

LUCY READING

TUMOR-BUSTING

A new technique called virotherapy harnesses viruses, those banes of humankind, to stop another scourge—cancer

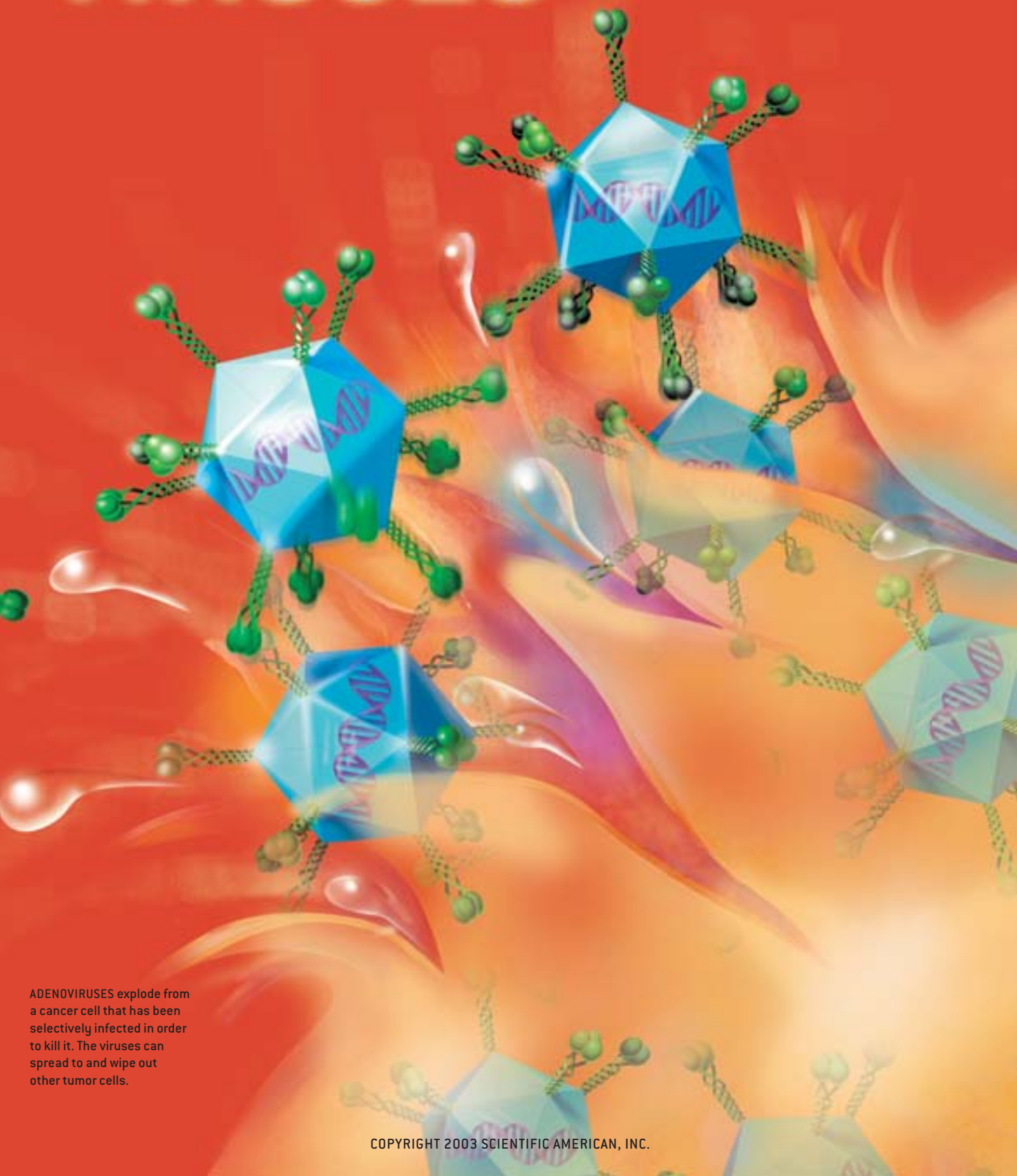
By Dirk M. Nettelbeck and David T. Curiel



Viruses are some of the most insidious creations in nature. They travel light: equipped with just their genetic material packed tightly inside a crystalline case of protein, they latch onto cells, insert their genes, and co-opt the cells' gene-copying and protein-making machinery, using them to make billions of copies of themselves. Once formed, the new viruses percolate to the cell surface, pinch off inside minuscule bubbles of cell membrane and drift away, or else they continue reproducing until the cell finally bursts. In any case, they go on to infect and destroy other cells, resulting in diseases from AIDS to the common cold.

Different viruses cause different diseases in part because each virus enters a cell by first attaching to a specific suction-cuplike receptor on its surface. Liver cells display one kind of receptor used by one family of viruses, whereas nerve cells display another receptor used by a different viral family, so each type of virus infects a particular variety of cell. Cancer researchers have envied this selectivity for years: if they could only target cancer therapies to tumor cells and avoid damaging normal ones, they might be able to eliminate many of the noxious side effects of cancer treatment.

VIRUSES



ADENOVIRUSES explode from a cancer cell that has been selectively infected in order to kill it. The viruses can spread to and wipe out other tumor cells.

Some scientists, including ourselves, are now genetically engineering a range of viruses that act as search-and-destroy missiles: selectively infecting and killing cancer cells while leaving healthy ones alone. This new strategy, called virotherapy, has shown promise in animal tests, and clinical trials involving human patients are now under way. Researchers are evaluating virotherapy alone and as a novel means for administering traditional chemotherapies solely to tumor cells. They are also developing methods to label viruses with radioactive or fluorescent tags in order to track the movement of the viral agents in patients.

Viruses to the Rescue?

ONE OF THE FIRST INKLINGS that viruses could be useful in combating cancer came in 1912, when an Italian gynecologist observed the regression of cervical cancer in a woman who was inoculated with a rabies vaccine made from a live, crippled form of the rabies virus. Physicians first injected viruses into cancer patients intentionally in the late 1940s, but only a handful appeared to benefit. Twenty years later scientists found that a virus that causes the veterinary disorder Newcastle disease shows a preference for infecting tumor cells and began to try to enhance that tendency by growing the viruses for generations in human cancer cells in laboratory culture dishes. Although critics countered that such viruses could be exerting only an indirect effect against cancer by generally activating an individual's immune system and making it more likely to detect and kill cancer cells, reports continued to pop up in the medical literature linking viral infection and cancer remission. In the early 1970s and 1980s two groups of physicians described patients whose lymphomas shrank after they came down with measles.

The modern concept of virotherapy began in the late 1990s, when researchers led by Frank McCormick of ONYX Pharmaceuticals in Richmond, Calif., and Daniel R. Henderson of Calydon in Sunnyvale, Calif., independently published reports showing they could target virotherapy to human cancer cells grafted into mice, thereby eliminating the human tumors. (ONYX is no longer developing therapeutic viruses, and Calydon has been acquired by Cell Genesys in South San Francisco, Calif.) Both groups used adenovirus, a cause of the common cold that has been intensively explored for virotherapy. (Other viruses under study include herpes simplex, parvovirus, vaccinia and reovirus.) Adenovirus is appealing in part because researchers understand its biology very well after years of trying to cure colds and of using the virus in molecular biology and gene therapy research. It consists of a 20-sided protein case, or capsid, filled with DNA and equipped with 12 protein "arms." These protrusions have evolved over millennia to latch onto a cellular receptor whose normal function is to help cells adhere to one another.

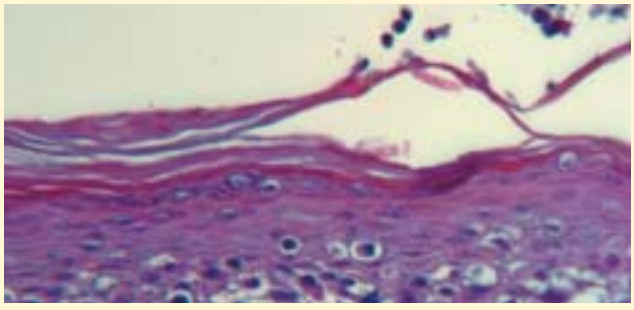
Adenoviruses are distinct from the types of viruses usually used in gene therapy to treat inherited disorders. Gene therapy traditionally employs retroviruses to splice a functioning copy of a gene permanently into the body of a patient in whom that gene has ceased to work properly. Unlike retroviruses, however, adenoviruses do not integrate their DNA into the genes of cells they infect; the genes they ferry into a cell usually work only for a while and then break down. Scientists have investigated adenoviruses extensively in gene therapy approaches to treat cancer, in which the viruses are armed with genes that, for example, make cancer cells more susceptible than normal ones to chemotherapy. In general, tests involving adenoviruses have been safe, but regrettably a volunteer died in 1999 after receiving an infusion of adenoviruses as part of a clinical trial to test a potential gene therapy for a genetic liver disorder [*see box on page 74*].

Gene therapists have been working to tailor adenoviruses and other viral vectors, or gene-delivery systems, to improve their safety and reduce the chances that such a tragedy might occur again. It is perhaps even more essential for researchers, such as ourselves, who are investigating virotherapy to develop safer, more targeted vectors, because virotherapy by definition aims to kill the cells the viruses infect, not just insert a therapeutic gene into them. Killing the wrong cells could be dangerous.

Adenoviruses bring with them characteristics that can make

TARGETING MELANOMA

THE SKIN CANCER melanoma is one of the most lethal cancers unless detected early; it arises from the uncontrolled growth and spread of pigmented cells in the skin called melanocytes. Scientists are using the new approach of virotherapy to selectively kill melanoma cells while leaving healthy cells alone. One technique for studying melanoma involves combining

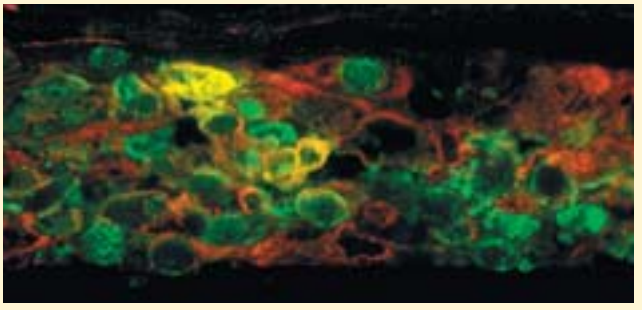


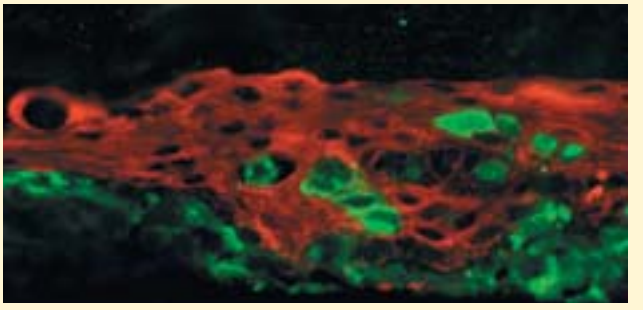
TERESE WINSLOW (*preceding pages*)

Overview/*Anticancer Viruses*

- Virotherapy is a new strategy to treat cancer by selectively infecting and killing tumor cells. Researchers are testing various approaches to target viruses—particularly adenoviruses—to cancer cells, leaving normal cells untouched.
- The viruses used in virotherapy can either kill tumor cells by bursting them open or deliver genes that make the cells more susceptible to traditional chemotherapies.
- The same types of viruses used in virotherapy can also be labeled with fluorescent or radioactive tags. Once delivered into the body, they home in on cancer cells. In the future, physicians might be able to use this imaging technique to detect the presence of tiny tumors.

melanoma cells (*dark dots in micrograph below left*) with normal skin cells called keratinocytes and collagen to make cancer-bearing artificial skin that can be grown in laboratory culture dishes. One of us (Nettelbeck) and colleagues have devised an adenovirus that can specifically reproduce in melanoma cells. In the center and right micrographs below, healthy keratinocytes appear red; cells infected with the virus show up green. The center micrograph was made using viruses that were not specifically targeted to melanomas. The viruses were able to grow in healthy cells, making those cells look yellow. In contrast, the targeted virus (*below right*) did not replicate in healthy cells, so none of the cells are yellow. *—D.M.N. and D.T.C.*





them riskier or safer, depending on the circumstances. Nearly everyone has been exposed at one time or another to adenoviruses, so almost all of us carry antibodies the immune system makes to target them for destruction. Accordingly, shots of adenoviruses as cancer therapies might cause severe, flulike symptoms if the body recognizes them as foreign and ramps up an immune response to eradicate them. (Wiping out the viruses would also squelch the therapy.) At the same time, recognition by the immune system ensures that the viruses do not reproduce out of control. Investigators are now designing various therapeutic approaches to optimize the efficacy of virotherapy and minimize the chances that adenoviruses will cause side effects. These strategies include giving immunosuppressive drugs at the time of virotherapy and modifying the adenoviruses so that they do not trigger a reaction by the immune system.

Homing In on the Target

VIROTHERAPISTS ARE DEVISING two main strategies to make sure their missiles hit their objectives accurately with no collateral damage. In the first approach, termed transductional targeting, researchers are attempting to adapt the viruses so that they preferentially infect, or transduce, cancer cells. The second method, called transcriptional targeting, involves altering the viruses so that their genes can be active, or transcribed, only in tumors [*see box on next two pages*].

Transductional targeting is particularly necessary because, unfortunately, adenoviruses bind more efficiently to the variety of normal tissues in the human body than they do to most tumor cells. We can reverse this pattern using specially generated adapter molecules made of antibodies that snap onto the arms of the virus like sockets on a socket wrench. By attaching carefully chosen antibodies or other molecules that selectively bind only to a specific protein found on tumor cells, we can render adenoviruses unable to infect any cells but cancerous ones. Once the antibody-bearing virus latches onto a targeted cell, the hapless cell engulfs it in a membrane sac and pulls it inside. As the sac disintegrates, the viral capsid travels to a pore in the cell's nucleus and injects its own DNA. Soon the viral DNA directs the cell to make copies of the viral DNA, synthesize viral proteins and combine the two into billions of new adenoviruses. When the cell is full to capacity, the virus activates a "death gene" and prompts the cell to burst, releasing the new viruses to spread to other cells.

The viruses can also be engineered more directly. In this regard, Curiel's group at the University of Alabama's Gene Therapy Center has designed adenoviruses that bind to cellular proteins called integrins. These molecules help cells stick to the network of connective tissue, called the extracellular matrix, that organizes the cells into cohesive tissues. Although integrins are also made by healthy cells, cancer cells produce them in abundance as they become metastatic and begin to squeeze through tissue layers and travel throughout the body. The University of Alabama research group has had encouraging results using the engineered viruses in mice bearing human ovarian cancers. The viruses homed in on the ovarian tumor cells

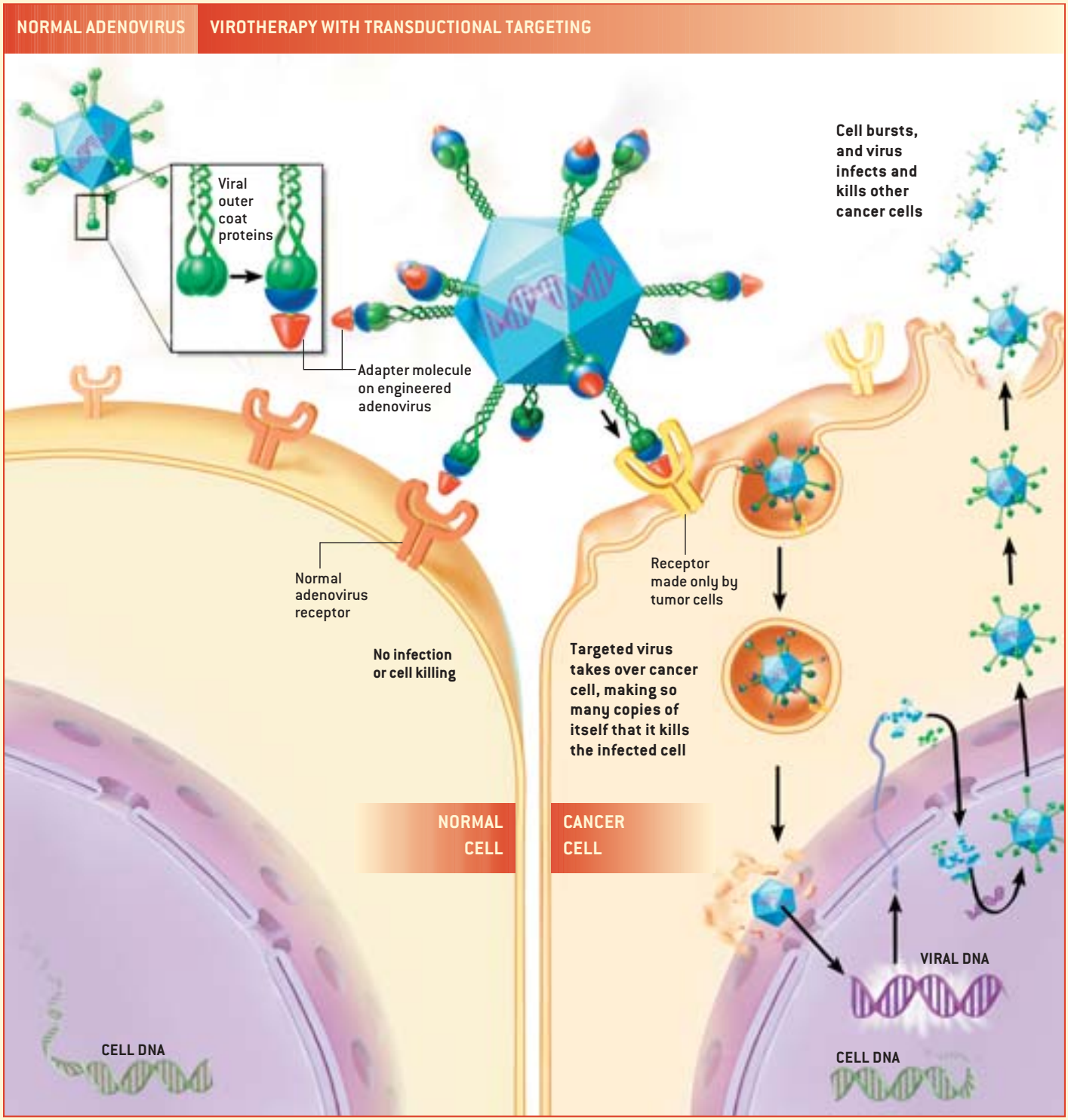
THE AUTHORS

DIRK M. NETTELBECK and *DAVID T. CURIEL* began their collaboration at the Gene Therapy Center of the University of Alabama at Birmingham (UAB), where Curiel is director of the division of human gene therapy. Curiel, who holds an M.D. and a Ph.D., is the Jeanne and Anne Griffin Chair for Women's Cancer Research at UAB and a professor of gene therapy at the Free University of Amsterdam. Nettelbeck—who is now heading a research group focusing on virotherapy for malignant melanoma in the department of dermatology at the University of Erlangen-Nuremberg in Germany—was a molecular biologist and postdoctoral fellow of the German Research Association at the University of Alabama from 2000 to 2003. He received his Ph.D. in 2000 from Philipps University in Marburg, Germany, and was honored with a graduation award from the Novartis Foundation for Therapeutic Research.

DIRK M. NETTELBECK *University of Erlangen* AND N. SANJIB BANERJEE, THOMAS R. BROKER AND LOUISE T. CHOW *University of Alabama at Birmingham*

ZAPPING CANCER CELLS WITH VIRUSES

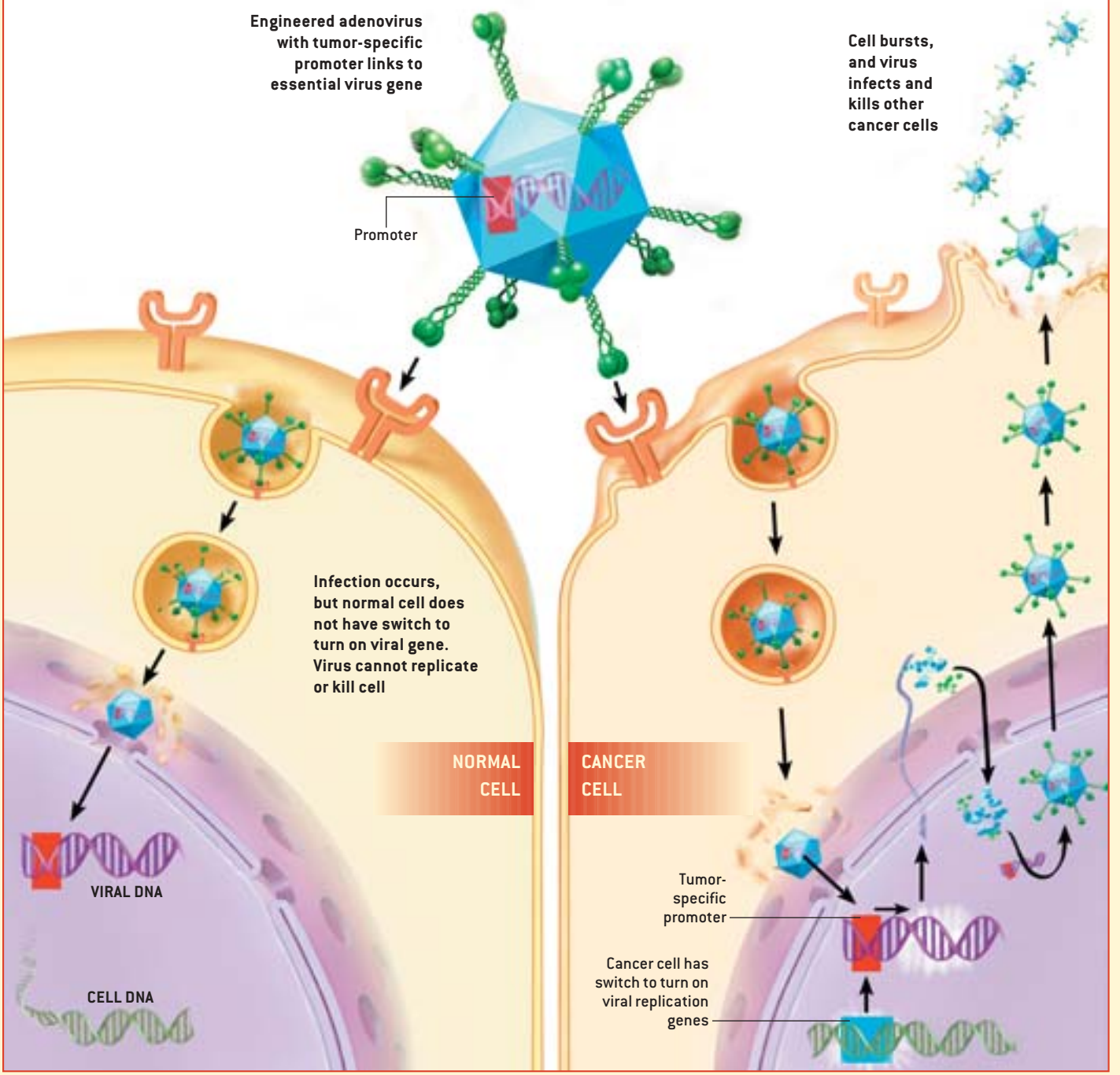
TWO MAIN STRATEGIES are being explored for virotherapy, which is the technique of using reproducing viruses to kill tumors. In the first method, dubbed transductional targeting (*below*), scientists are attempting to engineer viruses such as adenovirus—which normally causes respiratory infections—to selectively infect and destroy only cells that have turned cancerous. They are attaching adapter molecules onto the viral outer coat proteins or directly modifying these proteins to try to prevent the viruses from entering normal cells and instead prompt them to home in



TERESE WINSLOW

on tumor cells. The second approach (*below*) involves placing a snippet of DNA called a tumor-specific promoter next to one of adenovirus's essential genes. The promoter acts as an "on" switch that permits the gene to function only in cancer cells. The engineered viruses can enter normal cells, but they cannot reproduce and kill them. Once they enter cancer cells, however, the tumor-specific promoter lets them make millions of copies of themselves and ultimately burst the cancer cells. They can then spread to—and destroy—other tumors. *—D.M.N. and D.T.C.*

VIROTHERAPY WITH TRANSCRIPTIONAL TARGETING



and killed them, ridding the treated animals of the disease.

Transcriptional targeting generally takes advantage of genetic switches (promoters) that dictate how often a given gene is functional (gives rise to the protein it encodes) in a particular type of cell. Although each body cell contains the same encyclopedia of genetic information, some cells use different chapters of the encyclopedia more often than others in order to fulfill their specialized tasks. Skin cells called melanocytes, for instance, must make much more of the pigment melanin than liver cells, which have little use for the protein. Accordingly, the promoter for the key enzyme for making melanin gets turned on in melanocytes but generally is off in most other body tissues. In the deadly skin cancer melanoma, the gene encoding this enzyme is fully functional, making the tumors appear black. We, and others, have engineered adenoviruses that have a promoter for the enzyme adjacent to genes that are essential for the viruses' ability to replicate. Although these viruses might infect normal cells, such as liver cells, they can reproduce only inside melanocytes, which contain the special combination of proteins needed to turn on the promoter.

Researchers are currently tailoring adenoviruses with a variety of promoters that limit their activity to particular organs or tissues. In liver cancers, for example, the promoter for the gene α-fetoprotein—which is normally shut down after fetal development—becomes reactivated. Adenoviruses containing that same promoter hold promise for eradicating liver tumors. Scientists led by Jonathan W. Simons at Johns Hopkins University have tested the approach in men whose prostate cancer recurred following treatment with radiation. The researchers used adenoviruses that had been engineered by Cell Genesys to contain the promoter for prostate-specific antigen, a protein made in abundance by prostate tumors. They administered the virotherapy to 20 men who received varying doses of the adenoviruses. In 2001 Simons and his colleagues reported that none of the men experienced serious side effects and that the tumors of the five men who received the highest doses of the virotherapy shrank by at least 50 percent.

Other Strategies

VIROTHERAPISTS MIGHT END UP combining the transductional and transcriptional targeting strategies to ensure that the viruses kill only tumor cells and not normal ones. Adenoviruses engineered to contain the promoter for the enzyme that makes melanin, for instance, can also replicate in normal melanocytes, so on their own they might cause spots of depigmentation. And adenoviruses that are designed to bind to receptors on the surfaces of tumor cells can still invade a small proportion of healthy cells. But viruses altered to have several

But Is It Safe?

Many approaches to virotherapy use adenoviruses, which caused a death in a clinical trial of gene therapy four years ago

IN SEPTEMBER 1999 18-year-old Jesse Gelsinger died after receiving an infusion of adenoviruses into his liver. He had a mild form of an inherited liver disease called ornithine transcarbamylase deficiency (OTCD) and was participating in a clinical trial of a new gene therapy to use adenoviruses to ferry a corrected copy of the gene encoding OTCD into his liver cells. Unfortunately, four days after an infusion of the viruses, he died of acute respiratory distress syndrome and multiple organ failure, apparently caused by an overwhelming immune reaction to the large dose of adenoviruses he had been administered as part of the trial.

Although Gelsinger's death was part of a gene therapy trial, the tragedy also has ramifications for the new field of virotherapy. Gene therapy uses crippled versions of viruses such as adenovirus to introduce a new gene into cells; virotherapy employs actively replicating viruses (which may or may not contain added genes) to kill specific types of cells. Both, however, rely heavily on adenoviruses.

JESSE GELSINGER, who died in 1999 after receiving an infusion of adenoviruses, in a family photograph.

Gelsinger's autopsy showed that the engineered adenoviruses had spread to his spleen, lymph nodes and bone marrow, and an examination of his records revealed that his liver function was probably too impaired for him to be a volunteer in the trial. A number of scientists have also suggested that he might have mounted such an extreme immune reaction because he had previously been infected with a naturally occurring adenovirus.

Since Gelsinger's death, gene therapists and virotherapists alike have focused on refining adenoviruses to make them safer. But researchers are still unsure why Gelsinger reacted so violently to the adenoviral infusions: a second patient participating in the same clinical trial tolerated a similar dose of the viruses. And dozens of other people worldwide have been treated so far with adenoviruses with no serious side effects.

A National Institutes of Health report generated in the aftermath of Gelsinger's demise recommends that all participants in such clinical trials be monitored closely for toxic reactions before and after the infusion of therapeutic viruses. It also stipulates that volunteers be screened for any predisposing conditions that would increase their sensitivity for the viruses.

—D.M.N. and D.T.C.

MBR/KRT

SELECTED COMPANIES INVOLVED IN VIROTHERAPY

Company	Headquarters	Virus	Diseases	Viral Modifications	Clinical Trial Status
BioVex	Abingdon, Oxfordshire, U.K.	Herpes simplex virus (HSV)	Breast cancer and melanoma	Carries the gene for granulocyte-macrophage colony stimulating factor, an immune system stimulant	Phase I/II
Cell Genesys	South San Francisco, Calif.	Adenovirus	Prostate cancer	Targeted to prostate cancer cells using prostate-specific promoters	Phase I/II
Crusade Laboratories	Glasgow	HSV	Glioma (brain cancer), head and neck cancer, melanoma	Has a gene deletion that restricts it to actively dividing cells such as cancers	Phase II for glioma and head and neck cancer; Phase I for melanoma
MediGene	Martinsried, Germany	HSV	Glioma and colon cancer that has spread (metastasized)	Harbors two gene deletions that prevent it from reproducing in normal cells	Phase II for glioma; Phase I for colon cancer metastases
Oncolytics Biotech	Calgary, Alberta, Canada	Reovirus	Prostate cancer and glioma	Able to replicate only in cancer cells bearing the activated oncogene *ras*	Phase II for prostate cancer; Phase I/II for glioma

NOTE: Phase I tests are designed to evaluate safety in small numbers of patients. Phases II and III are intended to determine the appropriate dose and efficacy, respectively.

fail-safe mechanisms would be expected to be less likely to harm normal cells. There are no results at present, however, to demonstrate that a combination of approaches makes viruses more targeted.

A further strategy for targeting virotherapy makes the most of one of cancer's hallmarks: the ability of tumor cells to divide again and again in an uncontrolled manner. Healthy cells make proteins that serve as natural brakes on cell division—notably, the retinoblastoma (Rb) and p53 proteins. As cells turn cancerous, however, the genes that code for one or the other of these proteins become mutated or otherwise inactivated. Certain viruses, including adenovirus, interfere with the braking mechanisms of a normal cell by making proteins that stick to and inactivate Rb or p53. They do this because they can replicate only in cells that are preparing to divide.

Several research groups and biotechnology companies have engineered adenoviruses that fail to make the Rb or p53 blockers. Normal cells, which make these blockers, will stall the replication of these viruses by putting the brake on cell division. But these viruses will replicate in cells in which the Rb or p53 proteins are already disabled—cancer cells—and kill them. Curiel is planning clinical trials of the approach for ovarian cancer.

Researchers are also arming therapeutic viruses with genes that make the cells they infect uniquely susceptible to chemotherapy. The technique involves splicing into the viruses genes that encode enzymes that turn nontoxic precursors, or "prodrugs," into noxious chemotherapies. In one example, which was reported in 2002, André Lieber of the University of Washington and his co-workers designed adenoviruses to carry genes encoding the enzymes capable of converting innocuous prodrugs into the anticancer compounds camptothecin and 5-fluorouracil. The scientists engineered the viruses so that they could make the enzymes only in actively dividing cells, such as cancer cells. When they injected the viruses and the prodrugs into mice bearing implanted human colon or cervical cancer cells, they found that the viruses reproduced and spread in the tumors.

Such "smart" virotherapies are the vanguard of the future. But physicians will also need to track the activity of virotherapies in a patient's body to best assess how well the strategies are working and refine them further. Virotherapists are now teaming with radiologists to establish novel imaging technologies to easily measure how effectively a given virotherapy is replicating.

The imaging strategies involve inserting a gene that governs the production of a tracer molecule into a virus or virus-infected cell. The tracer can be either a fluorescent protein that can be observed directly or one that binds readily to the radionuclides used in standard radiological imaging techniques. The fluorescent protein might work best for cancers that are accessible by an endoscope, such as cancers of the larynx. Physicians could peer into the endoscope and see exactly where the viruses—and therefore, cancer cells—are by looking for fluorescence. So far the approach has worked best with viruses that do not kill cells, however. Nevertheless, we are convinced that such sophisticated imaging technologies will enable scientists to draw more meaningful conclusions from future clinical trials of virotherapy.

In 1995 gene therapy pioneer W. French Anderson of the University of Southern California School of Medicine predicted in this magazine that "by 2000 … early versions of injectable vectors that target specific cells will be in clinical trials." These trials indeed began on schedule, as well as some he could not have envisioned then. We envision a substantial role for viruses—that is, *therapeutic* viruses—in 21st-century medicine. SA

MORE TO EXPLORE

Gene Therapy: Designer Promoters for Tumour Targeting. D. M. Nettelbeck, V. Jérôme and R. Müller in *Trends in Genetics,* Vol. 16, pages 174–181; 2000.

Replicative Adenoviruses for Cancer Therapy. R. Alemany, C. Balagué and D. T. Curiel in *Nature Biotechnology,* Vol. 18, pages 723–727; 2000.

Vector Targeting for Therapeutic Gene Delivery. Edited by D. T. Curiel and J. T. Douglas. John Wiley & Sons, 2002.

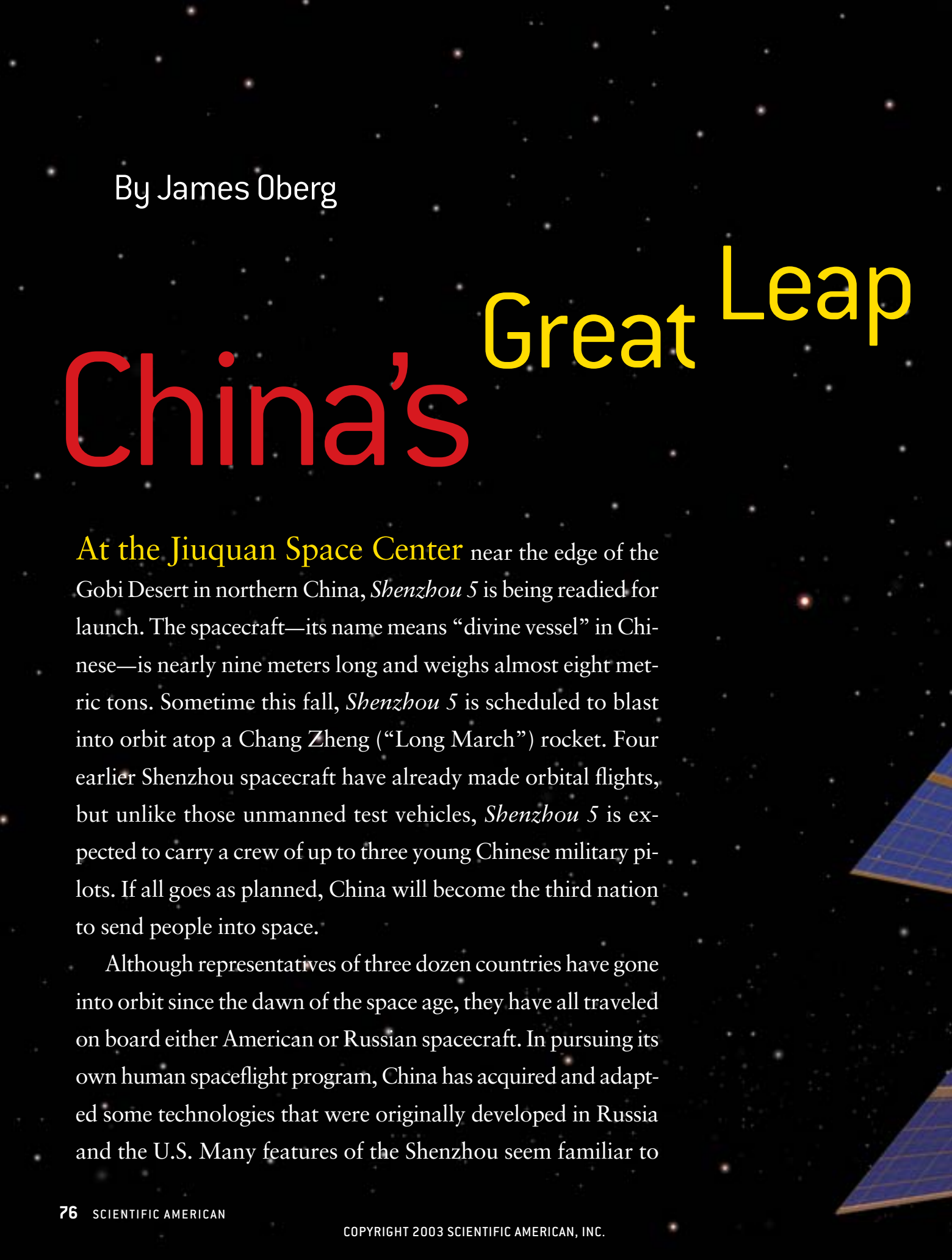
Cytolytic Viruses as Potential Anti-Cancer Agents. C.J.A. Ring in *Journal of General Virology,* Vol. 83, pages 491–502; 2002.

Gene therapy clinical trials database of *Journal of Gene Medicine:* **www.wiley.com/legacy/wileychi/genmed/clinical/**

American Society of Gene Therapy: **www.asgt.org**

By James Oberg

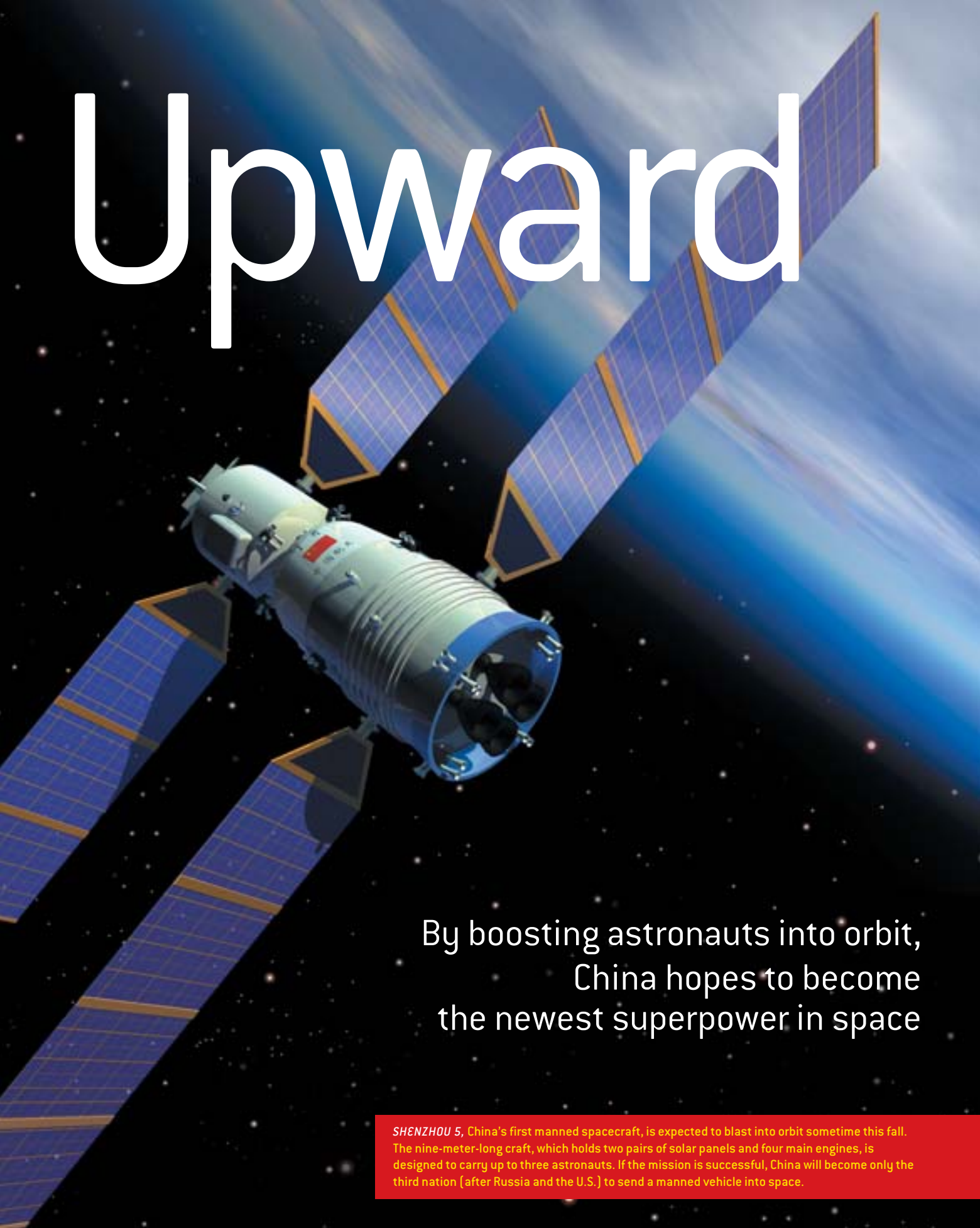
China's Great Leap



At the Jiuquan Space Center near the edge of the Gobi Desert in northern China, *Shenzhou 5* is being readied for launch. The spacecraft—its name means "divine vessel" in Chinese—is nearly nine meters long and weighs almost eight metric tons. Sometime this fall, *Shenzhou 5* is scheduled to blast into orbit atop a Chang Zheng ("Long March") rocket. Four earlier Shenzhou spacecraft have already made orbital flights, but unlike those unmanned test vehicles, *Shenzhou 5* is expected to carry a crew of up to three young Chinese military pilots. If all goes as planned, China will become the third nation to send people into space.

Although representatives of three dozen countries have gone into orbit since the dawn of the space age, they have all traveled on board either American or Russian spacecraft. In pursuing its own human spaceflight program, China has acquired and adapted some technologies that were originally developed in Russia and the U.S. Many features of the Shenzhou seem familiar to

Upward



By boosting astronauts into orbit, China hopes to become the newest superpower in space

SHENZHOU 5, China's first manned spacecraft, is expected to blast into orbit sometime this fall. The nine-meter-long craft, which holds two pairs of solar panels and four main engines, is designed to carry up to three astronauts. If the mission is successful, China will become only the third nation (after Russia and the U.S.) to send a manned vehicle into space.

space experts; at first glance, the craft looks like a slightly bigger version of the Russian Soyuz vessel. But China's space agency built several key systems on its own, and in some ways the Shenzhou is technologically superior to the Soyuz.

The first manned flight of the Shenzhou will have profound diplomatic echoes. In addition to garnering international prestige, China hopes that the program will stimulate advances in the country's aerospace, computer and electronics industries. If the initial missions are successful, China will probably establish its own space station in Earth orbit. Within a decade, China's space activities may well surpass those of Russia and the European Space Agency. And if China becomes the most important space power after the U.S., an entirely new "space race" may be in the offing.

The Upcoming Flight

CHINA'S SPACE AMBITIONS date back to 1970, when a Chang Zheng rocket boosted the country's first satellite into orbit. (The meter-wide craft transmitted the patriotic song "The East Is Red" as it circled the globe.) In the early 1990s President Jiang Zemin gave the go-ahead for a manned space program, and the decision set off a tremendous construction boom that was in some ways as impressive as NASA's buildup for the Apollo project in the 1960s. At the Jiuquan Space Center, the oldest and largest of the three locations where Chinese missiles and satellites are launched, a new rocket-assembly building and launch pad were completed in time for the first unmanned Shenzhou flight in 1999. Although the agency overseeing the program, the Chinese Academy of Space Technology, has been secretive about certain aspects—such as the training of its astronauts—many technical details about the spacecraft have appeared in the Chinese press.

The Shenzhou consists of three modules. The crew will ride in a coffee-cup-shaped unit—the command module, in NASA parlance—that contains their couches (on which the crew members lie horizontally during liftoff) and control panels. Attached to the front of the command module is a cylindrical pressurized unit (referred to as the orbital module) that will provide extra workspace for the astronauts. Behind the command module is a cylindrical equipment section (the service module) that holds rocket engines and an electrical power system. This module also sports a pair of solar power panels with a total area of 24 square meters.

The official designation of the launch vehicle is Chang Zheng 2F, but it is also poetically referred to as the Shen Jian ("divine arrow"). It is a version of the Chang Zheng 2E—a liquid-propellant rocket derived from Chinese military missiles—that has been upgraded to make it reliable enough for manned flight. Like other launch vehicles of this family that have been designed to carry heavier payloads, its original four-engine core is augmented by four strap-on boosters. All the rocket engines use liquid nitrogen tetroxide and hydrazine propellants, not the more powerful but troublesome liquid hydrogen fuel used by the main engines of NASA's space shuttle. The Shen Jian's launch thrust of 604,000 kilograms of force (kgf) puts it between the standard booster for the Soyuz (411,000 kgf) and the Saturn 1B (740,000 kgf) that carried several Apollo spacecraft into Earth orbit.

Shenzhou 5 will take the same flight path as its four unmanned predecessors. After liftoff, the booster will head east, crossing the Chinese coastline near Qingdao (site of a new tracking station), then move over the Yellow Sea at a maximum acceleration of about five g's. Passing just south of the Japanese island of Kyushu, the spacecraft will be in communication with a large tracking ship and soon achieve an orbital velocity of about eight kilometers a second. Its orbital inclination will be about 42 degrees, somewhat more equatorial than the 52-degree inclination of the International Space Station.

To return to Earth, the Shenzhou first jettisons the orbital module. The service module then fires a small braking engine to slow the spacecraft's velocity by about 200 meters a second. This maneuver allows Earth's gravity to tug the craft down into the upper atmosphere, where air braking slows it down further. Because the aim point is a landing zone in Inner Mongolia about 500 kilometers east of the Jiuquan Space Center, the braking engine fires while the spacecraft is still half a world away, over the South Atlantic Ocean. After the firing, the service and command modules separate; the former burns up in the atmosphere, whereas the latter is equipped with a heat shield on its base that protects the crew from the searing temperatures of reentry.

The landing sequence follows the pattern of the Soyuz and Apollo spacecraft. After a relatively high-g deceleration, the command module falls freely though the lower atmosphere, releasing a drogue parachute at an altitude of about 30 kilo-

Within a decade, China's space activities may well surpass those of Russia and the European Space Agency.

Overview/*China's Manned Space Program*

- Since 1992 China has been striving to put people into orbit. In addition to developing the Shenzhou spacecraft, the country has upgraded its launch vehicles, built new spaceflight facilities and trained a cadre of astronauts.
- Though superficially similar to the Russian Soyuz spacecraft, the Shenzhou is larger and, in some ways, technologically superior.
- China's government expects its manned space program to enhance the nation's world status and stimulate its high-tech industries.

ALFRED T. KAMAJIAN (*preceding pages*)

UNMANNED TEST FLIGHT of *Shenzhou 2* began with a night launch on January 10, 2001. The two-stage Chang Zheng 2F rocket gets added thrust from four strap-on boosters at its base.

WANG JIANMIN *Xinhua/Reuters*

meters and then a larger main parachute. The heavy heat shield is jettisoned. Touchdown on the open steppe is softened by the triggering of a set of rockets in the capsule's base. Recovery teams in the zone rush to the landing point and retrieve the crew.

Copy of Soyuz?

SHENZHOU 1, the first unmanned test vehicle, was launched on November 20, 1999, and spent less than a day in orbit. During each of the three subsequent flights (launched on January 10, 2001, March 25, 2002, and December 30, 2002), the spacecraft spent a week in orbit as progressively more sophisticated hardware was tested. *Shenzhou 2* reportedly carried a monkey, a dog and a rabbit to test the craft's life-support system. *Shenzhou 4* was described as a human-capable vessel, with all the equipment needed to accommodate three people. In fact, astronauts reportedly took part in the countdown inside the command module, leaving the craft only a few hours before liftoff.

Because of the superficial resemblance of the Shenzhou to the Soyuz, many observers concluded that the Chinese vehicle was nothing more than a knockoff with only minor modifications. David Baker, editor of the British publication *Jane's Space Directory*, said the Shenzhou is "a Soyuz to a very considerable extent—it's off the shelf, as it were." But Brian Harvey, author of *The Chinese Space Programme: From Conception to Future Capabilities*, has criticized this view. "There are a lot of misperceptions about the Chinese space program," Harvey says, referring to myths of low-technology spacecraft and widespread copying of foreign designs. "A lot of it reflects a Western cultural notion that the Chinese couldn't possibly master this kind of technology. I think it's more helpful to look at the way in which the Chinese have built their program up over the years—slowly, patiently, carefully, in a disciplined way, with the careful selection of choices, borrowing from elsewhere, but only to a limited extent."

Most other independent experts agree. Some reports indicate that although the Chinese had asked to purchase a fully functional Soyuz vehicle for study, the Russians demanded a price so high that the deal never took place. The capsule the Chinese bought was a stripped-down version missing many key space systems. And the Chinese may also have picked up at least one off-course, unmanned Soyuz-class capsule that crashed inside their country early in the space age.

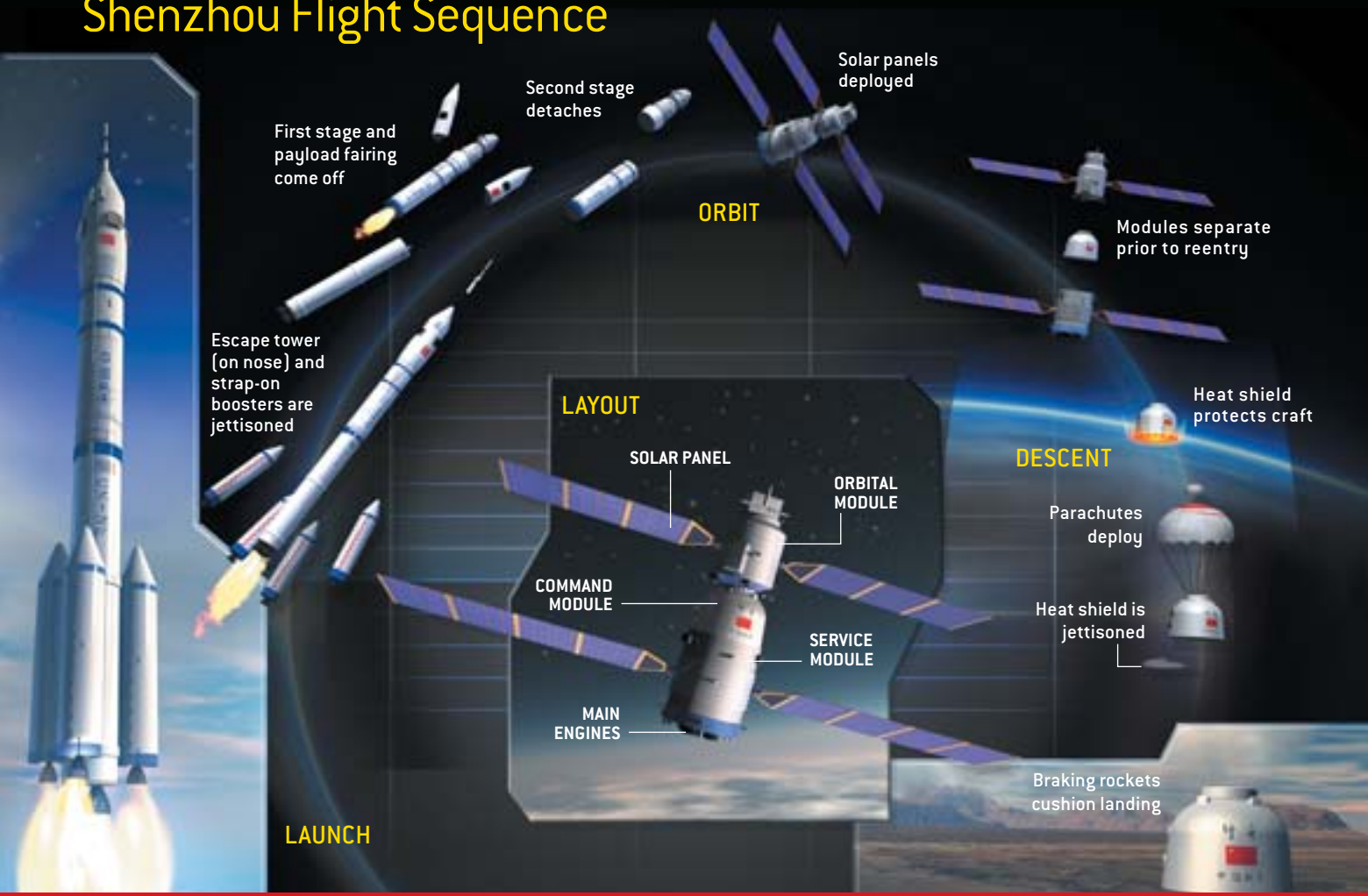
The three-module design of the Shenzhou is a logical arrangement, first developed independently by U.S. and Soviet teams more than 40 years ago. The Shenzhou looks less like today's Soyuz than like the preliminary designs for the Soyuz and Apollo. Although the Shenzhou's command module has some similarities to a Soyuz reentry capsule, the other modules are quite different from their Russian and American counterparts. The service module, for example, has four main engines, whereas Apollo's service module had only one and the Soyuz has one main and one backup engine. Also, the Shenzhou's large solar arrays generate several times more electrical power than the Russian system. And unlike the Soyuz, the Chinese orbital module carries its own solar panels and independent flight-control system, allowing it to continue as a free-flying, unmanned mini laboratory long after the reentry module has brought the crew back to Earth.

One example of outright Chinese copying is the pressure suits used to protect the astronauts in case of an air leak in the cabin during flight. The Russians introduced such a survival suit (called the

THE AUTHOR

JAMES OBERG spent 22 years at NASA Mission Control in Houston as a contractor employee specializing in orbital rendezvous for the space shuttle. He is now a consultant, lecturer and widely published author, having written a dozen books and countless magazine articles on the U.S. and Russian space programs. His interests include spaceflight safety and reliability, the contribution of space activities to national power, and the hidden histories of the world's space programs. Since February he has been the space consultant for NBC News. His Web site is available at www.jamesoberg.com

Shenzhou Flight Sequence



CHINA'S FIRST MANNED MISSION will most likely follow the same sequence that was practiced in the four unmanned test flights. As the rocket rises, its strap-on boosters, escape tower, first stage and payload fairing are cast off. Once in orbit, the rocket's second stage detaches from the Shenzhou spacecraft, which then deploys its solar panels. The Shenzhou's modules separate prior to reentry: the orbital module remains in space, the service module burns up in the atmosphere, and the command module returns the crew to Earth, using parachutes and braking rockets to soften the landing.

Sokol) in 1972 after three cosmonauts died in an accidental cabin depressurization during their return to Earth. (A much more sophisticated suit is used for spacewalks.) The Chinese needed a suit with similar functions, so after obtaining samples of the Sokol design they copied it exactly, right down to the stitching and color scheme.

Chinese officials have made no secret of such technology transfers. An article on Chinese space plans appearing in the Xinhua News Agency's magazine *Liaowang* in 2002 stated: "After China and Russia signed a space cooperation agreement in 1996, the two countries carried out very fruitful cooperation in docking system installations, model spaceships, flight control, and means of life support and other areas of manned spaceflight. Russia's experience in space technology development was and is of momentous significance as enlightenment to China."

The mention of docking systems is especially illuminating. Although Russia and the U.S. have used different types of docking mechanisms over the years to link spacecraft in orbit, photographs of the Shenzhou indicate that the Chinese have chosen a Russian variant called the APAS-89. The device consists of a pressurized tunnel 80 centimeters in diameter surrounded by sloping metal petals that allow any two units of the same design to latch together. Originally developed for the Russian Mir space station, the APAS-89 is used to dock NASA's space shuttles to the International Space Station (ISS). Although China is primarily interested in docking its spacecraft with its own small space stations, the decision to employ the APAS-89 mechanism allows the Shenzhou to link with both the space shuttles and the ISS.

Building the Booster

THE DEVELOPMENT of the launch vehicle for the Shenzhou also illustrates China's strategy. The rocket is an outgrowth of earlier versions used for scientific and commercial launches, but some very specific improvements have been made. According to Liu Zhusheng, the chief designer of the Chang Zheng 2F booster, the reliability of the vehicle is rated at 97 percent—that is, it is expected to fail in only three of 100 launches. (The rated reliability of the unmodified Chang Zheng booster is 91 percent.) Furthermore, the Shenzhou can separate from its booster in the event of a launch catastrophe. Speaking to Xinhua News Agency reporters last January, Liu described a new malfunc-

ALFRED T. KAMAJIAN

tion-diagnosis system designed to monitor the booster's status and, if needed, activate backup mechanisms or initiate the spacecraft's escape. He claimed that the escape system gives the crew a survival probability of 99.7 percent.

The Shenzhou's launch escape system is based on the "tractor rocket" scheme originally developed for NASA's Mercury flights in the early 1960s and later adopted by the designers of the Apollo and Soyuz spacecraft. A set of solid-propellant rockets mounted on a tower on the nose of the Shenzhou pulls the spacecraft clear of the booster in the event of a launch disaster. The spacecraft then parachutes to the ground. The solid-fuel rockets are ready to fire from 15 minutes before liftoff until 160 seconds afterward (at which point the spacecraft is at an altitude of 110 kilometers).

In an interview with *People's Daily*, the official newspaper of the Chinese Communist Party, launch vehicle manager Huang Chunping showed a reporter a model of the booster that had four wings on the payload fairing that surrounds the Shenzhou during liftoff. The wings, each resembling a lattice, are required to keep the Shenzhou stable during a launch escape. "This is the most difficult part of the escape system," he explained. "We once wanted to inquire about it from Russian experts, but they set the price at $10 million. Finally, we solved the problem on our own." This pattern of studying previous work but designing the actual flight hardware independently was followed on most other Shenzhou systems.

Astronauts or *Taikonauts?*

CHINA DID, HOWEVER, receive Russian help in training its astronauts. As part of the 1996 cooperation agreement between the two countries, two Chinese military pilots—Wu Jie and Li Qinglong—took courses at Russia's Gagarin Cosmonaut Training Center in Star City outside Moscow. They reportedly are now directing the training program for China's flight candidates. Twelve other pilots were later picked to join them in the first group of Chinese astronauts. Last January, Hong Kong newspapers named Chen Long as the leading candidate for the command of the first manned flight.

All the Chinese astronauts are young, university-educated jet pilots with more than 1,000 hours of flying experience. Because the Shenzhou, like the Soyuz, has a limited amount of cabin space, small stature is a requirement: each astronaut is under 170 centimeters (five feet, seven inches) tall and weighs less than 65 kilograms (143 pounds). What the Chinese astronauts should be called is still in dispute. One Chinese space enthusiast coined the term *taikonaut,* from the Chinese *tai kong,* or "outer space." Chinese officials and newspapers prefer *yuhangyuan,* which roughly translates as "space navigator."

In preparation for the Shenzhou missions, China vastly improved its facilities for spaceflight research and training. The new China Space Center in Aerospace City, a southwestern suburb of Beijing, is home to a collection of pressure chambers, space-vehicle simulators, a centrifuge and a landing-impact tower, along with classrooms and medical facilities for astronauts. Mission control for the Shenzhou flights is also located here. In addition, the new Space Technology Research and Test Center in Tangjialing, northwest of Beijing, has spacecraft-integration halls, space-environment chambers (including the world's fifth largest vacuum chamber) and vibration-test facilities.

What is more, China has four ocean-going ships to track its missiles and spacecraft. These Yuan Wang ("Long View") ships have been deployed in the Pacific Ocean to monitor military missile tests and in the Indian Ocean to control the maneuvering of satellites into geosynchronous orbit. The ships are sent into the South Atlantic, Indian and South Pacific oceans to support the Shenzhou flights. The Russians used to have a similar fleet but scrapped it in the 1990s because of budget constraints. Rather than purchase the Russian ships, China built its own.

Because some of the critical ground-control functions for the Shenzhou's return to Earth must be performed while the craft is over the South Atlantic, China signed an agreement with the African nation of Namibia in 2000 to build a tracking station near the town of Swakopmund. Construction started in early 2001 and was completed by year's end. Five permanent residents occupy the facility, and the staff expands to 20 during missions. The site lies under the reentry path of the Shenzhou, and because the craft's orbit has a different inclination than the International Space Station's, the Namibian base could not be used to track flights returning from there. This

GETTY/AFP

RECOVERY TEAMS rushed toward the command module of the *Shenzhou 3* after it landed in Inner Mongolia on April 1, 2002. The unmanned spacecraft had spent a week in orbit.

Trajectory of the "Divine Vessel"

AFTER LIFTOFF at the Jiuquan Space Center, the Shenzhou ascends over the Pacific Ocean, where it is tracked by one of China's Yuan Wang ("Long View") ships. To return to Earth, the craft fires a braking engine over the South Atlantic Ocean. This retro-burn is monitored by a tracking station near Swakopmund in Namibia. The landing zone is about 500 kilometers east of Jiuquan, which is the largest of China's three launch sites (*inset*).

arrangement suggests that despite the Shenzhou's compatible docking gear, the Chinese have no near-term interest in visiting the ISS.

China's Goals

THE FIRST UNMANNED test flight of the Shenzhou in 1999 prompted a great outpouring of national pride. "Landmark Launch Shakes the World!" thundered *People's Daily*. But to justify the expense of the Shenzhou program—officials say 19 billion yuan, or $2.3 billion, has been spent so far—China surely expects more than feel-good headlines. The government anticipates that a successful manned space program will enhance China's world status and the reputation of its high-tech exports, giving the country greater diplomatic and commercial power.

Furthermore, China sees space technology as critical to quickly achieving technological parity with Western nations and Japan. A white paper issued in 2000 by the Information Office of the State Council claimed that the space industry is "an integral part of the state's comprehensive development strategy." According to an article in the June 2000 issue of *Xiandai Bingqi*, the monthly journal of a military technology research institute, the human spaceflight program "will raise levels in areas such as computers, space materials, manufacturing technology, electronic equipment, systems integration and testing." The journal also notes that the experience of developing spacecraft navigation, propulsion, life support and other subsystems could be applied to "dual-use military/civilian projects."

China's plan is to build an impressive stand-alone space capability on a narrow technological base. Instead of developing a wide variety of aerospace technologies, as the U.S. has done, China will focus on specific areas where it can match and then outdo the accomplishments of other nations. In an article appearing in *People's Daily* in April 2000, Luan Enjie, director of China's Aerospace Bureau, explained: "With limited state financial resources and very weak industrial and technical foundations, we do not have the strength to comprehensively catch up with and surpass world advanced levels in all aerospace fields."

Although China is still far from challenging the space status of the U.S., it may have more attainable goals in mind. If there is a new space race under way, it is for second place. Russia's space program faded in the 1990s and now preserves its remaining capabilities only through massive commercial sales to Western customers. With a GNP and federal budget five times as great as Russia's, China can easily afford to outspend America's former chief rival in the space race. But the Chinese are not being profligate. The $2.3 billion that China has reportedly spent on the Shenzhou project over the past decade is a little more than

NINA FINKEL

half of what NASA spends on the space shuttle every year.

The burgeoning Chinese space program is already beginning to eclipse the European Space Agency (ESA), which has never sent astronauts into orbit but leads the world in launching unmanned commercial satellites. The ESA faces political disenchantment among its member nations and cuts in its space research budget. Harvey, the spaceflight expert and author, says China may match the ESA's launch rate (about 10 liftoffs annually) over the next five years. "Europe's launches will be mainly commercial and scientific, whereas China will concentrate on applications [including weather and reconnaissance satellites] and its manned program," Harvey notes. In 2001 Roger-Maurice Bonnet, the retiring director of the ESA Science Program, declared that European governments must make the political decision to spend what it takes to maintain the ESA's second-ranked position in space research (after the U.S.). Otherwise, Bonnet warned, China would overtake the Europeans within a decade.

Many space experts agree that China's drive to overtake the Europeans and Russians is entirely credible. "China certainly has the political will to forge ahead with its space program," says Joan Johnson-Freese, formerly of the Asia-Pacific Center for Security Studies and now with the Naval War College. "[China] recognizes all the internal and external prestige-related benefits of space that the U.S. and the FSU [former Soviet Union] did in the 1960s, as well as the technology-industrialization-economic benefits that pushed Europe into space later."

As evidence of their determination, China's space officials have already laid out the steps to follow the initial manned Shenzhou missions. Zhang Qingwei, a leading official of the Chinese agency overseeing the program, told *People's Daily* last January: "[The] orbital cabin [remains in space] to lay a foundation for China's second-step manned spaceflight project—forming a docking link between a spacecraft and another flight vehicle." Observers generally interpret this to mean that in the near future, a manned Shenzhou craft will rendezvous and dock with an unmanned orbital module left in space by a previous mission.

Once this capability has been demonstrated, China could proceed to docking the Shenzhou with a small space laboratory. The aforementioned 2002 article in *Liaowang* magazine described the development plan: "After it succeeds in manned spaceflight, China will very soon launch a cosmic experimental capsule capable of catering to astronauts' short stays." This capsule is elsewhere described as "a laboratory with short-term human presence," to be followed later on by a space station designed for long-term stays. Last January, unnamed officials provided further background to Xinhua News Agency reporters: "As the next step, China will endeavor to achieve breakthroughs in docking technology for manned spacecraft and space vehicles, and will launch a skylab. After that it will build a long-term manned space station to resolve problems related to large-scale space science experiments and applied technology, and to make contributions to mankind's peaceful development of outer space."

Phillip S. Clark, a British space consultant specializing in Russian and Chinese technology, expects China's space agency to launch a small 12- to 14-ton laboratory, perhaps within the next two years. Clark predicts that in 2006 or 2007 China will loft a larger station similar to the Russian Salyut stations launched in the 1970s and 1980s. Ultimately, Clark believes, China will begin the orbital assembly of a structure like the 130-ton Russian Mir station.

In addition, China has been accelerating its deployment of unmanned satellites for communications, weather, navigation and space research. Chinese officials have also discussed plans to land small remote-controlled rovers on the moon by 2010. Some Western media have reported that China wants to land astronauts on the moon as well, but these accounts have come from unofficial sources and may have been mistranslated. A manned mission to the moon would be many times more expensive than the Shenzhou project, whereas the payoffs might be only incrementally more.

According to space experts such as Harvey, boosting astronauts into orbit will be enough to make the world see China in a new light. "There will be a perception that the country has reached space superpower status," he says. "If China follows that with its own Salyut-class space station, it will impress the Asian region specifically and the world as a whole." China's goals for its space program are obviously not the same as America's, Russia's or Europe's. Judging from the hardware already built and the infrastructure in place, it seems clear that for the foreseeable future China intends to follow its own path in space. SA

China sees space technology as critical to quickly achieving parity with Western nations.

MORE TO EXPLORE

Thread of the Silkworm. Iris Chang. Basic Books, 1996.

The Chinese Space Program: A Mystery within a Maze. Joan Johnson-Freese. Krieger Publishing Company, 1998.

The Chinese Space Programme: From Conception to Future Capabilities. Brian Harvey. Wiley-Praxis Series in Space Science and Technology. John Wiley & Sons, 1998.

More information about China's space efforts can be found at **www.spacedaily.com/dragonspace.html, www.astronautix.com, www.fas.org/spp/guide/china/index.html, people.itu.int/~meens/Shenzhou/shenzhou.htm, www.cnsa.gov.cn** and **www.cgwic.com/index1.html**

The Economics of Child Labor

Campaigns against child labor are most likely to succeed when they combine the long arm of the law with the invisible hand of the marketplace

By Kaushik Basu

DHANBAD COAL MINES, Bihar, India (*opposite page*)

EAST WIND TRUCK FACTORY, Shiyan, China (*right*)

PHOTOGRAPHS BY SEBASTIÃO SALGADO

In the early days of the industrial revolution, inventors were often very forthright about the aims of their innovations. The co-inventor of the roller spinning machine, English mechanic John Wyatt, promoted it as a way for textile factories to downsize their labor forces. The contraption was so easy to run, Wyatt said, that businesses didn't need as many skilled craftspeople with spinning wheels; they could get by with children instead. "Adopting the machine, a Clothier formerly employing a hundred spinners might turn

off thirty of the best of them but employ an additional ten infirm people or children," he wrote in 1741. The British attorney general was won over and, in granting a patent, noted how "even Children of five or six Years of age" could operate the machine.

Commending an invention for facilitating child labor is now a matter of distant history. By the end of the 19th century, child labor was on the decline in most industrialized nations. But globally the problem has not come to an end. In 2000, according to the International Labor Organization (ILO), 186 million children between the ages of five and 14—roughly one in six children—were illegal laborers, mostly in developing nations. Of these, 111 million did hazardous work, such as mining, construction and hard farm labor, with lifelong consequences for their health. Some eight million were slave laborers, child soldiers or prostitutes.

These figures have to be treated with caution. Child labor is notoriously difficult to measure or even to define. The ILO is careful not to count ordinary household chores as child labor; nonetheless, in different ways, its estimates both overcount and undercount the problem. At times, a modest amount of work classifies a child as a laborer, which inflates the figures; on the other hand, girls' work around the home, which often comes at the expense of their education, is severely underestimated. Despite these caveats, researchers agree that the ILO estimates may not be far off the mark.

But what should be done about it? The answer depends critically on what gives rise to child labor and why it persists. With a large and growing number of researchers working on this topic, our understanding has deepened considerably in recent years. This scrutiny has called into question the zero-tolerance stance that many politicians and policymakers once took. It was common in the 1990s to hear calls for an immediate ban on imports of products made with child labor. Activists sought to involve the World Trade Organization in imposing trade sanctions against nations where child labor was prevalent. These appeals were an unfortunate product of genuine misunderstanding and, often, self-interested economic protectionism—an effort to protect jobs from foreign competition masquerading as a concern for impoverished children.

SUGAR CANE FIELDS, Zona da Mata, Pernambuco, Brazil

SEBASTIÃO SALGADO *Contact Press Images*

An example of what can go wrong happened in Nepal, as a study in 1995 by UNICEF described. In the 1990s opponents of child labor pushed for a global boycott of hand-knotted carpets made by children. Many Nepalese carpet makers had a simple response: they summarily fired their children. As a result, between 5,000 and 7,000 girls became prostitutes. A well-intentioned campaign ended up hurting the very people it sought to protect. With a little more sophisticated knowledge of economics, such failures could be avoided.

Overview/*Child Labor*

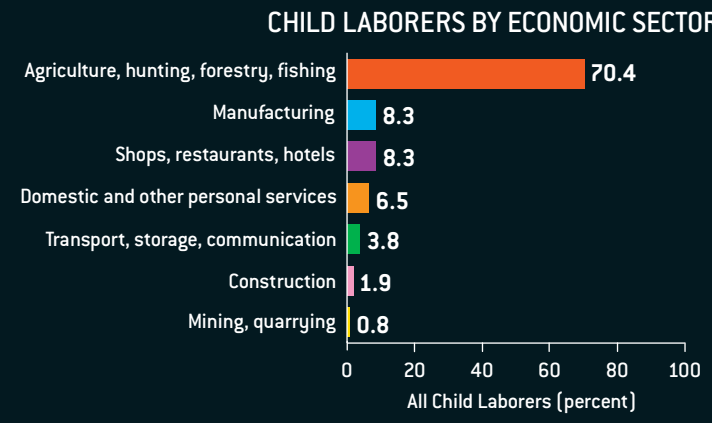
- Child labor, including some of its worst forms in factories, mines and brothels, continues to thrive—especially (though not only) in developing countries. Blanket bans on employing children in export industries can force kids into even worse situations, such as starvation. Policymakers need to take a more nuanced approach.
- Economic theory indicates that child labor can be self-reinforcing. It increases the pool of workers, keeping wages low—and ensuring that families must continue to send their children to work. Yet the abolition of child labor can also be self-reinforcing, by decreasing the labor supply, raising adult wages and eliminating the need for children to work. When these effects operate, selective bans and government assistance to families can tip an economy from the first condition to the second.

Hard Work

IN THE 19TH CENTURY commentators tended to blame child labor on parental sloth. Equating child labor with child abuse was a convenient assumption for justifying legislative action, and with repetition people came to believe it to be true. But most now recognize—and studies of working-class autobiographical writings and large data sets on household behavior in developing countries confirm—that the main cause of child labor is parental poverty. Few parents want to send their children to work unless forced by circumstance.

One of those household data sets, a 1991 study of rural Pakistan, illustrates how cause and effect operate. The mechanism is not always obvious. In some of the poorer areas, the study found that richer households were actually more likely to send their children to work than poorer ones. At first sight, the results seemed to contradict the poverty hypothesis. But in backward rural areas, labor markets often function inefficiently. Even if a household is impoverished and needs to supplement

Most people associate child labor with sweatshops, but manufacturing actually accounts for a fairly small fraction of economically active children. The majority work in agriculture, ranging from family farms to commercial plantations. In some countries, more than a quarter of farm workers are children under age 15.



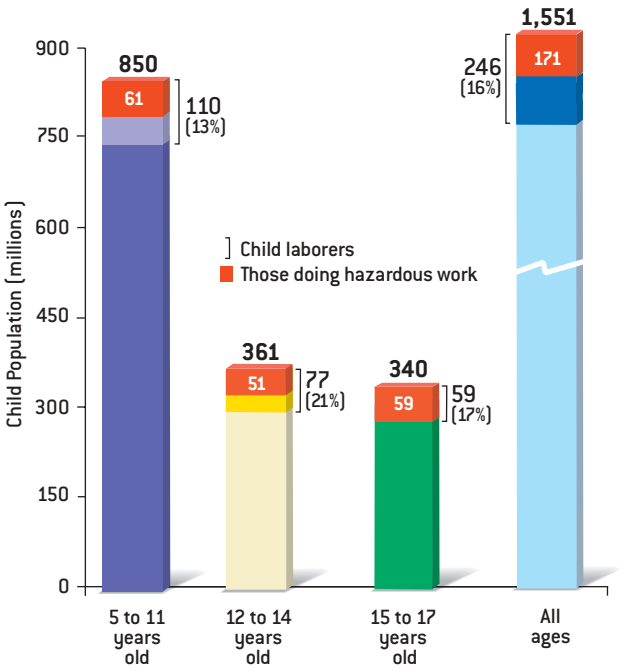
its income by putting the children to work, jobs may not be available. Only those households that own some land can avoid this problem: they can employ their children on their own property. Because land ownership is a form of wealth, it is not surprising to find that child labor is more common in richer households. Greater wealth does not cause the children to work but simply reflects land ownership, which creates opportunities to employ children.

Beyond a certain threshold of wealth, the incidence of child labor starts to go down. One can divide the Pakistani households into three categories: marginal (those owning less than one hectare of land), small (one to three hectares) and large (more than three). The percentage of children who worked increased from marginal to small households but declined for large households. Once a household becomes sufficiently rich, it has less need to make its children work.

The view that child labor declines with prosperity fits well with global time-series statistics. In China, for example, the percentage of children aged 10 to 14 who work—a number known as the child labor participation rate—steadily decreased from 48 percent in 1950 to 12 percent in 1995. The sharpest drop occurred in the 1980s, when the country's economic growth rate soared. Much the same has been true in Vietnam and India, for which Western economists have more reliable data. Conversely, in nations that have done worse economically, the decline in child labor has been marginal. In Cambodia, for instance, the child labor participation rate fell from 29 percent in 1950 to 25 percent in 1995.

The role of parental poverty illuminates why child labor is so difficult to eradicate. In Britain the practice continued to

Child Laborers by Age



CHILD LABOR is extremely common in the world today. The International Labor Organization counts 245.5 million laborers under the age of 17—nearly one in six young people. Of these, 170.5 million are doing hazardous work involving unsafe conditions, long hours or outright abuse (*red*). Below age 12, any child who works for pay is considered a laborer. From age 12 to 14, children can work up to 14 hours per week without being considered laborers. From age 15 and up, they can do any nonhazardous job.

SOURCE: INTERNATIONAL LABOR ORGANIZATION (1997) (*top*);
SOURCE: INTERNATIONAL LABOR ORGANIZATION (2000); UNITED NATIONS POPULATION DIVISION (*bottom*)

Easing into the Curve

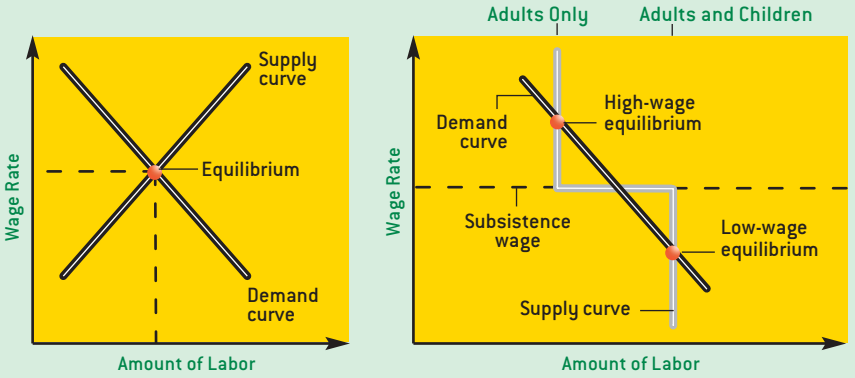
SUPPLY-DEMAND CURVES are one of the most basic conceptual tools in economics. The "curves" are two lines (which may or may not actually be curved) on a graph, symbolizing people's willingness to buy or sell a product depending on its price. Usually the curves cross at a single point, which indicates the price of the product in a free market of buyers and sellers (*left*).

But in labor markets where children are considered potential workers, the supply and demand curves can cross at more than one point. Households typically send their children to work when the income of the adults drops to intolerably low levels. The threshold for "intolerable" will vary from society to society and household to household, but let us ignore such variations for the sake of clarifying the basic principles. Let us also suppose that a child's labor is equivalent to some fraction of an adult's. Based on these assumptions, one can draw a supply curve for labor with multiple equilibria (*right*).

This curve shows how much labor is available to employers, depending on the wage they offer. It has a steplike shape. If the wage exceeds the subsistence level, households do not send their children to work, and each household supplies one unit of labor (provided by the adults). If the wage is lower than the subsistence level, each household supplies somewhat more than one unit of labor (provided by adults and kids together). The behavior of employers is described by a demand curve. Typically demand curves slope downward: for an employer to hire lots of workers, the wages must be low.

This market can settle into one of two equilibria. Either wages are high and adults work, or wages are low and both adults and children work. With more sophisticated assumptions, the supply curve could look different. For example, if children work only to make up the difference between their parents' income and the subsistence level, then the horizontal segment of the supply curve becomes a downward sloping curve (a hyperbola, to be precise). But the possibility of multiple equilibria remains.

When a market has only a single equilibrium, measures such as a legal ban on child labor have to fight against the natural tendencies of the marketplace. Even if these measures succeed, they come at the price of economic inefficiencies, which can create a different set of social problems. Multiple equilibria eliminate this dilemma. If a nation is caught in one equilibrium, a ban could deflect it to another equilibrium, at which point the marketplace would work with, rather than against, the policy. Economically, all these equilibria are equally efficient. —*K.B.*



IN ECONOMICS TEXTBOOKS, the labor-supply curve is a straight line slanting upward (meaning that high wages attract more workers); the demand curve slants downward (high wages discourage employers from hiring); and the two meet at a single point (*left*). But in poor countries, the supply curve can have a stair-step shape (*right*). Low wages actually attract more workers, because children must enter the work force for families to make ends meet. When this happens, the supply and demand curves meet at more than one point.

spread until around 1860, despite the adoption of new laws and policies throughout the first half of the 19th century. The effect of the laws was to impose a cost on firms that were found to be employing children. It is arguable that the added cost, by making children less attractive to hire, tended to lower their wages. But because children worked primarily to reach a minimal acceptable level of income for their households, a lower hourly wage induced them to work longer hours. Hence, paradoxically, the laws may have contributed to a worsening of child labor. The same risk is there with modern laws, such as India's Child Labor Act of 1986, that impose fines on firms that hire children.

Restoring Balance

SOME COMMENTATORS look at the data and jump to the opposite extreme, drawing the conclusion that there is no scope, or need, for governmental action against child labor. But that, too, is an overreaction. In certain situations, a legal ban can be extremely helpful in eliminating child labor while leaving the children and their parents better off. This argument emerges from analysis of demand and supply curves.

These curves show how a market comes to equilibrium, in which the price of a product adjusts to ensure that demand equals supply. In the case of labor markets, the product is the amount of work being performed, and the price is the wage rate. The wage rate determines the willingness of employees to work (the supply) and of employers to hire them (demand). In the standard textbook model, if the wage goes up, the supply of labor goes up but the demand goes down; if the wage drops, the opposite happens. There is only one wage rate at which the two match.

Some markets, however, differ from this textbook case and are characterized by multiple equilibria—that is, they have more than one wage rate at which demand equals supply. Labor markets in poor countries are an example. Consider a country where adult wages are low and children are (for that very reason) made

NINA FINKEL

to work. Suppose that child labor is banned and that the ban is enforced. Firms that were using child labor will seek adults to fill those gaps. Competing for a smaller pool of workers, they will have to pay a higher wage. It is entirely possible that if wages had been high to start with, parents would not have sent the children out to work. Suppose the law is then revoked. Earning enough money to get by, parents will have no need to send their children to work.

The law works simply as a mechanism for deflecting the economy from one equilibrium, in which wages are low and children have to work, to the other, in which wages are high and children can go to school instead [*see box on opposite page*]. Once the economy locks into the new equilibrium, the self-interest of employers and employees keeps it there, even through minor economic shocks such as recession. The law is a one-time effort. In a paper that I wrote with Pham Hoang Van of the University of Missouri, we named such laws "benign legislative interventions" to distinguish them from routine legislative interventions, which require continuous surveillance and a sustained threat of punitive action [*see box at right*].

In theoretical models, economists often treat the switch from one equilibrium to another as costless. In reality, of course, that is not the case. Firms used to employing children may have to make investments in technology to cope with the new situation. In the country as a whole, the government will need to build more schools. Nevertheless, the models have been substantiated by historical analyses of, for example, the role that legislation played in reducing child labor in late-19th-century America.

Trapped

MULTIPLE EQUILIBRIA occur in other ways, too. A household that sends its children to work has to face some stigma for doing so. The degree of stigma is inversely related to the amount of child labor that occurs in society. That is, if child labor is widespread, people will be used to it and the stigma will be minor. To see how this may lead to multiple equilibria, suppose that only a few children work. Parents who send their children to work face the disapproval of their friends and neighbors. Therefore, only those parents who have a very great need to send their children to work do so. Hence, few children work. The society is in equilibrium—the low prevalence of child labor reinforces itself.

THE AUTHOR

KAUSHIK BASU is professor of economics, the Carl Marks Professor of International Studies and the director of the Program on Comparative Economic Development at Cornell University. (Carl Marks was a venerable merchant banker, not an alternative spelling of Karl Marx.) He has written extensively on development economics, welfare economics, game theory and industrial organization. When he graduated from high school in Calcutta, he says, his father wanted him to study physics, and he wanted to study nothing; they settled on economics as a halfway compromise. Initially Basu found economics dry and unappealing, but while at the London School of Economics he was completely taken in by the lure of the free-floating deductive reasoning on which economics is founded and decided to make that his career.

More Than One Way to Run a Market

THE IDEA of multiple equilibria also applies to situations other than child labor. In general, economists today accept that market economies are more efficient than centrally controlled ones. As individuals pursue their self-interest, their actions collectively make life better for all. But it is possible to go overboard and believe that individual rationality is always sufficient to achieve some desired social end. Economic models that have multiple equilibria demonstrate this fallacy.

Consider the debate that occurred in the 19th century about whether to have statutory limits on work hours. The standard laissez-faire argument was that if a worker was willing to work for 14 hours a day and an employer was willing to pay for that, the state had no reason to intervene. To do so would be paternalistic meddling.

SHIH-YEN FOUNDRY, Hubei Province, China

A more sophisticated argument would look at the particulars of supply and demand. Suppose that there were 100 workers and several employers and that the subsistence income was $12 a day. Each worker preferred to put in eight hours a day, as long as he could earn at least $12, and worked longer only when that was necessary to reach the subsistence level. If the hourly wage was $2, workers stayed eight hours on the job and earned $16, for a total labor supply of 800 man-hours a day. If the wage was $1, workers had to toil for 12 hours to make the subsistence income, and the total labor supply was 1,200 man-hours.

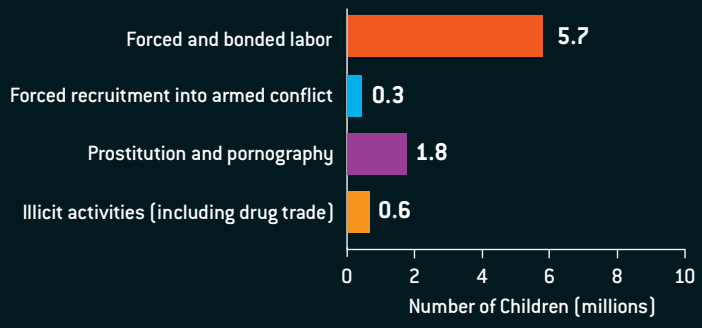
As for the employers, they were more willing to hire workers when the wage was low. Suppose that for a wage of $1 an hour, the aggregate demand for labor was 1,200 man-hours a day and that at $2 an hour, it was 800 man-hours.

In this scenario, the wage rates of $1 an hour and $2 an hour are both equilibria: at these values, demand equals supply. If the economy were caught at the $1 wage equilibrium, it is true that each worker would want to work for 12 hours a day. Nevertheless, a statutory limit of eight hours per workday could be justified in the workers' own interest. This would be a benign intervention. All the workers prefer this equilibrium and would choose it freely, but they have no way to reach it unless they act collectively. To be sure, the high-wage equilibrium is less favorable for the entrepreneurs. Each equilibrium has its winners and losers.

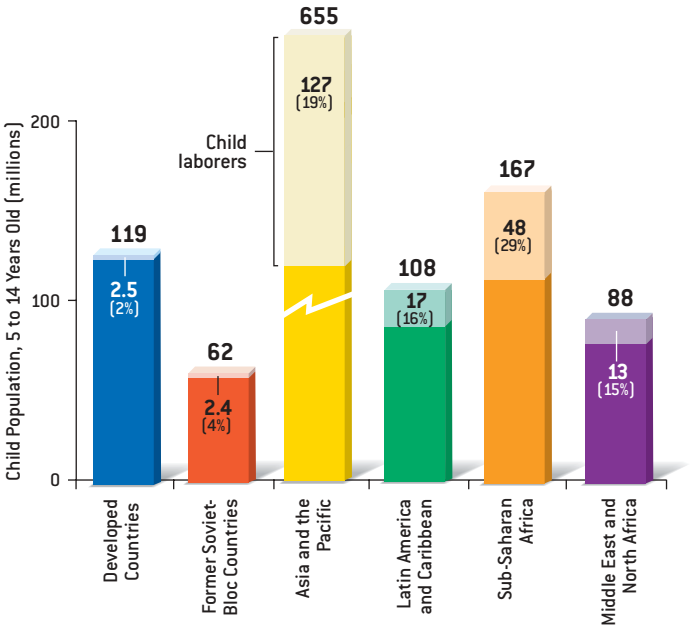
Policy debates in economics are often polarized between those who see an unhindered market as the sole instrument of economic progress and those who would entrust it all to government. The concept of multiple equilibria is one illustration of the importance of the middle way, which recognizes the role of markets as well as the need for government intervention in certain situations. —*K.B.*

More than eight million children are thought to be trapped in debt bondage, forced military service, sexual slavery or other abusive situations. Developed countries, including the U.S. and European Union members, account for about half a million of these children. These figures are consistent with estimates that 1.2 million children are trafficked across borders every year.

CHILDREN DOING THE WORST FORMS OF CHILD LABOR



Child Laborers by Region



IN ABSOLUTE TERMS, the largest number of child laborers live in the developing countries of South and East Asia. But in relative terms, the prevalence of child labor is highest in sub-Saharan Africa, where an estimated 29 percent of five- to 14-year-olds work for a living. Child labor is much less common in the developed world but has not disappeared altogether. (Technical note: The total child population figures here differ slightly from those on page 87 because of revisions in the population estimates.)

If, on the other hand, many children work, the stigma of child labor is lower. More parents are inclined to send their children to work, and again the society is in equilibrium. If the prevalence of child labor falls in between these levels, the society is out of equilibrium and will go through a period of change until it reaches one of the equilibria.

Another type of equilibrium relates to the vicious cycle of poverty. People who worked as children received less education and tend to be poorer as adults, so they are more likely to send their children to work. Families can be trapped in a perpetual child-labor cycle. Conversely, they might be trapped in a virtuous cycle of rising wealth. This theory has recently been tested and validated using a large data set on Brazilian households. Parents who had worked as children were indeed more likely to put their own children into the labor force.

Interestingly, this trend was true even when the experimenters controlled for adult incomes. People who had worked as children were more likely to send their own kids to work than were people of equal income who had not worked as children. Thus, although economics is the main factor in child labor, it is not the only one. Parents who have a history of child labor tend to have social norms and preferences that attach a lower value to their children's schooling. Just as the cause of the fire that destroys a house can be both the spilled kerosene on the floor and the discarded cigarette stub, child labor can have many triggering factors.

The tendency of child labor to reinforce itself, through either economic or social means, seems to be cause for pessimism. But the converse—the tendency of child labor, once eliminated, to stay that way—works in favor of reformers. Multiple equi-

SOURCE: INTERNATIONAL LABOR ORGANIZATION (2000)

STEEL PRODUCTION PLANT, Cairo, Egypt

libria are closely related to the phenomenon of "tipping," whereby a small change suddenly leads to a sharp movement, just as the jug that is tilted gently will at some point tip over. Consider a society that has multiple equilibria and is caught in the high-child-labor equilibrium. Now suppose that through some intervention, such as a law or slow shift in attitudes, child labor is cut down little by little. At some point, the labor market moves into the zone of attraction of the other equilibrium. Child labor then falls off rapidly, without any further intervention. In the U.S., child labor remained widespread until 1900 despite more than 70 years of attempts by state governments to ban it. But when it finally began to decline, the decline was extremely rapid. By 1930 it was almost gone.

Step by Step

A FRIEND OF MINE once tried to persuade me to take up regular jogging by claiming that every 10 minutes of jogging would increase my life expectancy by eight minutes. At first, that seemed incentive enough, but then it struck me that it all depended on what I wanted to maximize in life. If it was my nonjogging life span, then I needed to worry that every 10 minutes of jogging would decrease my nonjogging life by two minutes. Facetious though this example may sound, it points to the important and often overlooked fact that whether a particular policy is desirable or not depends on its ultimate objectives.

If controlling child labor is not an end in itself but an instrument for enabling children to grow up into productive and happy individuals, then policies have to be evaluated against this larger yardstick and not just the immediate one of whether it halts child labor. In the poorest regions, society may in fact have to permit children to work a few hours each day. Studies in Peru and Brazil have shown that children's labor is often the only way they can finance schooling for themselves or their siblings and thereby ensure an eventual escape from poverty for their progeny. Such findings raise troubling moral questions, but if policy is to be effective, it must account for the reality of people's lives.

Many commentators have argued—and I agree—that legislative action is not the best way to control child labor, barring some special cases, such as when we have reason to believe that there are multiple equilibria (and so a benign intervention will work). In general, policymakers should work to improve the conditions and earnings of adult laborers so as to diffuse the conditions that foster sending children to work. For instance, during an economic downturn, fluctuations in income may compel parents to withdraw their children from school. Even if the kids return to school later, they find it difficult to catch up and often drop out altogether. Providing parents with access to affordable credit and insurance can help them ride out hard times without resorting to child labor. Small incentives, such as providing children with a midday meal in school or giving parents a subsidy for sending their kids to school, can also sharply reduce child labor, as has been shown in Brazil and Bangladesh.

In tackling the problem of child labor, it is easy to fall into the trap of complacency, leaving it all to the markets, or into the trap of moral self-righteousness, trying to remove child labor in one stroke, with no concern for the well-being of the alleged beneficiary of such a policy. We now have enough information and understanding of the problem that we can aim to eliminate child labor. It will take restraint and a careful construction of nuanced policy interventions. SA

MORE TO EXPLORE

The Economics of Child Labor. Kaushik Basu and Pham Hoang Van in *American Economic Review*, Vol. 88, No. 3, pages 412–427; June 1998.

Child Labor: Cause, Consequence and Cure with Remarks on International Labor Standards. Kaushik Basu in *Journal of Economic Literature*, Vol. 37, pages 1083–1119; September 1999.

Hard at Work in Factories and Mines: The Economics of Child Labor during the British Industrial Revolution. Carolyn Tuttle. Westview Press, 1999.

Child Farm Labor: The Wealth Paradox. Sonia Bhalotra and Christopher Heady. July 2000. Available online at **www.bris.ac.uk/Depts/Economics/research/pdffiles/dp00492.pdf**

Does Globalization Increase Child Labor? Evidence from Vietnam. Eric Edmonds and Nina Pavcnik. NBER Working Paper No. w8760; January 2002. Available online at **papers.nber.org/papers/w8760**

A Social Stigma Model of Child Labor. Luis-Felipe López-Calva in *Estudios Económicos*, Vol. 17, No. 2; July–December 2002. Available online at **mailweb.udlap.mx/~llec_www/documentos/Lopez-Calva-revisado-stigma.pdf**

Is There a Child Labor Trap? Intergenerational Persistence of Child Labor in Brazil. Patrick M. Emerson and Andre P. Souza in *Economic Development and Cultural Change*, Vol. 51, No. 2, pages 375–398; January 2003. Available online at **www.vanderbilt.edu/Econ/wparchive/workpaper/vu02-w14.pdf**

For statistics on child labor, visit the International Labor Organization at **www.ilo.org/**

SMART FABRICS

Cool Shirt

Sweat cools your body as it evaporates from the skin, but clothing traps that moisture, raising body temperature and causing you to sweat even more. To help, garment makers are infusing the athletic-apparel market with "moisture management" fabrics that wick away sweat and dry quickly—and these are just the first of high-tech clothes to come.

Several factors enhance fabrics such as Coolmax from DuPont and Moistex from Asahi Kasei. Manufacturers are extruding advanced polyesters into fibers with a moisture content as low as 0.5 percent, versus 4 percent for nylon and 6 to 7 percent for cotton, so that they wick and dry more quickly. New extrusion techniques also allow makers to produce fibers with unusually shaped cross sections [*see illustration*] that channel away sweat. Crafting the coolest fabric "is a balancing act of many properties," says Michael Hunt, senior research chemist at DuPont Textiles and Interiors in High Point, N.C.

To make winter gear that dissipates moisture but holds in heat, manufacturers use specially extruded hollow fibers that retain insulating air. "The hair in polar bear fur is hollow," Hunt notes. CW-X, Under Armour and other makers of so-called compression garments, which help to hold muscles in place, are also combining the wicking fibers with the compressing strands (often Lycra) so that the snug fit doesn't cause athletes to overheat.

Scientists are vying to create smart fabrics that actually react to changing conditions, such as shirts that change color in sunlight and airy jackets that suddenly become waterproof when raindrops hit. Key are conductive fibers made of polymers doped with additives, such as camphorsulfonic acid, that can conduct charge. Rain would alter the fabric's conductivity, causing dopants to shrink, pulling closed a garment's pores.

"We have produced conductive yarns and have woven them into fabric," says Frank Ko, materials engineering professor at Drexel University. But products are at least several years away. "The challenge now," Ko adds, "is to make the fabric stable enough to survive sweat and tough enough to survive the washing machine." —*Mark Fischetti*

FABRIC that quickly wicks away sweat and dries is knit from polyester fibers.

KENT SNODGRASS *Precision Graphics [illustrations]*; DUPONT TEXTILES AND INTERIORS *[micrograph]*

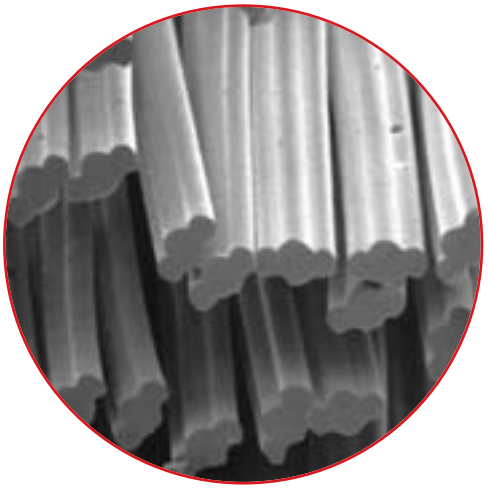
DID YOU KNOW . . .

➤ ANTI-CLING: Synthetic fibers are comfortable in part because they do not retain moisture. But it's hard for such fabric to disperse static charge, so it tends to cling. Some makers add antistatic agents.

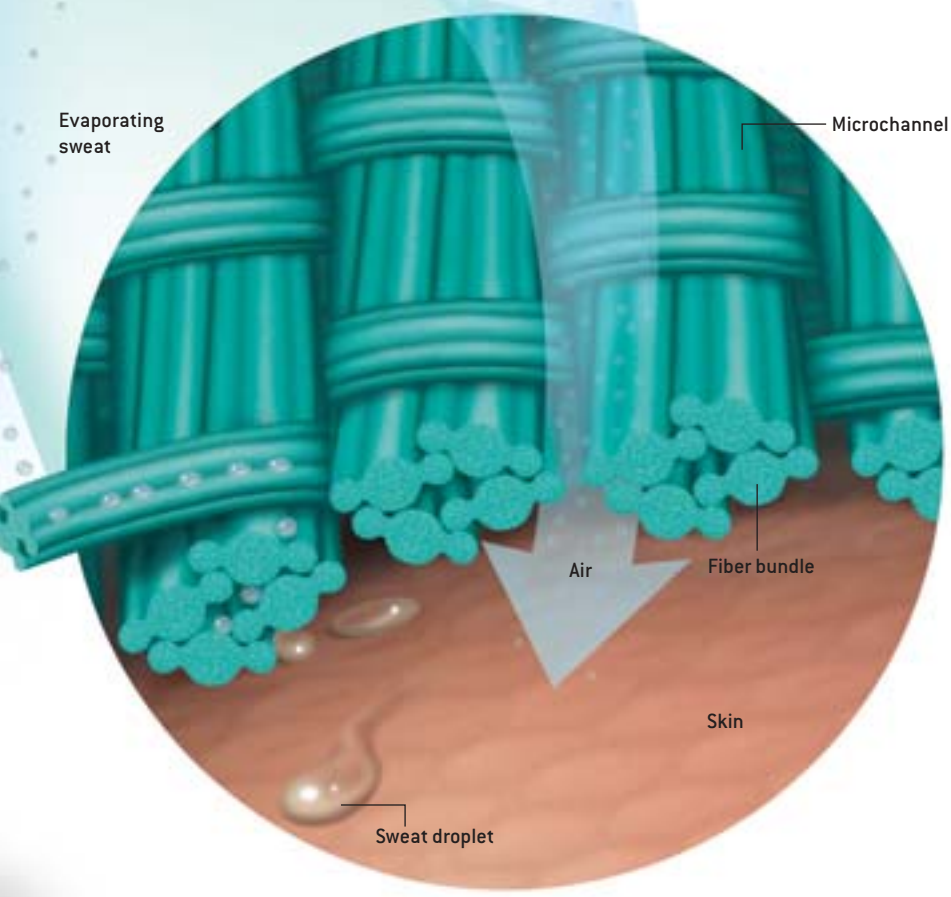
➤ ANTI-STAIN: In advertising its new Go Khakis line of pants, Dockers says it uses "nanotechnology" for superior stain resistance. But industry experts observe that the pants simply come coated with Teflon, which lowers surface tension so that liquids are less likely to adhere. Molecules can be dubbed "nanotech," but Teflon coatings are not new.

➤ ANTI-SUN: Yes, you can get sunburn through a shirt made with a loose construction. Manufacturers can produce a tighter weave to close the holes, but this may decrease airflow, making the shirt feel clammy to wear. Alternatively, makers can add a delusterant such as titanium dioxide to the fibers, which scatters ultraviolet rays.

➤ CHAMELEON: If chemists can devise conductive polymers doped with additives that can conduct charge (*see main text*), they could fashion a blouse that shifts from green to blue when a woman steps from shade into sunlight. Energy from the sun would alter the current, shifting the molecular orientation of dopants and thus the colors of light they absorb. The U.S. Army is already experimenting with this cloth for more versatile camouflage. (In so-called mood clothing, the dyes change color in response to changing heat.)



POLYESTER in Coolmax fabric is extruded into fibers with a scalloped oval cross section that provides microchannels, which help to draw beads of sweat away from skin through capillary action.



IN ONE DESIGN, yarn bundles against the skin contain fewer but larger filaments (and thus less surface area) than bundles along the exterior, which have more and smaller filaments (greater area). The gradient in surface area drives sweat from the inside out, like a one-way sponge, and spreads the moisture on the outside face, enhancing evaporation. Gaps in the knit allow cooling air in.

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

The Infinite Arcade Machine

BUILDING THE WORLD'S LARGEST VIDEO ARCADE—IN YOUR FAMILY ROOM **BY W. WAYT GIBBS**

Three steps into Aaron H. Mahler's house, there is no mistaking that this is a man who knows computers and who loves arcade games. His office, just off the foyer, hosts a desk with two computers and two PDAs, a Road King pinball machine, and a classic Pac-Man game rescued from the bistro at Sweet Briar College, where Mahler (pronounced "mailer") works as the director of network services. And in the middle of the room sits the toy that was worth flying to Virginia to see: an arcade machine with custom-built controls and a PC inside running software, called MAME, that allows it to play more than 2,200 classic video games with almost perfect fidelity. No quarters required.

I had read about MAME, the Multiple Arcade Machine Emulator software project that enables a PC to mimic dozens of microchips used in arcade games. First released in 1997 by Nicola Salmoria, a programmer in Siena, Italy, the open-source project has since attracted some 100 hackers to work on improving it. I had even gone to www.mame.net and downloaded a copy, along with a few of the public-domain ROM (read-only memory) files that contain programs for individual games. Yet standing here in front of Mahler's universal arcade machine, I am amazed by the thing.

To kids who were reared on PlayStations and Xboxes, the hulking video games that filled neighborhood arcades in the 1980s and 1990s may seem almost ludicrously simplistic. But that is not how I remember them. Like many in Generation X, I spent much of my youth twitching and shimmying to low-resolution graphics and low-fidelity sound at arcades, cinemas and pizza parlors. To this day, the sight of Centipede evokes memories of the defeats I suffered at that game to my junior high school girlfriend—whom, incidentally, I ended up marrying. And I still recall my awe upon climbing into the cockpitlike Star Wars, which in 1984 was the first space combat game you could play sitting down.

GAME NOT OVER: That's the motto for the MAME project, a program that enables home computers to emulate almost any video arcade machine built between 1975 and 1990.

JOHN FRASER

Much of my allowance went into that Star Wars machine, Tron and a select few others. You had to invest your quarters judiciously if you were ever to add your initials to the high-score chart. To be able to play any game, from Asteroids to Zaxxon, simply by picking it from a menu—well, it's downright decadent.

Building such a machine, however, re-

quires a lot of time and perhaps a touch of fanaticism. Mahler has made a hobby of collecting and restoring old video game machines. At one time he owned more than 30, in various states of disrepair. So he has more experience than most with how the devices are wired. He has a Millipede cabinet to house the MAME machine. And he has legal rights to use the ROM files for each of the dozens of arcade games he actually owned.

"You can get MAME ROM files from various places on the Internet," Mahler says. "In fact, I was able to restore some broken machines I owned by reprogramming them with downloaded ROMs. But technically it is illegal to use the ROMs with MAME unless you own the actual chips from which the software was extracted or if the copyrights have lapsed or have been waived by the author of the game." As with laws against sharing MP3 music files, these restrictions are infrequently observed and hardly ever enforced.

The first hurdle Mahler faced in his conversion project was getting his fiancée's permission to tinker with Millipede, at which she is a master. "I didn't like the idea of him messing with my game," she says, casting a possessive eye on the fancifully painted cabinet. But after scrolling through a number of the step-by-step guides for building MAME machines that enthusiasts have posted online, Mahler worked out a way to disconnect but not disable the original game hardware.

The monthlong process—documented in his own extensive guide at http://sparhawk.sbc.edu/MAME—began with an old computer running the Linux operating system. Mahler jury-rigged a cable to connect the PC to a refurbished arcade monitor. "You can play the games on a modern multisync PC monitor," Mahler avers. "But it just doesn't cut it for authenticity; you just don't get the right look."

Setting up the software was relatively straightforward. In addition to MAME, he installed AdvanceMAME, an add-on

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 6 one-hour episodes each season. Visit the Scientific American Frontiers Web site at www.pbs.org/saf

www.sciam.com

Our award-winning Internet resource, updated daily 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. $10.95 each. Elsewhere $13.95 each. Many issues are available. Fax your order with your Visa, MasterCard or AMEX information to: (212) 355-0408.

Reprints

$4 each, minimum order 25 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.

(also free) that puts the monitor into the same low-frequency, low-resolution modes used in the original arcade machines. AdvanceMAME also shrinks and rotates games that are designed for wide screens so that they appear normal (if a bit small) on the tall, narrow monitor mounted in the cabinet.

Most of the effort and cost went into the control panel. But it is a setup to make a gamer whistle. Three joysticks, a trackball, a spinner and a whopping 23 buttons—all placed exactly where Mahler wants them—sprout from the Formica-covered surface. Connecting each of these doohickeys to the computer took two interface boards, more than 200 feet of wire, and several late-night stripping and splicing sessions. Even with such ample controls, however, there are some games—such as Star Wars, alas—that just do not feel right without their special yokes, wheels or pedals.

BUTTONS GALORE adorn the custom-built control panel that Aaron Mahler attached to his old Millipede game. Inside, a Dell PC serves up any of 2,000-plus arcade classics.

So how well does it all work? I played Pac-Man as a test. On Mahler's original Pac-Man game, a bit faded now after 40,724 plays, I scored 21,150. (So I'm a little out of shape; it's been a while.) On the MAME machine, with its brighter colors, clearer sound and more responsive joysticks, I scored 21,700. The game play was indistinguishable.

But on the MAME machine, you can hit a button to pause the game when the phone rings. If the pixelated characters in Donkey Kong seem too jagged, you can pop up a menu and choose an option that smooths the pixels to look almost like a hand-drawn cartoon. And because the MAME ROMs mimic the original hardware, you can change those secret settings that arcade operators twiddled to make games harder or easier.

Mahler's infinite arcade game has been a big hit with his friends, he says. He is helping a buddy to build one now. "And I'm going to build a second machine from a Marble Madness game that's sitting in my office at Sweet Briar," Mahler adds. "That one has a horizontally mounted display, so it's better for the wider games."

Enough MAME enthusiasts have begun building dedicated cabinets that a cottage industry has sprung up to serve them. X-Arcade, for example, sells prewired control panels for $150 and a full machine cabinet (minus PC) for $1,100. But dabblers can make do with MAME on whatever computer they already own. The software runs on Windows, DOS, MacOS, Linux, PDAs, some digital cameras—almost anything with a microprocessor and a display.

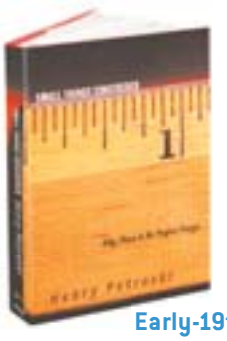
Who knows, perhaps MAME will allow Generation X to pass its infatuation with arcade games on to Generation Xbox. On an interminable flight back from Australia recently, I whittled at the time by playing Space Invaders, the game that started the video arcade craze (as well as a run on 100-yen coins) back in 1978. As the invaders marched relentlessly across my laptop screen, an 18-month-old boy sitting next to me watched, with rapt attention. SA

AARON H. MAHLER

Thinking inside the Box

THE REAL WORLD PLACES REAL CONSTRAINTS ON ENGINEERS

BY MICHAEL M. SOKAL



SMALL THINGS CONSIDERED: WHY THERE IS NO PERFECT DESIGN
by Henry Petroski
Alfred A. Knopf, New York, 2003 ($25)

Early-19th-century American steam engines were less fuel-efficient and more dangerous than their late-18th-century British counterparts. These details don't surprise most first-year engineering students. After all, didn't American technology lag Britain's for many years? They are often taken aback, however, to discover that these faults derived partly from explicit choices of American steam engineers. Historian George Basalla suggested in his 1988 book *The Evolution of Technology* (Cambridge University Press) that such choices can be understood as adaptations to the resource-rich and skills-poor American environment, in which heavy fuel consumption mattered less than the ease of design, construction and maintenance afforded by high-pressure operation. (These higher-pressure engines could also run at greater speeds, an important feature on a larger landmass.) More generally, like the Y2K problem, what later generations saw as a design flaw can be most richly seen as the result of a designer's attempt to work within the technological boundary conditions of a given time and place.

In *Small Things Considered,* Henry Petroski's approach to the question of "why there is no perfect design" is less evolutionary than Basalla's and reflects his own experience as a practicing engineer and a keen observer of the made world and of how people live in it. But like Basalla, Petroski continually emphasizes that all made things, both physical and social, are designed, whether consciously or unconsciously, and that anyone designing anything must work within a set of physical and social constraints. As he writes, in considering the design of chairs, "All designs must involve tradeoffs, if not in materials, then in function; if not in cost, then in fashion; if not in quality, then in proportion; if not in size, then in shape; if not in this, then in that." The design process is thus often labyrinthine, and successive compromises in response to specific constraints close off and open up different choice points later in the process. Indeed, even as designers "perfect" their creations, they usually both improve (in some ways) and impair (in others) what came before.

Petroski, a professor of civil engineering and of history at Duke University, illustrates his argument in a series of delightful chapters, many derived from his "Engineering" column in *American Scientist* magazine. He starts each with observations (some stimulated by reports in the *New York Times* "Patents" feature) about one or another made thing—staircases, paper cups and paper bags and the machinery to make them, duct tape and WD-40, automobile cup holders, and ink-jet printers. Many trace the origins and development of a particular device or way of doing things, and they all continually reassert Petroski's primary theme. These reflections lead him (often apparently by free association) to other examples—from the invention of single-lever faucets to the redesign of vegetable peelers, from the placement of doorknobs to that of light switches—that further demonstrate just how designers made their choices within constraints to achieve workable compromises. Although the regular echo of Petroski's the-

OUR OWN DEVICES: THE PAST AND FUTURE OF BODY TECHNOLOGY
by Edward Tenner. Alfred A. Knopf, New York, 2003 ($26)

If Henry Petroski is the engineer interpreter of everyday technology, Edward Tenner is its philosopher. This fascinating collection of essays delves into, mulls over, and teases apart eyeglasses, shoes, chairs and other innovations that have changed our bodies in unexpected ways. Tenner, the author of *Why Things Bite Back,* is a researcher at the Smithsonian's National Museum of American History.

NINETEENTH-CENTURY reclining chair

COURTESY OF ALFRED A. KNOPF

sis (not unexpected in articles that first appeared months apart) at times seems repetitious, the pleasure and excitement of seeing his playful mind at work more than make up for any annoyance.

One design constraint that often emerges is the need for designers to think *within* the box and to consider the "compatibility with the existing world," the oversight of which "can jeopardize years of development work and result in an ultimate design and financial disaster." This statement of principle comes as Petroski traces the design of easier-to-manipulate toothbrushes too thick for traditional bathroom racks and leads into discussions of the *Great Eastern* (a mid-19th-century steamship too large for most contemporaneous harbors) and of the Concorde.

Petroski also raises this point in considering the two different keypad arrangements found on telephones, on one hand, and on electronic calculators, on the other. In doing so, he reviews late-1950s Bell Labs studies of other designs that concluded that one alternative would be faster and that another was most preferred by those who tried it. He notes (more or less approvingly) that, despite these results, telephone manufacturers chose to keep the now conventional arrangement "'since it uses the available space efficiently and permits a simplified design in the initial application.'" Contrarily, however, he points out that few people have trouble shifting between telephone keypads and calculator keypads and that "we all seem to adapt easily to the machine before us." One wonders, however, how Petroski would view attempts of standard QWERTY keyboard users to type on Dvorak keyboards, and vice versa.

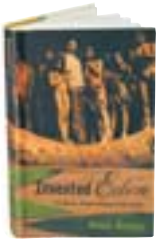
Petroski's Web site describes one of his primary interests as "the use of case histories to understand the role of human error and failure in engineering design." Yet, as his book well illustrates, most (or at least many) design failures are not caused by human errors but derive, perhaps inevitably, from the necessity of compromise and the impossibility of making choices that satisfy all constraints. Whatever their cause, such failures do occur, and these often have major consequences. One such failure that Petroski does not discuss is the design of the butterfly ballot, used in south Florida in the year 2000. Some argue that the use of this ballot determined the results of that year's presidential election and, in some ways, the fate of the nation. One need not accept this argument in full to realize just how significant and complicated the process of design can be, and *Small Things Considered* provides all sorts of penetrating and broadly interesting insights into the nature of this process. SA

Michael M. Sokal has taught history of technology to engineering undergraduates at Worcester Polytechnic Institute since 1970. He will serve as president of the History of Science Society in 2004 and 2005.

THE EDITORS **RECOMMEND**

INVENTED EDEN: THE ELUSIVE, DISPUTED HISTORY OF THE TASADAY
by Robin Hemley. Farrar, Straus and Giroux, New York, 2003 ($25)

"Elusive" and "disputed" are certainly the words to describe the history of the Tasaday, a group of about two dozen people discovered in 1971 living in an apparently primitive way in a remote Philippine jungle. The "lost tribe" was for a while an international sensation. In 1986, however, a Swiss reporter wrote that several Tasaday had told him that they had been coerced into pretending to be cavemen and were in fact farmers from a neighboring tribe. Hemley, professor of English at the University of Utah, has dug deeply into the Tasaday story. His conclusion: "The Tasaday were pseudo-archaics in Claude Lévi-Strauss's terminology, a small group that had fled into the forest to escape an epidemic of some sort. By 1986 they had become what anthropologist Richard Fox calls professional primitives."



RECENT BOOKS **BY STAFF MEMBERS AND CONTRIBUTORS**

Merchants of Immortality: Chasing the Dream of Human Life Extension
by Stephen S. Hall. Houghton Mifflin, Boston, 2003 ($25)

Cities of Gold: A Novel of the Ancient and Modern Southwest
by William K. Hartmann. Forge Books, New York, 2003 ($25.95)

The Moons of Jupiter
by Kristin Leutwyler. W. W. Norton, New York, 2003 ($39.95)

The New Killer Diseases: How the Alarming Evolution of Mutant Germs Threatens Us All
by Elinor Levy and Mark Fischetti. Crown Publishing Group, New York, 2003 ($24.95)

The Pathological Protein: Mad Cow, Chronic Wasting, and Other Deadly Prion Diseases
by Philip Yam. Copernicus Books, New York, 2003 ($27.50)

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

Strategic Bullying

BY DENNIS E. SHASHA

In schoolyards, the world of crime, and international affairs, certain entities try to take valuables from others by force. They attack most happily when they have overwhelming power and the target is wealthy. Sometimes individuals or groups join forces, forming a temporary coalition to gain wealth or to deter an attack.

Let's assume that the following rules hold. An entity's power is represented by a number; if the entity is a coalition, the coalition's power is equal to the sum of its members' powers; wealth is proportional to power. Entity A will attack B only if A is sure of winning (if A's power is greater than B's). When an entity is beaten, the winners gain its wealth but not its power. Every entity plays selfishly, striving to acquire riches yet avoid certain destruction. A configuration of power values is said to be stable if no fighting will occur—as will happen if a coalition can form that is strong enough to prevent fighting and if being loyal to the coalition is in every member's best interests.

Suppose the power values are 4, 2, 1. The configuration is unstable because 4 will fight 2 and 1 and will beat them whether they form a coalition or not. In contrast, if the power values are 4, 2, 1, 1, the configuration is stable: 2, 1 and 1 can form a coalition to prevent 4 from attacking. Moreover, if any coalition members were destroyed, the remaining ones would be, too, so each coalition member would be strongly motivated to stay in the coalition.

To warm up, consider the configuration 5, 4, 4. Is it stable? What about 5, 4, 3? The 5, 4, 4 configuration is unstable because the two 4s will certainly gang up on 5. After splitting 5's wealth, they will separate into a stable (in fact invincible) 4, 4 configuration. But 5, 4, 3 is stable because no one can be assured of winning with impunity (*below*).

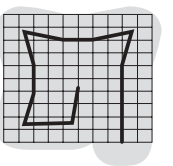
Your problem: What is the largest set of distinct power values (positive integers all different from one another) that is stable when the highest value is 21? What if the integers need not be distinct?

Here's another challenge. Certain entities can have a "bee-sting" capability—defenses able to destroy even a superior attacking power. Those using the sting, though, die in the battle along with the enemy. For example, if a 5 attacks a 4 having a bee sting, both the 5 and the 4 will perish. Thus, 5 won't attack. A bee-sting capability may seem inherently stabilizing because it should prevent attacks. But might the acquisition of this capacity by an entity render a configuration unstable? SA

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

Answer to Last Month's Puzzle

To find the missing hiker, begin two miles west of the square's southeastern corner and follow the route shown in the illustration below. The total length of the route is about 29.31 miles, so the search party will detect the hiker's signal within about 293 minutes.



Web Solution

For more discussion of last month's problem and a peek at this month's solution, visit www.sciam.com

CORNELIA BLIK

WARM-UP SOLUTION: A 5, 4, 3 configuration is stable because peace is in each group's best interests.

What's Wrong with This Picture?

FOR FOOTBALL FANS IN THE SUNSHINE STATE, IT'S SOMETIMES THE GAME OF THE NAME **BY STEVE MIRSKY**

In this world, nothing is certain but death and taxonomy. Everyone is interested in death, but few of us outside of the order Strigiformes give a hoot about taxonomy. Nevertheless, people can have taxonomy—which Merriam-Webster's defines as the "orderly classification of plants and animals according to their presumed natural relationships"—thrust upon them. Just ask the mortified folks at the University of Florida.

Two of Florida's major products are the fruits of *Citrus sinensis* trees and football. The University of Florida's football team is called the Gators, in honor of *Alligator mississippiensis*. Although the name might lead you to assume that the Gators would be the football team at the University of Mississippi, the gridiron gladiators—sorry, I was channeling melodramatic sportswriter Grantland Rice (*Oryza sativa*) for a second there—at Ole Miss are in fact the Rebels, an appellation that pays tribute to the halcyon days of yesteryear when some *Homo sapiens* fought for the right to own other *Homo sapiens*. Go figure.

The Linnean alligator moniker presumably derives from the fact that American alligators can be found along the Mississippi River as far north as Oklahoma. Despite the geographical misappropriation, since 1987 *Alligator mississippiensis* has been Florida's official state reptile. Buckle your alligator belt, because things are about to get even more confusing.

In late July the University of Florida football team released its media guide for the 2003 season. They mailed out some 13,000 copies of the guide, which school officials thought featured a large and aggressive-looking alligator on the cover. Except that, on closer inspection, the gator turned out to be a croc. Conspicuously unconsulted University of Florida crocodilian (both alligators and crocodiles are referred to as crocodilians—I told you it would be confusing) expert Kent Vliet told the *Palm Beach Post* that the photo-

graph was probably of the Nile crocodile, *Crocodylus niloticus,* a nasty beast indeed. Differences in the snout shapes and the appearance of the big, pointy teeth are dead, and I do mean dead, giveaways.

The pictured critter was not even the elusive American crocodile, *Crocodylus acutus*. Best estimates put the number of *acutus* individuals in Florida at somewhere between 500 and 1,200, a figure apparently arrived at via the same tallying methodology used in Florida elections.

Now, to the untrained eye, mistaking a crocodile for an alligator seems like a trifle. But it's a major deal taxonomically. The two species are grouped together within the family Crocodylidae. To put the relationship in perspective, human beings are in the family Hominidae, which we share with chimpanzees, gorillas and orangutans. Therefore, the Florida fumble is roughly the equivalent of using a photo of a group of gorillas to illustrate the faculty, who probably feel like monkeys' uncles over the entire affair.

Despite the preceding ribbing, I don't want to be accused of piling on. We all make mistakes, which is why pencils come with erasers—except for golf pencils, because lying in golf is not a mistake—and why this magazine has an erratum section. *Scientific American* recently corrected an April news story that contended that, in one study, cloned pigs had variable numbers of teeth. In fact, they had variable numbers of teats. The reporter can blame a poor phone connection for our own pigskin faux pas.

Besides, the Florida folks have probably become immune by now, the gator gaffe having no doubt prompted plenty of abuse from their fellow Southeastern Conference members. These schools include the universities of Arkansas (the Razorbacks, a fancy name for hogs, *Sus scrofa*), South Carolina (the Gamecocks, a belligerent kind of chicken, *Gallus domesticus*) and Alabama (the Crimson Tide, presumably a bloom of the algae species *Karenia brevis*). All told, I think I'd rather find myself in one of Florida's numerous gatored communities. See ya later. SA

ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

ASK THE EXPERTS

Q What causes insomnia?

—H. York, Builth Wells, Wales

Henry Olders, an assistant professor of psychiatry at McGill University who conducts sleep research, provides this answer:

People can lose sleep for a variety of reasons, including medications, alcohol, caffeine, stress and pain. When the underlying cause is removed, these bouts usually get better on their own. For some people, however, sleep problems turn into insomnia, the chronic inability to either fall asleep or stay sleeping. Research suggests that attitudes about sleep, and the resulting slumber patterns and behaviors, make certain individuals vulnerable to insomnia.

Many insomniacs feel they lack sufficient sleep, but evidence is mounting that they are getting at least as much as they require and possibly more. Insomniacs tend to go to bed early, stay there late and sleep during the day—all of which contribute to the problem.

Why would anyone spend more time asleep than he or she needs? Charles M. Morin of Laval University in Quebec found that insomniacs hold stronger beliefs than normal sleepers do about the detrimental effects of insomnia to physical and mental health and that they perceive their sleep as less controllable and predictable. Individuals with insomnia are more likely to be concerned about not sleeping and to think about problems, events of the day and noises in the environment before falling asleep. Simply put, if you are convinced that you need eight hours of sleep a night, you will arrange your bedtime and rising time so that you spend eight hours in bed. If you require only six hours of sleep, however, you will spend two hours tossing and turning.

How much sleep do you need? And how can you tell if you are getting the right amount? Although eight hours a night is a figure repeated so often that it has almost become an article of faith, the reality is that sleep need is highly individual. Large-scale epidemiological studies have demonstrated that sleeping seven hours a night is associated with the lowest mortality risk (for factors including heart disease, cancer and accidental death) compared with longer or shorter periods of shut-eye. In addition, it is probable that as we age, we need less sleep.

To help treat insomnia, practice "sleep hygiene." This includes adjusting the levels of noise, light and temperature so that you are comfortable; not reading or watching television in bed; avoiding excess food, alcohol, nicotine, caffeine and other stimulants before you turn in; completing exercise at least three hours before lights out; and then determining your optimum bedtime. The longer you are awake, the more slow-wave (delta) sleep you will have; slow-wave sleep is what leads to feeling rested and refreshed. Limiting the time you spend in bed may also help. Together these nonpharmacological approaches are more effective and longer-lasting than medications for insomnia.

Q Why is the sky blue?

—M. Nasrallah, Amman, Jordan

Anthony D. Del Genio of the NASA Goddard Institute for Space Studies and adjunct professor of earth and environmental sciences at Columbia University explains:

We can thank the scattering effect, which disperses nearly 10 times as much blue light in the atmosphere as light of longer wavelengths (such as red). Sunlight is a mixture of all colors. As sunlight passes through the atmosphere, it acts as a mixture of electromagnetic waves that causes the oscillation of charged particles (electrons and protons) in air molecules. This oscillation produces electromagnetic radiation at the same frequencies as the incoming sunlight, but the radiation is scattered in every direction.

The blue component of visible light has shorter wavelengths and higher frequencies than red. Thus, blue light makes charged particles oscillate faster than red light does. The result is that the scattered blue light is almost 10 times as prevalent as red light. Violet light is scattered even more than blue, but less violet light enters the atmosphere, and our eyes are more sensitive to blue.

A planet with no atmosphere cannot have a bright sky, because there is no scattering effect. Photographs taken by astronauts on the moon show a midnight-black sky. SA

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

MATT COLLINS