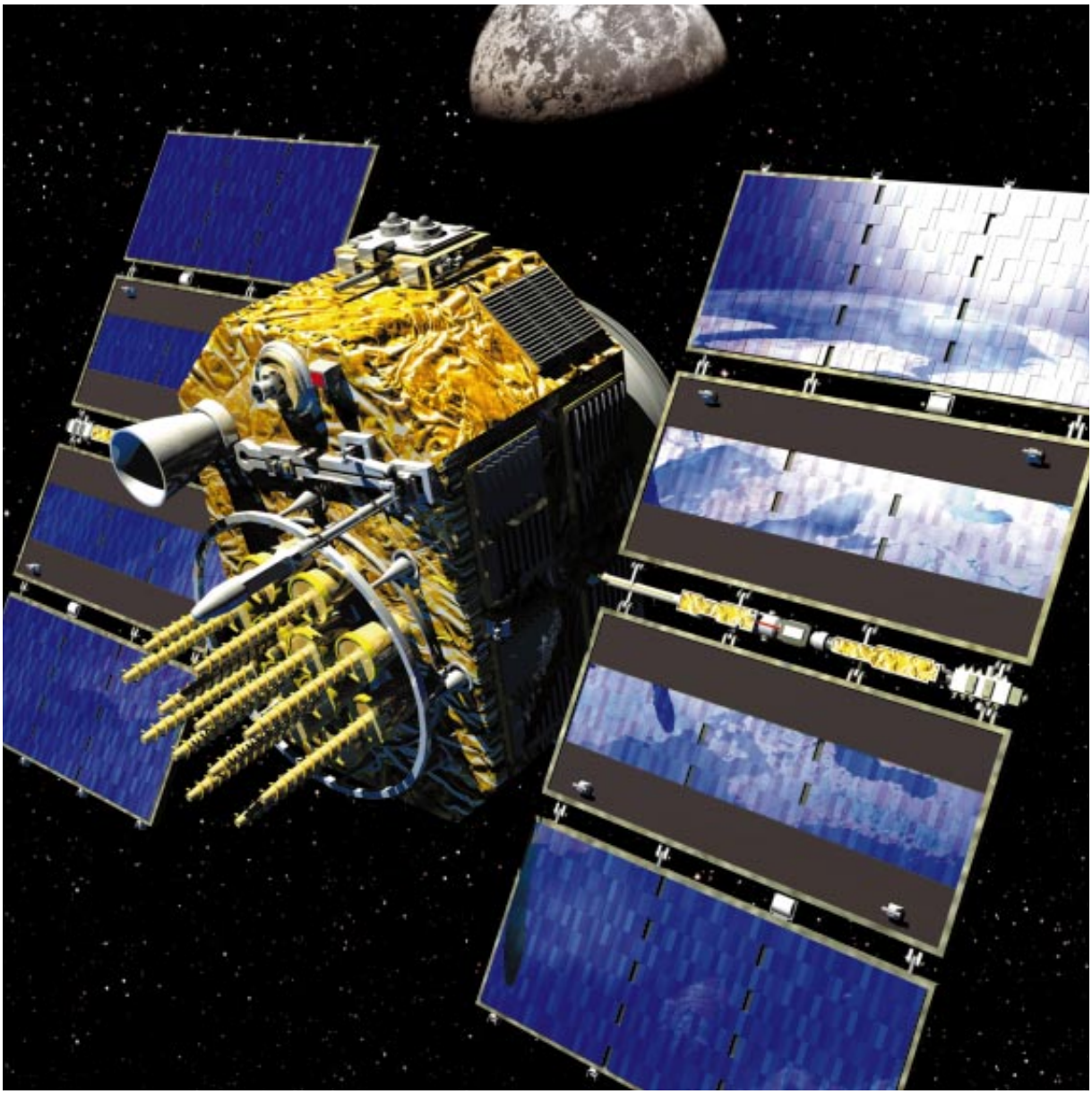
# SCIENTIFIC AMERICAN

FEBRUARY 1996
$4.95

*Saving malnourished minds.*
*Exploding galaxies.*
*Antibiotics against ulcers.*



***Global Positioning System** brings space-age navigation down to earth.*

SCIENTIFIC AMERICAN

February 1996 Volume 274 Number 2

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111.  Second-class postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian GST No. R 127387652; QST No. Q1015332537. Subscription rates: one year $36 (outside U.S. and possessions add $11 per year for postage). 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; fax: (212) 355-0408 or send E-mail to SCAinquiry@aol.com. **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.**



# DEPARTMENTS

## Letter from the Editor

Just in time for Valentine's Day, this issue offers "The Loves of the Plants," in which Londa Schiebinger examines the sexual classification system that 18th-century taxonomist Carl Linnaeus imposed on plants. From our, ahem, enlightened standpoint today, Linnaeus's fixation seems quaint and misguided. In the words of that social philosopher Tina Turner, "What's love got to do with it?"

History is peppered with investigators seeing what they were predisposed to see. Georges Cuvier and other supporters of the catastrophism school of geology looked at sedimentary deposits high in the mountains and saw evidence of world-drowning floods. (Their faith in a Biblical Flood may have been a factor.) The controversy over IQ and intelligence measurement has always been inflamed by fears about whether culture and prejudice skew the search for an "honest" answer. Stephen Jay Gould, in his book *Wonderful Life,* describes the exotic Cambrian fossils of the Burgess Shale, including *Hallucigenia*—a living nightmare with spikes for legs and tentacles on its back. Until the 1970s, it and other creatures had been improbably crammed into the known groups of arthropods because of assumptions about evolutionary progress.

***COVER art by Slim Films***

But sometimes even errata need errata. Three years ago paleontologists again reappraised *Hallucigenia* and concluded that people had been looking at the fossil upside-down. *Hallucigenia* walked on flexible legs and wore spikes on top, a member of the more mundane class of velvet worms.

Drawing the line between observation and interpretation is never easy. It's easy to lose track of your assumptions, to forget which keystones in the edifice of your theory are loosely packed sand. The triumph of the scientific method is that over time, through collective effort, mistakes can be overturned. Science accepts error as something to be corrected over time. What could be a more tolerant and humane philosophy?

On the subject of errors, here's one of mine. The first time I ever saw a global positioning system in practice was a few years ago, while cruising up the sidewalk of Fifth Avenue with our intrepid Phil Morrison. He was demonstrating a handheld unit, and as it chewed on the problem—the skyscrapers were making its satellite communications balky—he extolled the device's ability to find our location. "We're at the corner of Fifth and 46th Street," I said, glancing at a signpost. "What's the big deal?" Find out how big a deal it is on page 44.

JOHN RENNIE, *Editor in Chief*


**SCIENTIFIC AMERICAN**®

*Established 1845*

**EDITOR IN CHIEF:** John Rennie

**BOARD OF EDITORS:** Michelle Press, *Managing Editor;* Marguerite Holloway, *News Editor;* Ricki L. Rusting, *Associate Editor;* Timothy M. Beardsley; W. Wayt Gibbs; John Horgan, *Senior Writer;* Kristin Leutwyler; Madhusree Mukerjee; Sasha Nemecek; Corey S. Powell; David A. Schneider; Gary Stix; Paul Wallich; Philip M. Yam; Glenn Zorpette

**ART:** Edward Bell, *Art Director;* Jessie Nathans, *Senior Associate Art Director;* Jana Brenning, *Associate Art Director;* Johnny Johnson, *Assistant Art Director;* Carey S. Ballard, *Assistant Art Director;* Nisa Geller, *Photography Editor;* Lisa Burnett, *Production Editor*

**COPY:** Maria-Christina Keller, *Copy Chief;* Molly K. Frances; Daniel C. Schlenoff; Terrance Dolan; Bridget Gerety

**PRODUCTION:** Richard Sasso, *Associate Publisher/Vice President, Production;* William Sherman, *Director, Production; Managers:* Carol Albert, *Print Production;* Janet Cermak, *Makeup & Quality Control;* Tanya DeSilva, *Prepress;* Silvia Di Placido, *Graphic Systems;* Carol Hansen, *Composition;* Madelyn Keyes, *Systems; Ad Traffic:* Carl Cherebin; Rolf Ebeling

**CIRCULATION:** Lorraine Leib Terlecki, *Associate Publisher/Circulation Director;* Katherine Robold, *Circulation Manager;* Joanne Guralnick, *Circulation Promotion Manager;* Rosa Davis, *Fulfillment Manager*

**ADVERTISING:** Kate Dobson, *Associate Publisher/Advertising Director.* OFFICES: NEW YORK: Meryle Lowenthal, *New York Advertising Manager;* Randy James, Thom Potratz, Elizabeth Ryan, Timothy Whiting. CHICAGO: 333 N. Michigan Ave., Suite 912, Chicago, IL 60601; Patrick Bachler, *Advertising Manager.* DETROIT: 3000 Town Center, Suite 1435, Southfield, MI 48075; Edward A. Bartley, *Detroit Manager.* WEST COAST: 1554 S. Sepulveda Blvd., Suite 212, Los Angeles, CA 90025; Lisa K. Carden, *Advertising Manager;* Tonia Wendt. 235 Montgomery St., Suite 724, San Francisco, CA 94104; Debra Silver. CANADA: Fenn Company, Inc. DALLAS: Griffith Group

**MARKETING SERVICES:** Laura Salant, *Marketing Director;* Diane Schube, *Promotion Manager;* Susan Spirakis, *Research Manager;* Nancy Mongelli, *Assistant Marketing Manager;* Ruth M. Mendum, *Communications Specialist*

**INTERNATIONAL:** EUROPE: Roy Edwards, *International Advertising Manager,* London; Vivienne Davidson, Linda Kaufman, Intermedia Ltd., Paris; Karin Ohff, Groupe Expansion, Frankfurt; Barth David Schwartz, *Director, Special Projects,* Amsterdam. SEOUL: Biscom, Inc. TOKYO: Nikkei International Ltd.; TAIPEI: Jennifer Wu, JR International Ltd.

**ADMINISTRATION:** John J. Moeling, Jr., *Publisher;* Marie M. Beaumonte, *General Manager;* Constance Holmes, *Manager, Advertising Accounting and Coordination*

**SCIENTIFIC AMERICAN, INC.**
415 Madison Avenue,
New York, NY 10017-1111

**CHAIRMAN AND CHIEF EXECUTIVE OFFICER:** John J. Hanley

**CORPORATE OFFICERS:** John J. Moeling, Jr., *President;* Robert L. Biewen, *Vice President;* Anthony C. Degutis, *Chief Financial Officer*

**DIRECTOR, PROGRAM DEVELOPMENT:** Linnéa C. Elliott

**DIRECTOR, ELECTRONIC PUBLISHING:** Martin Paul

**PRINTED IN U.S.A.**

# LETTERS TO THE EDITORS

## Debating Darwin, Socially

I was disappointed that John Horgan, in his article "The New Social Darwinists" [SCIENTIFIC AMERICAN, October 1995], appears to share *Scientific American*'s long-standing inability to look at human sociobiology objectively. His worst lapse comes at the very beginning. Devendra Singh has helped isolate a major variable in the attribution of human attractiveness, the waist-hip ratio. Throughout the range of female human shapes, increasing waist-to-hip ratios are associated with increasing mortality, decreasing fertility, increasing free testosterone levels and decreasing attractiveness. But Horgan goes for the cheap laughs and misses all the content. His article reveals a recurring problem serious scientists have on the subject of human behavior.

ROBERT TRIVERS
Rutgers University

I would like to propose an evolutionary explanation for why "gentlemen prefer blondes" over brunettes in Western cultures. Several authors have suggested that florid displays of secondary sexual characteristics "inform" the female that the suitor is healthy and free of dermal parasites. I suggest that being blonde serves a similar purpose. Anemia (a common symptom of intestinal parasites), cyanosis, jaundice and skin infection are much easier to detect in fair-skinned individuals than in brunettes. Also, the skin of blondes "ages" faster and more visibly than that of brunettes. Fertility in women declines with age and with disease, so men may gravitate toward blondes, in whom such signs are easier to observe. I originally intended the above as a parody of ad hoc sociobiological theories of human mate selection but came to realize that this idea is at least as viable as many others currently in vogue, including those mentioned by Horgan.

V. S. RAMACHANDRAN
University of California at San Diego

## Darwin Strikes Again

I was fascinated by Madhusree Mukerjee's short feature on mating Australian Redback spiders, "Giving Your All" ["Science and the Citizen," SCIENTIFIC AMERICAN, October 1995], which showed a pair of the spiders doing the black widow thing and described their strategy: "This is for the kids." Ah, the uses of evolutionary psychology! In a fraction of a page, you not only helped me with my arachnophobia but also gave me a better understanding of my relationship with my first wife.

FRANK DURHAM
Tulane University

## Cancer Controversies

I would like to correct erroneous statements made by Devra Lee Davis and H. Leon Bradlow in their article "Can Environmental Estrogens Cause Breast Cancer?" [SCIENTIFIC AMERICAN, October 1995] concerning one of Ciba's products, atrazine herbicide. Atrazine is not a xenoestrogen. It has been shown to be devoid of estrogenic activity in the standard bioassays. The authors claim atrazine increases "breast cancer" in male rats. Ciba toxicologists are unaware of any data showing an increase in breast cancer in males of any species resulting from exposure to atrazine. Davis and Bradlow imply that atrazine is only slowly degraded. In fact, it has a half-life in humans of about 12 hours.

DONALD R. SAUNDERS
Ciba-Geigy Corporation
Greensboro, N.C.

Davis and Bradlow should not have limited their inquiry to synthetic substances. Naturally occurring estrogens abound (for example, in ginseng and in toxins produced by molds). If estrogenicity is the key to causing breast cancer, these naturally occurring substances may be equally hazardous.

DEAN O. CLIVER
University of California at Davis

***Davis and Bradlow reply:***

Our hypothesis holds that substances that increase estrogen levels—xenoestrogens—explain some of the incidence of breast cancer among women having no known risk factors. In one recently published test, we and our colleagues found that atrazine increases the production of bad estrogens (16-α-hydroxyestrone) about as potently as some known breast cancer–causing agents do. In contrast, several natural xenoestrogens, such as those in broccoli and fish oil, actually suppress production of bad estrogen.

Atrazine is one of the most widely used herbicides in the U.S. Where exposures are common, compounds having short half-lives can profoundly affect public health. Atrazine belongs to the class of triazine herbicides, which the EPA has placed under Special Pesticide Review because of their great exposure potential and because a number of studies suggest they may increase several types of cancer. A recent study links atrazine in particular with ovarian cancer in humans. Public interest is best served by resolving these issues, not by denying their existence.

## The Endless Frontier

Regarding Gerald Holton's essay "The Controversy over the End of Science" [SCIENTIFIC AMERICAN, October 1995], it is interesting to reread a lecture by Max Planck from December 1, 1924: "When I began my physical studies [in Munich in 1874] and sought advice from my venerable teacher Philipp von Jolly…he portrayed to me physics as a highly developed, almost fully matured science…. Possibly in one or another nook there would perhaps be a dust particle or a small bubble to be examined and classified, but the system as a whole stood there fairly secured, and theoretical physics approached visibly that degree of perfection which, for example, geometry has had already for centuries."

FRIEDRICH KATSCHER
Vienna, Austria

*Letters may be edited for length and clarity. Because of the volume of mail, we cannot answer all correspondence.*

ERRATUM

The 100,000 cholera deaths mentioned in James Burke's column "Connections: Top to Bottom" [SCIENTIFIC AMERICAN, December 1995] occurred around the world, not just in England as stated.

# 50, 100 AND 150 YEARS AGO

## FEBRUARY 1946

No matter whether one wants much power or little power from the fission of uranium, the same irreducible minimum of U-235 must be used. Any such power source must be heavy and awkward, and quite dangerous in case of mishap of one sort or another. These points of themselves rule out the pre-war dreams of the automobile with the lifetime power-supply built in. Any visions of individual atomic power units for every home or farm must also be abandoned, so long as uranium fission is to be the source."

"The extent of industrial growth in the production of amino acids is apparent when it is realized that ten years ago no amino acids were produced commercially; within the last few months, tons of the life-giving substances were flown to Europe to help restore the health of the half-starved inhabitants of war-devastated countries. Just as sulfa drugs, vitamins, and penicillin have outgrown their test tubes, so have amino acids progressed far beyond the laboratory stage."

"Those who still look askance upon synthetic resins as a reliable bonding agent for wood will find many of their questions answered by work going forward on the 200-ton flying boat nearly completed by the Hughes Aircraft Company. Here, a group of adhesives—liquid and film—are being used to make the millions of glued joints required in the mammoth aircraft. Tests extending over a three-year period have proved the efficacy of the bonding materials."

## FEBRUARY 1896

A large aerolite, or meteorite, exploded above the city of Madrid, Spain, at 9:30 A.M., February 10. The explosion was accompanied by a vivid flash of light and a loud report. The concussion was so severe that the partition wall of the United States legation building collapsed, and nearly all of its windows were broken. The officials of the Madrid Observatory state that the explosion occurred 20 miles above the earth. A general panic prevailed in the city."

"The London and Northwestern Railway Company, of England, is said to be the greatest corporation on earth. It has 2,300 engines, and employs 60,000 men. Everything is made by the company—bridges, engines, rails, carriages, and an innumerable lot of other things; even the coal scuttles and wooden limbs for the injured of its staff."

"Comparing housework with other industries, it is obviously belated in respect to using mechanical devices and labor-saving inventions. There are endless devices for saving steps, for avoiding dust, for transporting things up and down, which might be studied in the building of our homes and go in with the laths and plaster. Since the architect views the home chiefly as a social rather than an industrial center, and the carpenter is guided by other considerations than planning to save a woman's time, the perfection of a home plant for housework will only be known when the housewife has a head, if not a hand, in the building."

"Antwerp is rivaling London for the ivory trade of the world. The British consul general at Antwerp reports the large extent to which ivory is brought to Belgium from the Congo."

*An idea for 360-degree cinema, 1896*

"In our illustration, we give a general view of the electric cyclorama, or panorama, as conceived by the inventor, Mr. Chase of Chicago. The projection apparatus, suspended in the center of the panorama by a steel tube and guys of steel wire, is 8 feet in diameter. The operator stands within the apparatus and is surrounded by an annular table supporting eight double projectors, lanterns and all the arrangements necessary for imparting life to a panorama 300 feet in circumference and over 30 in height. It is possible at will to animate such or such a part of the view by combining this apparatus with the Edison kinetoscope or the Lumière kinematograph."

## FEBRUARY 1846

A Mr. Philips of London has introduced an apparatus for the instantaneous extinguishment of fires. The principle of his fire annihilator is to project upon the fire a gaseous vapor which has a greater affinity for the oxygen of the atmosphere than the burning combustibles, and consequently extinguishes the fire by depriving it of the element oxygen, on which combustion particularly depends."

"A new and excellent work, 'The Art of Weaving,' by C. G. Gilroy, presents a somewhat complicated French loom, or as it is termed, a Jacquard machine. The machine is capable of working an unlimited variety of figures and colors, as would appear from the fact that the night dress of Pope Boniface, which was woven in one of these machines, contained 276 different colors, so arranged and blended as to display the likeness of 276 heretics, each suffering under some species of torture different from any of the others. Thirty different colors, requiring as many different shuttles, were employed in the manufacture of the coronation dress of Queen Victoria."

"A newspaper, printed on silk, is published weekly in Pekin, China. Some of the issues measure more than 30 feet in length. Who says China is not ahead of us?"

## Keeping Vaccines Cold

*Travails of immunizing the world's children*

In Papua New Guinea they are transported over land for two weeks by porters who store them in gas refrigerators that are held on bamboo poles. In Nigeria, motor launches take them upriver. On the Afghan-Pakistan border, they are packed with ice and stored in caves.

Every year the World Health Organization assists national governments in delivering vaccines to more than 100 million children worldwide. Because these vaccines are perishable, this effort requires a logistical network of refrigeration and transport called a cold chain. The WHO began helping to construct the cold chain in the mid-1970s to ensure that vaccines for major childhood diseases—poliomyelitis, measles, tuberculosis, diphtheria, tetanus and pertussis (whooping cough)—survive for the up to two years it may take to get them from a European factory into a child's arm in Nepal. At the same time, the WHO and a companion organization have recently struggled to implement new vaccine technologies that could eliminate the cold chain entirely.

Because the cold chain passes through places where a supply of electricity is intermittent or nonexistent, the WHO helped to develop a refrigerator lined with water-filled tubes that can make enough ice in eight hours to store vaccines safely for up to a 16-hour lapse of electricity. The WHO has also promoted the replacement of inefficient kerosene refrigerators with solar-powered refrigerators. Some 5,000 of these solar refrigerators, based on designs that originated with the U.S. space program, are now in use.

To encourage broader use of solar technology, the WHO has tried to couple vaccine storage to broader economic benefits. The WHO organized a project earlier this year in which a town in a remote area of Colombia's Chocó district installed a solar-power-generating system—not just for vaccine storage but also for providing energy to the local health center and other facilities, including a community television set.

The imminent worldwide ban on the use of ozone-depleting chlorofluorocarbons poses a challenge to the cold chain. The phasing out of these chemicals, which in developing countries is scheduled to occur early in the next decade, means that new refrigerator designs will be needed. But new equipment that employs substitute refrigerant chemicals has performed poorly and requires further development. "There's a 30 percent drop in efficiency," says John S. Lloyd, a WHO technical officer.

On the other hand, a modest new technology introduced this year should make it easier for health care workers who administer vaccines to assess the effectiveness of the cold chain. Beginning in January, chemical monitors will start to appear on vials of highly heat-sensitive polio vaccines. A small dot on the vial changes color once a vaccine has been exposed to elevated temperatures for enough time to affect its potency.

A product-freshness dot falls far short of the WHO's ultimate technological goal: elimination of the cold chain itself. The Children's Vaccine Initiative (CVI), a program sponsored by the WHO and other leading international organizations, seeks to develop a supervaccine—a single-dose oral vaccine that works against multiple diseases and does not break down in the heat of the tropics.

The CVI's initial attempt to move toward this goal has foundered in a morass of bureaucratic confusion. The program was pushing a novel processing technique to create a heat-stable oral polio vaccine to assist in a multiagency campaign to eliminate that disease entirely by the year 2000. The newly formulated vaccine, which could withstand temperatures of 37 degrees Celsius for a period of a week, achieved its stability by the bonding of a live but attenuated polio virus to molecules of heavy water (deuterium oxide).

WHO officials involved with the CVI program proceeded eagerly to line up two vaccine manufacturers and a supplier of heavy water in preparation for clinical trials. At a meeting last summer,

ROGER LEMOYNE; UNICEF

***VACCINE DELIVERY** takes place by motorcycle, porter or boat, such as this one carrying a white "cold chain" box along a river in the Ogun state of Nigeria.*

however, managers of regional and national vaccine programs said that a new vaccine was not needed, because the current one had proved adequate so far for the polio eradication campaign.

At that gathering, concern also arose that a vaccine made with heavy water might generate unfounded rumors that it was radioactive, because deuterium oxide is used as a coolant and moderator in nuclear reactors. These worries effectively halted development of the vaccine, although a meeting this month was scheduled to determine whether to continue some research.

Vaccine manufacturers were left scratching their heads at the WHO's indecision. "We're very disappointed and frustrated," says Jacques F. Martin, chief executive of Biocine, the vaccine subsidiary of Chiron, a U.S.-based biologicals company. "Because of this experience, we won't be fond of going along with them next time," Martin remarks.

Also upset was D. A. Henderson, a professor at Johns Hopkins University's School of Public Health who led the successful WHO program to eradicate smallpox and who is now a member of the CVI's scientific advisory board. The decision to sideline the heat-stable vaccine program, he says, may cause the campaign against polio to run into difficulties when it tries to reach into outlying areas of Africa. Henderson characterized the CVI-WHO leadership as "uncertain, hesitant and weak."

Despite setbacks, development of heat-stable vaccines continues outside the confines of the WHO bureaucracy. In October, Quadrant Holdings, a British-based health care technology company, announced an agreement with Biocine to develop a heat-stable vaccine that provides protection against diphtheria, pertussis and tetanus. The stabilizing agent is a sugar, called trehalose. When the vaccines are dried in the presence of trehalose, they appear to suffer no molecular damage, and they resist degradation at tropical temperatures.

In similar research, two Japanese companies have developed lyophilized (freeze-dried) vaccines that are heat stable. Kaketsuken, based in Kumamoto, Japan, has nearly finished a clinical trial in Indonesia of a combination vaccine that provides protection against diphtheria, pertussis and tetanus.

A potentially effective form of delivery for such dried vaccines is direct injection of the powder into the skin, where it can elicit an immune response. Reconstitution of the vaccine into a highly perishable liquid would not be required. To get the vaccine into the skin, Oxford Biosciences in England has developed the makings of a needleless injection system. A supersonic jet of helium emitted from a penlike device would carry the vaccine through an outer skin layer. If this works, the cold chain might ultimately be supplanted, replacing the prick of a needle with a shot of dry powder. *—Gary Stix*

FIELD NOTES

## Bose Knows P.R.

Quantum mechanics rarely makes the morning headlines. But the subject has not daunted reporters who for weeks have trooped to a nondescript, second-floor laboratory at the University of Colorado at Boulder. There a long-sought state of matter first materialized last summer. By cooling 2,000 rubidium atoms to an unimaginably frigid 170 billionths of a degree above absolute zero, Eric Cornell, Carl Wieman and their colleagues produced the Bose-Einstein condensate, in which all the atoms act as a single giant atom.

KEN ABBOTT *University of Colorado at Boulder*

***HIGH SCHOOLERS and Carl Wieman* (right) *inspect the traps for the Bose-Einstein condensate.***

In beating out several other researchers, the two principal investigators became stars. Nearly every major university and research laboratory has invited them to present lectures. "We've been plodding away on this stuff for six years, then the media came," Wieman remarks. "I'm booked through '97, and Eric's not too different. I gave a talk last week—where was I? I can't remember anymore." The local papers have plastered color photographs of the two on the front pages, although reporters seemed to show limited interest in quantum mechanics. "They asked what Eric and I were like, what we did in our spare time, what our hobbies were, our taste in clothes," Wieman recalls.

Their celebrity status has also forced Cornell and Wieman to face a different kind of explanatory challenge: teenagers. On a brisk Saturday morning in November, on a campus jammed with college-football fanatics, the two physicists and their pool of graduate laborers found themselves immersed, armpit deep, in about 50 seventh and eighth graders. Intrigued by news accounts, students around the area have become curious about the discovery. This group came as part of a National Science Foundation program run by C. G. Mendez and Ernest Cisneros of Metropolitan State College of Denver.

It's time for down-to-earth analogies and hands-on demos. Several metal skewers piercing an apple represent the laser beams coming in from a number of directions to slow ever kinetic atoms. A bowl filled with small balls and sloshed about illustrates how the researchers' instruments eject the warmer atoms, thereby leaving behind the less active, cooler ones. Peeling back the aluminum foil on a piece of equipment, Wieman reveals strips cut from refrigerator magnets, which guide cold atoms down a narrow glass tube. Cornell patches through a live video feed of the trapped atoms, showing a nearly forming Bose-Einstein condensate that vanishes the moment he blocks a laser beam with his hand.

The young charges pay attention to the show but may be more curious about the laboratory paraphernalia. They sight along the laser paths and lean on the air-cushioned laboratory table (to Cornell's chagrin) and wonder why empty cans that once stored caffeine-rich soda are stacked in a corner.

Engaging in good public relations, though, eats into research time; the team has yet to glean any vital statistics from the condensate. "The setup was really optimized for getting the Bose condensate fast, not for doing things once we got there," Cornell explains. They are now redesigning the equipment so that they can increase the number of atoms in the condensate. That may keep them ahead of their competition. Two other groups have subsequently produced the condensate as well—one group at Rice University, using lithium atoms, and another at the Massachusetts Institute of Technology, with sodium. The bask in the limelight may soon end. *—Philip Yam*

## Frozen Assets

*U.S. officials question the value of Antarctic science programs*

It is by far the coldest, most hostile place on the planet. Antarctica is also a mecca of sorts for many American scientists, who have journeyed there for decades to study everything from fish with antifreeze in their veins to the microwave afterglow of the big bang. Even the most avid supporters of research in the Antarctic acknowledge that it is an expensive enterprise. Last year the U.S. National Science Foundation allocated $196 million, roughly 6 percent of its entire budget, to support research in Antarctica. Of that money, about 85 percent, or $167 million, paid for logistical and operational support; only research conducted in space requires more overhead.

Inevitably, some members of Congress have begun to question whether the money is well spent. Last fall the Senate Appropriations Committee inserted a provision into the NSF's budget calling for the Clinton administration to review the program with an eye toward shrinking it and possibly closing one or more of the three bases—namely McMurdo, a coastal town south of New Zealand that supports more than 1,200 people during the summer; South Pole Station, home to some 125 scientists and staff; and Palmer Station, below South America, which sustains about 40 personnel.

Even before the NSF's budget was passed, the White House's Office of Science and Technology Policy (OSTP) had begun forming a panel to review the program. One reason, says Gerald T. Garvey, the OSTP's assistant director for physical sciences, was that the NSF had been planning to ask for an increase in funds to replace its 20-year-old South Pole Station, which is sinking into the ice cap. The NSF has estimated that a new base could be built in eight years for about $200 million.

One staunch defender of the upgrade, and of the Antarctic program in general, is Louis J. Lanzerotti, an atmospheric scientist at AT&T Bell Laboratories who has visited Antarctica twice and has been conducting research there remotely for 25 years. Lanzerotti says that when he first became involved in Antarctic studies in the 1970s, he "didn't think all the research was of high quality." Since then, he notes, the science has improved "enormously" as a result of stricter oversight by the NSF and more rigorous peer review. Most of the studies done in Antarctica now cannot be done as well anywhere else, he says.

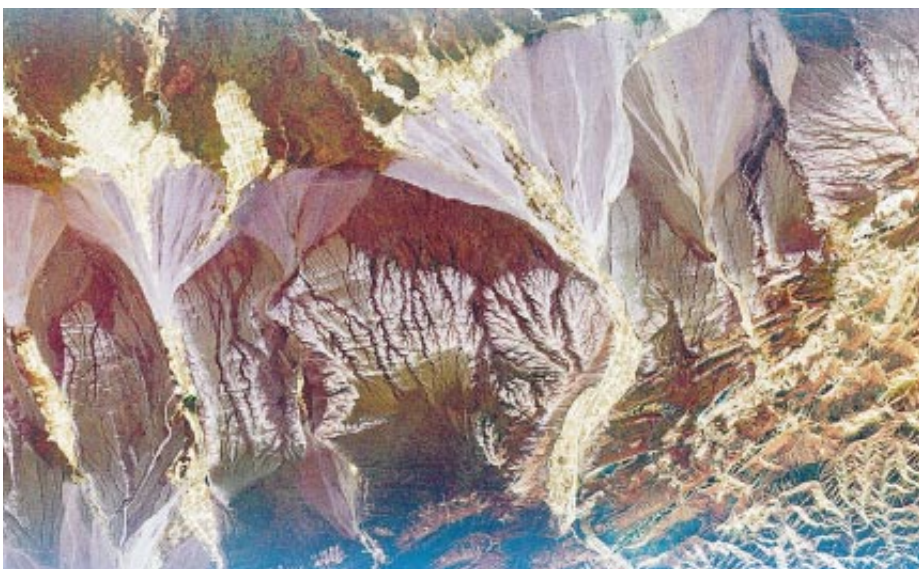
Indeed, the unusually dry, frigid conditions at the South Pole—and its long, sunless winter—make it an excellent place for astronomy and atmospheric science. The pristine continent also serves as a bellwether of environmental change. In 1986 and 1987 researchers at McMurdo linked the notorious ozone hole detected above the Antarctic to a buildup of chlorofluorocarbons; this work helped to convince leading industrial nations to sign a pact banning CFCs. Subsequent investigations at Palmer Station have monitored the effects of ozone depletion on organisms exposed to elevated levels of ultraviolet radiation. Scientists are now trying to determine whether global warming will trigger a precipitous collapse of the Antarctic ice sheet, which would cause sea levels worldwide to surge by as much as 60 meters.

Some lawmakers have suggested that the end of the cold war has eliminated a major reason for the U.S. Antarctic program: keeping pace with the former Soviet Union, which has long been active there. But the U.S. still has good reasons to remain, counters R. Tucker Scully, director of the Office of Oceans and Polar Affairs at the State Department. He notes that seven nations—the U.K., Norway, New Zealand, Chile, Argentina, France and Australia—have claimed sovereignty over parts of the Antarctic. The U.S. rejects those claims. By maintaining a vigorous presence on the continent, Scully says, the U.S. can ensure that Antarctica remains open both for its own scientists and for those from other countries. The U.S. can also help enforce treaties banning military and mining operations there, Scully adds.

## Radar Range

There is more to these pictures of the earth than meets the eye—literally—because they were taken with radar, not visible light. As part of the ambitious "Mission to Planet Earth" project, the National Aeronautics and Space Administration, in conjunction with the Italian and German governments, has created a sophisticated radar system that rides on board the space shuttle and compiles detailed maps of the earth below. These recently processed images showcase the system's ability to reveal subtle geologic and environmental details that are difficult, if not impossible, to detect from the ground. Radar has no inherent color, so researchers assigned arbitrary hues to the various wavelengths and degrees of polarization of the radar; colors were selected to highlight certain details and to make the images aesthetically pleasing.

A radar map of the region around Pishan, a town in northwest China that lies along the ancient Silk Route, captures the changing climate of the region (*below*). The bright lavender, fernlike branches that cut across the center of the image represent ancient gravel deposits, known as alluvial fans, that washed down from the surrounding mountains during an earlier time when the area received much more rainfall. Modern erosion features show up as broad, lavender triangular features located above the older fans. A massive irrigation project overcomes the current, dry climate, creating the vegetated



JET PROPULSION LABORATORY

GALEN ROWELL *Mountain Light*

***SOUTH POLE STATION is sinking into the ice and needs to be replaced, according to the National Science Foundation.***

The Senate Appropriations Committee has proposed that the NSF might reduce its costs by enlisting other nations as partners for, say, the upgrade of the South Pole Station. John Lynch, manager of polar aeronomy and astrophysics for the NSF Antarctic program, says that at first glance an international base "has some appeal." But he fears that if the U.S. enlists foreign partners, all the nations with claims on the Pole may demand to participate as well. Lynch worries, too, that the tiny station would become so overrun by visiting dignitaries that its resources would be strained, and scientists would be hard-pressed to get their work done.

Another way to reduce the costs of logistics would be to find private contractors who can provide transportation and other services more cheaply than the U.S. Navy does now, according to Erick Chiang, head of the NSF's polar-research support section. He estimates that costs could be cut by as much as $10 million annually by taking steps such as replacing military helicopters and pilots with civilian ones. But there is no substitute, Chiang adds, for the Defense Department's C5s and C141s, which lug heavy equipment to the coastal bases, and the ski-equipped C130s, which can land in regions lacking airstrips.

One factor that the NSF has in its favor as the debate unfolds is its adept handling of the media. Every year the agency ships a handful of reporters down to the continent for a red-carpet tour; the journalists need only pay their airfare to New Zealand or Chile. (This reporter made the trip in November 1992.) These jaunts, the costs of which the NSF has not estimated, have yielded a steady stream of generally favorable coverage of the Antarctic program. The NSF has no plans to eliminate its journalism program soon. —*John Horgan*

JPL

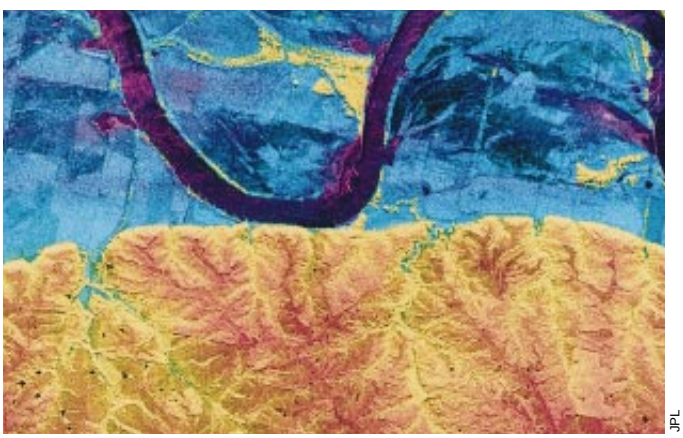
space shuttle with related radar views taken from a NASA DC-8 aircraft to evaluate the environmental effects of the floods that ravaged the Mississippi Valley in 1993. In this map (*below*), color represents elevation: blue is the lowest, orange the highest. The blue area at the top is part of the floodplain of the Missouri River. Dark bands and streaks denote areas that were severely eroded when levees gave way. From radar studies, Arvidson's group estimates that the flood dumped five million tons of sand into the floodplain and carried away some three million metric tons of soil. Images such as these are assisting scientists in the assessment of the total damage, the potential for future flooding and ways to preserve natural wetlands that help to control the fluctuations of the river, Arvidson explains. —*Corey S. Powell*

*These and many other radar images can be viewed on-line at http://www.jpl.nasa.gov/news/ They can also be obtained on CD-ROM; for more information, go to http://southport.jpl.nasa.gov/education.html or send an e-mail message to edc@eos.nasa.gov*

oases that appear yellow here; the grid patterns denote stands of poplar trees, planted as windbreaks.

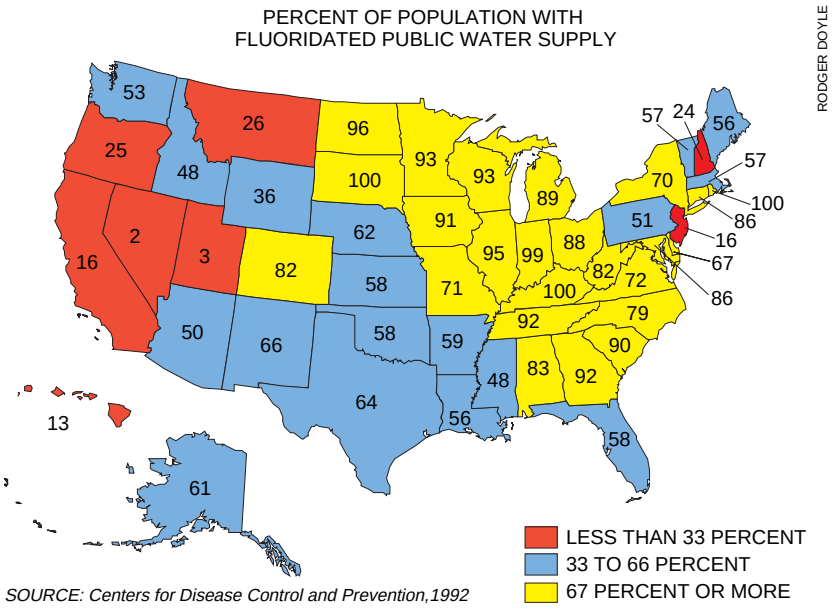
The dramatic view of the nearby Karakax Valley in western China (*above*) was created by a group led by Diane Evans of the Jet Propulsion Laboratory in Pasadena, Calif.; the researchers combined two radar pictures to convey the three-dimensional topography of the valley. Although this area is now fairly stable, the giant Altyn Tagh fault that runs through the entire scene (the diagonal line marking a change in slope and color of the side of the valley) testifies to powerful seismic disturbances in the past. An international team, guided by these radar views, has visited China and collected rock samples, hoping to gain new insight into the instabilities associated with giant faults. In this image, erosion channels and gravel deposits show up as gray areas. The manner in which the radar reflects indicates the ages of the channels and so offers further insight into how the region's climate has changed over the eons.

Devastation caused by too much rain was the subject of radar studies by Ray Arvidson of Washington University and his colleagues. The researchers combined radar data from the



JPL

## Fluoridation




*SOURCE: Centers for Disease Control and Prevention,1992*

In January 1945 Grand Rapids, Mich., became the first city in the world to have controlled water fluoridation. The results there and, subsequently, in scores of other cities were overwhelmingly positive: the rate of cavities in permanent teeth of schoolchildren was typically 50 to 60 percent lower in cities with fluoridation compared with control cities with no fluoridation. Adults as well as children benefited.

Public health officials were optimistic that most communities in the country would soon take advantage of this highly cost-efficient method of improving dental health. But unexpected opposition arose: in the late 1940s many communities decided against fluoridation, apparently because there was widespread ignorance and confusion about its benefits and because some opponents misrepresented the scientific facts. Beginning in the 1950s, opponents claimed that fluoridation caused a variety of diseases, including cancer, birth defects and kidney disease. Despite more than 40 years of research concerning these claims, there is no believable evidence that these or any other diseases are caused or promoted by fluoride in public water supplies.

Primarily because of the unexpected opposition, as late as 1992, according to the latest census taken by the Centers for Disease Control and Prevention, only 62 percent of Americans using public water supplies enjoyed the benefits of fluoride. (Ninety percent of the population uses public water.)

In the West the generally low proportion of homes with fluoridated water probably reflects, in part, mistrust of government and also skepticism about "mass medication." People in western states generally have greater access to citizen-initiated referendums than Americans elsewhere, and they have repeatedly used them to prevent fluoridation. This situation may be changing, however, at least in California, which in late 1995 enacted legislation mandating fluoridation for all water companies serving 25,000 or more people. In the East, New Jersey has an exceptionally low level of fluoridation because of a peculiar situation in which fluoridation of any large water system can be blocked by one community fed by the system. Eight states in addition to California—Illinois, South Dakota, Michigan, Ohio, Minnesota, Nebraska, Connecticut and Georgia—now mandate fluoridation.

Fluoridation has one well-documented drawback: it causes dental fluorosis, a condition that ranges from barely noticeable white specks to unattractive black staining of the teeth. Noticeable fluorosis is rare in areas with optimally fluoridated water (0.7 to 1.2 parts per million). In those few communities with high levels of naturally fluoridated water, fluorosis may be a cosmetic problem but not a health problem.

Of the 50 largest U.S. cities, only Los Angeles, San Antonio, San Jose, Portland (Oregon), Sacramento, San Diego, Honolulu and Tucson did not have fluoridated water as of January 1996. —*Rodger Doyle*

# The Doctor Glut

*Experts debate predictions of a physician surplus*

SAN ANTONIO, TX: "Board-certified OB/GYN with well-established solo practice ($500K+ annually) looking for an associate to share the work load." So reads a recent job advertisement in the *Journal of the American Medical Association.* Yet in the same publication, readers have lately been presented with copious analysis and commentary about an impending physician glut—a surplus that could exceed 165,000 by the turn of the century. It seems difficult to reconcile the ubiquity of physicians earning hundreds of thousands of dollars in salary with the image of hundreds of thousands of doctors pounding the streets in search of gainful employment. Could both sets of numbers possibly be correct?

Some analysts who have grappled with the problem have concluded that the danger of an oversupply is real. In November 1995 a report from the University of California's Pew Health Professions Commission stated, "There seems little reason to doubt the modest assumptions that have been used to generate the projections of a physician oversupply." That document also warns that "American medicine will soon face a dislocation of crisis proportions."

The National Research Council's Institute of Medicine will soon publish the results of a similar study entitled, "The Nation's Physician Workforce: Options for Balancing Supply and Requirements." Neal A. Vanselow, a professor of medicine at Tulane University who headed the study, explains that some of the warning signs are clear. He notes that "physician salaries are beginning to drop" and that "we're beginning to see physician bankruptcies."

Others find these assessments too dire. "I haven't actually seen a lot of unemployed physicians yet," says Steffie Woolhandler, a professor at Harvard Medical School and a founder of the Physicians for a National Health Program. "The thing I see is a sense of desperation—but not much destitution yet."

Woolhandler's view seems to be supported by recent statistics. Although salaries have reached a plateau, physicians average just over $145,000 annually. Even without consideration of other income, that figure still places the average physician in the top 3 percent of American families. Woolhandler expresses more concern for patients than for physicians. She also questions the fundamental premise of studies that base

their estimates of physician oversupply on the recent practices of health maintenance organizations.

Under some HMO plans, Woolhandler notes, a single physician may be responsible for as many as 800 patients. "It's not unusual for patients to be scheduled seven minutes apart," she says. The raw statistics from HMOs could automatically imply an oversupply of doctors because the patients seen in these settings are usually allocated less time than patients traditionally have been allowed. Other factors also make the use of HMO statistics problematic. The clientele of such organizations—with relatively few of the older, poorer and sicker patients enrolled—is not representative of the general population.

Carolyn Clancy, director of the Center for Primary Care Research at the U.S. Public Health Service, points out that no physician-supply study has adequately dealt with the issue of women in the medical workforce. Although women currently constitute only about 20 percent of practicing doctors, their proportion in medical schools tops 40 percent, and thus their ranks will grow to influence enormously the physician population. According to Clancy, female doctors tend to spend more time with patients and work fewer hours each week; hence, they appear to be less "productive" than the average doctor now in practice. With more women working in medicine, the anticipated surplus may not in fact turn out to be so large.

The question of whether the U.S. will eventually be awash in doctors no doubt concerns those students now contemplating medical school. But it is perhaps more keenly relevant to the future patients of doctors in training. Many teaching hospitals rely on the labor of modestly paid medical residents to care for what would otherwise be grossly underserved patient populations. In essence, the medical education system performs two distinct functions: it supplies the most needy with immediate care, and it trains doctors for the future.

The Pew Commission has recommended closing a substantial percentage of the nation's medical schools or limiting the number of foreign graduates of American medical schools who can become residents in the U.S. But these "solutions" to the feared oversupply problem, without other, compensating actions to help those now served by medical residents, could actually create a shortage of caregivers. For Woolhandler, the logic of such restrictive prescriptions, while there exists an acute need for health care by so many Americans, remains difficult to swallow. —*David Schneider*

# Schizophrenia Revisited

*New studies focus on malfunctions in the brain*

Schizophrenia may well be the most mysterious of all mental illnesses. Psychiatrists long blamed bad parenting, but the evidence now fingers faulty genes: four separate papers in the November 1995 issue of *Nature Genetics* singled out a suspicious DNA sequence on chromosome six. (Two reports in the same issue failed to confirm the link.)

Even so, the mystery is far from solved. Although the damage may be done early on, schizophrenic symptoms—which include paranoia, delusions, social withdrawal, auditory and visual hallucinations, and disorganized thoughts—often do not surface until adulthood. Patients bear none of the telltale tumors or lesions that lie behind some similarly disruptive disorders.

Having found little wrong in the form of a schizophrenic's brain, many workers are turning their studies toward function. David A. Silbersweig and Emily Stern of the New York Hospital–Cornell University Medical Center, working with colleagues at Hammersmith Hospital in London, devised a clever means for exposing the brain's blood flow in the act of hallucinating. They examined six schizophrenic men, all of whom heard voices and were unresponsive or unexposed to treatment; one had visual hallucinations as well. The team placed each subject in a positron emission tomography, or PET, scanning machine and asked him to press a button with his right thumb when he hallucinated.

In the November 9, 1995, issue of *Nature,* the group reported that tissues at the brain's surface needed for hearing were active in all the patients. In the man who had visual hallucinations, activity was also observed in the cortical areas involved in coordinating sights and sounds. All the men had increased blood flow in another set of structures, positioned deeper in the brain, including the hippocampus, parahippocampal gyrus, cingulate gyrus, thalamus and striatum. "All these interconnected regions are involved in integrating thoughts and emotion," Silbersweig says. Significantly, the hallucinators did not experience higher blood flows in an area used to distinguish between internal and external stimuli.

Carol Tamminga of the Maryland Psychiatric Research Center has also found evidence that hallucinations stem from processes deep within the brain. At the annual meeting of the Society for Neuroscience in San Diego last November, she described PET scans she took of 24 volunteers—half of them schizophrenic and the other half healthy—after they received ketamine, a drug that prompted psychotic symptoms in both groups. "The activity in the anterior cingulate gyrus greatly increased in both schizophrenics and normals," Tamminga notes. "But the schizophrenics showed a far greater increase."

Next, Tamminga analyzed the effects of haloperidol, an antipsychotic drug. Haloperidol blocks receptors in the brain that bind to dopamine, a neurotransmitter suspected of playing a key role in schizophrenia. As expected, the medication lowered neuronal activity in the frontal and cingulate regions of the cortex but increased blood flow in

# The Rainbow Majority

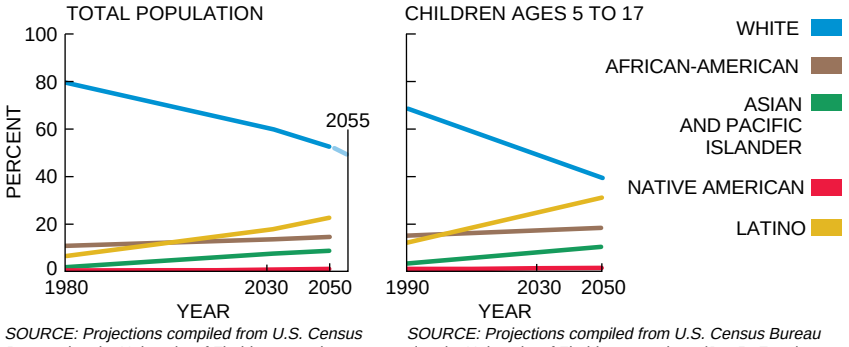
The conservative tide that has gripped America derides programs intended to foster ethnic or racial understanding as a threat to core social values. Conservative pundit Rush Limbaugh rails that multiculturalism is the "tool of revenge of many who have failed to assimilate and fit into mainstream American life." During coming decades, public debate will most likely focus on what exactly constitutes "mainstream American life." The Limbaughian backlash represents a reaction to an inexorable demographic trend. White European Americans face the threat of losing their majority status in U.S. society.

By the year 2055 groups now classified as minorities, taken together, will probably outnumber whites of European descent—a result of immigration and higher fertility. The crossover point will happen even sooner—in the year 2030—for school-age children, according to researchers from the University of Florida. By that year, moreover, Latinos will become more numerous than blacks (*graphs*).

Joe R. Feagin, a professor of sociology at the University of Florida, foresees that these groups may achieve newfound power by the strength of their votes. The weight of numbers could create a rainbow majority that would change the face of American politics. Whites will still be the largest single group, but Feagin believes that complex brokering among ethnic and racial coalitions may or may not leave whites on top. "Whites can still rule as a minority if the subgroups in a rainbow majority are at loggerheads with one another," Feagin says.

Divisiveness training classes may have already begun as a replacement for multicultural education. Feagin points to books such as *Alien Nation,* authored by *Forbes* senior editor Peter Brimelow. *Alien Nation* begins with the statement that current immigration policy may be Adolf Hitler's revenge on America—a human tidal wave that may destroy the country. "These writings are not so different than white supremacist publications analyzing the need to restore white, racial hegemony," Feagin remarks. If visions of an alien nation prevail, a rainbow majority may be supplanted by the monochrome realities of a neoapartheid. —*Gary Stix*

**U.S. RACIAL AND ETHNIC COMPOSITION**




*SOURCE: Projections compiled from U.S. Census Bureau data by University of Florida researchers. The blue trend line is extended to 2055, when whites may constitute less than 50 percent of the population.*

*SOURCE: Projections compiled from U.S. Census Bureau data by University of Florida researchers (Joe R. Feagin, Hernán Vera and Barbara A. Zsembik).*

JOHN AY JOHNSON

several regions in the middle of the brain. Tamminga also recorded abnormal activity taking place in the cingulate gyrus when schizophrenics attempted to discriminate between similar sounds. Although the schizophrenic patients performed as well as normal subjects did on these auditory tests, they used more of their brain in the process. "Schizophrenics activated all the correct areas and just a couple more," Tamminga states, "but they didn't use them in a normal fashion."

Neuropathology also indicates functional problems in some of the same cerebral circuits. The cingulate gyrus in schizophrenics typically contains fewer GABA, or inhibitory, neurons—cells that help to filter incoming stimuli—notes Francine Benes of McLean Hospital in Belmont, Mass., and Harvard Medical School. Benes has investigated whether excess dopamine might cause this decrease. She counted the contacts between dopamine-releasing fibers in the brain and other excitatory and inhibitory neurons in postmortem tissue samples taken from 10 schizophrenic and 15 unaffected people. In all the specimens, the dopamine fibers made more contacts with inhibitory neurons, but the pattern was most pronounced in the cingulate gyrus of the schizophrenics.

Like other researchers, Benes suspects that the inhibitory neurons in schizophrenics are impaired—a situation that would only be made worse by a deluge of incoming dopamine. "My model postulates that the medications we use to treat schizophrenia block the dopamine receptors on the inhibitory neurons, thus freeing them up to perform more efficiently," Benes explains.

The greatest number of contacts between inhibitory neurons and dopamine fibers in schizophrenics appeared in layer II of the cingulate cortex—a layer that is actively developing near the time of a normal birth. This discovery helps to corroborate the theory that obstetric complications may increase the likelihood of an infant acquiring schizophrenia later in life, Benes adds.

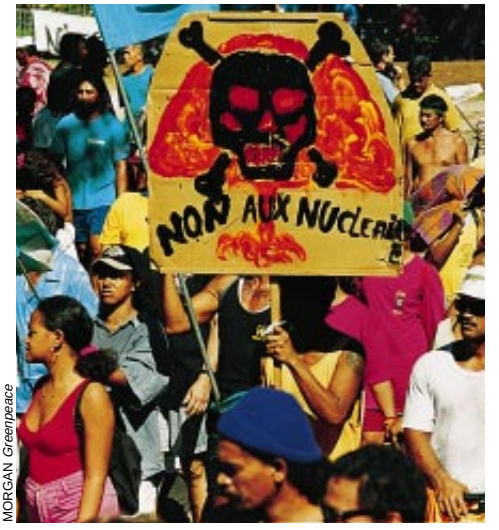
"The consensus is that schizophrenia is neurodevelopmental and that it very likely involves some abnormality in the way the structures that mediate information processing are connected," says Nancy Andreasen of the University of Iowa. For now, this theory cannot offer much help to the 1 percent of the population worldwide affected by schizophrenia. But serious investigations into the disease have only just begun: "Once we find the mechanism," Andreasen believes, "we will be able to find better treatments and maybe even means of prevention." —*Kristin Leutwyler*

## Going Out with a Bang

When the French government resumed testing nuclear weapons last September below the South Pacific atolls of Mururoa and Fangataufa, it provoked an international uproar of surprising intensity. Attempting to quell it, French officials reduced the number of planned tests from eight to six and portrayed most of them as a necessary preliminary step to the country's participation in a worldwide ban on testing, which negotiators hope to implement by the end of this year. To nuclear experts outside France, however, the official rationale for the tests makes little sense.

By all accounts, one of the tests was to make sure that the TN-75, a new warhead for submarine-launched ballistic missiles, worked well. This test took place on October 1 under Fangataufa. It is the other five tests (of which three had been carried out by mid-December) whose purposes are somewhat murky. According to Daniel LeRoy, counselor for nuclear affairs at the French Embassy in Washington, D.C., all these tests are to generate a last burst of data. The information, he says, is necessary to help the country's nuclear scientists adapt to a post–cold war world in which trials involving nuclear weaponry are limited to so-called aboveground experiments that do not entail nuclear blasts.

Last fall French president Jacques Chirac said a couple of tests in the series would be used mainly to study "mixing," a potentially problematic phenomenon during the detonation sequence that saps energy from the overall yield. But experts say it is hard to imagine French scientists, who have already conducted more than 200 tests, significantly advancing their understanding of the phenomenon in this short series.



MORGAN *Greenpeace*

***TAHITIANS protested the French resumption of nuclear testing on nearby Mururoa.***

In a recent statement to the press, Jacques Bouchard, director of the military division of the French Atomic Energy Commission, said French nuclear-weapons designs had become so sophisticated that officials could not have confidence the weapons would work in the absence of testing. Simpler designs that could be trusted to work without having been tested were needed, he noted, adding that the validation of such designs was "one objective of the final test campaign."

"These tests are being conducted to increase French confidence in their nuclear deterrent under a CTB [comprehensive test ban]," says Christopher E. Paine of the Natural Resources Defense Council (NRDC), a Washington, D.C., legal and research organization.

Puzzlingly enough, however, in interviews in 1994 with U.S. nuclear experts, both Bouchard and his predecessor, Roger Baléras, asserted that the country would need 10 to 20 more tests to produce the robust weapons suited to deployment under a CTB. Suzanna van Moyland, a researcher at the Verification Technology Information Center, suspects that the French are testing a new warhead variant in this final series. "They'd have the data already if it wasn't a new variant," she reasons. "It would be a laborious, expensive and politically risky venture just to get the same results again."

Whatever the true purpose of this series, it has already accomplished something unexpected, if unintended, according to Robert S. Norris of the NRDC. As recently as last summer, French officials had been insisting that a test-ban agreement allow tests with yields up to several hundred tons. But "in an ironic way, Chirac's decision to resume testing, and the outcry against it, has forced the French to retract their insistence on these permissible yields and adopt a true, zero-yield test ban," Norris declares. "That's the most important thing. And nobody could have predicted it would happen." —*Glenn Zorpette*

## A Harebrained Scheme

*Experimental rabbit-killing virus runs amok in Australia*

Despite many efforts to control them, rabbits have overrun Australia and are causing blight, erosion and extinction wherever they go. On an island off the Australian coast, scientists had been testing a deadly rabbit virus to see whether it could serve as a weapon against the long-eared pests. The virus escaped from the island and is now spreading, uncontrolled, through the Australian outback. The unplanned viral outbreak—along with a Scottish experiment involving pollen—raises concerns about the ability of scientists to control the spread of organisms that they introduce into the environment.

The offending bug causes rabbit calicivirus disease (RCD). First identified in China in 1984, the virus surfaced in Europe in 1986. The only animals affected by the virus are European rabbits, *Oryctolagus cuniculus,* which usually die within two days of contracting the disease. Australians and New Zealanders soon noticed RCD's potential as an anti-rabbit agent, and the Australia and New Zealand Rabbit Calicivirus Disease Program was born. Funded by the Commonwealth Science and Industry Research Organization (CSIRO), among other agencies, the program set up a test site at Wardang Island off the coast of South Australia to evaluate the disease for an eventual release, probably in 1997.

Officials foresaw little possibility of the virus's escaping from the island, but escape it did. In October, CSIRO announced that RCD had spread beyond the quarantine area despite "strict surveillance" and "comprehensive microsecurity measures." Within a week a rabbit had died of RCD at Point Pearce on Yorke Peninsula near the island. Scientists speculate that bush flies may have transmitted the virus to the mainland.

A frantic attempt to contain the virus ensued. Scientists were stunned when the virus appeared at Yunta, 260 kilometers from Point Pearce, and then at Blinman, another 200 kilometers away. Officials suspect that humans may have been responsible for the disease's spread. A reporter for the *Sydney Morn-*

CSIRO

***RESEARCHER examines rabbits in an Australian laboratory of the Commonwealth Science and Industry Research Organization.***

ANTI GRAVITY

## Sinister Movements

The difference between a violin and a viola, a disgruntled musician once noted, is that a viola burns longer. One similarity between violins and violas—and cellos and basses for that matter—is that they each demand a vastly unequal division of manual labor. The left hand subtly flutters, each finger dancing independently over the strings; the right hand could be sawing through a 2-by-4.

MICHAEL CRAWFORD

The brain, of course, is controlling that left hand, while also desperately calculating how long to hold the dotted 16th note in cut time that the eyes have just spotted up ahead. The virtuosity of the southpaw digits is maintained through an oddity already suspected by anyone with violinist friends—their brains are different, according to a report last October in *Science.*

The authors did magnetic imaging of the brains of six violinists and three other string players, then compared them with six people who, like Jack Benny, can't play the violin. The musicians' cortical regions associated with the left digits were larger than both the regions corresponding to the right hand and either region in controls. The musicians' brains also showed greater responses to tactile stimulation of the sinister, dextrous digits.

Of course, the violin is merely a convenient marker for asymmetrical digital stimuli. Other studies have shown comparable adaptations in the brains of owl monkeys that had one or two digits stimulated over long periods. (No evidence supports rumors that these reports neglect to mention violins only because the monkeys attempted to blow into them.)

The researchers acknowledged that their experiment doesn't prove that playing the violin makes the brain grow bigger. It might be the other way around. "It could be argued," they wrote, "that individuals with a genetically determined large representation of the left-hand digits make superior string players and therefore are more likely to continue with musical training once they have begun."

On the other hand (it had to be said somewhere in this piece), the investigation also showed that the cortical differences were largest in the musicians who began their studies youngest. So chances are that playing indeed trains the brain. All of which means that the conductor George Szell was more right than he ever could have guessed when he said, "In music one must think with the heart and feel with the brain."

*—Steve Mirsky*

*ing Herald* may have unintentionally carried the virus from Point Pearce to Yunta; farmers, whose animosity toward rabbits runs deep, may also be helping the virus along, possibly by releasing infected animals.

By the end of October, officials gave up trying to control the spread of RCD and instead concentrated on measuring the disease's effect. The virus spread rapidly, reaching Broken Hill in New South Wales by early December. "In broad terms, where the virus has been present for four to five weeks, there is over a 95 percent kill rate," said Niall Byrne of CSIRO. This is enough to reduce the population, despite rabbits' legendary fecundity.

Rough estimates put the current rabbit population at 200 to 300 million. At these numbers, they can devastate the flora of the countryside and drive other plant-eaters into starvation. "The scale of rabbit damage is just vast," Byrne says. "Twenty-five percent of native mammals in New South Wales have been rendered locally extinct, and the major factor is rabbits."

RCD is not the first antirabbit virus released in Australia. In 1950 myxomatosis was introduced, and the initial success was astounding. A premyxomatosis rabbit population of about 600 million crashed to about 50 million. But the virus failed to control the rabbit population for long.

"There was coevolution between the virus and rabbits," says Edward M. Berger, a biologist at Dartmouth College. "A less virulent strain of virus evolved, and the [rabbit] survivors became more resistant," enabling both the virus and host to survive. "It's a classic example of coevolution." Although myxomatosis occasionally has local kill rates of 70 to 95 percent, the overall rate has dropped to about 40 percent.

RCD is not the only organism that is difficult to control. A recent Scottish study shows that a previously used pollen dispersion model—which would probably have been used to predict how fast genes from a transgenic crop leak into the environment—badly underestimated the amount of pollen that spreads from large oilseed rape fields.

The study's authors discovered that pollen can disperse much farther than the model predicts; pollen levels that had been expected no more than 100 meters away were observed at distances up to 2.5 kilometers. The study thus demonstrates the principle that any genes that scientists introduce into a crop can quickly spread into wild populations. A herbicide-resistant strain might cause "superweeds" that would be difficult to contain. This incident, like the one in Australia, should serve as a warning: it is easy to release an organism into a new environment but hard to control its spread. —*Charles Seife*

## THE ANALYTICAL ECONOMIST

# Reaching an Economic Event Horizon

A communist government facing economic ruin makes deals with Western businesses, promising market reforms and a docile workforce to attract investment. It sounds like yet another story from Eastern Europe—right down to $100 million from financier George Soros to help build a new billion-dollar petrochemical complex. Instead it's Calcutta, capital of the state of West Bengal in India, where last fall government authorities working to modernize the former gem of the British Empire finally banned rickshas.

After Maoist insurgents in the state were brutally suppressed by India's central government, during the late 1960s and early 1970s, a more moderate Marxist-Communist Party won at the polls in 1977 and has stayed in power in West Bengal ever since. Former foreign investors remember the early years of the party's accession as a time of constant strikes and almost nonexistent electricity. The conservative Congress Party now governs most of the rest of India, and West Bengal has an economic growth rate a little better than half that of the rest of the country.

Unable to generate funds internally for development—or to get substantial sums from the central government—West Bengal officials have begun courting the same foreign investors they made unwelcome almost 20 years ago.

Many Western observers have commented on the irony of a communist government seeking capitalist succor, seeing in the Bengali turnabout an echo of the fall of the Soviet Union and the Eastern Bloc. Despite massive state ownership of businesses ranging from power plants to hotels, however, the vast bulk of the region's productive capacity has remained in private hands, according to government figures. The power of the people has been reflected instead in a workforce where general strikes may be called at least once a year and where government employees (according to local newspaper accounts) often do not show up at their posts until noon.

If anything, the region's economic organization still harks back to the colonial bureaucracy that dominated the area from the mid-1700s until 1911, when the British moved their capital to Delhi. Since independence and partition, however, Calcutta has far less territory to administer. One World Bank official suggests that communism has failed to raise living standards because its development requires a prior stage of capitalism—which Bengal has yet to undergo. (Communists in Chile in the 1960s preached a similar creed, promising to out-market the capitalists on their way to a socialist future, but they never had a chance to realize their theories.)

Thus far efforts toward market reforms have been mixed. During the past year, the state has approved about $4 billion in foreign investments, and Chief Minister Jyoti Basu has tried to reform the civil service by threatening to fine workers who do not show up by 10:30 A.M. Yet less than half of those investments have actually been made; the rest continue to wend their way through the ancient bureaucracy.

The future of Basu's initiatives (or retrenchments, depending on where you stand) may depend far more on pragmatic considerations than on ideology or economic theory. Although it came to power because of popular dissatisfaction with the widespread corruption of the Congress Party, the Communist Party has maintained its position at least in part by intimidation—newspapers regularly report beatings of opposition-party leaders. The gangs who helped to win past elections have now become an embarrassment to party leaders, although at least this class strife appears less severe than the Muslim-Hindu religious violence that plagues the rest of the country.

A number of social scientists have reported that the Communist Party's land reforms and other redistributive efforts have failed to close the gap between rich and poor in West Bengal, and that disparities may actually be growing as economic development picks up. If that trend continues, the party may find itself pushed out of office after having been co-opted by the same capitalist system that it sought to turn to its own purposes. —*Paul Wallich*

## Green Policies

*Insurers warm to climate change*

When it comes to environmental issues, the insurance industry may be best remembered for its broadsides against Superfund, the program that mandates cleanup of hazardous-waste sites. During the past few years, though, some members of the $650-billion global property and casualty industry have started to take a decidedly progressive stance on the hottest environmental issue of the decade.

Take a full-page advertisement that appeared in the *Financial Times* last October. It entreats the reader to "stop and think: giant storms are triggered by global warming; this is caused by the greenhouse effect; which is, in turn, accelerated by man." The advertisement in the salmon-colored broadsheet could have been paid for by Greenpeace. Instead the bill was picked up by Swiss Re, one of the world's largest reinsurance companies (reinsurers absorb risk of losses from other insurers).

A sense of pragmatism has turned a number of insurers into true believers. Traditionally, companies predict the risk of incurring future losses based on past experience. Until 1987, the single largest loss the industry had experienced from a natural disaster amounted to less than $1 billion. Then came a $2.5-billion European storm. That was just an appetizer. The Big One, Hurricane Andrew in 1992, left insurers with losses of $16.5 billion. Suddenly, estimates of future losses—what would have happened if Andrew had directly hit Miami and New Orleans—reached $85 billion. That number could have forced major insurance companies out of business. In fact, Andrew did push nine smaller companies into bankruptcy.

Insurers acknowledge that the reasons for large losses may have less to do with climate change than they do with insuring homes in exposed coastal areas and an increase in property values. Nevertheless, the mounting evidence of a warming trend—and its possible impact on worldwide weather patterns—has caught the attention of an industry that survives on its ability to estimate the financial impact of future events.

If some insurers have begun to sound a bit like Greenpeace, perhaps it's because they have had a little coaching. Greenpeace International activist Jeremy Leggett has worked for several years to encourage the insurance industry to hop on the global-warming bandwagon. Insurers could serve as a counterbalance to a powerful oil and coal lobby that ardently combats measures to reduce emissions of greenhouse gases.

MARK WILSON *AP Photo*

***HURRICANE HAVOC, such as that experienced in St. Thomas last September, has caused some insurers to look for better climate data and some to worry about global warming.***

Leggett organized a conference in Berlin in March 1995, held on the eve of the United Nations's climate summit there, at which insurance officials voiced strong opinions on a changing climate. "I'm personally convinced that global warming is taking place and showing itself in the frequency and severity of natural disasters," says Gerhard A. Berz, a meteorologist who attended the conference representing Munich Re, the world's largest reinsurer.

The insurance industry has kept a high profile on this issue during the past year. Members of the industry contributed to a recent report on measures to be taken to mitigate the impact of climate change, put out by the U.N.'s Intergovernmental Panel on Climate Change and its World Meteorological Organization. Last fall, in a statement coordinated by the U.N.'s Environment Program, a group of some 50 insurers from outside the U.S. pledged to adhere to environmentally sound principles. And last winter, U.S. insurers met with Vice President Al Gore to discuss how the industry might become involved with climate change issues.

Talk has translated into action. Since Andrew, insurers have, in fact, provoked the ire of some homebuilders by asking for installation of window glass that can resist hurricane-force winds. The industry has also sought various forms of government help to cover losses from catastrophes that exceed the resources of individual companies. But insurers have yet to set their sights on the environmentalists' goal of disinvesting from coal or oil companies in favor of investments in alternative-energy companies. "The financial department here says it has to get as much profit as it can," Berz states. "As soon as the environmental sector starts producing profits, we'll start investing."

The insurers' main response to drastic changes in the weather is a familiar one: when the going gets tough, run for the hills. Before Andrew, Florida insurers engaged in a pitched competition to sign on new customers. After the storm, Allstate attempted to drop 300,000 homeowners' policies in the state until the legislature placed strict limits on the amount of coverage insurers could withdraw. Worries about covering property losses in hurricane-prone Florida have yet to abate. State Farm this year

decided to stop the writing of most new homeowners' policies there.

Insurers' views on global warming are by no means monolithic. Many U.S.- and Bermuda-based insurers have shied away from making any statements about the likelihood of climate change and the consequences of global warming. "It's not a good practice to raise people's fears unless you have solid science; otherwise people don't believe you the next time," says Charles L. Kline, president of Centre Cat, a Bermuda reinsurer whose second largest shareholder is General Motors's pension fund. Although a consensus on the expectations of global warming has continued to build, the debate on how this trend would affect the frequency and intensity of hurricanes is by no means settled. Kline also suggests that speculating about the dire effects of global warming could be a ploy by some insurers to lay the groundwork for raising premiums.

But even those companies that do not purchase full-page newspaper advertisements would like a better idea over the short run of when it is time to pack their bags. So they have begun to develop closer contacts with climate-modeling scientists. Twelve companies from the insurance industry have begun funding the Risk Prediction Initiative (RPI), a program within the Bermuda Biological Station for Research, which itself does climate-related investigations.

The RPI backs major researchers from around the world in projects that analyze data from global climate modeling—for instance, the impact of El Niño, a periodic oceanic disturbance that affects weather, on hurricanes and typhoons. Insurers can then crunch the data into the statistical models they use to estimate future losses. During 1996 the companies will put up $750,000 for more than 10 projects. In one, researchers at Florida State University are probing the effectiveness of a supercomputer-based atmospheric model for predicting future hurricanes.

Some companies have even brought these studies in-house. Richard T. Gordon, a physicist employed by the Chubb Group of Insurance Companies, often works in chaos modeling. Following the pattern set by Wall Street's effort to discern market trends through new types of modeling, the insurance industry is seeking to predict the vagaries of climate fluctuations by employing the full array of modern soothsaying tools, from chaos theory to neural networks to fuzzy set theory. Gordon's most recent writing is a chapter entitled "A Hybrid Neural Network Genetic Algorithm Model for El Niño"; it appeared in a book published last November by the American Society of Mechanical Engineers Press.

In coming years, climate researchers may find a good prediction lands them squarely on the money. Centre Cat, the Bermuda reinsurer, planned to announce this month two annual $25,000 research prizes, one for the best prediction on the formation and landfall of Atlantic hurricanes, the other for Pacific typhoons. If hurricane divining improves, the industry may be able literally to place bets on the weather.

In 1992 catastrophe futures began to be traded on the Chicago Board of Trade as a hedge against the losses that may be incurred from big storms or earthquakes. A futures market in acts of God means that it may become increasingly lucrative to distinguish when a prediction is worth following or when it is just so much hot air. —*Gary Stix*

## A New Way to Spell Relief: V-e-n-o-m

*A toxin from killer sea snails promises a better painkiller*

The only thing worse than a pain-filled life is a painful death. Both are altogether too common. The World Health Organization estimates that on any given day over three million people struggle with chronic pain from cancer alone. In the U.S., a study published last November in the *Journal of the American Medical Association* revealed that of thousands of terminal patients interviewed at five major hospitals, about half spent their final days in agony. The study's authors laid much of the blame at the feet of a medical culture that chases miracles to the bitter end rather than doing its best to assuage suffering.

BALDOMERA OLIVERA

***CONE SNAILS stun their prey by pumping a toxin cocktail through a tooth-tipped tube* (right).**

But there is another reason that medicine so often fails to offer relief: doctors still have few weapons against pain, and each has major drawbacks. Aspirin and other over-the-counter remedies are far too weak for most chronic conditions. Narcotics such as morphine often work for a while, but many patients quickly become physically dependent and require ever stronger infusions. Eventually, says William G. Brose, director of the Pain Clinic at the Stanford University School of Medicine, "opiate-tolerant patients feel no effect even from a dose thousands of times stronger than that needed to kill you or me." As a result, many sufferers can obtain comfort only at the cost of their faculties.

A better alternative may be on the way. At a recent meeting of the American Pain Society, Brose presented encouraging results from small-scale clinical trials of a novel drug that he says could represent "the first of a new and exciting class of compounds to treat severe pain." Brose administered the drug, known as SNX-111, to seven of his toughest cases. Five were dying of cancer; two others had shoulder nerves torn from their spinal cord. All had stopped responding to opiates and were beyond the help of conventional medicine.

Within three days of starting on the new drug, five of the seven patients reported that their pain had disappeared—in the case of one amputee, for the first time in 25 years. Only one patient did not respond at all. More remarkable, the most serious side effects reported were mild eye jitters and a slight drop in blood pressure. None of the fuzzy-headedness and lethargy typically caused by narcotics was seen. Nor, after as much as nine months' treatment, did any patients show signs of tolerance or addiction. Quite the contrary: one woman was able to trim the cost of her daily dose of opiate-based drugs from $6,000 to $100 after starting on SNX-111; the new drug seemed to have reversed much of her tolerance for the old.

Veteran doctors might easily dismiss these results as so much snake oil and voodoo, because none of the early clinical trials were controlled against placebo effect. In fact, snake oil is not too far from the truth: the new painkiller was isolated from the venom of cone snails.

SCIENTIFIC AMERICAN

CORRESPONDENCE

**Reprints** are available; to order, write Reprint Department, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, or fax inquiries to (212) 355-0408.

**Back issues:** $8.95 each ($9.95 outside U.S.) prepaid. Most numbers available. Credit card (Mastercard/Visa) orders for two or more issues accepted. To order, fax (212) 355-0408.

**Index of articles since 1948** available in electronic format. Write SciDex®, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, fax (212) 980-8175 or call (800) 777-0444.

***Scientific American*-branded products** available. For free catalogue, write Scientific American Selections, P.O. Box 11314, Des Moines, IA 50340-1314, or call (800) 777-0444.



**Editorial correspondence** should be addressed to The Editors, Scientific American, 415 Madison Avenue, New York, NY 10017-1111. Manuscripts are submitted at the authors' risk and will not be returned unless accompanied by postage.

**Advertising correspondence** should be addressed to Advertising Manager, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, or fax (212) 754-1138.

**Subscription correspondence** should be addressed to Subscription Manager, Scientific American, P.O. Box 3187, Harlan, IA 51537. The date of the last issue of your subscription appears on each month's mailing label. For change of address notify us at least four weeks in advance. Please send your old address (mailing label, if possible) and your new address. We occasionally make subscribers' names available to reputable companies whose products and services we think will interest you. If you would like your name excluded from these mailings, send your request with your mailing label to us at the Harlan, IA address. E-mail: SCAinquiry@aol.com.

But evolution, not voodoo, seems to explain why SNX-111 and similar compounds might be ideal for blocking certain types of pain.

There are more than 500 species of carnivorous cone snails lurking in the oceans. All use chemical weapons—a cocktail of neurotoxins fired through a harpoonlike tube tipped with a barbed tooth—to catch the worms, mollusks or fish on which they feed. More than a decade ago Baldomero Olivera and his colleagues at the University of the Philippines in Manila observed that several varieties of cone snail had evolved a sting potent enough to completely paralyze a fish within seconds—no doubt to keep it from flailing like a hooked bass as the snail reeled the prey into its distensible stomach. Encounters with one of the species, *Conus geographus,* have even killed at least 20 humans.

Analyzing these snails' venom, Olivera discovered two keys to their lethal efficiency. First, the toxic protein fragments they brew, known as conotoxins, are smaller than almost any others seen in nature. They are as small as peptides can be, in fact, while still retaining particular shapes. Those shapes, perfectly suited to block critical openings on nerve cells, are their second advantage. Small and specific, the conotoxins blitz straight to their targets and shut down the doomed fish's nerves.

The same characteristics happen to be ideal in a drug. Chemicals that are too large or indiscriminate tend to break up and bind to things they shouldn't, causing side effects. Opiates, for example, affect many parts of the brain besides those that only perceive pain. And any drug that works on nerve cells directly, rather than by stimulating an intermediate chemical (as morphine does), is less likely to produce tolerance.

From the witches' brew of peptides in cone snail venom, researchers have isolated one class, called ω-conotoxins, that may ease human suffering by interrupting pain signals as they travel through the spinal cord to the brain. These signals, which in many patients arise from pressure or injury to a major nerve, depend on calcium ions moving within cells. A few promising ω-conotoxins stop up particular pain-associated channels through which these ions pass without interfering with other, very similar channels that convey normal sensations. As a consequence, pain is blocked without numbness.

SNX-111, which is a synthetic ω-conotoxin developed by Neurex in Menlo Park, Calif., is 100 to 1,000 times more potent than morphine. It must be administered through a small tube directly to the spinal cord. Although many opiates have the same limitation, the inconvenience was great enough to move the same amputee who had finally found relief from 25 years of pain to discontinue using the new drug.

In a large-scale controlled trial of SNX-111 that was scheduled to begin in January at 30 medical centers, Neurex planned for miniature pumps to be implanted directly into 150 to 300 patients. The refillable pumps, which are made by Medtronic in Minneapolis and are regularly used to deliver narcotics, can administer a few nanograms of the drug an hour for two to three months. If the results of the trial, due in 1997, prove the substance to be safe and effective, life for a few hundred thousand anguished souls may become a bit more bearable. *—W. Wayt Gibbs*

## Teleconomics

*Information marketeers focus on content rather than access*

Think for a minute about what happens if the Internet really does change forever the business of telecommunications. Unlikely as it seemed a few years ago, this do-it-yourself research network, cobbled together by academic amateurs, seems to have gained an economic momentum far greater than most commercial alternatives. Interactive television has lost its sparkle, and with it has dimmed the idea that the telecommunications future belongs to information turnpikes built and operated by giant cable and telecommunications companies. Commercial on-line services are also falling by the info-wayside.

Even Bill Gates, the most powerful man in computing, is bowing to the strength of the Internet by ostensibly transforming Microsoft Network from a traditional on-line service, which prospers on the quality of the connections it provides, to a publishing center on the Internet, which will live or die on the quality of its information. As Microsoft goes, so goes the rest of the business world. For if the Internet model does triumph—and there are good reasons to think that it might—all companies will increasingly come to depend for survival on their ability to get the attention of information-glutted consumers.

The sheer ability to distribute—by owning a printing press, a television station or even a global chain of shops—

## More Rules of the Road

You can listen to it on the radio or watch it on television late at night, but you may not be allowed to read, see or hear "indecency" on the Internet if congressional censors have their way. Recent U.S. House and Senate actions could impose a two-year jail sentence and a $100,000 fine on anyone who publishes, in electronic form, material that lawmakers deem indecent even though the Supreme Court has struck down attempts to ban indecent language (around which comedian George Carlin built his famous "Seven Words You Can't Say on TV" monologue) from broadcast and print media.

Civil liberties organizations and Internet companies (which could be criminally liable if they do not monitor and censor their users) stand ready to challenge the federal legislation. In the meantime, however, many states have already leapt into the arena with even stricter legislation of their own. It is not clear whether federal law will preempt local rules, says Ann Beeson of the American Civil Liberties Union, so anyone who ventures into cyberspace should probably take notice.

Illinois is one of several states hoping to protect minors. It is now a felony in that state to send an e-mail proposition to anyone under the age of 13, even if you reasonably believe them to be older; it is also a felony for 17-year-olds and their elders to proposition anyone younger than 17. In Kansas, meanwhile, it is illegal to possess or transmit any digital data that depicts or simulates sex involving teenagers under 16.

In Oklahoma it may soon be illegal, depending on the context, to allow unmarried persons under the age of 18 access to electronic depictions or descriptions of "nudity"—definition of which includes "buttocks with less than a full opaque covering." Georgia sees a different danger: it has criminalized any instruction about explosives or other dangerous weapons if the speaker has reason to believe that someone somewhere in the audience might use the knowledge for illicit purposes.

And if your e-mail, Usenet posting or World Wide Web page might be read by someone in Connecticut, be aware that it is a felony there to transmit text that contains threats with the intent to harass, annoy or alarm. (Any civil libertarians who send e-mail to the Connecticut legislature should probably avoid statements like "I would like to punch your face in for passing such a stupid law.") —*Paul Wallich*


COURTESY OF SÃO PAULO MUSEUM OF ART; *Seated Bather Drying Her Leg*, by Auguste Renoir

***FEDERAL CRIME? An electronic version of this magazine cover might fall afoul of any U.S. law barring "indecent" material from the Internet (courts have judged some nudes decent, others not). Local statutes forbidding computer dissemination of nudity are both more strict and less ambiguous.***

will no longer matter much. But getting attention will require businesspeople to cultivate a deeper understanding of both the economics of information networks and the nature of information itself. A quick examination of how existing information networks have grown illuminates both issues.

A lesson from the world of banking should give businesspeople a profound respect for the power of the economic forces underlying the Internet's growth. Citibank, the retail banking arm of Citicorp, had the biggest network of retail banks in New York City in the 1970s. When the new technology of the automated teller machine (ATM) came on the horizon, Citibank tried to capitalize on that advantage by rushing to install ATM machines outside each branch. It hoped for a virtuous circle. With more branches than any of its rivals, Citibank's ATMs would be more convenient. More convenient ATMs would lure more customers, which would enable Citibank to open more branches and lure yet more customers.

Instead Citibank found itself in a vicious circle. Citibank's rivals formed the New York Cash Exchange (NYCE), an internetwork of ATMs. Customers of any member of NYCE could use the ATMs of any other. By pooling technology to create an open network, the banks decreased the unit costs of the machines and increased the value to customers of joining their network. Anyone seeking to build a proprietary network removed from the Internet should study Citibank's failed strategy.

Airlines provide a glimpse of the new competitive issues that crop up in open networks. Almost as soon as U.S. airlines employed computers to manage reservations, they, like Citicorp, sought to use them to lock in their customers—and they, too, were frustrated. Each airline hoped to put onto the desks of travel agents a terminal linked directly to its reservation system. Because it would be easiest for the travel agents to book, say, American Airlines's flights, they would book more of them. Nice theory. But in practice the scheme was thwarted by the same kind of self-escalating openness that quashed Citibank's scheme to dominate ATMs.

With American and United Airlines vying for desktop dominance, agents opted for the terminal that offered the most convenience—and the search for convenience, in turn, forced airlines to open up their reservation systems to list the flights of rivals. With all reservation systems compelled by competitive pressures to offer more or less the same content, airlines tried to differentiate themselves by maintaining control over the context in which that information was viewed.

Executives from American, whose ticketing symbol is AA, explained, with

only a slight smile, that their flights always appeared first on the reservations screen because of their commitment to alphabetization, for easy viewer navigation through the data. Executives at United used a different excuse to put their flights first. In the end, complaints of airlines pushed to the bottom of the last screen forced the U.S. government to specify exactly how reservations systems could display the data they made so ubiquitously available.

The Internet is unlikely to face such regulation. For unlike the users of dumb reservations terminals linked to airline mainframes, the machines on the Internet have enough horsepower and versatility to let the consumers of information create their own layouts—regardless of what the producers would like to foist on them. And if users do not wish to do it for themselves, a small legion of companies is queuing up to sell them search engines or intelligent agents or whatever other technology promises to make comprehensible the sprawling web of information that stretches across the Internet. But the lessons of content and context are worth paying attention to nonetheless—for they point to two very different methods of directing the focus of customers' attention.

The obvious way is to provide answers to the questions that consumers are already curious about—that is, to provide content. But as the airlines discovered as soon as they began electronic delivery of flight times and seat availability, it is hard to keep a competitive grip on content in the interlinked world of computer networks. Like it or not, in a realm in which there are no barriers to the distribution of information, no enterprise can prevent customers from having a peek at the competition's wares. A company can, however, try to manipulate the questions customers ask, and the context in which answers are offered, in such a way as to put the competition into the shade.

Already the Internet is creating ways to manipulate context more cleverly, and more profitably, than the airlines' crude attempts to bias reservations toward their own flights. Some of the tricks involved are variations on themes familiar to advertisers and marketers. Travel companies have long advertised in the pages of magazines such as *Condé Nast Traveler* because they are an obvious place to get in front of the eyes of people interested in travel. But those companies should soon be able to deliver their services via the electronic *Traveler* by taking bookings directly from its Web pages. *Traveler* itself will morph from magazine to mall, much to the advantage of the companies that are savvy enough to rent shop space in its virtual corridors.

A more subtle approach for companies looking to use the Internet to get a grip on their customers, however, is to try to engage them with the same kinds of questions the company itself is asking. Instead of keeping the R&D department locked away in the back room, it can be put on display via the Internet. The obvious hope is that the technological vision exciting the people in the laboratory will also sway the views of those roaming the Internet.

Sun Microsystems has worked this trick magnificently with Java, its object-oriented programming language for the Internet. As well as sending out press releases, Sun made available from its Web site briefing documents and copies of the software—even early versions that, in addition to containing potentially off-putting bugs, would also alert potential commercial competitors to forthcoming features. In the new world of the Internet, competitive advantage will lie not in who you know but how you know them. —*John Browning*

## Bill Gates's Apocryphal History

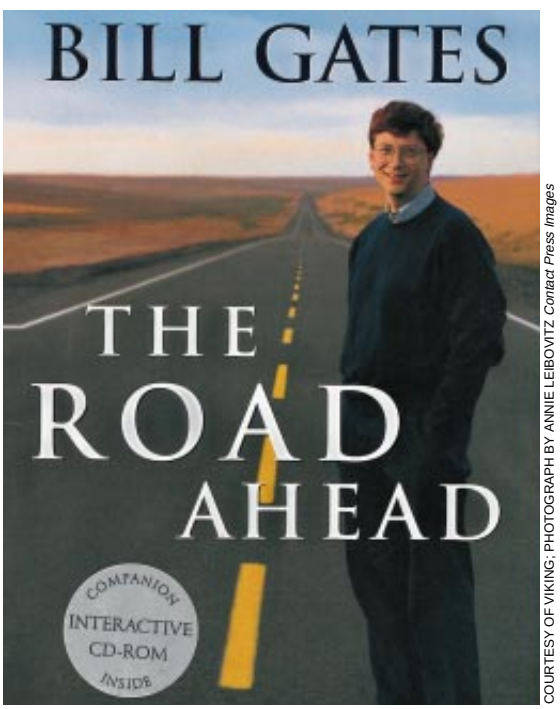
### The Road Ahead *perpetuates a myth about the road behind*

If Bill Gates's grasp of the past is any guide, readers should take his visions of the future with a dose of skepticism. In his instant best-seller *The Road Ahead,* the mega-entrepreneur ruminates (along with two co-authors) about where the computer revolution will take us. On page xiii of his preface, Gates mentions past prophets whose prognostications "look silly today." Among these, he says, is "the commissioner of U.S. patents who in 1899 asked that his office be abolished because 'everything that can be invented has been invented.' "

This anecdote is both ancient and apocryphal, according to science historian Morgan Sherwood of the University of California at Davis. The story was widespread more than 50 years ago, Sherwood says, when a scholar named Eber Jeffery did an exhaustive investigation of it. In a 1940 article, entitled "Nothing Left to Invent," published in the *Journal of the Patent Office Society,* Jeffery traced the tale to testimony delivered before Congress by Henry L. Ellsworth, the commissioner of patents, in 1843. Ellsworth told lawmakers that the rapid pace of innovation "taxes our credulity and seems to presage the arrival of that period when human improvement must end."

But this statement was "a mere rhetorical flourish," Jeffery wrote, "intended to emphasize the remarkable strides forward in inventions then current and to be expected in the future." Indeed, Commissioner Ellsworth asked Congress to provide him with extra funds to cope with the flood of inventions he expected in transportation, communications and agriculture. Ellsworth stepped down from his post two years later. But in his resignation letter, far from recommending that his office be closed, he expressed pride at having expanded it.

Bill Gates is hardly alone in perpetuating the myth of the nearsighted patent official. In the March 17, 1995, issue of *Science,* Daniel E. Koshland, Jr., then editor-in-chief of the journal, repeated the apocryphal anecdote in an essay on science's future. In Koshland's telling, however, the legendarily myopic commissioner made his gaffe in the 1860s rather than in 1899. —*John Horgan*




COURTESY OF VIKING; PHOTOGRAPH BY ANNIE LEIBOVITZ *Contact Press Images*

***BEST-SELLER by cyberseer Gates tells the tale of the 19th-century patent commissioner who thought everything had been invented.***

PROFILE: DANIEL C. DENNETT

# Dennett's Dangerous Idea

The effect is baffling. I am scrutinizing an image of an everyday scene flashing on and off on a video screen several times a second. In fact, two almost identical images, differing only in one significant feature, are rapidly alternating. My task is simply to spot the feature.

It sounds like child's play, but, inexplicably, I am blind to what is in plain view for 10, 20, even 30 seconds. When the scales finally fall from my eyes, the chimney jumping from one end of a roof to the other is suddenly so obvious as to be comic. It's the same with a long series of pairs of images. I feel humiliated, but Daniel C. Dennett, who has lured me into this perceptual pickle, is exultant. The effect, discovered recently by Ronald A. Rensink of Cambridge Basic Research, "allows me to say 'I told you so!' " Dennett exclaims.

My miserable performance is actually par for the course. And according to Dennett, a philosopher at Tufts University who has some singular ideas about the mind, the sensory snafu created by the flickering images reveals a good deal about awareness. Specifically, our powerful subjective impression that we are conscious of sensory perceptions in real time is in truth an illusion. "We don't take the world in at a glance and keep a copy of it to compare to the next glance. We take in a very limited amount of material, and that's all we keep," Dennett explains. Hence my helplessness when called on to compare images flashed a split second apart.

Philosophers tumbled long ago to the realization that things are not how they seem. Dennett has pushed this insight a mind-bending step further: even our illusions are not what they seem, because they are built of still more illusions. With this mental nutcracker he claims to have split open the conceptual chestnut of consciousness, which he sees as the product of a "virtual machine" running in the brain.

Based on psychophysical observations like Rensink's, Dennett's theory has divided thinkers about the mind. His brashly titled 1991 book *Consciousness Explained* has sold some 100,000 copies, an unprecedented number for a hard-core philosophical treatment. Last year Dennett published a take-no-prisoners manifesto, *Darwin's Dangerous Idea*, aimed at establishing natural selection as "the single best idea anyone has ever had." And currently he is advising on the construction of a robot at the Massachusetts Institute of Technology that "will be conscious if we get done all the things we've got written down in the calendar."

Dennett's book jackets proclaim that he directs the Center for Cognitive Studies. In fact, he virtually—as Dennett himself might put it—*is* the Center for Cognitive Studies. The center comprises, when I visit, just the director, an administrative assistant and a cognitively challenged dog. Dennett, who holds the title of Distinguished Arts and Sciences Professor, finds the situation perfect, as he has plenty of time to write. "What more could anyone want?" he asks.

*Dennett thinks Darwinism, properly understood, has no "intolerable moral implications."*

Demystifying the mind has been the main focus of his campaign to bring philosophy up to date with biology. Dennett gave up hopes of becoming a full-time sculptor to study philosophy at the University of Oxford in the 1960s, but the supposed golden age of that field had, he now realizes, already turned decadent. "It drove me crazy—we were shooting our mouths off about the mind without knowing anything about the brain," he recalls. He found himself feeling "more and more like an outsider," but the late Gilbert Ryle, a preeminent figure in the school known as materialism, encouraged Dennett's eccentric interest in psychology. Today there are some philosophers who wish Ryle had not been so solicitous.

Materialists, including Dennett, hold that consciousness and will are wholly due to, well, material. The material is designed by evolution (or people) to serve useful functions. But Dennett perceives a deep reluctance, even among prominent intellectuals, to accept this account fully. The reluctance is based on "fears that the 'wrong' answer would have intolerable moral implications," he writes in *Darwin's Dangerous Idea.* A consistent theme in Dennett's writing is that these fears are based on ignorance.

Dennett has written, so far, six (and a half) books and dozens of papers. In his 1984 book *Elbow Room,* he explored the perennial conundrum of whether humans have free will. His firm conclusion was that we do; our deliberations form part of the "critical nexus" leading to a course of action, so the power to decide is ours even if the universe operates along deterministic lines. "The main problem for free will is political," declares Dennett, who identifies himself as a liberal and a member of the American Civil Liberties Union.

Dennett's most dangerous idea is that the key to understanding behavior, human and otherwise, is adopting the "intentional stance." Just as the best way of predicting a chess-playing computer's next move is to—rather than study its circuitry—assume it "intends" to win, so the best way to predict the behavior of an animal is to suppose that it has beliefs and desires. Dennett sees beliefs as virtual properties of brains similar to software in a computer.

In everyday life, we assume unthinkingly the intentional stance when dealing with people. The behaviorist movement in psychology, however, had attempted to expunge notions such as desire and expectation from the scientific lexicon. The intentional stance does not deny explanations of behavior based on nerve-cell activity. But Dennett says his approach has stimulated work in animal behavior and substantially influenced child developmental psychology.

Applying the intentional stance to humans led Dennett to his theory of consciousness. Several intriguing psychophysical experiments demonstrate that, contrary to intuition, the order in which we perceive things is not always the order in which the sensory data arrive in the brain. One of Dennett's favorite examples is called the color-phi phenomenon. If an experimenter arranges for two separate but stationary red and yellow lights to flash in rapid alternation—with a short period of darkness between flashes—viewers believe they see a single moving light that changes color en route between two positions. Dennett accepts this report of conscious experience—that's the intentional stance—even though the subjects retrospectively believe the color of the illusory moving light to be changing before the new-colored light is turned on.

Dennett's explanation is that at any instant there is no dividing line between sensory data that people are conscious of and those that are unconscious. In

fact, we are not conscious of anything at precisely the time we imagine. What we experience, Dennett maintains, is generated a little after the fact, as the result of a competition among multiple patterns of mental activity propagating within the brain. Awareness comprises the small fraction of those mental events whose influence will persist and so alter beliefs about what just happened. In the color-phi experiment, data about the new color are combined with the false idea of a moving light and wrongly referred backward in time.

The idea of consciousness as largely illusion is too much for some more conservative philosophers. Dennett has for some years been engaged in an acrimonious running battle with John R. Searle of the University of California at Berkeley. Searle insists (over Dennett's denials) that Dennett does not believe in consciousness as everyone else understands it. In a recent issue of the *New York Review of Books,* Searle accused Dennett of representing an "intellectual pathology." Dennett in turn charges Searle with "egregious misrepresentation." Dennett judges that his own work on the mind "has shown how to turn the mysteries into puzzles." Other philosophers agree that Dennett's contributions have been substantial. Owen Flanagan, a philosopher at Duke University, says the intentional stance represents "a tremendous contribution."

Just as behavior makes sense only with the intentional stance, Dennett holds, so evolution makes sense only with the parallel stance known as adaptationism. ("So our intentionality," Dennett once wrote, "is derived from the intentionality of our 'selfish' genes!") This idea, that features of organisms can be understood as natural selection's solutions to evolutionary problems, is hardly new, but some biologists have cast doubt on its value.

*Darwin's Dangerous Idea* is a hard-hitting pitch for adaptationism as a "universal acid" explaining not merely organic life but also culture and such perplexities as meaning and morality. "Feelings may get hurt," Dennett warns: the book is for "those who agree that the only meaning of life worth caring about is one that can withstand our best efforts to examine it."

Dennett takes vigorous issue with the idea that adaptationism should be relegated to a minor role. That notion he lays at the door of Harvard University paleontologist Stephen Jay Gould. Dennett devotes entire chapters of *Darwin's Dangerous Idea* to deconstructing Gould and Roger Penrose, the University of Oxford mathematician whose books argue that thought depends on effects explicable only by quantum physics. Both thinkers, Dennett believes, are forlornly searching for "skyhooks": higher-order explanations that will seem to give human beings a special meaning.

Dennett finds meaning in evolution itself. And despite being no man of the church, his reverence for Darwinism borders on the religious. The tree of life is not something he could pray to, he declares in *Darwin's Dangerous Idea,* but he insists that "this world is sacred." (He also sings sacred music with the New England Classical Singers.) And Dennett maintains that evolution can inform morality.

He steers well clear of the classic "naturalistic fallacy" of supposing that how things are in nature is how they should be in society. Moreover, he criticizes sociobiologists, including Harvard biologist Edward O. Wilson, for oversimplifying human behavior; sociobiologists forget "again and again" that people can think for themselves and learn from others.

BRUCE DAVIDSON *Magnum*

***DANIEL C. DENNETT says consciousness is created by a "virtual machine" in the brain.***

Yet Dennett believes that understanding the creative power of natural selection can enhance appreciation for its unique products, including cultures and endangered species and individual human beings. That kind of appreciation could form the basis of a practical morality, he suggests. Absolutist ethical principles, in contrast, often generate conflicts between competing rules.

By now Dennett and I have spent several hours talking about the books and Rensink's images. But I want to know what Dennett thinks about the really Big Questions: it seems fair to ask a philosopher how we find happiness and what is evil. For happiness, according to Dennett, you have to find a project that is "bigger than yourself." And absolutism extended into fanaticism is the biggest evil facing the world: "There's a growing sense of desperation and futility—until we have a more equitable economic and political situation, people are going to be strongly motivated to that sort of behavior."

Ultimately, though, Dennett is an optimist. He can't resist pointing out that his five-year-old grandson knows more about volcanoes than anyone did 100 years ago. Humans can assimilate science's bewildering insights, he asserts, and "education is the key." Writing with the energy of a Fury in his uniquely educational fashion, Dennett seems to be doing his share to keep the forces of darkness at bay. —*Tim Beardsley*

# Malnutrition, Poverty and Intellectual Development

*Research into childhood nutrition reveals that a poor diet influences mental development in more ways than expected. Other aspects of poverty exacerbate the effects*

by J. Larry Brown and Ernesto Pollitt

The prevalence of malnutrition in children is staggering. Globally, nearly 195 million children younger than five years are undernourished. Malnutrition is most obvious in the developing countries, where the condition often takes severe forms; images of emaciated bodies in famine-struck or war-torn regions are tragically familiar. Yet milder forms are more common, especially in developed nations. Indeed, in 1992 an estimated 12 million American children consumed diets that were significantly below the recommended allowances of nutrients established by the National Academy of Sciences.

Undernutrition triggers an array of health problems in children, many of which can become chronic. It can lead to extreme weight loss, stunted growth, weakened resistance to infection and, in the worst cases, early death. The effects can be particularly devastating in the first few years of life, when the body is growing rapidly and the need for calories and nutrients is greatest.

Inadequate nutrition can also disrupt cognition—although in different ways than were previously assumed. At one time, underfeeding in childhood was thought to hinder mental development solely by producing permanent, structural damage to the brain. More recent work, however, indicates that malnutrition can impair the intellect by other means as well. Furthermore, even in cases where the brain's hardware is damaged, some of the injury may be reversible. These new findings have important implications for policies aimed at bolstering achievement among underprivileged children.

Scientists first investigated the link between malnutrition and mental performance early in this century, but the subject did not attract serious attention until decades later. In the 1960s increasing evidence of undernutrition in industrial nations, including the U.S., along with continuing concern about severe malnutrition in developing countries, prompted researchers to examine the lasting effects of food deprivation. A number of studies in Latin America, Africa and the U.S. reported that on intelligence tests children with a history of malnutrition attained lower scores than children of similar social and economic status who were properly nourished. These surveys had various experimental limitations that made them inconclusive, but later research has firmly established that undernutrition in early life can limit long-term intellectual development.

### Worry over Brain Damage

For many years, scientists considered the connection between nutrition and intellectual development to be straightforward. They assumed that poor nutrition was primarily a worry from conception to age two, when the brain grows to roughly 80 percent of its adult size. In this critical period, any degree of malnutrition was thought to halt the normal development of the brain and thereby to inflict severe, lasting damage.

Gradually, though, investigators recognized that the main-effect model, as we have termed this view, was too simplistic. For instance, the emphasis on

STEVEN LEHMAN *SABA*

**HEALTHY BREAKFAST provided to schoolchildren helps them avoid malnutrition and its attendant problems. A growing consensus indicates that meeting nutritional requirements throughout childhood is essential to full intellectual development. The program providing food to this kindergarten in Central Falls, R.I., is funded by the National School Breakfast Program.**

the first two years of life proved somewhat misguided. Brain growth in that period is not always terminated irreversibly in undernourished children. Rather it may be put on hold temporarily; if diet improves by age three or so, growth of the brain may continue at close to a normal pace. Conversely, injury to the brain can occur even when a child suffers malnutrition after the first two years of life—a sign that providing adequate nutrition throughout childhood is important to cognitive development. Focusing exclusively on the first two years of life is thus inadequate.

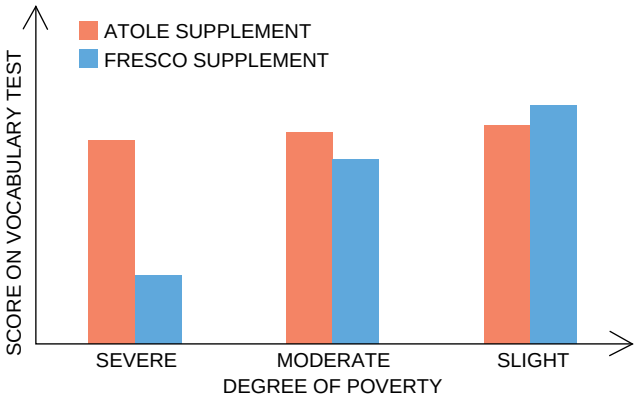
Furthermore, although severe underfeeding in infancy can certainly lead to irreparable cognitive deficits, as the main-effect model predicts, the model cannot fully account for intellectual impairment stemming from more moderate malnutrition. This flaw became apparent in the 1960s, when researchers showed that mildly undernourished children from middle- or upper-income families (whose nutrient deficits stemmed from medical conditions) did not suffer the same intellectual troubles as did mildly underfed children in impoverished communities. If poor nutrition impaired cognition only by structurally altering the brain, the two groups should have performed alike. Something else had to be at work as well. In other words, factors such as income, education and other aspects of the environment could apparently protect children against the harmful effects of a poor diet or could exacerbate the insult of malnutrition.

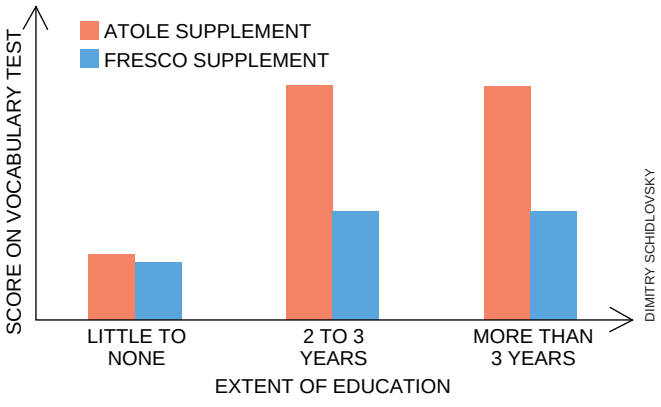
### No Energy to Learn

In the 1970s research by David A. Levitsky and Richard H. Barnes of Cornell University helped to clarify how malnutrition might hinder cognitive development in ways other than injuring the brain. Levitsky and Barnes studied rodents to examine the effects of malnutrition. Levitsky concluded that the malnourished animals performed less well on tests of mental ability, such as maze running, not because they suf-

## Effects of Poverty and Malnutrition: The Guatemalan Project

In a project carried out by the Institute of Nutrition of Central America and Panama, children and young adults in Guatemala who had received nutritional supplements in infancy were studied to assess the influence of early diet and poverty on later intellectual development. Subjects, including the boys at the right, were given a battery of cognitive tests. Individuals who regularly consumed a highly nutritious supplement called Atole before the age of two performed at about the same level on most tests, such as tests of vocabulary skills, regardless of economic status (*bottom left*). But the performance of those given a less nutritious supplement called Fresco varied with poverty level. Evidently, good nutrition early in life can help counteract the destructive effects of poverty on intellectual development. Among individuals who had more than two years of formal education, those who consumed Atole scored significantly higher than those who received Fresco (*bottom right*)—an indication that poor nutrition in infancy can subsequently undermine the benefits of schooling. —*E.P.*



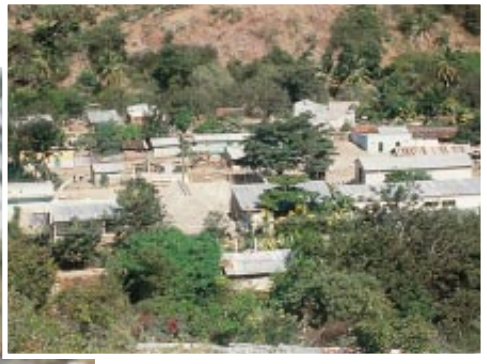




fered brain damage but mostly because, lacking energy, they essentially withdrew from contact with their peers and the objects in their environment. In addition, mothers coddled the less mobile infants, further hindering their growth and independence.

By extrapolation, the findings implied that cognitive disability in undernourished children might stem in part from reduced interaction with other people and with their surroundings. This fundamental shift in understanding produced increased optimism about the prospects for remediation; if decreased social interaction was partly at fault for cognitive impairment, then social and intellectual remediation could presumably help make up for deficits in the youngsters' experiences.

Although the new ideas were compelling, scientists did not have much human evidence to buttress the changing views. A recent study by one of us (Pollitt) and several collaborators adds strong support to the notion that malnutrition affects intellectual development in part by compromising many different aspects of a child's development. The research also provides added insight into how poor diet and economic adversities during childhood combine to impede intellectual functioning later in life. Pollitt's collaborators included Reynaldo Martorell of Emory University, Kathleen S. Gorman of the University of Vermont, Patrice L. Engle of California Polytechnic State University and Juan A. Rivera of the Institute of Nutrition of Central America and Panama.

The project was an extensive follow-up of Guatemalan children who were studied by other scientists many years earlier. In 1969 the Institute of Nutrition of Central America and Panama, with the help of various U.S. government agencies and private foundations, began a massive effort to examine the value of nutritional supplements in preventing the health problems of malnutrition. For



GUATEMALAN VILLAGE (*above*) was one of four towns where children and pregnant women received nutritional supplements.

PHOTOGRAPHS BY MARGARITA DE MARTINEZ *INCAP*

COGNITIVE TESTS given as part of the Guatemalan project assessed the participants' abilities in areas such as vocabulary, reading comprehension, general knowledge and arithmetic.

eight years, residents of four villages in Guatemala received one of two nutritional supplements. When this phase of the study was being planned, researchers felt that protein was the most important nutrient missing from diets in developing countries. Therefore, project workers looked specifically at how children would respond to added protein in their diets. The mothers and children in two of the villages received a high-protein supplement called Atole (the Guatemalan name for a hot maize gruel). Inhabitants of the other two villages—who constituted the control group—received Fresco, a sweet, fruit-flavored drink, which contained no protein. Both supplements provided vitamins and minerals as well as calories; Fresco provided a third the calories of Atole.

When the study began, all pregnant women, and all children under the age of seven in the villages, were invited to participate. During the course of the study, children under seven who moved into the villages and women who became pregnant were also asked to join the project. More than 2,000 children and mothers participated between 1969 and 1977. Regular medical exams of the children revealed that both supplements improved the health of the participants, but Atole performed more impressively. For instance, in all four villages, the rate of infant mortality decreased. But in the villages that received Atole, infant mortality decreased 69 percent, whereas in villages receiving Fresco, the rate went down by just 24 percent. Also, only Atole improved growth rates in children under three.

### Gains in Guatemala

In the follow-up study, carried out in 1988 and 1989, Pollitt and his colleagues visited the villages to assess how these early nutritional supplements affected intellectual development over the long term. More than 70 percent of the original participants—by then, ranging in age from 11 to 27 years old—agreed to take part in the follow-up. In particular, the team's analysis concentrated on the group of roughly 600 people who were exposed to Atole or Fresco both prenatally and for at least two years after birth. These adolescents and young adults took literacy, vocabulary and reading comprehension tests, a general knowledge exam, an arithmetic test and a standard nonverbal intelligence test. The researchers then determined how education and economic status (measured by house quality, father's occupation and mother's education) correlated with test scores.

The subjects who received Atole in early life performed significantly better on most tests of cognition than those who received Fresco. The strongest effects of Atole were observed among those at the low end of the social and economic ladder: these children performed as well as the more privileged children in their villages [*see box on these two pages*]. Atole thus served as a kind of social equalizer, helping children from low-income families achieve at the same level as their slightly more economically advantaged peers within the village. But the children of this study all lived in extreme poverty and did not perform at the same level as, say, a child from a middle-income household in a more prosperous area of Guatemala. Hence, adequate nutrition by itself could not fully compensate for the negative effects of poverty on intellectual growth.

In addition, Atole appeared to have increased the advantage of education. With every additional year of schooling, the differences in achievement between the adolescents who received Atole and those who consumed Fresco increased. This result indicates that poor nutrition can essentially negate some typical benefits of education. In separate but related studies, Pollitt and his collaborators, working in Peru, and Sally Grantham-McGregor of the University of the West Indies, working in Jamaica, have demonstrated that learning capabilities are affected by how recently one has eaten. So breakfast every day before school is indeed important, particularly among children at risk for undernutrition.

The better long-term effects in the Atole group can largely be explained by the differences in the children's motor skills, physical growth, and social and emotional development. The youngsters who received Fresco in their early life suffered more physical disadvantages—a slower rate of growth and a slower rate of recovery from infection, for example—compared with those who received Atole. Because development was hindered, these children also learned to crawl and walk slightly later on average than the infants who received Atole. Pollitt and his colleagues speculate that for the infants who took Fresco, this limitation delayed the acquisition of the cognitive skills that children develop when they explore their social and physical environment.

Furthermore, because these undernourished toddlers remained small for their age, adults might have tended to treat them as if they were younger than their actual age. Such a response would very likely slow cognitive development, if the toddlers were not challenged—to talk, for instance—in the same way that other children their age were. Children who consumed Atole, in contrast, avoided malnutrition, grew up faster and were presumably exposed to more challenges in their social environment. Of course, the results do not rule out the possibility that the Fresco recipients may have suffered some degree of brain damage that impeded their later functioning. The findings, however, imply that additional factors, such as the child's social environment, played a major role as well.

The results in Guatemala are also consistent with the prevailing understanding of the interactions between poor

nutrition, poverty and education. Nutritional supplements combat the effects of poverty, but only somewhat. A well-nourished child may be better able to explore the environment, but an impoverished community may offer little to investigate. And although schools can provide much of the stimulation children need, early malnutrition can undermine the overall value of education. Most important, this study demonstrates that poor nutrition in early childhood can continue to hinder intellectual performance into adulthood.

Because the early planners of the Guatemalan study chose to examine protein, these results emphasize protein's importance to intellectual growth. The supplements also included calories, vitamins and minerals; consequently, their role should be taken into account, but the arrangement of this particular study makes isolating the effects difficult.

Other work links essential vitamins and minerals to mental ability. For example, in one study in West Java, Pollitt and his colleagues showed a close association between iron-deficiency anemia (the most common consequence of malnutrition) and poor mental and motor skills in children. The researchers gave iron supplements to babies between 12 and 18 months old who were suffering from iron-deficiency anemia. The mineral significantly improved the infants' scores on mental and motor skills tests. Sadly, children with iron-deficiency anemia are more susceptible to lead poisoning, which produces its own set of neurological disorders that interfere with proper cognition. Consequently, poor children face a double jeopardy: they are more likely to be anemic and more likely to live where lead poisoning is widespread.

### Correcting and Preventing Impairment

Studies such as the one in Guatemala have prompted many scholars, including one of us (Brown), to suggest that when the social and economic aspects of a child's environment cannot be easily changed, providing adequate nutrition during infancy and later will at least lessen the cognitive deficits engendered by poverty. Nutritional supplements cannot by themselves reverse the long-term adverse effects of earlier undernutrition, however. The ideal would be to provide additional support, such as tutoring, opportunities to develop new social skills and guidance

## Avoiding Malnutrition

Opinions on what constitutes malnutrition—and recommendations for avoiding the problem—have been refined over time. Early studies considered lack of protein to be the most troubling deficiency in the diets of underfed children, especially in developing countries. Ingested protein is broken down into amino acids, which are then recycled to build the specific proteins needed by the individual at any given time. Proteins form many structural elements of the body and carry out most cellular processes. By the 1970s, though, investigators had begun to worry about calories, too. When faced with a lack of calories, the body breaks down amino acids for energy instead of using them to make new proteins.

RODICA PRATO

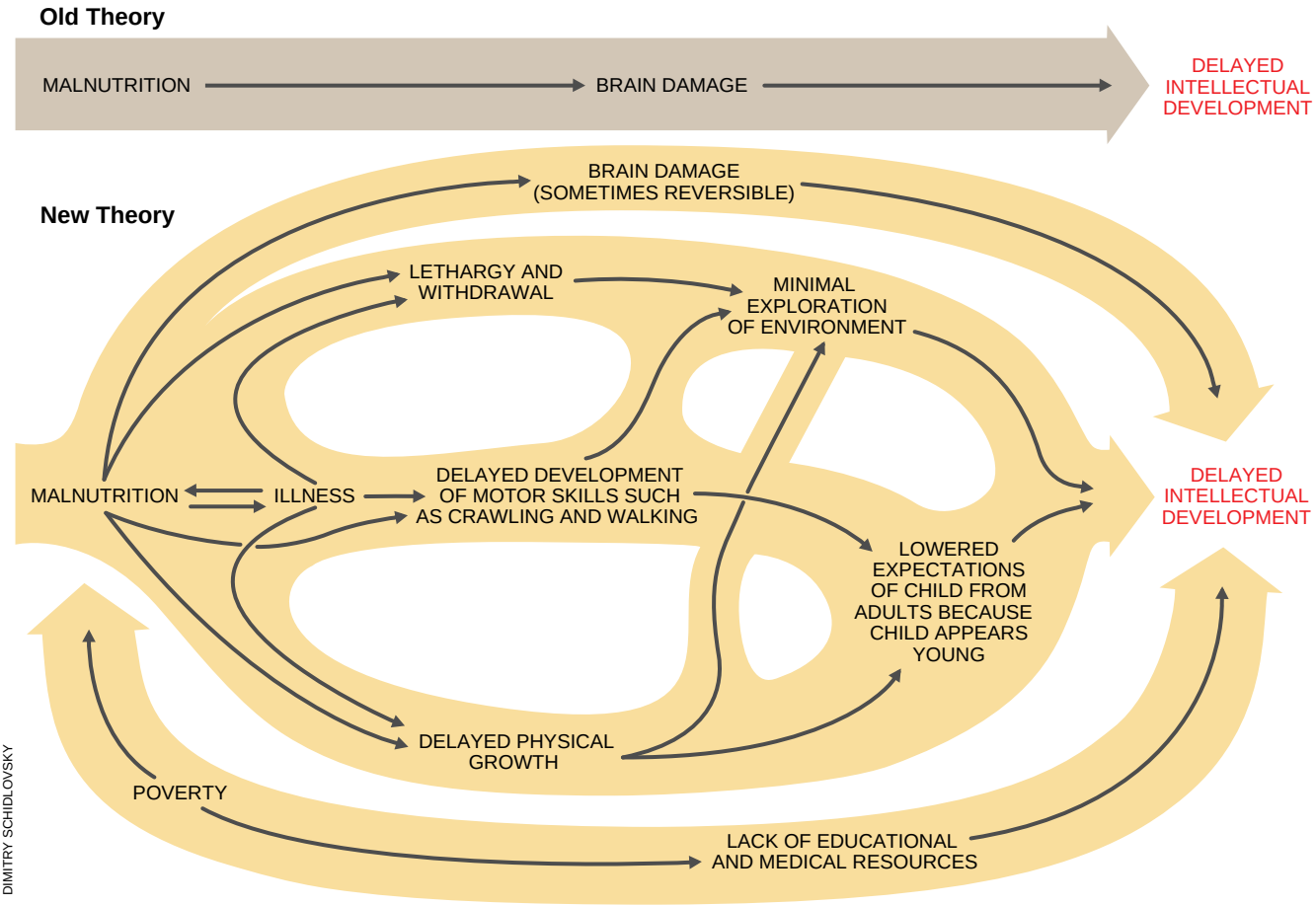
**DIETARY requirements for children can be met by eating several servings a day from each of five food categories.**

In more recent years, nutrition research has emphasized that shortages of vitamins and minerals—particularly vitamin A, iodine and iron—contribute to significant health problems. Vitamin A is important for good vision, bone growth, tooth development and resistance to infection. Iodine, which tends to be scarce in developing countries, is needed for proper operation of the central nervous system. Iron is a constituent of hemoglobin, which transports oxygen to tissues. Iron also helps the body fight infections; levels of the mineral are low in diets of many poor children in the U.S. Hence, most investigators now believe malnutrition is best avoided by a diet that supplies enough protein, calories, vitamins and minerals to ensure normal growth. Some standard guidelines for optimal nutrition in children are listed below. *—J.L.B.*

| FOOD CATEGORY | SERVINGS PER DAY | SERVING SIZE* | | |
|---|---|---|---|---|
| | | AGE 1 TO 3 YEARS | 4 TO 6 YEARS | 7 TO 10 YEARS |
| WHOLE-GRAIN OR ENRICHED BREADS, CEREALS, RICE, PASTA | 6 OR MORE | 1/2 SLICE BREAD OR 1/4 CUP RICE OR NOODLES | 1 SLICE BREAD OR 1/2 CUP RICE OR NOODLES | 1 TO 2 SLICES BREAD OR 1/2 TO 1 CUP RICE OR NOODLES |
| VEGETABLES | 3 OR MORE | 2 TO 4 TBSP OR 1/2 CUP JUICE | 1/4 TO 1/2 CUP OR 1/2 CUP JUICE | 1/2 TO 3/4 CUP OR 1/2 CUP JUICE |
| FRUITS | 2 OR MORE | 2 TO 4 TBSP OR 1/2 CUP JUICE | 1/4 TO 1/2 CUP OR 1/2 CUP JUICE | 1/2 TO 3/4 CUP OR 1/2 CUP JUICE |
| LEAN MEATS, FISH, POULTRY, EGGS, NUTS, BEANS | 2 OR MORE | 1 TO 2 OZ | 1 TO 2 OZ | 2 TO 3 OZ |
| MILK AND CHEESE | 3 TO 4 | 1/2 TO 3/4 CUP MILK OR 1/2 TO 3/4 OZ CHEESE | 3/4 CUP MILK OR 3/4 OZ CHEESE | 3/4 TO 1 CUP MILK OR 3/4 TO 1 OZ CHEESE |

** Data from "Growth and Nutrient Requirements of Children." P. M. Queen and R. R. Henry in* Pediatric Nutrition, *edited by R. J. Grand et al. Butterworth, 1987*

LISA BURNETT




**MALNUTRITION HINDERS COGNITIVE ABILITIES through several interacting routes, according to recent research. Early models of malnutrition considered cognitive deficiencies to result only from damage to the brain (*top*). Now scientists also believe (*bottom*) that malnutrition alters intellectual development by interfering with overall health as well as the child's energy level, rate of motor development and rate of growth. In addition, low economic status can exacerbate all these factors, placing impoverished children at particular risk for cognitive impairment later in life.**

from an involved parent or another concerned adult. Recent studies have shown that enriched education programs for children in economically impoverished communities can often ameliorate some of the problems associated with previous malnutrition.

To have the best chance at being useful, such intervention should be comprehensive and sustained. Most undernourished children face persistent challenges that can exacerbate the effects of underfeeding. They frequently live in areas with substandard schools and with little or no medical care. Their parents are often unemployed or work for very low wages. And the children may suffer from illnesses that sap energy needed for the tasks of learning.

On balance, it seems clear that prevention of malnutrition among young children remains the best policy—not only on moral grounds but on economic ones as well. The U.S., for example, invests billions of dollars in education, yet much of this money goes to waste when children appear at the school door intellectually crippled from undernutrition. The immediate expense of nutrition programs and broader interventions should be considered a critical investment in the future. Malnutrition alters educational preparedness and, later, workforce productivity, making it an unacceptable risk for its victims as well as for a nation's strength and competitiveness. Steps taken today to combat malnutrition and its intellectual effects can go a long way toward improving the quality of life—and productivity—of large segments of a population and thus of society as a whole.

## *The Authors*

J. LARRY BROWN and ERNESTO POLLITT have collaborated for several years on the policy implications of childhood nutrition. Brown, director of the Center on Hunger, Poverty and Nutrition Policy at Tufts University, is also professor of nutrition and health policy at the School of Nutrition Science and Policy and at the School of Medicine. Pollitt is professor of human development in the department of pediatrics at the School of Medicine at the University of California, Davis, and is also a member of the Program of International Nutrition.

## *Further Reading*

NUTRITION AND BEHAVIOR. Edited by Janina R. Galler. Plenum Press, 1984.

THE EFFECTS OF IMPROVED NUTRITION IN EARLY CHILDHOOD: THE INSTITUTE OF NUTRITION OF CENTRAL AMERICA AND PANAMA (INCAP) FOLLOW-UP STUDY. Edited by Reynaldo Martorell and Nevin S. Scrimshaw. Supplement to *Journal of Nutrition,* Vol. 125, No. 4S; April 1995.

THE RELATIONSHIP BETWEEN UNDERNUTRITION AND BEHAVIORAL DEVELOPMENT IN CHILDREN. Edited by Ernesto Pollitt. Supplement to *Journal of Nutrition,* Vol. 125, No. 8S; August 1995.

# The Global Positioning System

*Two dozen satellites hovering thousands of miles out in space are allowing people to locate themselves on the earth's surface with remarkable precision*

by Thomas A. Herring

Turning onto the final approach, the Boeing 737 airliner responded smoothly to the command of its computerized autopilot, setting up for what looked to be another perfect landing. Although automatic approaches are routinely performed in bad weather, this particular aircraft was not using the normal navigational signals beamed up from the airport to complete its so-called Category IIIA landing—the kind used when the pilot cannot see the runway until after the airplane touches down. The jet's occupants were relying instead on satellites of the U.S. Department of Defense's Global Positioning System (GPS) high in orbit overhead. These modern navigational benchmarks, floating in space at an altitude of more than 20,000 kilometers, were supposed to guide the swiftly moving aircraft safely to the ground.

As the 737 neared the runway, the GPS signals indicated that the ground loomed only 300 feet below the landing gear, and the airliner slowed its descent. Having completed numerous landings that day, the engineers on board had grown confident in the craft's satellite-guided abilities. But on this attempt the autopilot suddenly sounded an alarm: the GPS equipment had lost contact with a critical satellite. The airplane's human pilot quickly took over control from the computer and throttled up the engines to abort a potentially disastrous landing.

After later analysis, the engineers who had developed the airplane's guidance equipment realized that the temporary loss of signal had been caused by a software "bug" in the GPS satellite. The flaw had not been detected earlier, because no one before had relied on GPS navigation for such a demanding task. The GPS system had, in fact, been designed with a built-in uncertainty in position of 100 meters—longer than a football field. Several years ago few people would have dared to imagine that the GPS could lead an airplane all the way to the ground. But in the intervening time some clever tinkering has brought about a surprising level of navigational precision, and the GPS has evolved into something even its makers had not envisioned when they launched the first of its satellites.

The Department of Defense began construction of its sophisticated satellite positioning system in the mid-1970s to allow military ships, aircraft and ground vehicles to determine a location anywhere in the world. Although the designers of the GPS had meant it primarily for classified operation, they made provision for civilians to use the satellite signals to locate themselves—but far less well than their military counterparts. A reduction in accuracy seemed necessary for the unclassified signals; otherwise an enemy could easily use the GPS broadcasts, and the elaborate satellite system would not give the U.S. any military advantage. Yet remarkably, scientists and engineers working outside the armed forces have devised ways to circumvent the purposeful degradation of the GPS signals, and ordinary citizens are now able to achieve much better results than the Department of Defense had ever expected.

Such refinements allow GPS radio receivers to guide pleasure boats into foggy harbors or passenger cars along unfamiliar roads. This satellite positioning system is now even used to keep track of the placement of cargo containers as they are shuffled around the holding yards of Singapore's busy port. More impressively, with GPS surveying equipment, geologists can measure the subtle shifts in the earth's crust—movements of just a few millimeters—that show the motion of the planet's tectonic plates and help to define the location and extent of earthquake-prone zones.

The glitches that sometimes emerge during such exercises are not so much failures as they are indications that nonmilitary scientists and engineers are pushing GPS instrumentation to limits never intended by the system's originators. How did the Department of Defense intend to deny their signal's inherent precision to civilians? Why have so many people succeeded in circumventing the prohibitions on accuracy? The answers to these questions require a broad understanding of the mechanics of satellite navigation in general and of the GPS in particular.

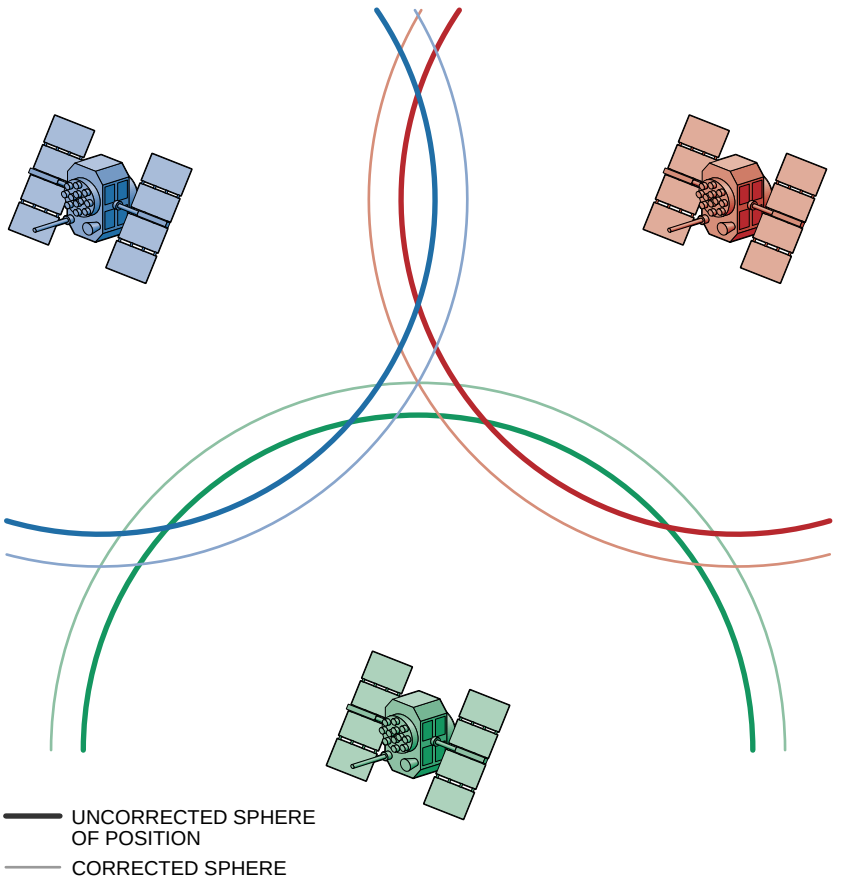
### A New Star to Steer By

Soon after the Soviets launched *Sputnik* in 1957, some scientists and engineers realized that radio transmissions from a satellite in a well-defined orbit could indicate the position of a receiver on the ground. The procedure uses the Doppler shift of radio signals as the satellite passes overhead. (A similar Doppler shift accounts for the sudden change in the tone of a train whistle as a locomotive speeds by.) Using this method, the U.S. Navy pioneered the "Transit" satellite positioning system during the 1960s.

The technique the navy employed

**SPHERES OF POSITION show the geometric basis for the operation of the Global Positioning System (GPS). By receiving coded broadcasts, a person on the earth can determine his or her distance to each of several satellites that orbit in precisely known patterns. Each distance measurement coincides with a set of possible locations that form an imaginary sphere (*purple* or *green*) centered on the satellite emitting the signal. The intersection of several spheres with the surface of the earth marks the person's exact location.**

ALFRED T. KAMAJIAN




JARED SCHNEIDMAN DESIGN

**CLOCK ERROR in the receiving electronics typically causes the distance measurements to the GPS satellites to be somewhat incorrect. With such inaccuracies, the corresponding spheres of position (*thick lines*) will not intersect neatly at a single point. Adjusting the receiver's clock slightly forward or back does, however, correct the ranges and allows all spheres to meet precisely (*thin lines*). This method appears here to require measurements from only three satellites, but in three dimensions, four satellite ranges are necessary.**

was, unfortunately, rather cumbersome. It required expensive electronic equipment on the ground and usually demanded reception of signals from two separate passes of the satellite overhead, which necessitated a wait of more than 100 minutes. Even under the best circumstances, with several days available for collecting signals, one could not hope to determine a location that was more accurate than about a single meter, and so Doppler positioning proved rather limited for precise land surveying.

But even before the deployment of the first Transit satellite, the Department of Defense had begun contemplating a more sophisticated approach that might, for example, allow a pilot flying a jet fighter instantaneously to determine his exact position. In particular, the U.S. Air Force was planning a navigation system that utilized "ranging"—the measurements of distances to several satellites—rather than the Doppler shift in radio frequency.

The determination of position by ranging is straightforward in concept. Suppose, for example, one is able to ascertain that a particular satellite is 20,000 kilometers away. Then the person's position must be somewhere on a huge sphere 20,000 kilometers in radius (40,000 kilometers in diameter) that surrounds that satellite. Because satellites travel in stable, predictable orbits, the location of the satellite, and the imaginary sphere surrounding it, is known exactly.

If at the same instant that the first range is taken the person can also measure the distance to a second satellite, a second "sphere of position" can be determined. A third range to a third satellite gives a third sphere, and so forth. In general, there will be few places where all the spheres meet. For example, two spheres can intersect along a circle; three spheres can coincide only at two points. Because one of these points typically represents an unreasonable solution to the navigation problem (it may be deep within the earth or far out in space), three satellite ranges are sufficient to give one's exact position.

### Synchronize Your Watches

The question the military planners first faced in designing the GPS satellite positioning system was how exactly to make the necessary range measurements. At that early juncture, there were many choices. For example, radar equipment could transmit a radio pulse and receive the echo after the signal had propagated up to a satellite and reflected back down again. A computer could then easily calculate the distance to the satellite from the measured delay and the known velocity of the radio pulse, the speed of light.

But such a system would force anyone using it to broadcast a stream of powerful radar bursts—not an ideal activity for soldiers, sailors or pilots who are trying to avoid being detected by their enemies. So the Department of Defense considered an alternative strategy. The navigation satellites could transmit radio pulses at specific, known times, and by measuring the exact instant when the pulses arrived, the receiving equipment could determine the distance to the satellite. That procedure demanded, however, that the receiver's clock be synchronized with the one on the satellite. This concept formed the basis for what became the Global Positioning System.

Exact synchronization may at first seem a rather severe requirement; a mismatch of as little as a millionth of a second would translate to an error of about 300 meters. Although the navigation satellites themselves could each carry a highly accurate "atomic clock," it would be prohibitively complicated and expensive for each receiver to be so equipped. But there was a way to avoid the need for such perfect timepieces on the ground: one need only establish how much the receiver's inferior clock had drifted from the correct time.

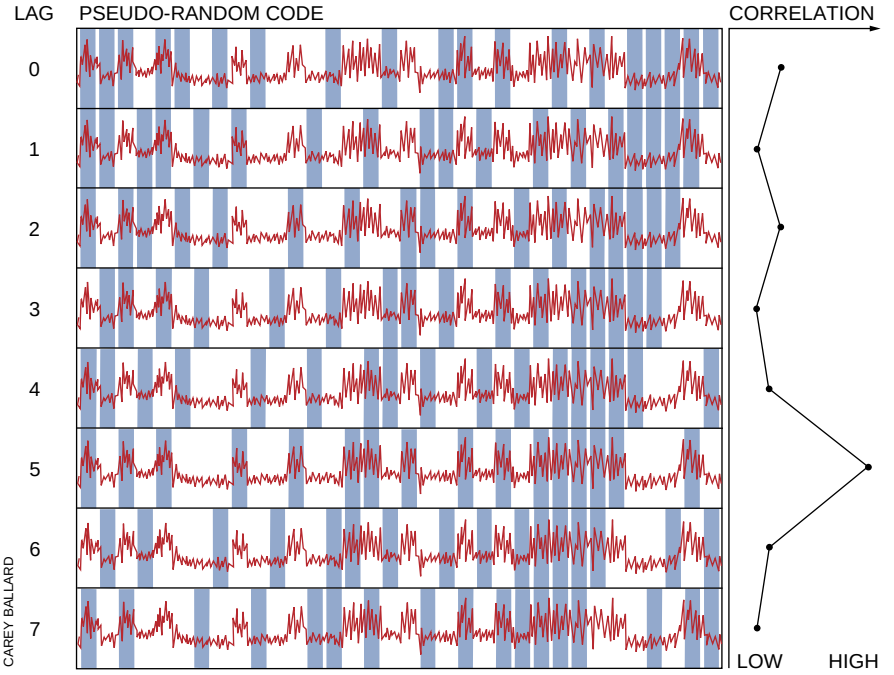
This task is not particularly difficult. One starts by assuming that the receiver's clock is approximately correct in calculating the ranges to four satellites. Because the receiver's clock is not in fact running exactly on time, the distances calculated, called pseudo-ranges, will not be entirely correct. The four pseudo-ranges will correspond to four imaginary spheres surrounding the satellites. These four spheres should ideally intersect at a single point—the receiver's location—but will not meet exactly, because the satellite and receiver clocks are not absolutely synchronized. All four spheres will be just a little too

large (if the receiver's clock is running fast) or too small (if its clock is slow). But there is one value for the amount of clock error that makes all four spheres meet perfectly, and so a few algebraic computations can determine the necessary adjustment. Thus, even a simple receiver, with an electronic clock that is no more complicated or expensive than an ordinary digital wristwatch, can be synchronized with the atomic clocks whizzing past high in the sky.

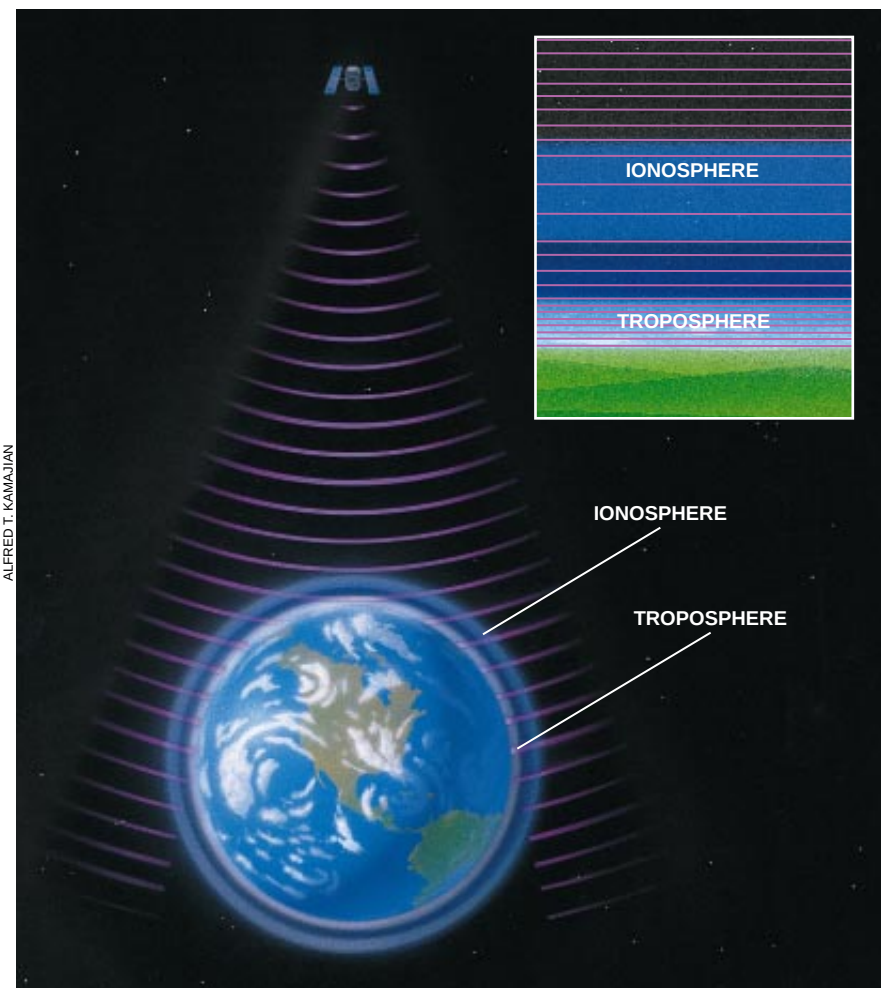
### Pseudo-this, Pseudo-that

Next, the military engineers who designed the system needed to decide how exactly to transmit the signals from the GPS satellites. They borrowed a technique that had been employed, strangely enough, by astronomers, among others, since the 1950s. Those scientists had been examining other planets by sending out radar pulses from their giant radio telescopes at what might have seemed to be random moments, but they were in reality following a carefully formulated code. The astronomers called these special cadences pseudo-random sequences. With them the researchers were able to measure the time delay in the weak radar reflections from the surface of a distant planet by finding the instant when the received signals and the transmitted pseudo-random sequence seemed to match most closely. In essence, the radar astronomers found the travel time (and hence the range to the radar target) by measuring when the two signals were most closely correlated.

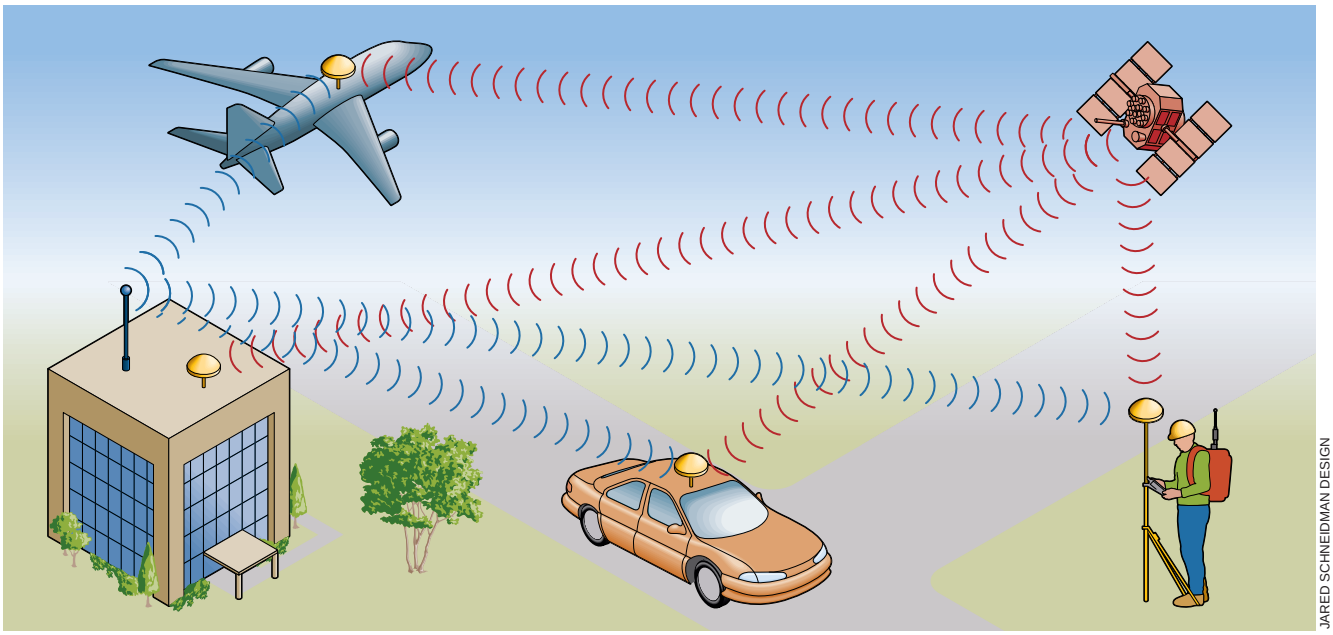
Noting the astronomers' success, the military engineers opted to use similar pseudo-random sequences for their new space-based positioning system. They decided, however, that the GPS satellites would emit high-frequency radio waves continuously rather than beam discrete radar pulses down to the earth. The use of pseudo-random sequences to code the radio emissions offered many advantages. One that would be greatly appreciated by boaters and hikers years




**PSEUDO-RANDOM SEQUENCES are broadcast from the GPS satellites at known times. The delay in the arrival of the radio emissions is found by comparing the noisy signal received (*red*) with versions of the known sequence (*blue*) that are shifted in time. A high correlation between the code sequence and the signal (*right*) indicates the time lag between transmission and reception of the signals.**




**ATMOSPHERIC LAYERS alter the GPS signals and can introduce significant errors. The most important effects arise as the radio waves pass through the earth's charged ionosphere and water-laden troposphere. Whereas the radio wave fronts (*inset*) tend to stretch out in the ionosphere, they bunch together in the troposphere. Because these disturbances to the GPS signals can be measured with a fixed receiver, scientists at the National Weather Service can now measure atmospheric water content this way.**



JARED SCHNEIDMAN DESIGN

**DIFFERENTIAL GPS circumvents the clock errors imposed for military security. A fixed receiver at a known location determines the clock errors in the satellite signals (*red*) and broadcasts the appropriate corrections (*blue*) to mobile receivers in use nearby. This method can decrease the uncertainty in GPS positioning from 100 meters to as little as one meter.**

later was that it allowed inexpensive GPS receivers to be built. All satellites could then transmit on the same frequency without creating a garbled mess of radio interference. Because each GPS satellite would transmit a unique code, an inexpensive, single-frequency radio receiver could easily separate the different signals.

The final decision that the military designers had to make concerned where to put the satellites. Nearly all space-bound hardware is placed in one of two types of orbit—either circling relatively close to the earth (in so-called low-earth orbit) or fixed at some 36,000 kilometers above the equator in a 24-hour-long geosynchronous orbit. Low orbits would cost relatively little for each launch and would demand only modest power from the satellites' transmitters because they would not have to broadcast their signals for any great distance. But such placement would necessitate that hundreds of separate satellites be swarming around the planet to provide global coverage. Lofty geosynchronous orbits, on the other hand, would require far fewer satellites, but each would have to carry a more powerful transmitter, and these signals would have difficulty reaching the earth's polar regions.

The GPS planners chose a compromise solution, launching the satellites into orbits that were neither particularly low nor high; the satellites were set to orbit at an altitude of about 20,000 kilometers. At that altitude, 17 satellites would be sufficient to ensure that at least four of them—the minimum number needed to fix a position—would always be available from any location on the earth's surface. The final configuration adopted for the GPS has 21 primary satellites and three spares in orbit.

### Selective Service

Because the U.S. defense forces intended to achieve a tactical advantage with the new satellite navigation system, from the outset they encoded the radio emissions to prevent adversaries from also gaining the ability to determine locations precisely. But the Department of Defense anticipated permitting ordinary people to use the GPS, at least in a coarse fashion. So the system's designers faced the question of how to limit civilian accuracy while still allowing the U.S. military to use the system to its full potential. There were several ways this dual operation could be accomplished. One method was to transmit incorrect information to unau-

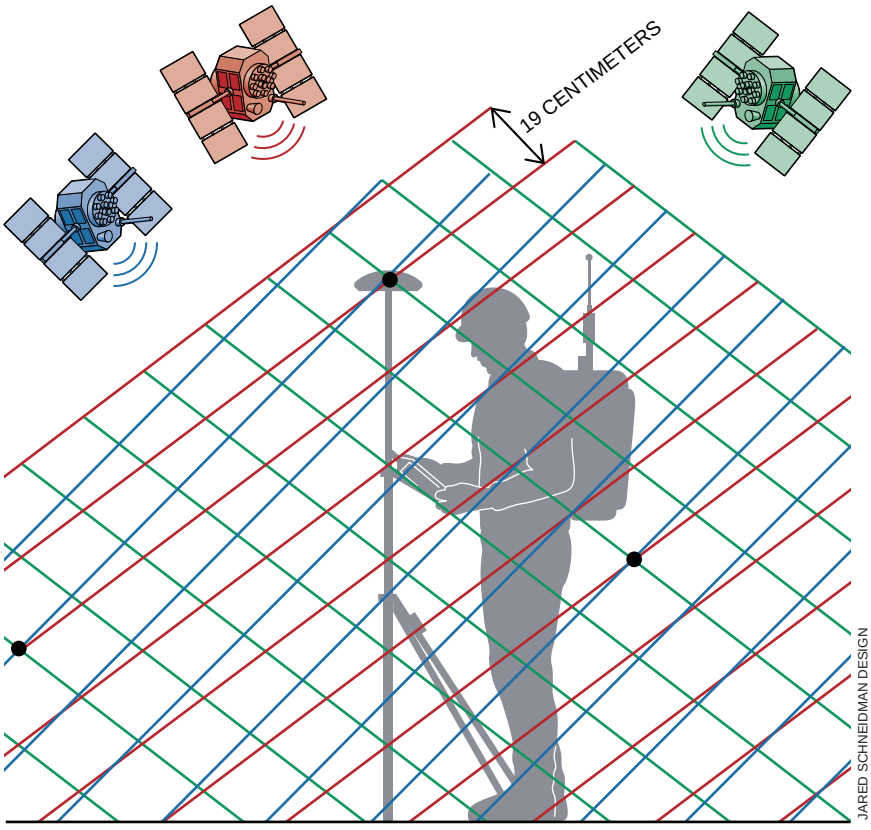
LOUIS PSIHOYOS *Matrix*

**BLIND NAVIGATION is no longer restricted to ships in fog. Visually impaired people may be able to use GPS to get around outdoors. This experimental application, and others, would be far easier without the confounding effects of military security measures.**

## Catching the Waves

Geophysicists have applied the GPS since the mid-1980s to help monitor the slow but relentless deformation of the earth's crust in geologically active regions—changes that can eventually cause the ground to rupture in an earthquake. For such investigations, they seek a maximum in precision from the GPS and often employ a technique called carrier tracking. Compared with differential GPS (which can locate a position to within about a meter or so), carrier tracking allows locations to be determined to within a few millimeters.

Carrier tracking gets its name from the satellite broadcasts that convey GPS signals on a set of so-called radio carrier waves. It works by determining which part of the radio wave strikes the antenna at a given instant—the "phase" of the received emission. Like an ocean swimmer sensing whether he or she is positioned at the crest of a wave, in a trough or somewhere in between, carrier tracking evaluates where on the 19-centimeter-long GPS radio waves the receiving antenna sits.




Carrier tracking allows a resolution that is a tiny fraction of a wavelength. The primary difficulty is in determining which of many identical waves the antenna is sensing. There are, however, a number of ways to overcome such ambiguity. The simplest is to track the carrier phase from several satellites simultaneously. If, for example, one found that the receiving antenna was located at the start of the waves (at zero phase) sent from three different satellites, there would be a limited number of spots where that coincidence was possible (*solid points on diagram*). With enough satellites, allowable points are spaced over a meter apart. Hence, by knowing the approximate position of the antenna (using the differential GPS technique), one can determine which of the points located by carrier tracking marks the correct location.

In practice, carrier tracking proves to be a rather delicate undertaking. The passage of the waves and the motion of the satellites need to be accounted for. Some uses of carrier tracking, such as the aircraft landing system developed by Bradford W. Parkinson and his colleagues at Stanford University, require special hardware to ensure the integrity of the navigation solutions. These efforts are made more difficult by a security feature introduced by the Department of Defense called anti-spoofing. Like the encoding of clock errors (called selective availability), anti-spoofing makes many marvelous high-precision civilian applications of the GPS harder and more costly to accomplish. —*T.A.H.*

thorized parties about exactly when the satellites had sent their signals. The GPS timing could be forced off slightly by altering the satellites' atomic clocks according to a specific code. Such "dithering" of the clocks appears to be what the Department of Defense has employed to keep the GPS secure, a procedure they term "selective availability." The modified signals allow all citizens to locate themselves reasonably well; navigational fixes will be off by no more than 100 meters. Military receivers that are equipped to interpret the classified code can readily work out a more refined position by removing the clock errors that have been added.

Civilian scientists and engineers interested in the GPS did not take long, however, to work out ways to get around the limitations of selective availability. Soon after the first group of these navigation satellites was launched, scientists had managed to find ways to reduce GPS errors—sometimes to as little as a few millimeters—achieving a level of accuracy that was many thousands of times finer than the system's military designers had thought possible. The first demonstration of such capability, by Charles C. Counselman III of the Massachusetts Institute of Technology and his colleagues, was performed rather unceremoniously in the parking lot of Haystack Observatory in Westford, Mass., during the fall of 1980.

To achieve a substantial improvement in accuracy, the scientists at Haystack and M.I.T. needed to correct the errors in the atomic clocks on the GPS satellites. The technique they employed was in fact quite simple: at a fixed point on the ground, they measured signals from several satellites. Knowing the exact location of the receiving antenna and the satellite positions, the scientists could then easily compare the pseudo-ranges (which they had measured) with the actual ranges (which they could calculate). The difference between the two numbers represented the error in the satellite clock, plus any inaccuracy in the clock used by the receiving equipment on the ground. The procedure of

examining several satellites simultaneously allowed the scientists to determine the clock error on the ground, and hence they could work out exactly how much each of the space-borne timekeepers was off.

The same method can be employed for circumventing selective availability today. The amount of clock dithering can be determined at a fixed ground station, and the corrections can be broadcast by radio. Mobile GPS apparatus operating nearby can use the information to calculate accurate locations. This scheme of "differential GPS" offers people outside the U.S. military the means to work out their whereabouts to within about a meter using surprisingly inexpensive equipment. (More specialized GPS receivers can achieve precision to about a centimeter.) There are currently a multitude of sources for differential GPS corrections. Many of them, curiously, are run by the U.S. government itself. The Federal Aviation Administration, for example, is starting to provide these services for aircraft. The U.S. Coast Guard, too, transmits corrections near major harbors. In addition, several commercial companies sell GPS corrections for most parts of the U.S. and for some other regions of the world as well.

THOMAS A. HERRING

**DEFORMATION OF THE EARTH'S CRUST in geologically active areas such as the Tien Shan of central Asia can be measured using precise GPS surveys of benchmarks. The GPS technique thus serves as a research tool to help monitor the accumulation of strain that can eventually cause devastating earthquakes.**

The widespread availability of differential GPS has sparked considerable debate as to why the U.S. military continues to spend money to encode the GPS during peacetime, forcing other branches of government to expend yet more resources to decode the errors and broadcast the results. Ironically, during two recent military operations, the Persian Gulf War and the occupation of Haiti, the Department of Defense turned off the security features of the GPS. They did so because there was not enough classified GPS equipment to go around, whereas civilian models were relatively easy to come by. (Many U.S. troops obtained this equipment, if in no other way, by telephoning home and purchasing GPS sets with their credit cards. The U.S. military had counted on their adversaries' lack of GPS-guided missiles and poor access to mail-order shopping.)

Moreover, the Russian government is now in the final stages of completing a satellite positioning system called GLONASS (for "Global Navigation Satellite System") that is largely similar to GPS. The Russian navigation system, however, does not encode its broadcasts, and thus anyone with the proper equipment can use it to full advantage. The existence of unencoded GLONASS, along with the widespread availability of GPS corrections, seems to negate any military advantage that might have accrued from purposeful degradation of the satellite clocks. A study recently conducted by the National Academy of Sciences has advised that selective availability is ineffective and should be discontinued. But so far the Department of Defense is still dithering—on the earth and in space.

### Where Next?

With each passing week, people seem to find clever new applications for satellite positioning. Meteorologists are measuring the delays in GPS signals caused by the atmosphere to aid in weather forecasting. Farmers are using this equipment to survey the condition of each square yard of their fields so that they can distribute fertilizer most effectively. And, reasonably enough, the GPS is more and more helping to guide ships, airliners, helicopters, satellites and even passenger cars. Experimental systems carried in backpacks may eventually lead blind people about. Indeed, the commercial applications now far outnumber the military uses of the system, and by the turn of the millennium the sale of GPS services should bring about $1 billion into the U.S. economy every year. With unprecedented speed, what was born as a military system has become a national economic resource. In this rapidly changing world, one must seriously wonder: Who should control the GPS?

## *The Author*

THOMAS A. HERRING is an associate professor in the department of earth, atmospheric and planetary sciences at the Massachusetts Institute of Technology. During his undergraduate studies at the University of Queensland, he spent his summers working underground in Australian mines and wrestling his way through the jungles of Papua New Guinea. These experiences convinced Herring that it was safer to pursue graduate studies in front of a computer at M.I.T., where he learned how to apply precise measurement systems to geophysical problems. In 1983 he received his Ph.D. and then joined the Harvard College Observatory. Herring remained there for six years before returning to M.I.T. After many years of respite from the rigors of fieldwork, Herring once again finds himself involved in field studies in such far-flung places as the central Asian republic of Kyrgyzstan and the Paraná River in Argentina.

## *Further Reading*

MEASUREMENT OF CRUSTAL DEFORMATION USING THE GLOBAL POSITIONING SYSTEM. Bradford H. Hager, Robert W. King and Mark H. Murray in *Annual Review of Earth and Planetary Sciences,* Vol. 19, pages 351–382; 1991.

THE NAVSTAR GLOBAL POSITIONING SYSTEM. Tom Logsdon. Van Nostrand Reinhold, 1992.

*GPS WORLD.* Bimonthly magazine published by Advanstar Communications, 859 Williamette Street, Eugene OR 97401.

International GPS Service for Geodynamics site on the World Wide Web at http://igscb.jpl.nasa.gov/

The University Navstar consortium site on the World Wide Web at http://www.unavco.ucar.edu/

# Seeing Underwater with Background Noise

*With a technique called acoustic-daylight imaging, sounds in the sea can "illuminate" submerged objects, thereby creating moving color pictures without sonar*

by Michael J. Buckingham, John R. Potter and Chad L. Epifanio

"Ping…ggg." The sound of a sonar transmission is familiar from classic films on submarine warfare, such as *Das Boot* and more recently *The Hunt for Red October.* An echo provides the submariner with the clue to a target's presence and position. Alternatively, one can passively listen for the sound generated by the target itself. In both techniques, however, the acoustic noise that permeates the oceans compromises the integrity of the signals. Breaking waves, passing ships, falling rain and even sea creatures such as snapping shrimp all contribute to this cacophony. It is only to be expected that sonar operators have traditionally regarded background noise as a nuisance and, accordingly, have directed great efforts to suppress the effects of ambient noise.

Yet that approach is gradually changing, as researchers have begun to recognize that the noise itself can be useful. Noise surrounds any object immersed in the ocean; the object, in turn, modifies this noise field in ways that depend on the object's shape, composition and position. Ambient noise has a familiar optical analogue: daylight in the atmosphere. We can see and photograph outdoor objects because they scatter, reflect and otherwise modify the light in the air. Likewise, noise that permeates the ocean acts as a kind of "acoustic daylight." Recent experiments have shown that we can indeed create images of underwater objects by using ambient noise as a source of illumination. Our results are sufficiently encouraging that we believe acoustic-daylight imaging should prove useful for a variety of purposes, from harbor security to underwater mine detection.

To be sure, at present the resulting pictures lack a certain aesthetic appeal. The image resolution is no match for that achieved with optical light. The acuity of human vision stems from the fact that the dilated pupil is 10,000 times the size of the wavelength of visible light, enabling the eye to "collect" a great number of light waves. Achieving a similar resolution with sound would demand an impractically large receiver 600 meters wide. But because seawater strongly absorbs light and all other forms of electromagnetic radiation, sound has become the favored—and in many cases, the only—means of acquiring information about the ocean depths.

Humanity's interest in sound in the ocean dates back to antiquity. Aristotle and Pliny the Younger wondered if fish could hear. Fishermen in ancient China located shoals of fish by using a bamboo stick as an underwater listening device, placing one end in the water. Leonardo da Vinci further developed the idea, noting in his studies of the properties of water that "if you cause your ship to stop, and place the head of a long tube in the water and place the outer extremity to your ear, you will hear ships at a great distance from you."

It was not until early in the 20th century, however, that inventors fashioned the first underwater sonic location systems, in order to counter the submarine threat during World War I. As rudimentary as those early devices were, they formed the basis of all subsequent sonar, the development of which accelerated rapidly during World War II. Current sonar systems, which have found widespread military, commercial and scientific application, have evolved to a high degree of sophistication. Still, they operate on much the same principles as their predecessors: they either actively transmit sounds or passively receive sounds produced by a target.

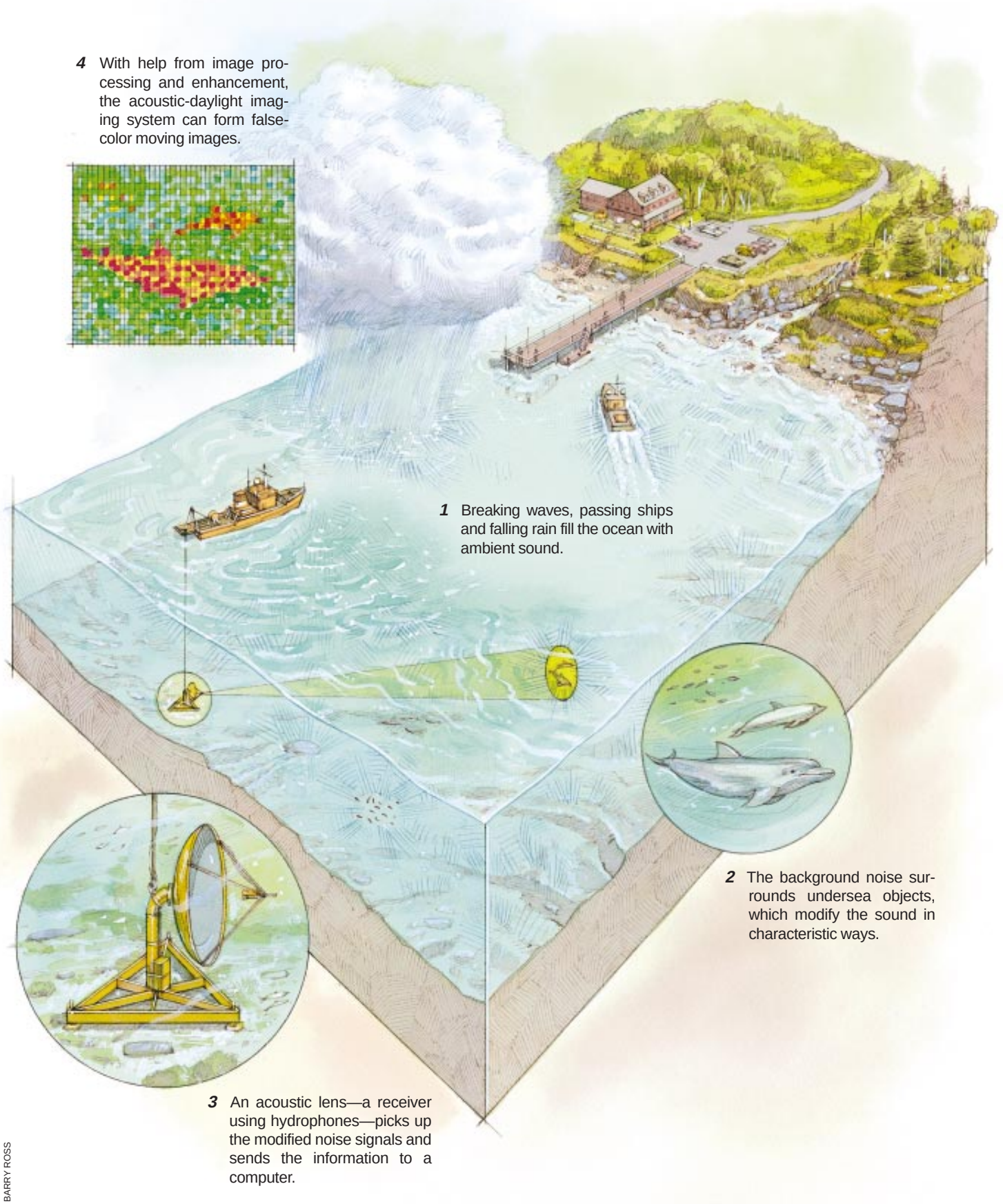
In view of the historical emphasis on active and passive techniques, it is not surprising that the notion that noise might provide an entirely new way of "seeing" in the ocean evolved only recently. In the mid-1980s one of us (Buckingham) recognized that visual imaging as performed by the eye is neither active nor passive. That is to say, the eye functions in a manner that differs fundamentally from the conventional ways of using acoustics in the ocean. Once this idea had registered, it became natural to speculate on the possibility of creating an underwater acoustic analogue of visual imaging. On a practical level, acoustic-daylight imaging would avoid the main drawbacks of conventional undersea detection techniques: echolocation unavoidably reveals the presence of the operator, and passive detection, though entirely covert, fails with quiet or silent targets.

### The First Experiment

In mid-1991 we conducted the first acoustic-daylight experiments in the Pacific Ocean off Scripps Pier at Scripps Institution of Oceanography in La Jolla, Calif. Working for his master of science degree at Scripps was a young navy lieutenant, Brodie Berkhout, who constructed and deployed the equipment. The main device was an acoustic receiver in the form of a simple parabolic reflector, 1.2 meters in diameter, with a single hydrophone (underwater microphone) at the focus. In effect, the reflector played the role of an acoustic lens.

The purpose of the experiment was to answer a simple question: Does the perceived noise level at the receiver change when an object is placed in its "beam," that is, its listening field? A rectangular plywood board, 0.9 by 0.77 meter and faced with neoprene rubber—a good reflector and scatterer of sound—served as the target. We found that for frequencies between five and 50 kilohertz (within the range produced by breaking waves, which are often the main source of ambient noise in the ocean),

# Acoustic Daylight in Action




BARRY ROSS


MICHAEL J. BUCKINGHAM

**KILLER WHALE at Sea World park in San Diego served as a moving target for ADONIS, the first acoustic-daylight imaging system.**

the noise intensity nominally doubled when the target was placed in the listening field of the reflector. This result persisted when we moved the target from seven to 12 meters from the receiver. Moreover, the target strongly reflected some frequencies and absorbed others, a phenomenon that can be interpreted as acoustic "color." This development suggested that we could translate the reflected acoustic signature into optical hues to create acoustic-daylight images in false color.

Spurred on by this success, we began thinking about the next stage of development. The parabolic reflector with a hydrophone at its focus "looks" only in a single direction, corresponding to just one pixel of an image. To create a more complete picture, more pixels are necessary, which means more receiver "beams" are needed (rather like the compound eye of a fly). The noise in each receiver beam could then be converted to a certain level of brilliance in a pixel on a video monitor, with the intensity of the noise governing the degree of the brightness. As in a newspaper photograph, the contrast between pixels would enable the eye to interpret the result as a more or less granular pictorial image.

With the success of the initial test, we became convinced of the feasibility of achieving genuine acoustic-daylight images that would contain 100 or more pixels. In mid-1992 we began designing a new acoustic lens, which came to be known as ADONIS, for acoustic-daylight, ambient-noise imaging system. Working in conjunction with EDO Acoustics in Salt Lake City, which produced an elliptical array of 128 hydrophones for ADONIS, we constructed a spherical reflector three meters in diameter and placed the hydrophones at the focus of the dish. This system formed a total field of view of approximately six degrees (horizontal) by five degrees (vertical), which is about one tenth the angular view afforded by a typical camera.

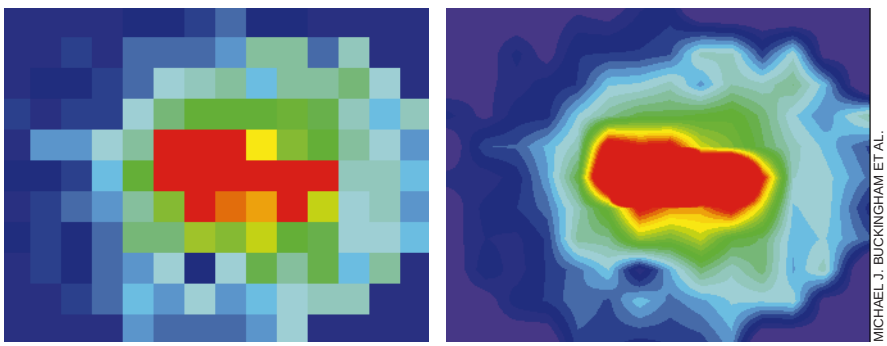
We lowered ADONIS, looking rather like a satellite dish, onto the seabed for the first time in August 1994. ADONIS was deployed from one of Scripps's research platforms, *R/P ORB,* moored off Point Loma in southern California. Square panels (one meter per side) of aluminum sheeting faced with neoprene rubber formed the targets to be imaged. The panels were mounted in various configurations on a square tic-tac-toe-type frame set on the seabed. Roiled-up sediment in the busy harbor made visibility through the water extremely poor during most of the experiment. On one occasion the turbidity was so bad that Hélène Vervoort, one of our divers, collided with the target frame.

An electronics package housed in a sealed pressure canister rested alongside the mast supporting the spherical dish. Among other processing tasks, the electronic equipment, designed by our colleague Grant B. Deane, would convert the ambient noise data acquired by ADONIS into digital form. The data would then be transmitted to the surface and rendered into real-time, false-color images on the screen of a Macintosh desktop computer. An immense amount of time and effort hung in the balance as ADONIS was lowered into the sea for its first deployment.

### To See or Not To See?

The air of hushed expectancy that hung over our group as ADONIS disappeared below the ocean surface was soon dispelled—not, however, because of an initial, resounding success. Almost immediately, the gauges monitoring several onboard power supplies surged—a strong indication that seawater was flooding into the electronics canister. Sure enough, when ADONIS was hauled up and the canister opened, saltwater gushed out. As a reflex reaction, we removed the delicate circuit boards and soaked them in deionized water, although nobody really believed they could be salvaged. But with help from a number of quarters, we flushed the boards with alcohol, tested all the electronic components of the complex 128-channel system, replaced them where necessary and sealed the leak in the canister. Twenty-four hours later ADONIS was again lowered into the water.

This time the tension on *ORB* was tangible as the divers made last-minute checks on the equipment. When the data started to flow, the laboratory became quiet. We had set three panels in the frame to form a simple horizontal target, one meter high by three meters wide, at a distance of 18 meters from



MICHAEL J. BUCKINGHAM ET AL.

**BAR TARGET (*left*) was imaged by ADONIS as a vaguely elongated form and artificially colored red (*upper left*). Each "pixel" represents the signals from a hydrophone. Computer processing enhanced the image (*upper right*).**

ADONIS. As we gathered around the screen, we realized that a faint rectangular shape was visible, almost filling the elliptical image space. We were watching the first acoustic-daylight picture.

Within minutes our confidence in the imaging system had soared. Divers had placed a sound source in the center of the target to help us align ADONIS with the target frame. But the source proved unnecessary: we could see where the targets were just from the ambient noise. We then extended the space between ADONIS and the target from 18 to 38 meters, as far as we could go without interfering with shipping traffic. At the greater range we expected perhaps a slight degradation in performance, but astonishingly the target became far clearer. Of course, the image was also smaller than it had been previously, but as a result, the surrounding ocean formed a nice, contrasting background that made the rectangular target stand out dramatically. As these raw images continued to appear on the screen, refreshed 30 times a second, we knew that acoustic-daylight imaging worked.

There was still much to be done during this deployment, however. We wanted to know if ADONIS could detect moving objects. A hydraulic motor mounted within the mast supporting the dish could slowly rotate the spherical receiver in azimuth, taking 12 minutes or so to complete a full 360-degree sweep. As the dish panned around, we watched the target appear on one side of the screen, creep to the center and finally drop off the far side. There was no doubt that we could create moving images.

One more test, the most demanding of all, remained. Divers replaced the bar-shaped target with four panels in the frame, forming a cross with vertical and horizontal arms and a one-meter-square hole in the center. Resolving the hole was the challenge: at a range of 38 meters, the size of the hole would be close to the resolution limit of ADONIS.

The first raw images of the cruciform target were indistinct. We could see the shape of the cross, but the appearance of the central hole fluctuated from instant to instant. Since then, we have reexamined the data and applied some computer processing. It turns out that the power spectrum of the noise—the intensity of the sound at different frequencies—serves a discriminatory function. It is essentially the acoustic version of color. By using the power spectrum, the four empty corners and the hole in the cruciform target could easily be identified and the edges of the panels located. The panels in the target frame showed a distinctly different "color" from the empty regions, including the central hole. It was as if the frame looked "red," and the hole appeared "blue." Currently we are exploring this technique as a means of enhancing acoustic-daylight images.

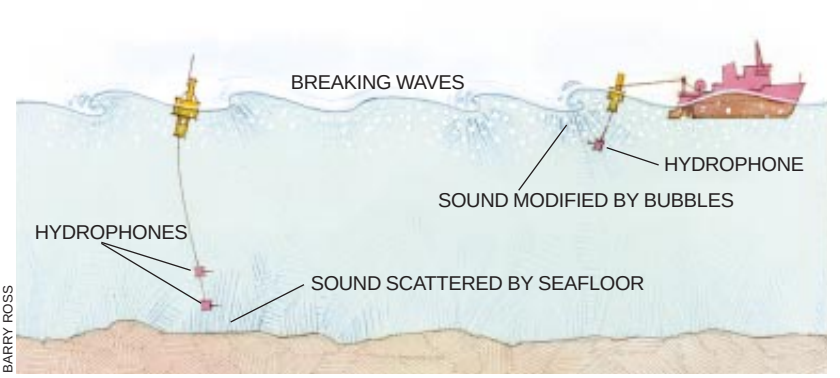
## Sounding Out New Uses for Noise

Acoustic-daylight imaging is just one form of remote-sensing technology that relies on the background noise in the seas. Oceanographers have recently demonstrated other examples of similar techniques. One is to use ambient noise to determine the acoustic properties of the seabed and hence to determine its composition to some extent. In the shallow waters over the continental shelves, where the depth is less than about 200 meters, the noise reflects off the seafloor. The manner in which the sound bounces off indicates the speed with which vibrations move in the floor. That, in turn, reveals the composition of the bottom: sound travels at different speeds through bedrock than it does through sand, for instance.

To carry out such measurements, one can deploy a fleet of hydrophone-dangling buoys to map the seabed using ambient noise. The hope is that this technique will offer a cost-effective alternative to conventional methods, such as the often slow and laborious practice of bouncing sonar signals off the sea bottom.

Background sounds may also prove beneficial in the study of processes occurring at the sea surface. In particular, they can reveal the amount of atmospheric gas the oceans are absorbing. Crucial for models of global warming and the greenhouse effect, the extent of gas exchange has been difficult to quantify. Ambient noise may help, because the phenomenon mostly responsible for the sound also happens to govern the transfer of gas from the air to water—namely, wave breaking. In driving air into the water, the process creates a layer of bubbles immediately below the surface. These bubbles modify the sound of the breaking waves in a characteristic way, leaving an acoustic signature for hydrophones below the bubbles to detect.

From such a simple acoustic measurement, it may be possible to infer the amount of air in the bubble layer and the depth to which the bubbles extend. Both quantities are related to the amount of gas entering the ocean. Some preliminary testing suggests the idea is feasible; major experiments are currently under way. —*M.J.B.*

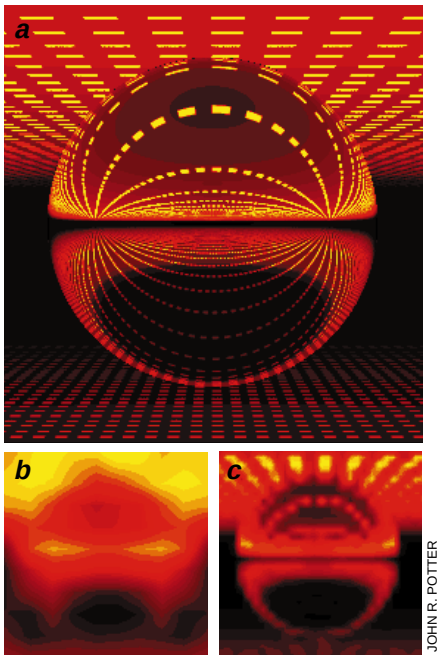



**AMBIENT NOISE could also measure the acoustic properties of the ocean bottom (*left*) and the amount of gas absorbed by the sea (*right*).**

### Imaging at Sea World

Static targets served us well in demonstrating that acoustic-daylight imaging is a workable technique. Inspired by our results, we were anxious to try a more difficult target: killer whales (*Orcinus orca*). Through the good offices of Ann Bowles, a research biologist at Hubbs Sea World Research Institute in San Diego, we were invited to deploy ADONIS in the outdoor killer-whale tank at Sea World. We could try to image highly mobile marine mammals while Bowles conducted behavioral studies on the response of the animals to a strange object in "their" tank; the whales, it seems, feel that anything placed in the tank, by definition, belongs to them.

In February 1995, working between the killer whales' public performances,




JOHN R. POTTER

**SIMULATED IMAGES of a steel sphere hint at the promise of acoustic daylight. The ambient noise comes from breaking waves, represented by yellow dashes. A system with 90,000 pixels would yield the highest resolution (*a*) but is probably not practical. Technology now uses about 100 pixels (*b*), and systems using 900 pixels are planned (*c*).**

we set up ADONIS in one corner of the tank in rather unpleasant weather conditions. Rain lashed down most of the time; to protect our computers and recording equipment, we rigged up makeshift tarpaulins, but even so water seeped everywhere.

Meanwhile, as we set up the system, the killer whales swam freely in the tank, taking as much interest in us as we did in them. Cautious at first, they quickly grew accustomed to the large reflecting dish. The whales became curious on finding that because of the focusing effect of the dish, sound reflected intensely back to them when they "pinged" in front of it. A young male, Splash, grew more adventurous, taking one of the oil-filled electronics cables (crunchy on the outside and chewy on the inside) into his mouth to try some exploratory mastication. Another animal swam fast toward the dish and breached over the top—at this point we felt that something had to be done. The trainers moved the whales to another tank, where they could play with their own toys until we were ready for them.

After one false start (all the electronics boards in the underwater housing had shaken free of their connectors during transportation), we switched on the equipment again, and data started flowing. We were not sure what to expect. Pumps and other machinery bring the noise in the Sea World tanks to quite high levels, comparable to those in the ocean. Despite some minor damage that the electronics boards had sustained when they were flooded by seawater, signals from all but two of the 128 channels were received and displayed as real-time moving images.

As we watched the raw data (that is, with no image enhancement) on the screen, a shadowy form suddenly appeared and remained in sight for several seconds. At the same time, we could see (with our own eyes) one of the whales move into the field of view of ADONIS, where it stayed as it swam directly away from the dish. Hydrophone monitors and the trainers confirmed that the whales themselves were not transmitting sound, indicating that the images we saw were a direct result of acoustic daylight. We still have to examine the images of the killer whales carefully and correlate them with the video recordings that were made simultaneously to confirm whether we actually imaged the whales. But the preliminary observations and the *ORB* experiment off Point Loma support the analogy between conventional photography with daylight and underwater imaging with ambient sound.

The results immediately suggest several potential applications. One is the detection of underwater mines, some of which can be rigged to detonate on receiving a sonar signal. An acoustic-daylight system might be able to locate these devices without triggering them. Imaging with ambient noise could provide vision for autonomous underwater vehicles, enabling them to steer around obstacles without help from a human operator on a surface ship and to monitor the structural integrity of oil rigs and other large maritime platforms. The inherently covert nature of acoustic-daylight imaging also makes it suitable for monitoring harbors—just as video cameras keep vigilance in shopping malls—and for counting marine mammals, because there would be no sonic interference with the animals themselves. (That, in turn, raises the question of whether marine mammals themselves use acoustic daylight to acquire information.)

Conceivably, we can take acoustic-daylight imaging further, for it is still a nascent concept. In recent tests, ADONIS successfully imaged plastic floats, titanium spheres and polyvinyl chloride oil drums containing wet sand and foam. Preliminary analysis indicates that the barrels can be seen even when they are on the seafloor. We have reached a stage rather like the earliest days of television: what is important is not the quality of the images but the fact that there are images at all. In the months ahead, we plan to replace the spherical reflector with a phased array containing as many as 1,000 hydrophones. At the same time, we shall be developing dedicated algorithms to provide image enhancement and automatic image recognition. These efforts will, we hope, improve the quality of acoustic-daylight images significantly and perhaps make the successors to ADONIS the underwater video cameras of the future.

## *The Authors*

MICHAEL J. BUCKINGHAM, JOHN R. POTTER and CHAD L. EPIFANIO developed acoustic-daylight imaging at the Scripps Institution of Oceanography in La Jolla, Calif. Buckingham is professor of ocean acoustics there and holds a visiting professorship at the University of Southampton in England. He received his Ph.D. in physics from the University of Reading and has written and edited numerous articles and books on acoustics. An itinerant yachtsman, Potter sailed across the Pacific Ocean last fall to direct the Acoustic Research Laboratory of the National University of Singapore. After spending four summers on the Antarctic peninsula, he received his Ph.D. from the Council for National Academic Awards and the University of Cambridge. Epifanio is close to completing his Ph.D. on acoustic-daylight research at Scripps. He received his B.S. in electrical engineering from Bucknell University in 1991. The authors are grateful to Sea World in San Diego, to Hubbs Sea World Research Institute and to the U.S. Office of Naval Research for their research support.

## *Further Reading*

IMAGING THE OCEAN WITH AMBIENT NOISE. Michael J. Buckingham, Broderick V. Berkhout and Stewart A. L. Glegg in *Nature,* Vol. 356, pages 327–329; March 26, 1992.

THEORY OF ACOUSTIC IMAGING IN THE OCEAN WITH AMBIENT NOISE. Michael J. Buckingham in *Journal of Computational Acoustics,* Vol. 1, No. 1, pages 117–140; March 1993.

ACOUSTIC IMAGING USING AMBIENT NOISE: SOME THEORY AND SIMULATION RESULTS. John R. Potter in *Journal of the Acoustical Society of America,* Vol. 95, No. 1, pages 21–33; January 1994.

ACOUSTIC DAYLIGHT IMAGING: VISION IN THE OCEAN. Michael J. Buckingham and John R. Potter in *GSA Today,* Vol. 4, No. 4, pages 97–102; April 1994.

# Telomeres, Telomerase and Cancer

*An unusual enzyme called telomerase acts on parts of chromosomes known as telomeres. The enzyme has recently been found in many human tumors and is being eyed as a new target for cancer therapy*

by Carol W. Greider and Elizabeth H. Blackburn

Often in nature things are not what they seem. A rock on the seafloor may be a poisonous fish; a beautiful flower in a garden may be a carnivorous insect lying in wait for prey. This misleading appearance extends to certain components of cells, including chromosomes—the strings of linear DNA that contain the genes. At one time, the DNA at the ends of chromosomes seemed to be static. Yet in most organisms that have been studied, the tips, called telomeres, are actually ever changing; they shorten and lengthen repeatedly.

During the past 15 years, investigation of this unexpected flux has produced a number of surprising discoveries. In particular, it has led to identification of an extraordinary enzyme named telomerase that acts on telomeres and is thought to be required for the maintenance of many human cancers. This last finding has sparked much speculation that drugs able to inhibit the enzyme might combat a wide array of malignancies. The research also opens the possibility that changes in telomere length over time may sometimes play a role in the aging of human cells.

Modern interest in telomeres and telomerase has its roots in experiments carried out in the 1930s by two remarkable geneticists: Barbara McClintock, then at the University of Missouri at Columbia, and Hermann J. Muller, then at the University of Edinburgh. Working separately and with different organisms, both investigators realized that chromosomes bore a special component at their ends that provided stability. Muller coined the term "telomere," from the Greek for "end" (*telos*) and "part" (*meros*). McClintock noted that without these end caps, chromosomes stick to one another, undergo structural changes and misbehave in other ways. These activities threaten the survival and faithful replication of chromosomes and, consequently, of the cells housing them.

It was not until the 1970s, however, that the precise makeup of the telomere was determined. In 1978 one of us (Blackburn), then working with Joseph G. Gall of Yale University, found that the telomeres in *Tetrahymena,* a ciliated, single-cell pond dweller, contained an extremely short, simple sequence of nucleotides—*TTGGGG*—repeated over and over. (Nucleotides are the building blocks of DNA; they are generally denoted as single letters representing the chemical bases that distinguish one nucleotide from another. The base in *T* nucleotides is thymine; that in *G* nucleotides is guanine.)

Since then, scientists have characterized the telomeres in a host of creatures, including animals, plants and microorganisms. As is true of *Tetrahymena,* virtually all telomeres—including those of mice, humans and other vertebrates—contain repeated short subunits often rich in *T* and *G* nucleotides [see "The Human Telomere," by Robert K. Moyzis; SCIENTIFIC AMERICAN, August 1991]. For instance, human and mouse telomeres feature the sequence *TTAGGG;* those of roundworms feature *TTAGGC*. (*A* stands for adenine, *C* for cytosine.)

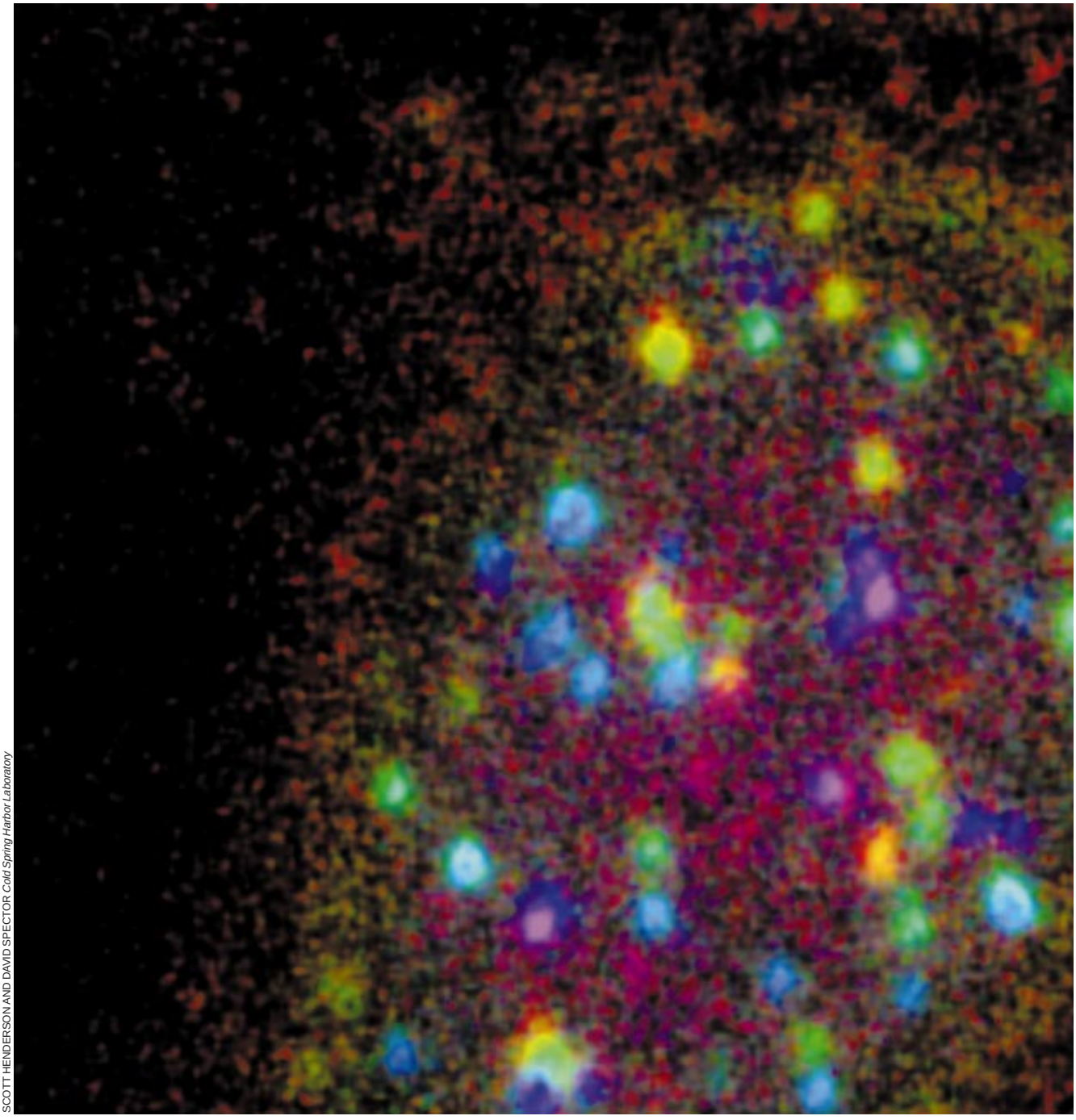
### In Search of Telomerase

The telomerase enzyme that is the object of so much attention today was found when comparisons of telomere length suggested such an enzyme could resolve a long-standing puzzle in biology. By the early 1980s investigations had revealed that, for some reason, the number of repeated subunits in telomeres differs between organisms and even between different cells in the same organism. Moreover, the number can fluctuate in a given cell over time. (Every species, however, has a characteristic average. In *Tetrahymena,* the average telomere has 70 repeats; in humans, 2,000.) The observed heterogeneity led Blackburn, who had moved to the University of California at Berkeley, Jack W. Szostak of Harvard University and Janis Shampay of Berkeley to propose a new solution to what has been called the end-replication problem.

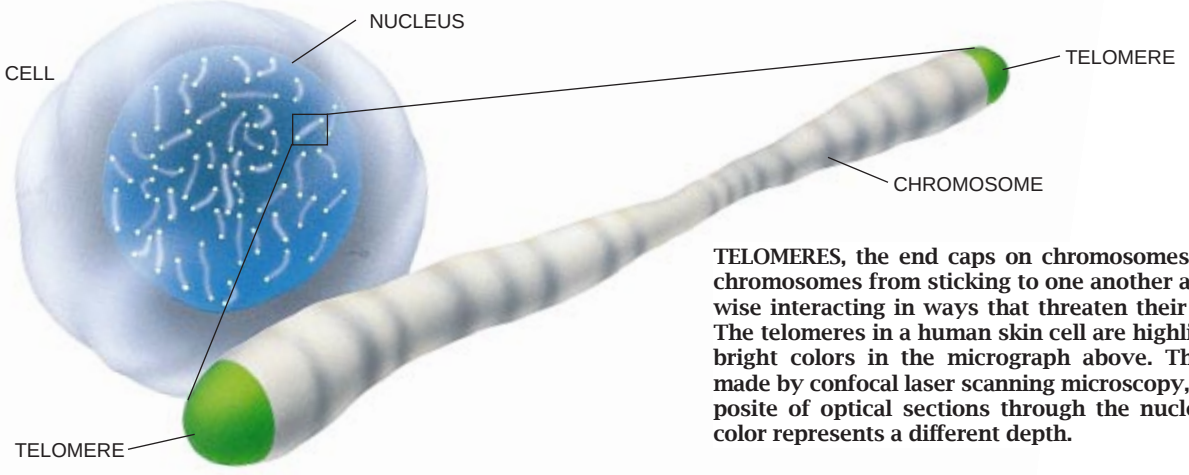
The problem has to do with the fact that cells must replicate their genes accurately whenever they divide, so that each so-called daughter cell receives a complete set. Without a full set of genes, a daughter cell may malfunction and die. (Genes are those sequences of nucleotides that give rise to proteins and RNA, the molecules that carry out most cellular functions. The genes in a chromosome are scattered throughout the large expanse of DNA that is bounded by the chromosome's two telomeres.)

In 1972 James D. Watson, working at both Harvard and Cold Spring Harbor Laboratory, noted that DNA polymerases, the enzymes that replicate DNA, could not copy linear chromosomes all the way to the tip. Hence, the replication machinery had to leave a small region at the end (a piece of the telomere) uncopied [*see box on page 94*]. In theory, if cells had no way to compensate for this quirk, chromosomes would shorten with each round of cell division. Eventually, the erosion would eliminate the telomeres and critical genes in some generation of the cells. These cells would thus perish, spelling the end of that cellular lineage. Clearly, all single-cell species subject to such shortening manage to counteract it, or they would have vanished long ago. So do germ-line cells (such as the precursors of sperm and eggs), which perpetuate the species in multicellular organisms. But how do such cells protect their telomeres?

For Blackburn, Szostak and Shampay, the observed fluctuations in telomere



SCOTT HENDERSON AND DAVID SPECTOR *Cold Spring Harbor Laboratory*




TOMO NARASHIMA

**TELOMERES, the end caps on chromosomes, prevent chromosomes from sticking to one another and otherwise interacting in ways that threaten their stability. The telomeres in a human skin cell are highlighted by bright colors in the micrograph above. The image, made by confocal laser scanning microscopy, is a composite of optical sections through the nucleus; each color represents a different depth.**

length were a sign that cells attempt to maintain telomeres at a roughly constant size. Yes, telomeres do shorten during cell division, but they are also lengthened by the attachment of newly synthesized telomeric subunits. The researchers suspected that the source of these additional repeats was some undiscovered enzyme capable of a trick that standard DNA polymerases could not perform.

When cells replicate their chromosomes, which consist of two strands of DNA twisted around each other, they begin by separating the double helix. The polymerases use each of these "parent" strands as a template for constructing a new partner. The special enzyme the workers envisioned would be able to build extensions to single strands of DNA from scratch, without benefit of an existing DNA template.

In 1984 the two of us, working in Blackburn's laboratory at Berkeley, set out to discover whether this putative telomere-lengthening enzyme—telomerase—actually existed. To our delight, we found it did. When we mixed synthetic telomeres with extracts of *Tetrahymena* cells, the telomeres gained added subunits, just as would be expected if the proposed enzyme were present.

Within the next several years we and our colleagues learned much about how telomerase works. Like all polymerases and virtually all enzymes, it consists mainly of protein, and it requires that protein to function. Uniquely, though, it also includes a single molecule of RNA (close cousin to DNA) that contains the critical nucleotide template for building telomeric subunits. Telomerase places the tip of one strand of DNA on the RNA, positioning itself so that the template lies adjacent to that tip. Then the enzyme adds one DNA nucleotide at a time until a full telomeric subunit is formed. When the subunit is complete, telomerase can attach another by sliding to the new end of the chromosome and repeating the synthetic process.

### Telomerase and Human Aging

In 1988 Greider left Berkeley for Cold Spring Harbor Laboratory, and later our groups and others found telomerase in ciliates distinct from *Tetrahymena,* as well as in yeast, frogs and mice. In 1989 Gregg B. Morin of Yale also discovered it, for the first time, in a human cancer cell line—that is, in malignant cells maintained for generations in culture dishes. Today it is evident that telomerase is synthesized by nearly all organisms with nucleated cells. The precise makeup of the enzyme can differ from species to species, but each version possesses a species-specific RNA template for building telomeric repeats.

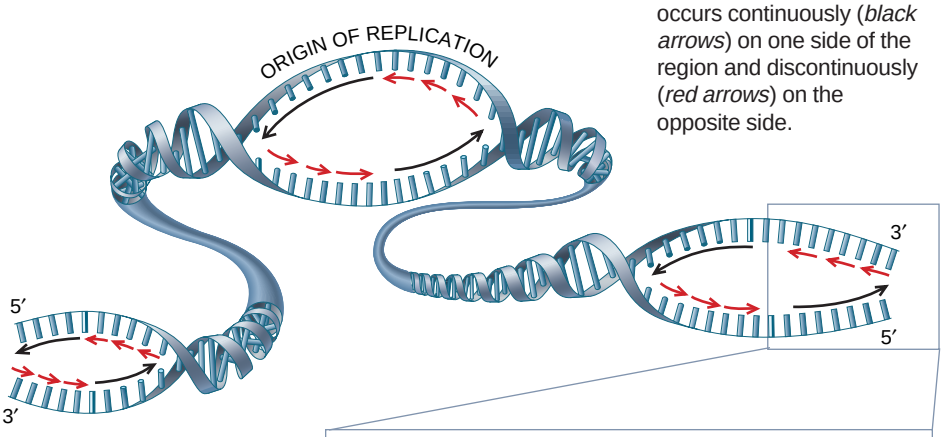
The importance of telomerase in many single-cell organisms is now indisputable. Such organisms are immortal in that, barring accidents or geneticists meddling in their lives, they can divide indefinitely. As Guo-Liang Yu in Blackburn's research group demonstrated in 1990, *Tetrahymena* needs telomerase in order to retain this immortality. When the enzyme is altered, telomeres shrink and cells die. Blackburn's team and others have similarly demonstrated in yeasts that cells lacking telomerase undergo telomere shortening and perish. But what role does telomerase play in the human body, which consists of a myriad of cell types and is considerably more complex than *Tetrahymena* or yeast?

Surprisingly, many human cells lack telomerase. Greider and others made this discovery in the late 1980s, when

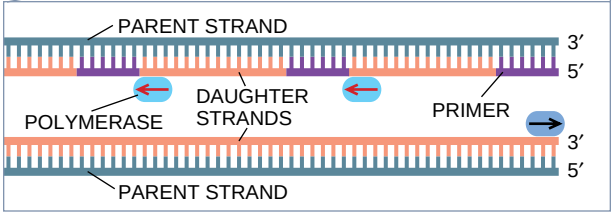
## The End-Replication Problem

The so-called end-replication problem arises from the inability of standard mechanisms for replicating chromosomes to do a complete job. When enzymes known as DNA polymerases copy the two original, or "parent," DNA strands in a chromosome, they leave each new "daughter" strand shortened at one tip (at the end traditionally labeled 5′). If cells did not compensate for this flaw in the replication mechanism, chromosomes would inexorably shrink. The existence of telomerase, the enzyme that elongates telomeres, was proposed in the early 1980s as a solution to the end-replication problem.

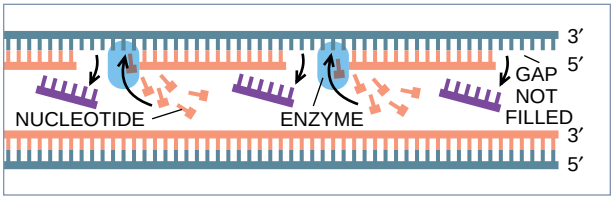
**1** Copying of chromosomal DNA, shown highly schematically, proceeds from regions known as origins of replication, at which the parent strands separate. Synthesis of each daughter strand occurs continuously (*black arrows*) on one side of the region and discontinuously (*red arrows*) on the opposite side.



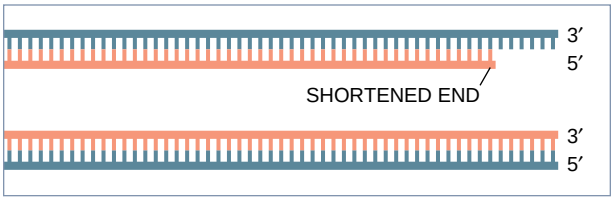

**2** Separated parent strands serve as templates on which polymerases synthesize daughter strands. The parts constructed as fragments extend from bits of RNA called primers.




**3** Other enzymes remove primers and fill the gaps between adjacent fragments.

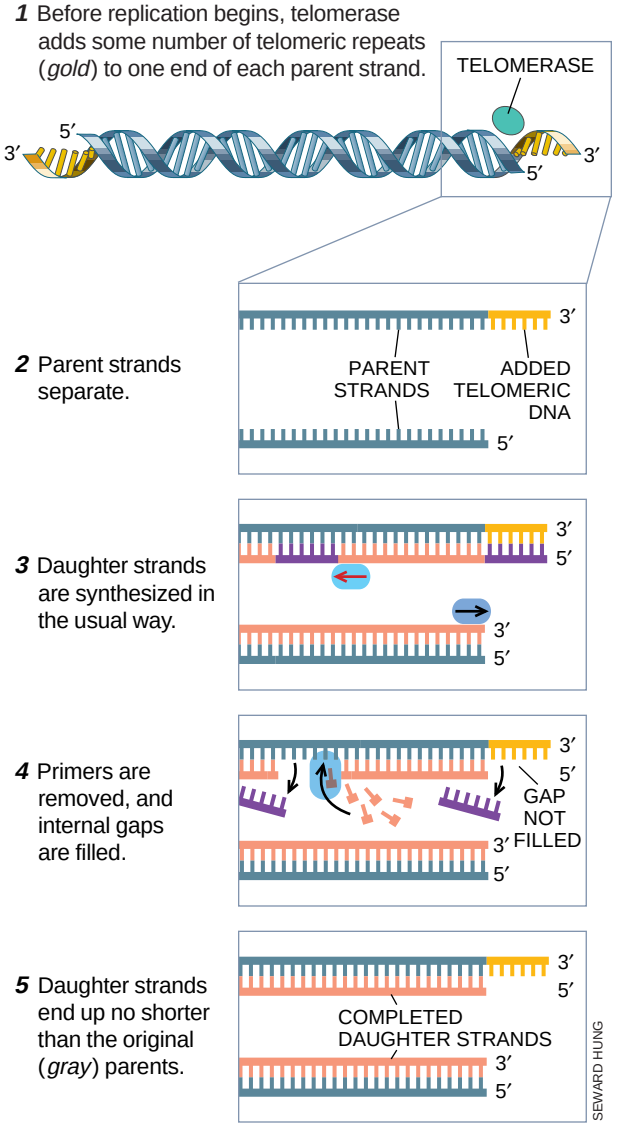


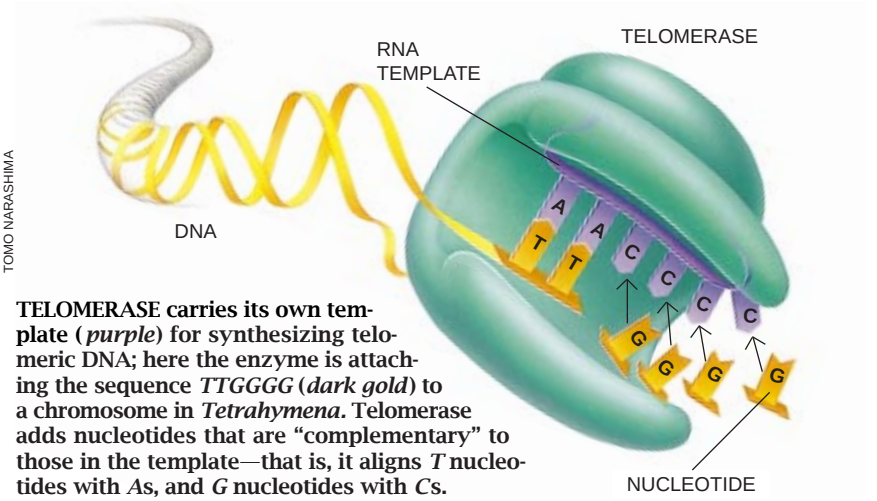

**4** But the enzymes cannot fill the gap remaining at one end (the 5′ tip) of each



## How Telomerase Solves the Problem

One scheme for how telomerase solves the end-replication problem proposes that the enzyme adds DNA to chromosomes before replication begins. The added DNA consists of one or more telomeric subunits: the short sequences of nucleotides that are repeated over and over in telomeres. The addition ensures that a daughter strand will be at least as long as its parent.




**TELOMERASE carries its own template (*purple*) for synthesizing telomeric DNA; here the enzyme is attaching the sequence *TTGGGG* (*dark gold*) to a chromosome in *Tetrahymena*. Telomerase adds nucleotides that are "complementary" to those in the template—that is, it aligns *T* nucleotides with *A*s, and *G* nucleotides with *C*s.**

they pulled together the threads of research that investigators in Philadelphia had initiated more than 25 years earlier. Before the 1960s, human cells that replicated in the body were thought to be capable of dividing endlessly. But then Leonard Hayflick and his co-workers at the Wistar Institute demonstrated unequivocally that this notion was incorrect. Today it is known that somatic cells (those not part of the germ line) derived from human newborns will usually divide 80 to 90 times in culture, whereas those from a 70-year-old are likely to divide only 20 to 30 times. When human cells that are normally capable of dividing stop reproducing—or, in Hayflick's words, become "senescent"—they look different and function less efficiently than they did in youth, and after a while they die.

In the 1970s a Soviet scientist named A. M. Olovnikov linked this programmed cessation of cell division to the end-replication problem. He proposed that human somatic cells might not correct the chromosomal shortening that occurs when cells replicate their DNA. Perhaps division ceased when cells discerned that their chromosomes had become too short.

We were unaware of Olovnikov's ideas until 1988, when Calvin B. Harley, then at McMaster University, brought them to Greider's attention. Intrigued, Greider, Harley and their collaborators decided to see if chromosomes do get shorter in human cells over time.

Sure enough, most normal somatic cells they examined lost segments of their telomeres as they divided in culture, a sign that telomerase was not active. Similarly, they and Nicholas D. Hastie's group at the Medical Research Council (MRC) in Edinburgh found that telomeres in some normal human tissues shrink as people age. (Reassuringly, Howard J. Cooke, also at the MRC in Edinburgh, had shown that telomeres are kept intact in the germ line.) These results indicated that human cells might "count" divisions by tracking the number of telomeric repeats they lose, and they might stop dividing when telomeres decline to some critical length. But definitive proof for this possibility has not yet been obtained.

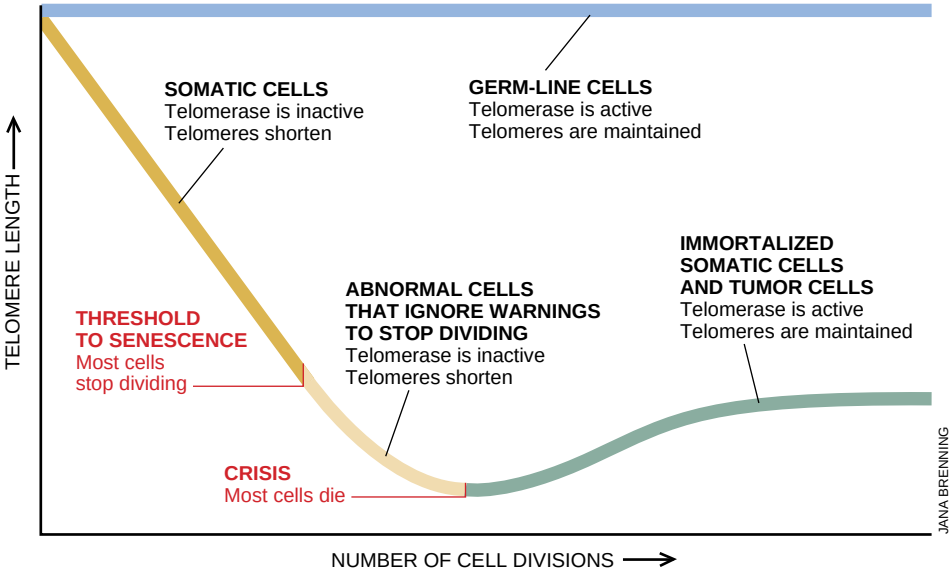
Could the reduction of telomeres and of proliferative capacity over time be a cause of human aging? It is probably not the main cause. After all, cells can usually divide more times than is required in a human life span. Nevertheless, the functioning of the older body may at times be compromised by the senescence of a subset of cells. For instance, local wound healing could be impaired by a reduction in the number of cells available to build new skin at a site of injury, and a reduction in the number of certain white blood cells could contribute to age-related declines in immunity. Further, it is known that atherosclerosis typically develops where blood vessel walls have been damaged. It is conceivable that cells at repeatedly injured sites could finally "use up" their replicative capacity, so that the vessels ultimately fail to replace lost cells. Then damage would persist, and atherosclerosis would set in.

### The Cancer Connection

Some investigators suspect that the loss of proliferative capacity observed in human cells lacking telomerase may have evolved not to make us decrepit but to help us avoid cancer. Cancers arise when a cell acquires multiple genetic mutations that together cause the cell to escape from normal controls on replication and migration. As the cell and its offspring multiply uncontrollably, they can invade and damage nearby tissue. Some may also break away and travel to parts of the

body where they do not belong, establishing new malignancies (metastases) at distant sites. In theory, a lack of telomerase would retard the growth of tumors by causing continually dividing cells to lose their telomeres and to succumb before they did much damage. If cancer cells made telomerase, they would retain their telomeres and would potentially survive indefinitely.

The notion that telomerase might be important to the maintenance of human cancers was discussed as early as 1990. But the evidence did not become compelling until recently. In 1994 Christopher M. Counter, Silvia Bacchetti, Harley and their colleagues at McMaster showed that telomerase was active not only in cancer-cell lines maintained in the laboratory but in ovarian tumors in the human body. Later that year groups led by Harley, who had moved to Geron Corporation in Menlo Park, Calif., and by Jerry W. Shay of the University of Texas Southwestern Medical Center at Dallas detected telomerase in 90 of 101 human tumor samples (representing 12 tumor types) and in none of 50 samples of normal somatic tissue (representing four tissue types).




Even before such evidence was obtained, however, researchers had begun exploring some of the details of how telomerase might contribute to cancer. That work suggests telomerase probably becomes active after a cell has already lost its brakes on proliferation.

The first clue was an initially mystifying discovery made independently by Titia de Lange, now at the Rockefeller University, and by Hastie's group. In 1990 these investigators reported that telomeres in human tumors were shorter than telomeres in the normal surrounding tissue—sometimes dramatically so.

Studies by Greider's, Bacchetti's and Harley's laboratories explained why the telomeres were so small. The teams had induced normal cells from humans to make a viral protein causing cells to ignore the alarm signals that usually warn them to stop dividing. The treated cells continued to proliferate long after they would normally enter senescence. In most of the cells, telomeres shortened drastically, and no telomerase was detected; eventually death ensued. Some cells, however, persisted after their siblings died and became immortal. In these immortal survivors, telomeres were maintained at a strikingly short length, and telomerase was present.

These outcomes imply that telomeres in cancer cells are small because cells synthesize telomerase only after they have already begun to replicate uncontrollably; by then, the cells have presumably lost a substantial number of telomeric subunits. When the enzyme is finally activated, it stabilizes the severely clipped telomeres, allowing overly prolific cells to become immortal.

These findings and others have led to an attractive but still hypothetical model for the normal and malignant activation of telomerase by the human body. According to this model, telomerase is made routinely by cells of the germ line in the developing embryo. Once the body is fully formed, however, telomerase is repressed in many somatic cells, and telomeres shorten as such cells reproduce. When telomeres decline to a threshold level, a signal is emitted that prevents the cells from dividing further.

If, however, cancer-promoting genetic mutations block issuance of such safety signals or allow cells to ignore them, cells will bypass normal senescence and continue to divide. They will also presumably continue to lose telomeric sequences and to undergo chromosomal alterations that allow further, possibly carcinogenic mutations to arise. When telomeres are completely or almost completely lost, cells may reach a point at which they crash and die.

But if the genetic derangements of the pre-crisis period lead to the manufacture of telomerase, cells will not completely lose their telomeres. Instead the shortened telomeres will be rescued and maintained. In this way, the genetically disturbed cells will gain the immortality characteristic of cancer.

This scenario has generally been borne out by the evidence, although, once again, things may not be entirely as they seem. Some advanced tumors lack telomerase, and some somatic cells—notably the white blood cells known as macrophages and lymphocytes—have recently been found to make the enzyme. Nevertheless, on balance, the collected evidence suggests that many tumor cells require telomerase in order to divide indefinitely.

### Prospects for Cancer Therapy

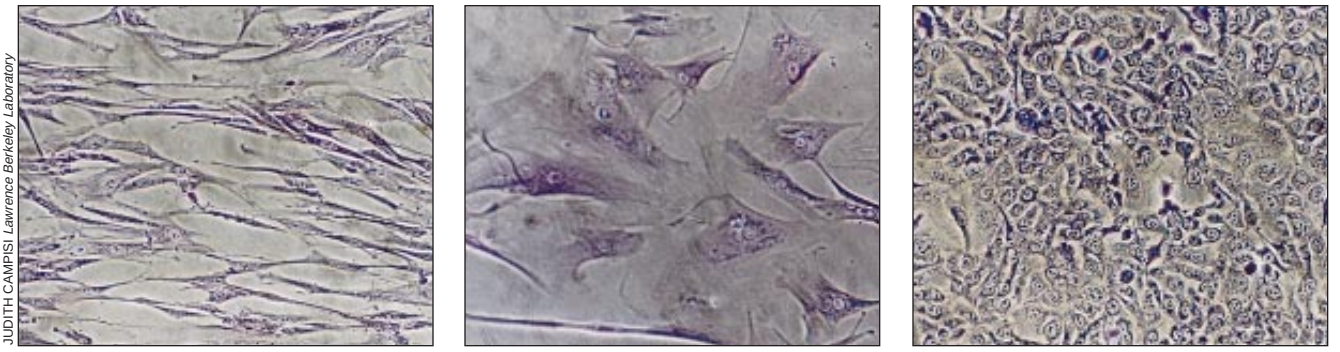
The presence of telomerase in various human cancers and its absence in many normal cells mean the enzyme might serve as a good target for anticancer drugs. Agents able to hobble telomerase might kill tumor cells (by allowing telomeres to shrink and disappear) without disrupting the functioning of many normal cells. In contrast, most existing anticancer therapies disturb normal cells as well as malignant ones and so are often quite toxic. Further, because telomerase occurs in numerous cancers, such agents might work against a broad array of tumors.

These exciting possibilities are now being actively explored by pharmaceutical and biotechnology companies. Nevertheless, a number of questions must be answered. For instance, researchers need to determine which normal cells (beyond the few already identified) make telomerase, and they need to assess the importance of the enzyme to those cells. If telomerase is crucial, drugs that interfere with it might in fact prove unacceptably toxic. The shortness of telomeres in certain tumor cells may obviate this problem, however. Telomerase-inhibiting agents might cause cancer cells to lose their telomeres and die well before normal cells, with their much longer telomeres, lose enough of

YOUNG CELLS

SENESCENT CELLS

TUMOR CELLS



JUDITH CAMPISI *Lawrence Berkeley Laboratory*

**MODEL OF TELOMERE LENGTH REGULATION in humans suggests that germ-line cells (*blue line in graph at left*) make telomerase and maintain telomeres at a long length. In contrast, many somatic, or nongerm, cells (*dark gold curve*) presumably lack telomerase, and so their telomeres shrink over time. Eventually, most somatic cells enter senescence: they stop dividing and undergo other changes. Aberrant cells (*light gold curve*) might continue to divide, however. Many will lose telomeric repeats until they reach a crisis point and die. But if the deranged cells begin to make telomerase, they may replicate indefinitely (*green curve*). This property, found in many tumor cells, is known as cellular immortality. The micrographs above show young somatic cells (*left*), senescent cells (*center*) and immortal tumor cells (*right*).**

their telomeres to suffer any ill effects.

Investigators must also demonstrate that inhibition of telomerase can destroy telomerase-producing tumors as expected. Last September, Harley, Greider and their co-workers showed that an inhibitory agent could cause the telomeres of cultured tumor cells to shrink; the affected cells died after about 25 cycles of cell division. Blackburn, now at the University of California at San Francisco, and her group have found, however, that cells sometimes compensate for the loss of telomerase. They repair their shortened ends by other means, such as by a process called recombination, in which one chromosome obtains DNA from another. If activation of alternative, "telomere-salvaging" pathways occurs frequently in human tumors, therapy targeted to telomerase would fail.

Studies of animals should help resolve such concerns. They should also help reveal whether inhibitors of telomerase will eliminate tumors in the living body and whether they will do so quickly enough to prevent cancers from injuring critical tissue.

To develop agents that will block telomerase in the human body, investigators must also have a sharper picture of exactly how the enzyme functions. How does it attach to DNA? How does it "decide" on the number of telomeric subunits to add? DNA in the nucleus is studded with all manner of proteins, including some that specifically bind to the telomere. What part do telomere-binding proteins play in controlling the activity of telomerase? Would altering their activity disrupt telomere elongation? Within the next 10 years we expect to learn a great deal about the interactions among the various molecules that influence telomere length.

Research into the regulation of telomere size could also yield benefits beyond new therapies for cancer. A popular approach to gene therapy for various diseases involves extracting cells from a patient, inserting the desired gene and then returning the genetically corrected cells to the patient. Frequently, though, the extracted cells proliferate poorly in the laboratory. Perhaps insertion of telomerase alone or in combination with other factors would temporarily enhance replication capacity, so that larger numbers of therapeutic cells could be delivered to the patient.

Modern research into telomeres has come a long way from the initial identification of repetitive DNA on the ends of chromosomes in a unicellular pond dweller. Elongation of telomeres by telomerase, initially considered to be merely a "cute" mechanism by which some single-cell creatures maintain their chromosomes, has proved, as ever, to be other than it seemed. Telomerase is, in fact, the predominant means by which nucleated cells of most animals protect their chromosomal end segments. And, now, study of this once obscure process may lead to innovative strategies for fighting a range of cancers.

In the early 1980s scientists would not have set out to identify potential anticancer therapies by studying chromosome maintenance in *Tetrahymena*. The research on telomerase reminds us that in studies of nature one can never predict when and where fundamental processes will be uncovered. You never know when a rock you find will turn out to be a gem.

## *The Authors*

CAROL W. GREIDER and ELIZABETH H. BLACKBURN began collaborating in 1983, when Greider joined Blackburn's laboratory at the University of California, Berkeley. Greider, who earned her Ph.D. in molecular biology from Berkeley in 1987, is senior staff scientist at Cold Spring Harbor Laboratory. Blackburn holds a 1975 doctorate in molecular biology from the University of Cambridge. She has been a professor of microbiology and immunology at the University of California, San Francisco, since 1990 and department chair since 1993.

## *Further Reading*

IDENTIFICATION OF A SPECIFIC TELOMERE TERMINAL TRANSFERASE ACTIVITY IN *TETRAHYMENA* EXTRACTS. Carol W. Greider and Elizabeth H. Blackburn in *Cell,* Vol. 43, Part 1, pages 405–413; December 1985.

*IN VIVO* ALTERATION OF TELOMERE SEQUENCES AND SENESCENCE CAUSED BY MUTATED *TETRAHYMENA* TELOMERASE RNAS. G.-L. Yu, J. D. Bradley, L. D. Attardi and E. H. Blackburn in *Nature,* Vol. 344, pages 126–132; March 8, 1990.

TELOMERASE, CELL IMMORTALITY AND CANCER. C. B. Harley, N. W. Kim, K. R. Prowse, S. L. Weinrich, K. S. Hirsch, M. D. West, S. Bacchetti, H. W. Hirte, C. W. Greider, W. E. Wright and J. W. Shay in *Cold Spring Harbor Symposia in Quantitative Biology,* Vol. LVIX, pages 307–315; 1994.

TELOMERES. Edited by E. H. Blackburn and C. W. Greider. Cold Spring Harbor Laboratory Press, 1995.

# Colossal Galactic Explosions

*Enormous outpourings of gas from the centers of nearby galaxies may ultimately help explain both star formation and the intergalactic medium*

by Sylvain Veilleux, Gerald Cecil and Jonathan Bland-Hawthorn

Millions of galaxies shine in the night sky, most made visible by the combined light of their billions of stars. In a few, however, a pointlike region in the central core dwarfs the brightness of the rest of the galaxy. The details of such galactic dynamos are too small to be resolved even with the *Hubble Space Telescope.* Fortunately, debris from these colossal explosions—in the form of hot gas glowing at temperatures well in excess of a million degrees—sometimes appears outside the compact core, on scales that can be seen directly from the earth.

The patterns that this superheated material traces through the interstellar gas and dust surrounding the site of the explosion provide important clues to the nature and history of the powerful forces at work inside the galactic nucleus. Astronomers can now determine what kind of engines drive these dynamos and the effects of their tremendous outpourings on the intergalactic medium.

Furthermore, because such cataclysms appear to have been taking place since early in the history of the universe, they have almost certainly affected the environment in which our own Milky Way galaxy evolved. Understanding how such events take place today may illuminate the distribution of chemical elements that has proved crucial to formation of stars like the sun.

Astronomers have proposed two distinctly different mechanisms for galactic dynamos. The first was the brainchild of Martin J. Rees of the University of Cambridge and Roger D. Blandford, now at the California Institute of Technology. During the early 1970s, the two sought to explain the prodigious luminosity—thousands of times that of the Milky Way—and the spectacular "radio jets" (highly focused streams of energetic material) that stretch over millions of light-years from the centers of some hyperactive young galaxies known as quasars. They suggested that an ultramassive black hole—not much larger than the sun but with perhaps a million times its mass—could power a quasar.

A black hole itself produces essentially no light, but the disk of accreted matter spiraling in toward the hole heats up and radiates as its density increases. The inner, hotter part of the disk produces ultraviolet and x-ray photons over a broad range of energies, a small fraction of which are absorbed by the surrounding gas and reemitted as discrete spectral lines of ultraviolet and visible light. In the years since Rees and Blandford proposed their model, astronomers have come to understand that similar black holes may be responsible for the energy output of nearer active galaxies.

As the disk heats up, gas in its vicinity reaches temperatures of millions of degrees and expands outward from the galactic nucleus at high speed. This flow, an enormous cousin to the solar wind that streams away from the sun or other stars, can sweep up other interstellar gases and expel them from the nucleus. The resulting luminous shock waves can span thousands of light-years—comparable to the visible sizes of the galaxies themselves—and can be studied from space or ground-based observatories. Some of these galaxies also produce radio jets: thin streams of rapidly moving


U.S. INFORMATION SERVICE


HALE OBSERVATORIES

**M82 (*a, b*), about 10 million light-years away from the earth, is distinguished by an outpouring of incandescent gas from the area around its core (*c*). Astronomers have deduced that the upheaval is caused by the rapid formation of stars near the galactic nucleus. The resulting heat and radiation cause dust and gas from the galactic disk to rush into intergalactic space. The galaxy's activity may have been triggered by interaction with its neighbor, M81.**

gas that emit radio waves as they traverse a magnetic field that may be anchored within the accretion disk.

Black holes are not the only engines that drive violent galactic events. Some galaxies apparently undergo short episodes of rapid star formation in their cores: so-called nuclear starbursts. The myriad new stars produce strong stellar winds and, as the stars age, a rash of supernovae. The fast-moving gas ejected from the supernovae strikes the background interstellar dust and gas and heats it to millions of degrees.

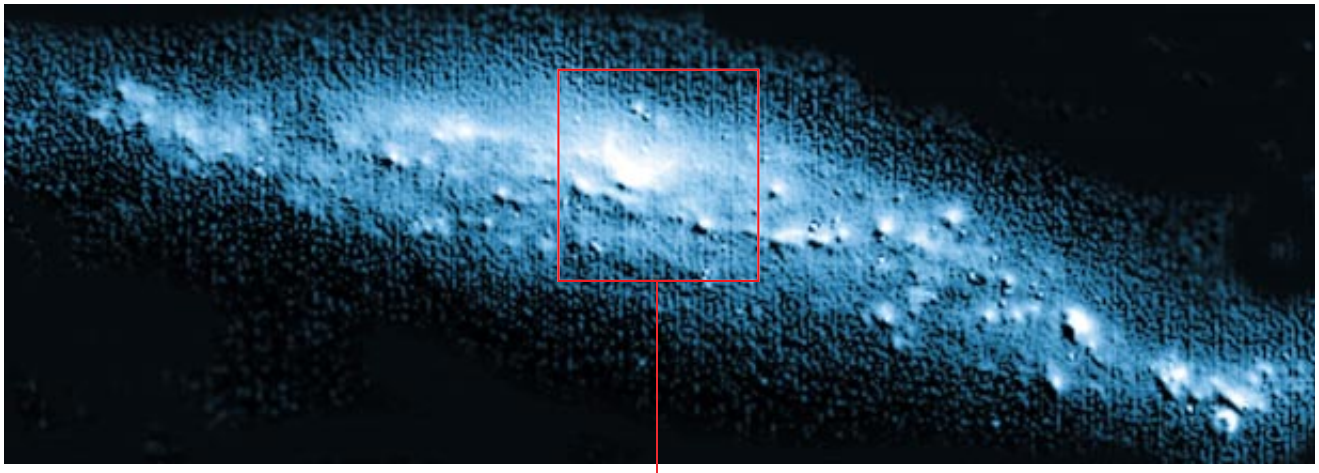
The pressure of this hot gas forms a cavity, like a steam bubble in boiling water. As the bubble expands, cooler gas and dust accumulate in a dense shell at the edge of the bubble, slowing its expansion. The transition from free flow inside the bubble to near stasis at its boundary gives rise to a zone of turbulence that is readily visible from the earth. If the energy injected into the cavity is large enough, the bubble bursts out of the galaxy's gas disk and spews the shell's fragments and hot gas into the galaxy halo or beyond, thousands of light-years from their origins.

Roberto Terlevich of the Royal Greenwich Observatory and his collaborators have led the most recent research aimed at determining whether starbursts alone can drive the outpourings of hot gas characteristic of active galaxies. In 1985 Terlevich and Jorge Melnick, now at the European Southern Observatory, argued that many such galaxies contain unusual stars they dubbed "warmers"—extremely hot stars with temperatures higher than 100,000 degrees and very powerful stellar winds. Such stars, the two scientists proposed, arise naturally when a starburst occurs in a region enriched in heavy chemical elements from previous supernovae. Terlevich and his colleagues contend that their model explains the spectra and many other properties of certain active galaxies.

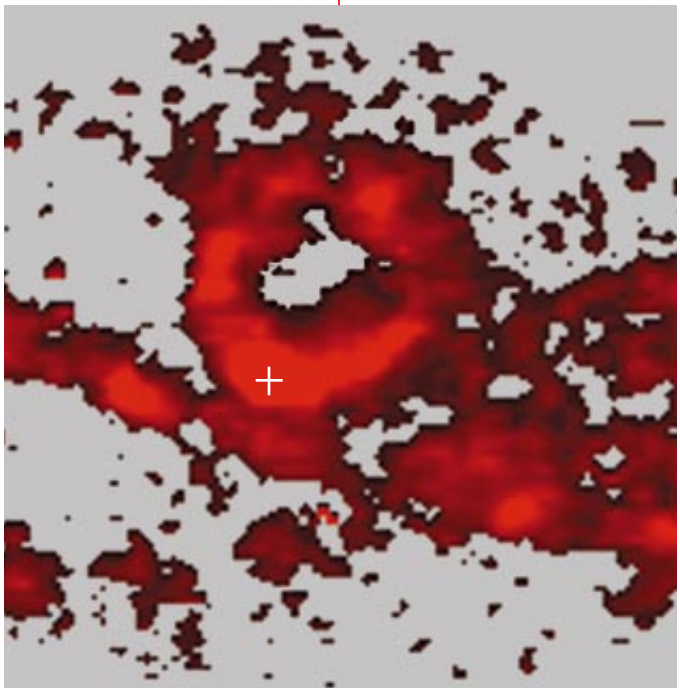
### Identifying the Engine

Both the starburst and the black-hole explanations appear plausible, but there are important differences between the two that may reveal which one is at work in a given galaxy. A black hole can convert as much as 10 percent of the infalling matter to energy. Starbursts, in contrast, rely on nuclear fusion, which can liberate only 0.1 percent of the reacting mass. As a result, they require at least 100 times as much matter, most


COURTESY OF GERALD CECIL



COURTESY OF GERALD CECIL



COURTESY OF GERALD CECIL

**STARBURST, a sudden pulse of star formation, may be responsible for the activity of NGC 3079 (*top*) even though the galaxy has a black hole at its center. A close-up view of the area near the nucleus (*white cross*) reveals the outlines of an enormous bubble that has been blown into the interstellar medium by the heat of the stars forming at the galaxy's center.**

of which accumulates as unburned fuel. Over the lifetime of a starburst-powered quasar, the total mass accumulated in the nucleus of the galaxy could reach 100 billion times the mass of the sun, equivalent to the mass of all the stars in the Milky Way galaxy.

The more mass near the nucleus, the more rapidly the orbiting stars must move. Ground-based optical observations, which are limited by atmospheric blurring, have not placed tight constraints on the concentration of mass in galactic centers. Recent radio-telescope findings, however, have revealed an accretion disk with an inner radius of half a light-year spinning rapidly around a mass 20 million times that of the sun at the center of a nearby spiral galaxy called NGC 4258.

Several research groups are now measuring the distributions of stellar motions across galactic nuclei using the spectrograph on board the *Hubble* telescope. The discovery that gas in the inner core of the active galaxy M87 is moving in a manner consistent with a black-hole accretion disk has demonstrated the promise of such techniques, and the instrument will become far more efficient after astronauts upgrade it in 1997.

Starbursts and black holes also differ in the spectra of the most energetic photons they produce. Near a black hole, the combination of a strong magnetic field and a dense accretion disk creates a soup of very fast particles that collide with one another and with photons to generate x-rays and gamma rays. A starburst, in contrast, produces most of its high-energy radiation from collisions between supernova ejecta and the surrounding galactic gas and dust. This impact heats gas to no more than about a billion degrees and so cannot produce any radiation more energetic than x-rays. The large numbers of gamma rays detected recently from some quasars by the *Compton Gamma Ray Observatory* imply that black holes are at their centers [see "The Compton Gamma Ray Observatory," by Neil Gehrels, Carl E. Fichtel, Gerald J. Fishman, James D. Kurfess and Volker Schönfelder; SCIENTIFIC AMERICAN, December 1993].

A final difference between black holes and starbursts lies in the forces that focus the flow of outrushing gas. The magnetic-field lines attached to the accretion disk around a black hole direct outflowing matter along the rotation axis of the disk in a thin jet. The material expelled by a starburst bubble, in contrast, simply follows the path of least resistance in the surrounding environment. A powerful starburst in a spiral galaxy will spew gas perpendicular to the plane of the galaxy's disk of stars and gas, but the flow will be distributed inside an hourglass-shaped region with a wide opening. The narrow radio jets that extend millions of light-years from the core of some active galaxies clearly suggest the presence of black holes.

All that we know about galaxies—active or otherwise—comes from the radiation they emit. Our observations supply the data that astrophysicists can use to choose among competing theories. The three of us have concentrated on visible light, from which we can determine the temperatures, pressures and concentrations of different atoms in the gas agitated by galactic explosions. We compare the wavelength and relative intensities of emission lines from excited or ionized atoms with those measured in terrestrial laboratories or derived from theoretical calculations.

### Spectral Signatures

Thanks to the Doppler shift, which changes the frequency and wavelength of light emitted by moving sources, this analysis also reveals how fast

the gas is moving. Approaching gas emits light shifted toward the blue end of the spectrum, and receding gas emits light shifted toward the red end.

Until recently, astronomers unraveled gas behavior by means of two complementary methods: emission-line imaging and long-slit spectroscopy. The first produces images through a filter that selects light of a particular wavelength emitted by an element such as hydrogen. Such images often dramatically reveal the filamentary patterns of explosions, but they cannot tell observers anything about the speed or direction of the gases' motions, because the filter does not discriminate finely enough to measure redshifts or blueshifts. Long-slit spectrometers, which disperse light into its constituent colors, provide detailed information about gas motions but only over a tiny region.

For almost a decade, our group has used an instrument that combines the advantages of these two methods without the main drawbacks. The Hawaii Imaging Fabry-Perot Interferometer (HIFI) yields detailed spectral information over a large field of view. Named after the turn-of-the-century French inventors Charles Fabry and Alfred Pérot, such interferometers have found wide-ranging applications in astronomy. At their heart are two glass plates that are kept perfectly parallel while separated by less than a twentieth of a millimeter. The inner surfaces of the plates are highly reflecting, so light passing through the plates is trapped into repeated reflections. Light of all but a specific wavelength—determined by the precise separation—is attenuated by destructive interference as the light waves bounce back and forth between the plates. By adjusting the separation between the plates, we can produce a series of images that are essentially a grid of spectra obtained by the interferometer at every position over the field of view.

The HIFI takes its pictures atop the 14,000-foot dormant volcano Mauna Kea, using the 2.2-meter telescope owned by the University of Hawaii and the 3.6-meter Canada-France-Hawaii instrument. The smooth airflow at the mountaintop produces sharp images. Charge-coupled devices, which are very stable and sensitive to faint light, collect the photons. In a single night, this powerful combination can generate records of up to a million spectra across the full extent of a galaxy.

### Nearby Active Galaxies

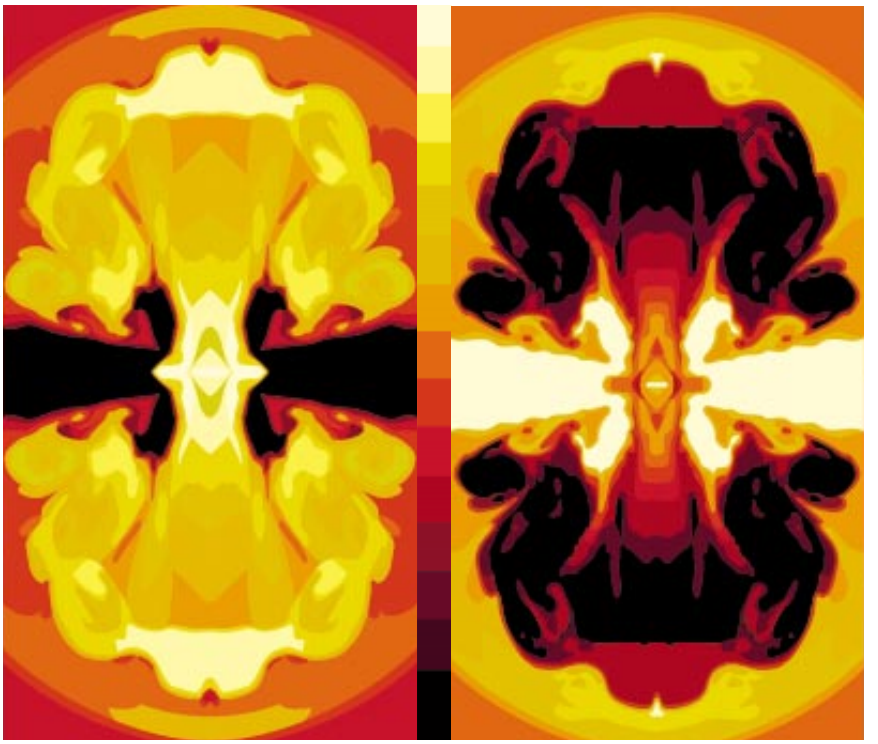
We have used the HIFI to explore NGC 1068, an active spiral galaxy 46 million light-years away. As the nearest and brightest galaxy of this type visible from the Northern Hemisphere, it has been studied extensively. At radio wavelengths, NGC 1068 looks like a miniature quasar: two jets extend about 900 light-years from the core, with more diffuse emission from regions farther out. Most likely, emission from gaseous plasma moving at relativistic speeds creates the radio jets, and the "radio lobes" arise where the plasma encounters matter from the galactic disk. As might a supersonic aircraft, the leading edge of the northeast jet produces a V-shaped shock front.

The same regions also emit large amounts of visible and ultraviolet light. We have found, however, that only 10 percent of the light comes from the nucleus. Another 5 percent comes from galaxy-disk gas that has piled up on the expanding edge of the northeast radio lobe. All the rest comes from two fans of high-velocity gas moving outward from the center at speeds of up to 1,500 kilometers per second.

The gas flows outward in two conical regions; it is probably composed of dense filaments of matter that have been swept up by the hot wind from the accretion disk. The axis of the cones of outflowing wind is tilted above the plane of the galaxy but does not point toward the poles.

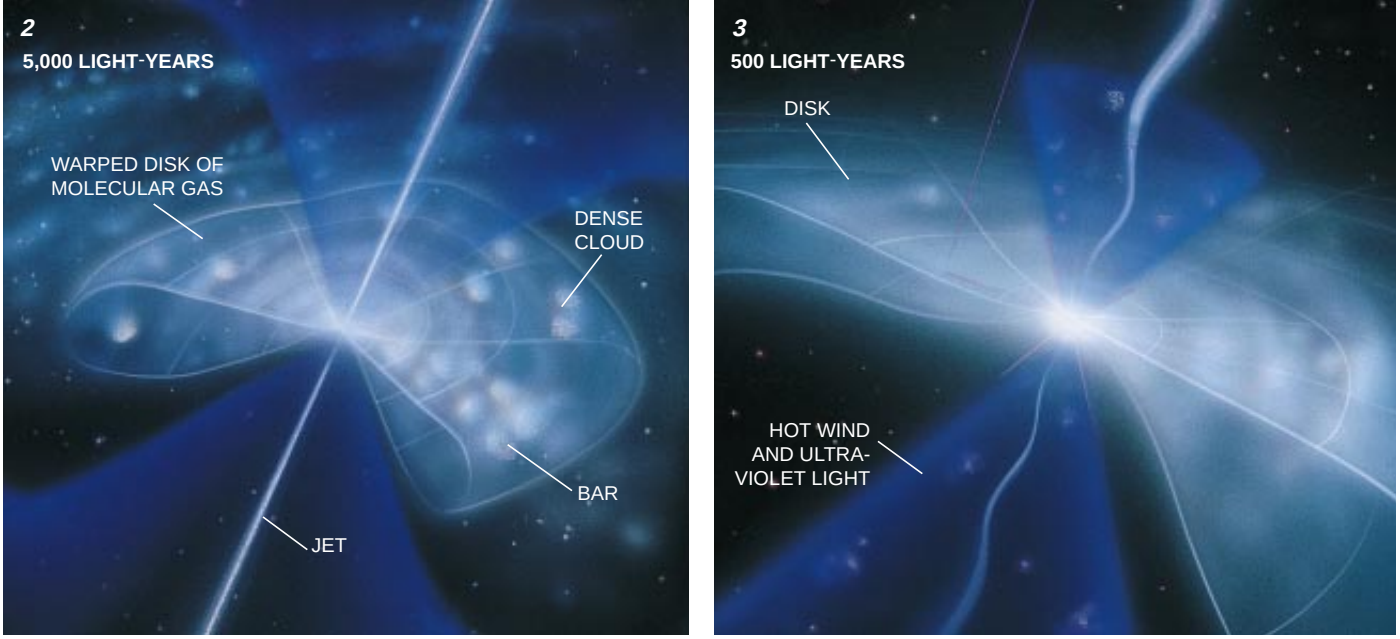
The effects of the activity within the nucleus reach out several thousand light-years, well beyond the radio lobes. The diffuse interstellar gas exhibits unusually high temperatures and a large fraction of the atoms have lost one or more electrons and become ionized. At the same time, phenomena in the disk appear to influence the nucleus. Infrared images reveal an elongated bar of stars that extends more than 3,000 light-years from the nucleus. The HIFI velocity measurements suggest that the bar distorts the circular orbit of the gas in the disk, funneling material toward the center of the galaxy. This inflow of material may in fact fuel the black hole.

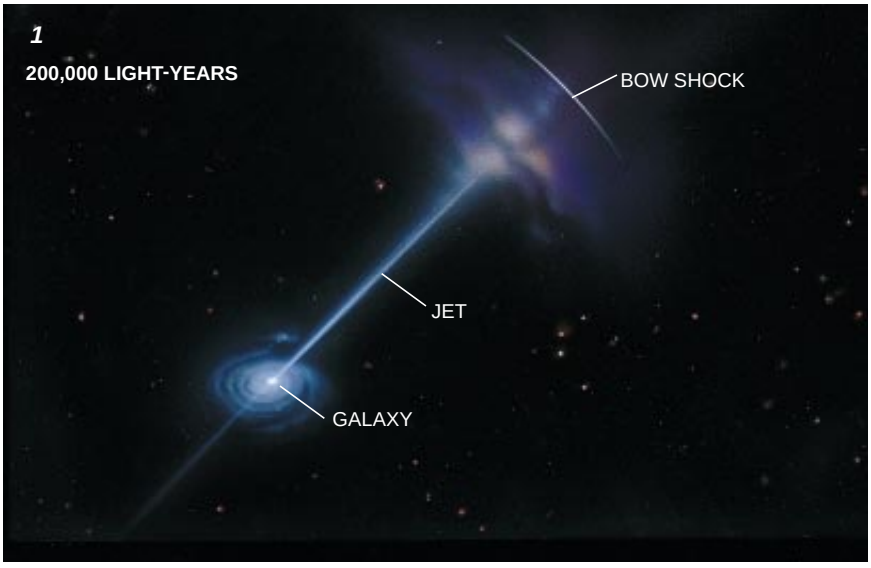
Another tremendous explosion is occurring in the core of one of our nearest neighbor galaxies, M82, just a few million light-years away. In contrast to NGC 1068, this cataclysm appears to be an archetypal starburst-driven event. Images exposed through a filter that passes the red light of forming hydrogen atoms reveal a web of filaments spraying outward along the galactic poles. Our spectral grids of emission from filaments perpendicular to the galactic disk reveal two main masses of gas, one receding and the other approaching. The difference in velocity between the two increases as the gas moves outward from the core, reaching about 350 kilometers per second at a



COURTESY OF GERALD CECIL, AFTER ANATOLY A. SUCHKOV ET AL.

**OUTPOURING OF GAS rapidly becomes turbulent in this computer simulation of a starburst-driven active galaxy. A temperature map (*left*) shows how the hot gas emanating from the nucleus displaces the cooler galactic gas around it. The resulting shock appears clearly in a map of gas density (*right*).**




ALFRED T. KAMAJIAN

**COLOSSAL FORCES at work in the center of an active galaxy can make themselves felt half a million light-years or more away as jets of gas moving at relativistic speeds plow into the intergalactic medium and create enormous shock waves (*1*). Closer to the center of the galaxy (*2, 3*), a dense equatorial disk of dust and molecular gas feeds matter to the active nucleus while hot gas and radiation spill out along the poles. The high density of the infalling gas within a few dozen light-years of the center causes a burst of star formation (*4*). Even closer to the center (*5*), the disk, glowing at ultraviolet and x-ray wavelengths, tapers inward to feed what astronomers believe is a black hole containing millions of stellar masses but still so small as to be invisible on this scale.**

distance of 3,000 light-years. At a distance of 4,500 light-years from the core, the velocity separation diminishes.

The core of M82 is undergoing an intense burst of star formation, possibly triggered by a recent orbital encounter with its neighbors M81 and NGC 3077. Its infrared luminosity is 30 billion times the total luminosity of the sun, and radio astronomers have identified the remnants of large numbers of supernovae. The filamentary web visible from the earth results from two elongated bubbles oriented roughly perpendicular to the disk of M82 and straddling the nucleus. X-ray observatories in space have detected the hot wind that inflates these bubbles; their foamy appearance probably arises from instabilities in the hot gas as it cools.

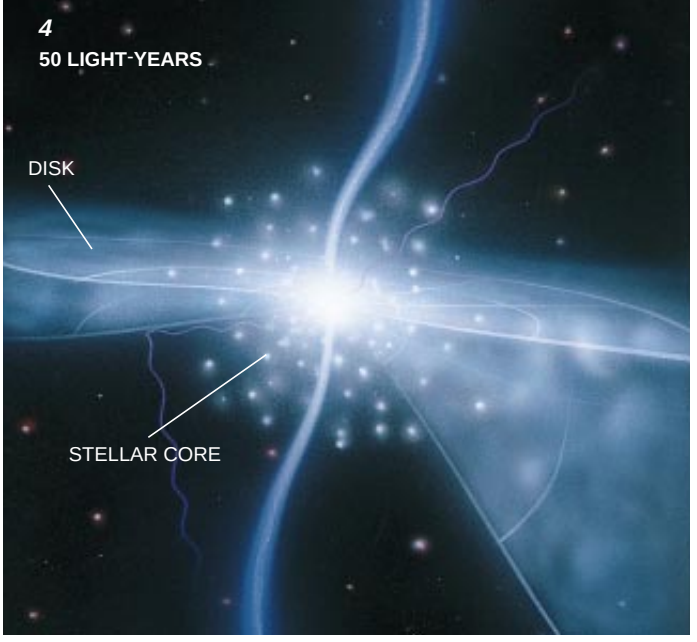
### Ambiguous Activity

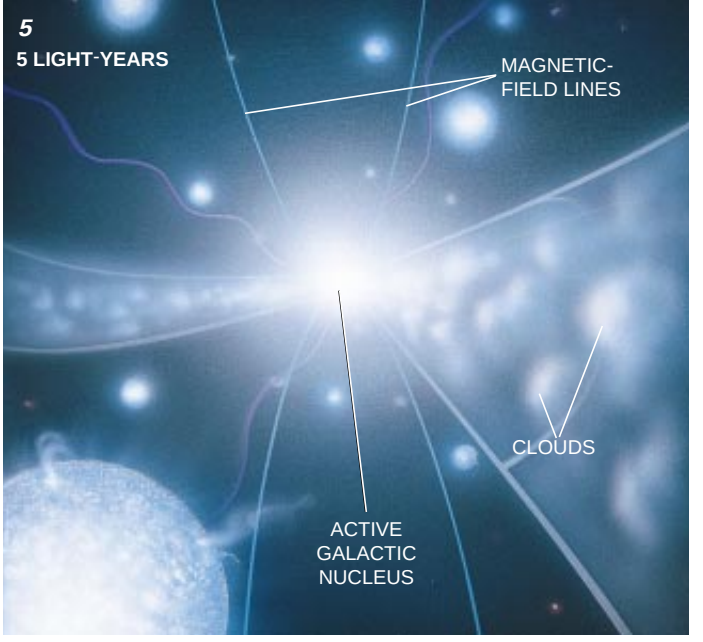
Unfortunately, the identity of the principal source of energy in active galaxies is not always so obvious. Sometimes a starburst appears to coexist with a black-hole engine. Like M82, many of these galaxies are abnormally bright at infrared wavelengths and rich in molecular gas, the raw material of stars. Radio emission and visual spectra resembling those of a quasar, however, suggest that a black hole may also be present.

Such ambiguity plagues interpretations of the behavior of the nearby galaxy NGC 3079. This spiral galaxy appears almost edge-on from the earth—an excellent vantage point from which to study the gas expelled from the nucleus. Like M82, NGC 3079 is anomalously bright in the infrared, and it also contains a massive disk of molecular gas spanning 8,000 light-years around its core. At the same time, the core is unusually bright at radio wavelengths, and the linear shape of radio-emitting regions near the core suggests a collimated jet outflow. On a larger scale, the radio-emission pattern is complex and extends more than 6,500 light-years from either side of the galactic disk.

Images made in red hydrogen light show a nearly circular ring 3,600 light-years across just east of the nucleus; velocity measurements from the HIFI confirm that the ring marks the edge of a bubble as seen from the side. The bubble resembles an egg with its pointed extremity balanced on the nucleus and its long axis aligned with the galactic pole. There is another bubble on the west side of the nucleus, but most of it is hidden behind the dusty galaxy disk.

Our spectral observations imply that the total energy of this violent outflow is probably 10 times that of the explosions in NGC 1068 or M82. The alignment of the bubble along the polar axis of the host galaxy implies that galactic dust and gas, rather than a central black hole, are collimating the outflow. Nevertheless, the evidence is fairly clear that NGC 3079 contains a massive black hole at its core.




Is the nuclear starburst solely responsible for such a gigantic explosion? We have tried to answer this question by analyzing the infrared radiation coming from the starburst area. Most of the radiation from young stars embedded in molecular clouds is absorbed and reemitted in the infrared, so the infrared luminosity of NGC 3079's nucleus may be a good indicator of the rate at which supernovae and stellar winds are injecting energy at the center of the galaxy. When we compare the predictions of the starburst model with our observations, we find that the stellar ejecta appears to have enough energy to inflate the bubble. Although the black hole presumed to exist in the core of NGC 3079 may contribute to the outflow, there is no need to invoke it as an energy source.

### How Active Galaxies Form

Although astronomers now understand the basic principles of operation of the engines that drive active galaxies, many details remain unclear. There is a vigorous debate about the nature of the processes that ignite a starburst or form a central black hole. What is the conveyor belt that transports fuel down to the pointlike nucleus? Most likely, gravitational interactions with gas-rich galaxies redistribute gas in the host galaxy, perhaps by forming a stellar bar such as the one in NGC 1068. Computer simulations appear to indicate that the bar, once formed, may be quite stable [see "Colliding Galaxies," by Joshua Barnes, Lars Hernquist and François Schweizer; SCIENTIFIC AMERICAN, August 1991]. (Indeed, the bar must be stable, because NGC 1068 currently has no close companion.)

Researchers are also divided on which comes first, nuclear starburst or black hole. Perhaps the starburst is an early phase in the evolution of active galaxies, eventually fading to leave a dense cluster of stellar remnants that rapidly coalesce into a massive black hole.

The anomalous gas flows in the galaxies that we and others have observed are almost certainly only particularly prominent examples of widespread, but more subtle, processes that affect many more galaxies. Luminous infrared galaxies are common, and growing evidence is leading astronomers to believe that many of their cores are also the seats of explosions. These events may profoundly affect the formation of stars throughout the galactic neighborhood. The bubble in NGC 3079, for instance, is partially ruptured at the top and so probably leaks material into the outer galactic halo or even into the vast space between galaxies. Nuclear reactions in the torrent of supernovae unleashed by the starburst enrich this hot wind in heavy chemical elements. As a result, the wind will not only heat its surroundings but also alter the environment's chemical composition.

The full impact of this "cosmic bubble bath" over the history of the universe is difficult to assess accurately because we currently know very little of the state of more distant galaxies. Images of distant galaxies taken by the *Hubble* will help clarify some of these questions. Indeed, as the light that left those galaxies billions of years ago reaches our instruments, we may be watching an equivalent of our own galactic prehistory unfolding elsewhere in the universe.

## *The Authors*

SYLVAIN VEILLEUX, GERALD CECIL and JONATHAN BLAND-HAWTHORN met while working at observatories in Hawaii and were drawn to collaborate by a shared interest in peculiar galaxies. Veilleux, now an assistant professor of astronomy at the University of Maryland, received his Ph.D. from the University of California, Santa Cruz. Cecil, an associate professor of astronomy and physics at the University of North Carolina at Chapel Hill and director of the Morehead Observatory there, received his doctorate from the University of Hawaii. Bland-Hawthorn received his Ph.D. in astronomy and astrophysics from the University of Sussex and the Royal Greenwich Observatory. He is now a research astronomer at the Anglo-Australian Observatory in Sydney.

## *Further Reading*

LARGE-SCALE BIPOLAR WIND IN M82. Jonathan Bland and R. Brent Tully in *Nature,* Vol. 334, pages 43–45; July 7, 1988.

IMAGING SPECTROPHOTOMETRY OF IONIZED GAS IN NGC 1068, Part 1: KINEMATICS OF THE NARROW-LINE REGION. G. Cecil, J. Bland and R. B. Tully in *Astrophysical Journal,* Vol. 355, pages 70–87; May 20, 1990.

RELATIONSHIPS BETWEEN ACTIVE GALACTIC NUCLEI AND STARBURST GALAXIES. Proceedings of the Taipei Astrophysics Workshop, 1991. Edited by Alexei V. Filippenko. Astronomical Society of the Pacific, 1992.

THE NUCLEAR SUPERBUBBLE OF NGC 3079. S. Veilleux, G. Cecil, J. Bland-Hawthorn, R. B. Tully, A. V. Filippenko and W.L.W. Sargent in *Astrophysical Journal,* Part 1, Vol. 433, pages 48–64; September 20, 1994.

DISSECTING COSMIC EXPLOSIONS. R. B. Tully, J. Morse and P. Shopbell in *Sky and Telescope,* Vol. 90, No. 1, pages 18–21; July 1995.

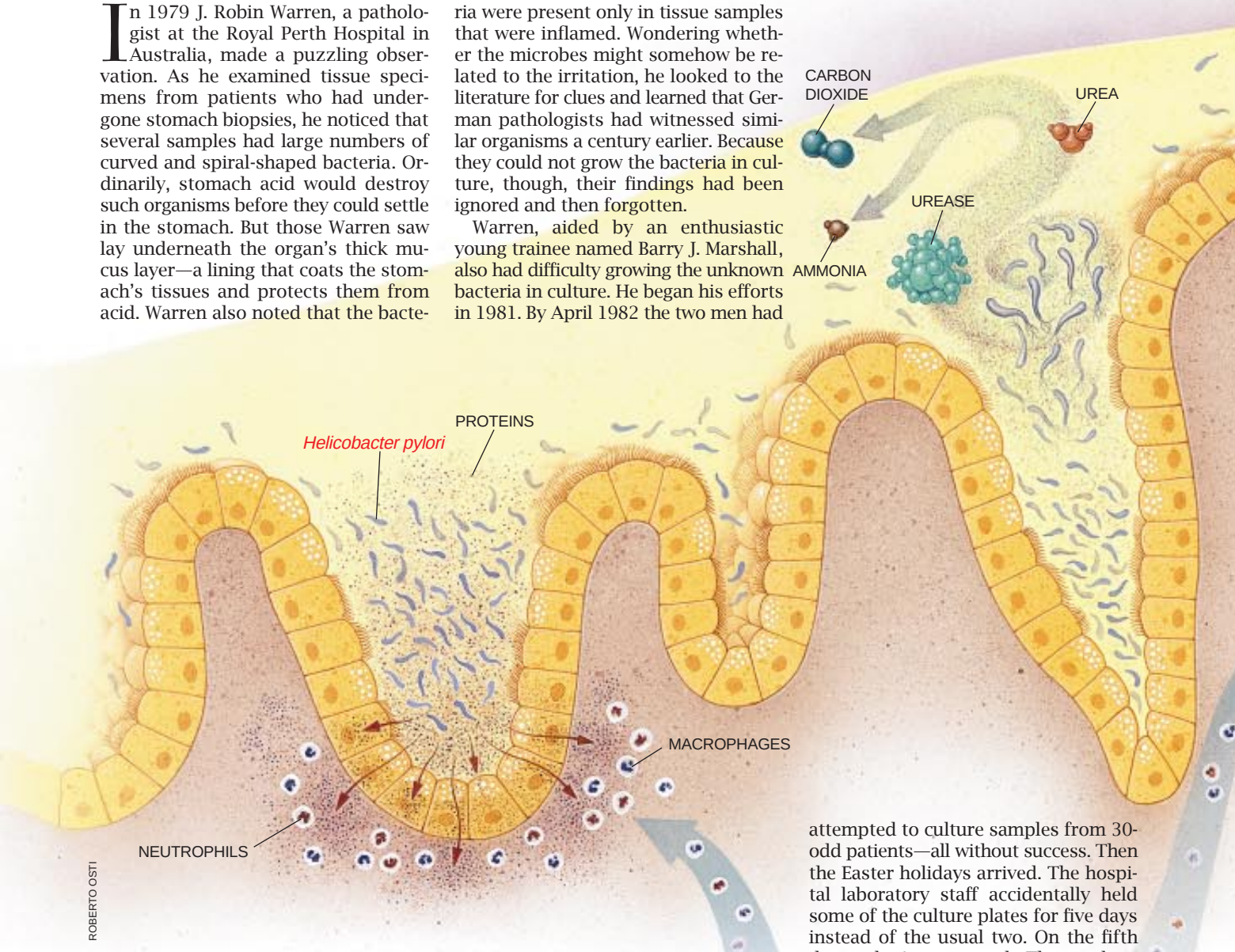
# The Bacteria behind Ulcers

*One half to one third of the world's population harbors* Helicobacter pylori, *"slow" bacteria that infect the stomach and can cause ulcers and cancer there*

by Martin J. Blaser

In 1979 J. Robin Warren, a pathologist at the Royal Perth Hospital in Australia, made a puzzling observation. As he examined tissue specimens from patients who had undergone stomach biopsies, he noticed that several samples had large numbers of curved and spiral-shaped bacteria. Ordinarily, stomach acid would destroy such organisms before they could settle in the stomach. But those Warren saw lay underneath the organ's thick mucus layer—a lining that coats the stomach's tissues and protects them from acid. Warren also noted that the bacteria were present only in tissue samples that were inflamed. Wondering whether the microbes might somehow be related to the irritation, he looked to the literature for clues and learned that German pathologists had witnessed similar organisms a century earlier. Because they could not grow the bacteria in culture, though, their findings had been ignored and then forgotten.

Warren, aided by an enthusiastic young trainee named Barry J. Marshall, also had difficulty growing the unknown bacteria in culture. He began his efforts in 1981. By April 1982 the two men had attempted to culture samples from 30-odd patients—all without success. Then the Easter holidays arrived. The hospital laboratory staff accidentally held some of the culture plates for five days instead of the usual two. On the fifth day, colonies emerged. The workers christened them *Campylobacter pylori-dis* because they resembled pathogenic bacteria of the *Campylobacter* genus found in the intestinal tract. Early in 1983 Warren and Marshall published their first report, and within months scientists around the world had isolated




ROBERTO OSTI

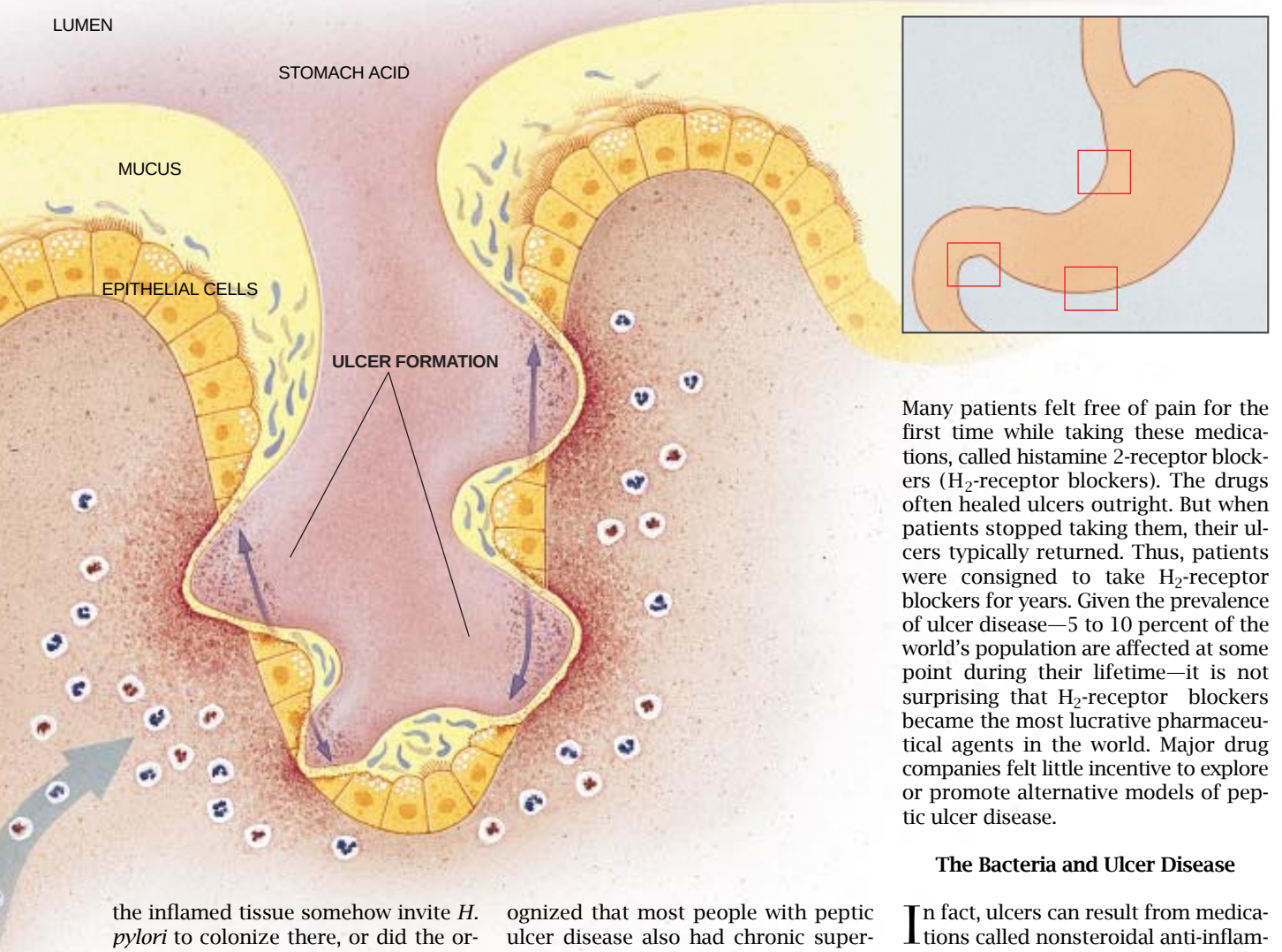
**ULCER-CAUSING BACTERIA (*Helicobacter pylori*) live in the mucus layer (*pale yellow*) lining the stomach. There they are partially protected from the stomach's acid (*pink*). The organisms secrete proteins that interact with the stomach's epithelial cells and attract macrophages and neutrophils, cells that cause inflammation (*left*). The bacteria further produce urease, an enzyme that helps to break down urea into ammonia and carbon dioxide; ammonia can neutralize stomach acid (*center*). *H. pylori* also secrete toxins that contribute to the formation of stomach ulcers (*right*). The microbes typically collect in the regions shown in the diagram at the far right.**

the bacteria. They found that it did not, in fact, fit into the *Campylobacter* genus, and so a new genus, *Helicobacter,* was created. These researchers also confirmed Warren's initial finding, namely that *Helicobacter pylori* infection is strongly associated with persistent stomach inflammation, termed chronic superficial gastritis.

The link led to a new question: Did the inflamed tissue somehow invite *H. pylori* to colonize there, or did the organisms actually cause the inflammation? Research proved the second hypothesis correct. In one of the studies, two male volunteers—Marshall included—actually ingested the organisms [*see box on next page*]. Both had healthy stomachs to start and subsequently developed gastritis. Similarly, when animals ingested *H. pylori,* gastritis ensued. In other investigations, antibiotics suppressed the infection and alleviated the irritation. If the organisms were eradicated, the inflammation went away, but if the infection recurred, so did the gastritis. We now know that virtually all people infected by *H. pylori* acquire chronic superficial gastritis. Left untreated, both the infection and the inflammation last for decades, even a lifetime. Moreover, this condition can lead to ulcers in the stomach and in the duodenum, the stretch of small intestine leading away from the stomach. *H. pylori* may be responsible for several forms of stomach cancer as well.

More than 40 years ago doctors recognized that most people with peptic ulcer disease also had chronic superficial gastritis. For a variety of reasons, though, when the link between *H. pylori* infection and gastritis was established, the medical profession did not guess that the bacteria might prompt peptic ulcer disease as well. Generations of medical students had learned instead that stress made the stomach produce more acid, which in turn brought on ulcers. The theory stemmed from work carried out by the German scientist K. Schwartz. In 1910, after noting that duodenal ulcers arose only in those individuals who had acid in their stomachs, he coined the phrase "No acid, no ulcer." Although gastric acidity is necessary for ulcers to form, it is not sufficient to explain their occurrence—most patients with ulcers have normal amounts of stomach acid, and some people who have high acid levels never acquire ulcers.

Nevertheless, the stress-acid theory of ulcers gained further credibility in the 1970s, when safe and effective agents to reduce gastric acid were introduced.




Many patients felt free of pain for the first time while taking these medications, called histamine 2-receptor blockers ($H_2$-receptor blockers). The drugs often healed ulcers outright. But when patients stopped taking them, their ulcers typically returned. Thus, patients were consigned to take $H_2$-receptor blockers for years. Given the prevalence of ulcer disease—5 to 10 percent of the world's population are affected at some point during their lifetime—it is not surprising that $H_2$-receptor blockers became the most lucrative pharmaceutical agents in the world. Major drug companies felt little incentive to explore or promote alternative models of peptic ulcer disease.

### The Bacteria and Ulcer Disease

In fact, ulcers can result from medications called nonsteroidal anti-inflammatory agents, which include aspirin and are often used to treat chronic arthritis. But all the evidence now indicates that *H. pylori* cause almost all cases of ulcer disease that are not medication related. Indeed, nearly all patients having such ulcers are infected by *H. pylori,* versus some 30 percent of age-matched control subjects in the U.S., for example. Nearly all individuals with ulcers in the duodenum have *H. pylori* present there. Studies show that *H. pylori* infection and chronic gastritis increase from three to 12 times the risk of a peptic ulcer developing within 10 to 20 years of infection with the bacte-

MICHAEL BAILEY *University of Virginia*

## Don't Try This at Home

Barry J. Marshall (*left*) of the Royal Perth Hospital in Australia made headlines after he announced in 1985 that he had ingested *Helicobacter pylori.* Marshall hoped to demonstrate that the bacteria could cause peptic ulcer disease. Marshall did in fact develop a severe case of gastritis, but the painful inflammation vanished without treatment.

Two years later Arthur J. Morris and Gordon I. Nicholson of the University of Auckland in New Zealand reported the case of another volunteer who wasn't so lucky. This man, a healthy 29-year-old, showed signs of infection for only 10 days, but his condition lasted much longer. On the 67th day of infection, the volunteer started treatment with bismuth subsalicylate (Pepto-Bismol). Five weeks later a biopsy indicated that the medication had worked. But a second biopsy taken nine months after the first showed that both the infection and the gastritis had recurred. Only when the subject received two different antibiotics as well as bismuth subcitrate was his infection finally cured three years later. *—M.J.B.*

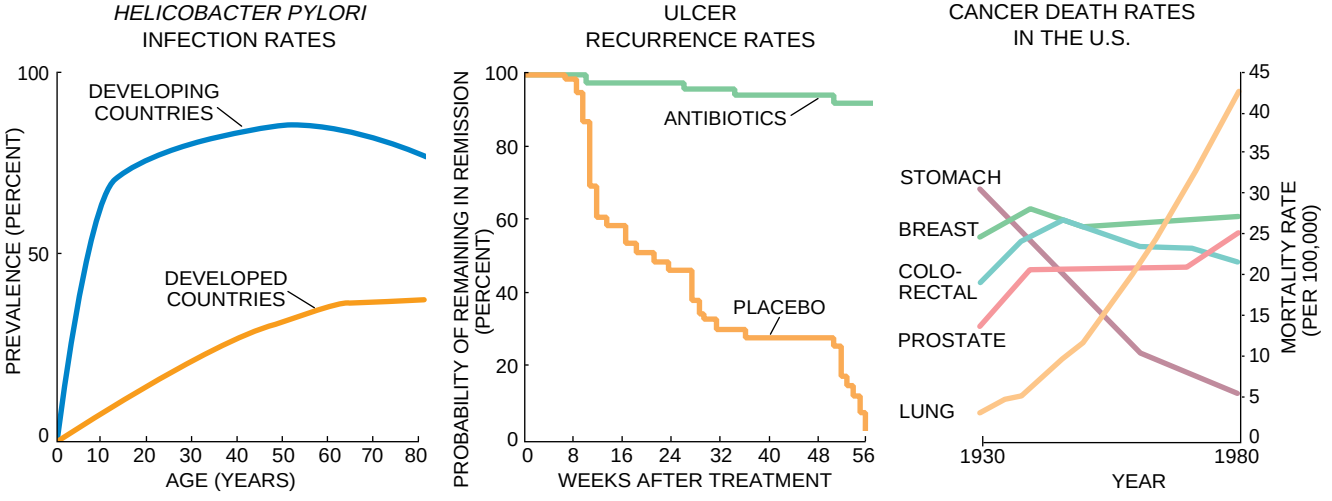
ria. Most important, antimicrobial medications can cure *H. pylori* infection and gastritis, thus markedly lowering the chances that a patient's ulcers will return. But few people can overcome *H. pylori* infection without specific antibiotic treatment.

When someone is exposed to *H. pylori,* his or her immune system reacts by making antibodies, molecules that can bind to and incapacitate some of the invaders. These antibodies cannot eliminate the microbes, but a blood test readily reveals the presence of antibodies, and so it is simple to detect infection. Surveys consistently show that one third to one half of the world's population carry *H. pylori.* In the U.S. and western Europe, children rarely become infected, but more than half of all 60-year-olds have the bacteria. In contrast, 60 to 70 percent of the children in developing countries show positive test results by age 10, and the infection rate remains high for adults. *H. pylori* infection is also common among institutionalized children.

Although it is as yet unclear how the organisms pass from one person to another, poor sanitation and crowding clearly facilitate the process. As living conditions have improved in many parts of the world during the past century, the rate of *H. pylori* infection has decreased, and the average age at which the infection is acquired has risen. Gastric cancer has also become progressively less common during the past 80 years. At the start of the 20th century, it was the leading cause of death from cancer in the U.S. and many other developed countries. Now it is far down on the list. The causes for its decline are not well understood, but we have reason to believe that the drop in *H. pylori* infection rates deserves some credit.

### A Connection to Cancer

In the 1970s Pelayo Correa, now at Louisiana State University Medical Center, proposed that gastric cancer resulted from a series of changes in the stomach taking place over a long period. In Correa's model, a normal stomach would initially succumb to chronic superficial gastritis for unknown reasons. We now know that *H. pylori* are to blame. In the second step—lasting for perhaps several decades—this gastritis would cause more serious harm in the form of a lesion, called atrophic gastritis. This lesion might then lead to fur-




DIMITRY SCHIDLOVSKY

**RATES OF INFECTION with *H. pylori* vary throughout the world. In developed countries, the infection is rare among children, but its prevalence rises with age. In developing countries, far more people are infected in all age groups (*left*). Supporting the fact that such infections cause ulcer disease, Enno Hentschel and his colleagues at Hanusch Hospital in Vienna found that antimicrobial therapy dramatically decreased the chance that a duodenal ulcer would recur (*center*). As infection rates have declined during the past century in the U.S., so, too, have the number of deaths from stomach cancer (*right*)—suggesting that *H. pylori* infection can, under some circumstances, cause that disease as well.**

ther changes, among them intestinal metaplasia and dysplasia, conditions that typically precede cancer. The big mystery since finding *H. pylori* has been: Could the bacteria account for the second transition—from superficial gastritis to atrophic gastritis and possibly cancer—in Correa's model?

The first real evidence linking *H. pylori* and gastric cancer came in 1991 from three separate studies. All three had similar designs and reached the same conclusions, but I will outline the one in which I participated, working with Abraham Nomura of Kuakini Medical Center in Honolulu. I must first give some background. In 1942, a year after the bombing of Pearl Harbor, the selective service system registered young Japanese-American men in Hawaii for military service. In the mid-1960s medical investigators in Hawaii examined a large group of these men—those born between 1900 and 1919—to gain information on the epidemiology of heart disease, cancer and other ailments. By the late 1960s they had assembled a cohort of about 8,000 men, administered questionnaires and obtained and frozen blood samples. They then tracked and monitored these men for particular diseases they might develop.

For many reasons, by the time we began our study we had sufficient information on only 5,924 men from this original group. Among them, however, 137 men, or more than 2 percent, had acquired gastric cancer between 1968 and 1989. We then focused on 109 of these patients, each of whom was matched with a healthy member of the cohort. Next, we examined the blood samples frozen in the 1960s for antibodies to *H. pylori*. One strength of this study was that the samples had been taken from these men, on average, 13 years before they were diagnosed with cancer. With the results in hand, we asked the critical question: Was evidence of preexisting *H. pylori* infection associated with gastric cancer? The answer was a strong yes. Those men who had a prior infection had been six times more likely to acquire cancer during the 21-year follow-up period than had men showing no signs of infection. If we confined our analysis to cancers affecting the lower part of the stomach—an area where *H. pylori* often collect—the risk became 12 times as great.

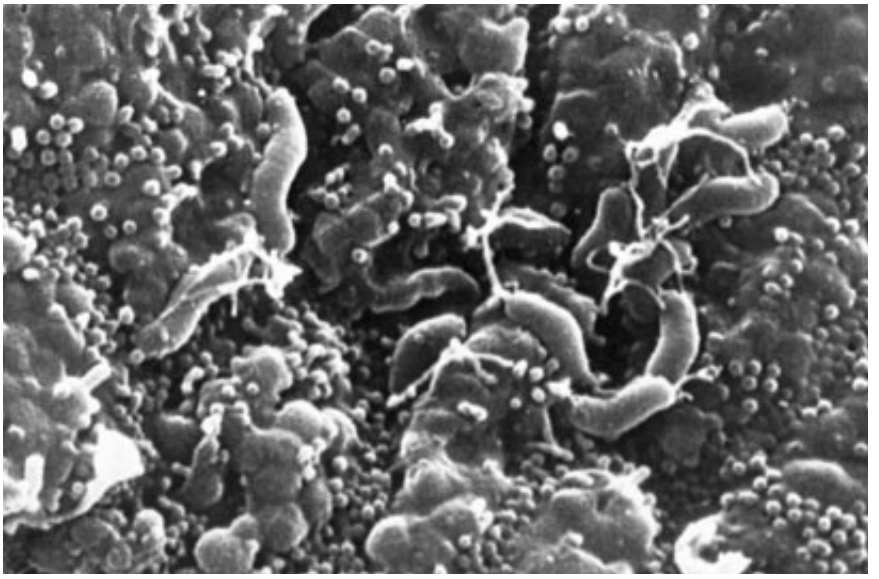
The other two studies, led by Julie Parsonnet of Stanford University and by David Forman of the Imperial Cancer Research Fund in London, produced like findings but revealed slightly lower risks. Over the past five years, further epidemiological and pathological investigations have confirmed the association of *H. pylori* infection and gastric cancer. In June 1994 the International Agency for Research in Cancer, an arm of the World Health Organization, declared that *H. pylori* is a class-1 carcinogen—the most dangerous rank given to cancer-causing agents. An uncommon cancer of the stomach, called gastric lymphoma, also appears to be largely caused by *H. pylori*. Recent evidence suggests that antimicrobial treatment to cure *H. pylori* infection may bring about regression in a subset of tumors of this kind, which is an exciting development in both clinical medicine and cancer biology.

### How Persistence Takes Place

Certainly most bacteria cannot survive in an acidic environment, but *H. pylori* are not the only exception. Since that bacteria's discovery, scientists have isolated 11 other organisms from the stomachs of other primates, dogs, cats, rodents, ferrets and even cheetahs. These bacteria, for now considered to be members of the *Helicobacter* family, seem to have a common ancestor. All are spiral-shaped and highly motile (they swim well)—properties enabling them to resist the muscle contractions that regularly empty the stomach. They grow best at oxygen levels of 5 percent, matching the level found in the stomach's mucus layer (ambient air is 21 percent oxygen). In addition, these microbes all manufacture large amounts of an enzyme called urease, which cleaves urea into ammonia and carbon dioxide. Fostering the production of ammonia may be one way helicobacters neutralize the acid in their local environment, further securing their survival.

An interesting puzzle involves what *H. pylori* eat. There are two obvious guesses: the mucus in which it lives and the food its human host ingests. But Denise Kirschner of Texas A&M University and I constructed a mathematical model showing that *H. pylori* would not be able to persist for years relying on those nutrient sources. In our model, the mathematics of persistence in the stomach requires some regulated interaction between the host cells and the bacteria. Inflammation provides one such interaction, and so I have proposed that *H. pylori* might trigger inflammation for the purpose of acquiring nutrients. An apparent paradox in *H. pylori* biology is that although the organisms do not invade the gastric tissue, they can cause irritation there. Rather, as we and others have found, the microbes release chemicals that the stomach tissue absorbs. These compounds attract phagocytic cells, such as leukocytes and macrophages, that induce gastritis.

The host is not entirely passive while *H. pylori* bombard it with noxious substances. Humans mount an immune response, primarily by making antibodies to the microbe. This response apparently does not function well, though, because the infection and the antibodies almost inevitably coexist for decades. In essence, faced with a pathogen that cannot be easily destroyed, humans had two evolutionary options: we could



COURTESY OF ARTHUR J. MORRIS, GORDON I. NICHOLSON AND MARTIN J. BLASER

**CURVED ORGANISMS, magnified 8,000 times, are *H. pylori* in the stomach of the second human volunteer to ingest the bacteria. The man had consumed the microbes 463 days earlier and developed chronic superficial gastritis as a result.**

have evolved to fight *H. pylori* infection to its death, possibly involving the abrogation of normal gastric function, or we could have become tolerant and tried to ignore the organisms. I believe the choice was made long ago in favor of tolerance. The response to other persistent pathogens—such as the microbes responsible for malaria and leprosy—may follow the same paradigm, in which it is adaptive for the host to dampen its immune reaction.

Fortunately, it is not in *H. pylori*'s best interest to take advantage of this passivity, growing to overwhelming numbers and ultimately killing its host. First, doing so would limit the infection's opportunity to spread. Second, even in a steady state, *H. pylori* reaches vast numbers (from $10^7$ to $10^{10}$ cells) in the stomach. And third, further growth might exhaust the mechanisms keeping the immune system in check, leading to severe inflammation, atrophic gastritis and, eventually, a loss of gastric acidity. When low acidity occurs, bacteria from the intestines, such as *Escherichia coli,* are free to move upstream and colonize the stomach. Although *H. pylori* can easily live longer than *E. coli* in an acid environment, *E. coli* crowds *H. pylori* out of more neutral surroundings. So to avoid any competition with intestinal bacteria, *H. pylori* must not cause too much inflammation, thereby upsetting the acid levels in the stomach.

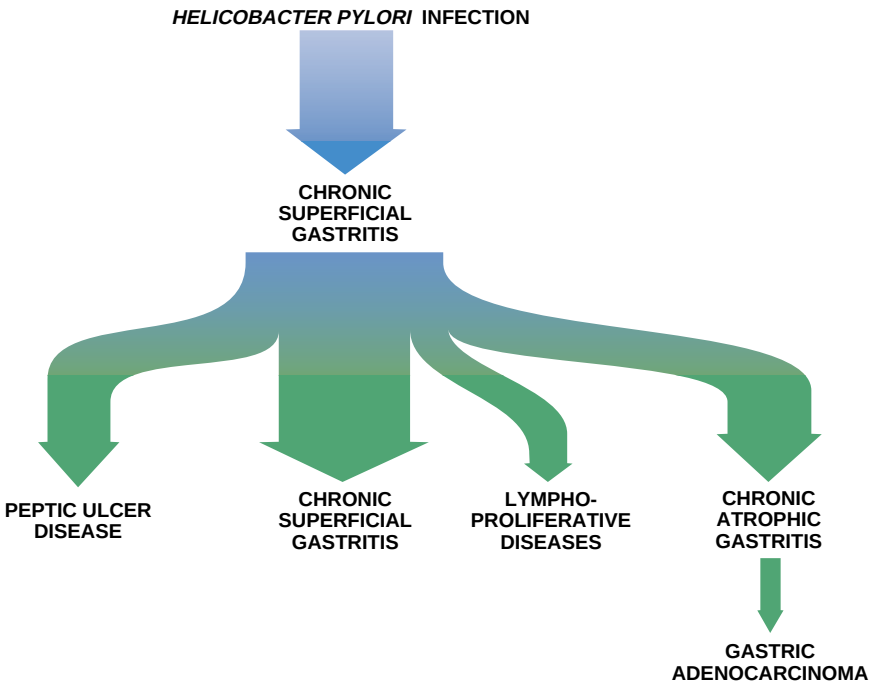
Are *H. pylori* symbionts that have only recently evolved into disease-causing organisms? Or are they pathogens on the long and as yet incomplete road toward symbiosis? We do not yet know, but we can learn from the biology of *Mycobacterium tuberculosis,* the agent responsible for tuberculosis. It, too, infects about one third of the world's population. But as in *H. pylori* infection, only 10 percent of all infected people become sick at some point in their life; the other 90 percent experience no symptoms whatsoever. The possible explanations fall into several main categories. Differences among microbial strains or among hosts could explain why some infected people acquire certain diseases and others do not. Environmental cofactors, such as whether someone eats well or smokes, could influence the course of infection. And the age at which someone acquires an infection might alter the risks. Each of these categories affects the outcome of *H. pylori* infection, but I will describe in the next section the microbial differences.

### Not All Bacteria Are Created Equal

Given its abundance throughout the world, it is not surprising that *H. pylori* are highly diverse at the genetic level. The sundry strains share many structural, biochemical and physiological characteristics, but they are not all equally virulent. Differences among them are associated with variations in two genes. One encodes a large protein that 60 percent of all strains produce. Our group at Vanderbilt University, comprising Murali Tummuru, Timothy L. Cover and myself, and a group at the company Biocine in Italy, led by Antonello Covacci and Rino Rappuoli, identified and cloned the gene nearly simultaneously in 1993 and by agreement called it *cagA*. Among patients suffering from chronic superficial gastritis alone, about 50 to 60 percent are infected by *H. pylori* strains having the *cagA* gene. In contrast, nearly all individuals with duodenal ulcers bear cagA strains. Recently we reexamined the results of the Hawaiian study and found that infection by a cagA strain was associated with a doubled risk of gastric cancer. Research done by Jean E. Crabtree of Leeds University in England and by the Vanderbilt group has shown that persons infected by cagA strains experience more severe inflammation and tissue injury than do those infected by strains lacking the *cagA* gene.

The other *H. pylori* gene that seems to influence disease encodes for a toxin. In 1988 Robert D. Leunk, working for Procter & Gamble—the makers of bismuth subsalicylate (Pepto-Bismol)—reported that a broth containing *H. pylori* could induce the formation of vacuoles, or small holes, in tissue cultures. In my group, Cover had clearly shown that a toxin caused this damage and that it was being made not only by *H. pylori* grown in the laboratory but also by those residing in human hosts. In 1991 we purified the toxin and confirmed Leunk's finding that only 50 to 60 percent of *H. pylori* strains produced it. Our paper was published in May 1992 and included a brief sequence of some of the amino acids that encode for the mature toxin. Based on that scanty information, within the next year four groups—two in the U.S., including our own, one in Italy and one in Germany—were able to clone the gene, which we all agreed to name *vacA*. The race to publish was on. Each of our four papers appeared in separate journals within a three-month period.

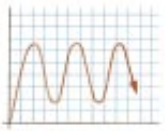
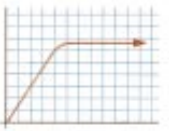
Lest this sounds like duplicated labor, I should point out that each team had in fact solved a different aspect of the problem. We learned, for example, that virtually all *H. pylori* strains possess *vacA,* whether or not they produce the toxin when grown in culture. We also discovered that there is an extraordinary amount of strain-to-strain variability in *vacA* itself. In addition, broth from toxin-producing strains inoculated directly into the stomach of mice brought




DIMITRY SCHIDLOVSKY

***H. PYLORI* INFECTION PROGRESSES to chronic superficial gastritis within months. Left untreated, the condition persists for life in most people. A small fraction, however, may develop peptic ulcer disease, lymphoproliferative diseases or severe chronic atrophic gastritis, leading to adenocarcinoma of the stomach.**

## Which Treatment Strategy Should You Choose?

| | OLD MODEL | NEW MODEL |
|---|---|---|
| **CAUSE**  | **Excess stomach acid** eats through tissues and causes inflammation | ***Helicobacter pylori*** bacteria secrete toxins and cause inflammation in the stomach, bringing about damage  |
| **TREATMENTS**  | **Bland diet**, including dairy products every hour, small meals, no citrus or spicy foods and no alcohol or caffeine<br>**$H_2$-receptor blockers** lessen blood levels of histamine, which increases the production of stomach acid<br>**Surgery** to remove ulcers that do not respond to medication or that bleed uncontrollably. In the 1970s, it was the most common operation a surgical resident learned. Now it is increasingly rare | **Antibiotic regimen.** In February 1994 an NIH panel endorsed a two-week course of antibiotics for treating ulcer disease: amoxicillin or tetracycline, metronidazole (Flagyl) and bismuth subsalicylate (Pepto-Bismol). In December 1995 an FDA advisory committee recommended approval of two new four-week treatments, involving clarithromycin (Biaxin) with either omeprazole (Prilosec) or ranitidine bismuth citrate (Tritec). One-week therapies are also highly effective  |
| **SUCCESS**  | Patients who stop taking $H_2$-receptor blockers face a **50 percent chance** that their ulcers will recur within six months and a 95 percent chance that they will reappear within two years | **No recurrence** after the underlying bacterial infection is eliminated  |
| **COST**  | $H_2$-receptor blockers cost from $60 to $100 per month, adding up to **thousands of dollars** over decades of care. Surgery can cost as much as $18,000 | **Less than $200** for a standard one-week therapy  |

ROBERTO OSTI (*drawings*); LISA BURNETT

about substantial injury. Strains that produce the toxin are some 30 to 40 percent overrepresented in ulcer patients compared with those having gastritis alone. And toxigenic strains usually but not always contain *cagA,* which is located far away from *vacA* on the chromosome.

### Slow-Acting Bacteria and Disease

Over the past 15 years, researchers and physicians have learned a good deal about *H. pylori.* This knowledge has revolutionized our understanding of gastritis, formerly thought to represent the aging stomach, and of peptic ulcer disease and gastric cancer. It has made possible new treatments and screening methods. In addition, a new field of study has emerged—the microbiology and immunology of the human stomach—that will undoubtedly reveal more about persistent infections within mucosal surfaces.

But let us extrapolate from these findings. Consider that slow-acting bacteria, *H. pylori,* cause a chronic inflammatory process, peptic ulcer disease, that was heretofore considered metabolic. And also keep in mind that this infection greatly enhances the risk of neoplasms developing, such as adenocarcinomas and lymphomas. It seems reasonable, then, to suggest that persistent microbes may be involved in the etiology of other chronic inflammatory diseases of unknown origin, such as ulcerative colitis, Crohn's disease, sarcoidosis, Wegener's granulomatosis, systemic lupus erythematosus and psoriasis, as well as various neoplasms, including carcinomas of the colon, pancreas and prostate. I believe *H. pylori* are very likely the first in a class of slow-acting bacteria that may well account for a number of perplexing diseases that we are facing today.

### *The Author*

MARTIN J. BLASER has been the Addison B. Scoville Professor of Medicine and director of the Division of Infectious Diseases at Vanderbilt University and at the Nashville Veterans Affairs Medical Center since 1989. He has worked at the Rockefeller University, the University of Colorado, the Denver Veterans Administration Medical Center, the Centers for Disease Control and Prevention, and St. Paul's Hospital in Addis Ababa, Ethiopia. He received a B.A. with honors in economics from the University of Pennsylvania in 1969 and an M.D. from New York University in 1973. He holds several patents and is a member of numerous professional societies and editorial boards. He has written more than 300 articles and edited several books.

### *Further Reading*

UNIDENTIFIED CURVED BACILLI IN THE STOMACH OF PATIENTS WITH GASTRITIS AND PEPTIC ULCERATION. Barry J. Marshall and J. Robin Warren in *Lancet,* No. 8390, pages 1311–1315; June 16, 1984.

HELICOBACTER PYLORI INFECTION AND GASTRIC CARCINOMA AMONG JAPANESE AMERICANS IN HAWAII. A. Nomura, G. N. Stemmermann, P.-H. Chyou, I. Kato, G. I. Perez-Perez and M. J. Blaser in *New England Journal of Medicine,* Vol. 325, No. 16, pages 1132–1136; October 17, 1991.

HUMAN GASTRIC CARCINOGENESIS: A MULTISTEP AND MULTIFACTORIAL PROCESS. Pelayo Correa in *Cancer Research,* Vol. 52, No. 24, pages 6735–6740; December 15, 1992.

EFFECT OF RANITIDINE AND AMOXICILLIN PLUS METRONIDAZOLE ON THE ERADICATION OF HELICOBACTER PYLORI AND THE RECURRENCE OF DUODENAL ULCER. Enno Hentschel et al. in *New England Journal of Medicine,* Vol. 328, No. 5, pages 308–312; February 4, 1993.

REGRESSION OF PRIMARY LOW-GRADE B-CELL GASTRIC LYMPHOMA OF MUCOSA-ASSOCIATED LYMPHOID TISSUE TYPE AFTER ERADICATION OF HELICOBACTER PYLORI. A. C. Wotherspoon et al. in *Lancet,* Vol. 342, No. 8871, pages 575–577; September 4, 1993.

PARASITISM BY THE "SLOW" BACTERIUM HELICOBACTER PYLORI LEADS TO ALTERED GASTRIC HOMEOSTASIS AND NEOPLASIA. Martin J. Blaser and Julie Parsonnet in *Journal of Clinical Investigation,* Vol. 94, No. 1, pages 4–8; July 1994.

# The Loves of the Plants

*Carl Linnaeus classified plants according to their reproductive parts, endowing them as well with sex lives reflecting 18th-century values and controversies*

by Londa Schiebinger

> The flowers' leaves…serve as bridal beds which the Creator has so gloriously arranged, adorned with such noble bed curtains, and perfumed with so many soft scents that the bridegroom with his bride might there celebrate their nuptials with so much the greater solemnity. When now the bed is so prepared, it is time for the bridegroom to embrace his beloved bride and offer her his gifts…
>
> Carl Linnaeus, *Praeludia Sponsaliorum Plantarum,* 1729

From Aristotle through Darwin and beyond, observers have infused nature with sexuality and gender. The great Swedish taxonomist Carl Linnaeus was thus not alone in imagining that plants have vaginas and penises and reproduce on "marriage beds." Although naturalists tended to believe that gender was a given of nature, the traits they ascribed to organisms have changed with shifting notions of masculinity and femininity in Western culture. For Aristotle, mares were sexually wanton, going "a-horsing" to satisfy their unbridled appetites. But in later centuries, females throughout nature—with the exception of Linnaeus's lusty flowers—were said to evince a patient modesty. Even among insects, females were observed to "repel the first [sexual] attacks of the males" and, in so doing, to win the respect of their paramours.

Since the Enlightenment, science has stirred hearts and minds with its promise of a neutral and privileged viewpoint, above and beyond the rough and tumble of political life. With respect to women, however, science is not a neutral culture. Gender—both the real relations between the sexes and cultural renderings of those relations—shaped European natural history and, in particular, botany. Crucial to this story is that Europeans who wrote about nature in this era were almost exclusively male.

It is ironic in this context that botany was long considered especially suited to women. Botany's reputation as "unmanly," an ornamental branch of science appropriate only for "ladies and effeminate youths," was such that it was sometimes questioned whether able-bodied young men should pursue it at all. (Hegel once compared the minds of women to plants because, in his view, both were essentially placid.) Eighteenth-century society condoned botany as a fitting pastime for middle-class ladies because it took them out into the air and taught them a certain intellectual discipline. This attitude rested in part on botany's historical affiliations with herbal healing—a realm in which women had traditionally been active in their role as midwives.

Then, too, botany among all the sciences was considered least offensive to the delicate spirit. As Rousseau pointed out, the student of anatomy was faced with oozing blood and stinking cadavers, entomologists with vile insects, geologists with dirt and filth. After Linnaeus, the study of plants seemed to call for more attention to sexuality than might seem suitable for ladies. Still, botany continued to be advocated for women, especially in England, as the science leading to the greatest appreciation of God and His universe.

There was, of course, a caveat. The most important directive issued to women was that their ambitions in botany should not transcend those of the amateur. The English botanist Thomas Martyn warned women away from mastering "long files of Latin words" and encouraged his "fair countrywomen" merely to amuse themselves with natural history. Serious science in any field was to be reserved for men. Women, Jean-Jacques Rousseau and others taught, lacked a certain genius—that "celestial flame" that sets fire to the soul—required for true innovation in science. Women of Rousseau's day were formally barred from universities and scientific academies. The few prominent women scientists in early modern Europe had maneuvered themselves into mathematics, physics, astronomy or scientific illustration through less formal routes, such as the salons of the elite and the craft guilds. Yet, even these women learned to recognize certain limits. Maria Sibylla Merian, the adventurous botanical illustrator and entomologist who traveled to Dutch Surinam in the late 17th century in search of exotic caterpillars, presented her observations in exquisitely crafted volumes but left classification to her male colleagues.

Despite their forays into botany, women were long to remain on the margins of intellectual life. As we shall see, parochial European notions of sex and sexual hierarchy became potent principles organizing 18th-century natural history—a matter of consequence in an age that looked to nature as the guiding light for social reform.

### The Private Lives of Plants

Not until the 17th century did European naturalists widely recognize that plants reproduce sexually. The ancient world, it is true, had some knowledge of sexual distinctions among plants. The Greek naturalist Theophrastus knew the age-old practice of fertilizing date palms by bringing male flowers to the female tree. Peasants working the land also recognized sexual distinctions in trees such as the pistachio. Plant sexuality, however, was not a focus of interest among naturalists of the ancient world. Eighteenth-century observers commonly charged that the ancients were ignorant of the basics: they sometimes called the seed-bearing plant "the male" and the barren plant "the female."

Between 1550 and 1700, as wonders from the voyages of discovery and the new colonies flooded Europe, the num-

**CUPID inspires the plants with love in this plate from Robert Thornton's *Temple of Flora* (London, 1805). The "romantic lives" of vegetation aroused much interest after Linnaeus's sexual system for classifying plants gained currency.**

COURTESY OF PENNSYLVANIA STATE UNIVERSITY

P. DE BAY Courtesy of the Trustees of the Victoria and Albert Museum

**LIFE CYCLE OF THE SAFFRON CROCUS is depicted by John Miller in *An Illustration of the Sexual System of Linnaeus* (1770). The hand-colored engraving dissects the parts of the plant, especially the flower, with its three stamens (male organs) and one pistil (female organ). The number of sexual parts leads to the plant's classification as *triandria monogynia*.**

ber of known plants quadrupled. Botanists were led to search for new methods of organizing this multitude of new specimens. As they sought simple principles that would hold universally, emphasis shifted from the medical uses of plants to more general and theoretical issues of pure taxonomy.

When plant sexuality exploded onto the scene, interest in assigning sex to plants outran understanding of botanical fertilization, or the "coitus of vegetables," as it was sometimes called. The English naturalist Nehemiah Grew, who developed his theory of plant sexuality from his knowledge of animals, first identified the stamen as the male organ in flowers in his 1682 treatise, *The Anatomy of Plants:*

> The blade (or stamen) does not unaptly resemble a small penis, with the sheath upon it, as its praeputium [prepuce]. And the…several thecae, are like so many little testicles. And the globulets [pollen] and other small particles upon the blade or penis…are as the vegetable sperme. Which as soon as the penis is erected, falls down upon the seed-case or womb, and so touches it with a prolific virtue.

By the early 1700s, the analogy between animal and plant sexuality was fully established. In his *Praeludia Sponsaliorum Plantarum,* Linnaeus related the terms of comparison: in the male the filaments of the stamens are the vas deferens, the anthers are the testes and the pollen that falls from them is the seminal fluid; in the female the stigma is the vulva, the style becomes the vagina, the tube running the length of the pistil is the fallopian tube, the pericarp is the impregnated ovary and the seeds are the eggs. The French physician Julien Offray de La Mettrie, along with other naturalists of the time, even claimed the honey reservoir found in a plant's nectary gland to be equivalent to mother's milk in humans.

Most flowers, however, are hermaphroditic, with both male and female reproductive organs. As one 18th-century botanist put it, there are two sexes—male and female—but three kinds of flowers: male, female and hermaphrodite, sometimes called androgyne. Although most botanists enthusiastically embraced sexual dimorphism, conceiving of plants as hermaphroditic was more difficult: they could not or would not recognize an unfamiliar sexual type. Even 40 years later, when Linnaeus's system was in wide use, William Smellie, chief compiler of the first edition of the *Encyclopaedia Britannica* (1771), rejected the whole idea of sexuality in plants and distanced himself from the term "hermaphrodite," noting when using the word that he merely spoke "the language of the system."

### The Bridal Bed

Still, the majority of European botanists gave undue primacy to sexual reproduction. Linnaeus was thoroughly taken with heterosexual coupling, confessing that "the singular structure and remarkable office of the stamens and pistil enticed my mind, to inquire what Nature had concealed in them. They commended themselves by the function they perform." He thus attributed sexual reproduction even to his *cryptogamia*—"plants that marry secretly"—by which he meant ferns, mosses, algae and fungi, whose reproductive habits were then not understood. The very fact that nonsexual reproduction is called asexual reveals the normative preference given sexual reproduction.

Not only were Linnaeus's plants sexed, they actually became "husbands" and "wives." Introducing new terminology to describe flowers, Linnaeus rejected the increasingly standard terms "stamen" and "pistil," for *andria* and *gynia*—suffixes that he derived from the Greek for husband (*aner*) and wife (*gyne*). Linnaeus's "Key to the Sexual System," published in *Systema Naturae* in 1737, was built upon the *nuptiae plantarum:* the marriages of plants. If male and female flowers occurred on the same plant, they shared the same house (*monoecia*) but not the same bed; if on separate plants, they lived in two houses (*dioecia*). Hermaphroditic flowers contained husbands and wives in one bed (*monoclinia*).

Linnaeus's classes of plants, based on

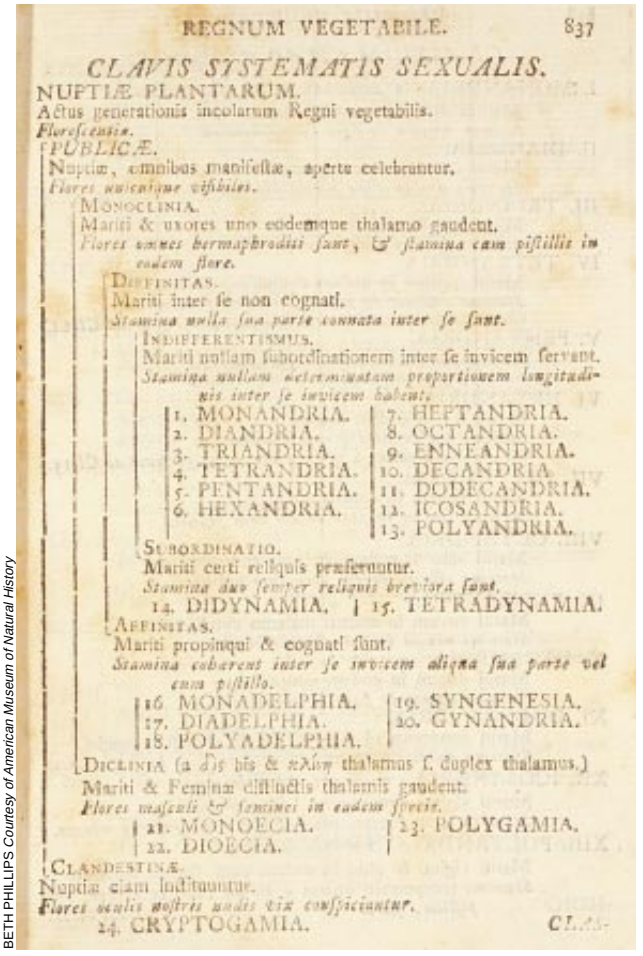
the number, proportion and position of the male stamens, end in *andria*. *Monandria* ("having only one man") signified one stamen, or husband, on a hermaphroditic flower; *diandria* signified two stamens and so on. The classes were subdivided into roughly 65 orders, based on the number, proportion and position of the female pistils: *monogynia, digynia, trigynia* and so forth. Thus, a saffron crocus, having three stamens and one pistil, would be called *triandria monogynia*. The orders were further divided into genera, based on the calyx, flower and other parts of the fruit; then again into species, based on the leaves or some other characteristic of the plant; and finally into varieties.

Linnaeus emphasized the "nuptials" of living plants as much as their sexual relations. Before their "lawful marriage," trees and shrubs donned "wedding gowns." Flower petals opened as "bridal beds" for a verdant groom and his cherished bride, while the curtain of the corolla lent privacy to the amorous newlyweds. Linnaeus divided the plant world according to the type of marriage each plant contracted—whether, for example, it had been wed "publicly" or "clandestinely." (The latter group consisted of the *cryptogamia*.) These two types of marriage, in fact, characterized custom in much of Europe at that time; only in 1753 in England did Lord Harwicke's Marriage Act do away with clandestine marriages by requiring a public proclamation (banns).

It is significant that Linnaeus, a Swedish country parson's son, focused on marriage when he thought of sexuality. European marriage customs were undergoing rapid change as the traditional anchors of the old order began to give way. Upper-class parents and even well-off peasants less often arranged marriages for their children out of property considerations alone. Increasingly love and affection became legitimate reasons to wed. Linnaeus's own marriage followed this pattern. He courted with tender expressions of love one Sara Moraea, the daughter of a wealthy physician. Linnaeus then left the running of his house entirely to his wife, while he concerned himself with the workings of nature.

Linnaeus's plants may have celebrated their nuptials, yet the majority did not engage in "lawful marriages." Only one class of plants, the *monandria*—exemplified by the tropical genus *Canna*—practiced "monogamy." Plants in other classes joined in liaisons consisting of two, three, 20 or more "husbands" who shared their marriage bed (that is, the petals of the flower) with one wife. The common iris, for example, enjoyed three husbands.

Erasmus Darwin, one of Linnaeus's many popularizers in England (and the grandfather of Charles), did not limit sexual relations to the bonds of holy matrimony. In his *Loves of the Plants* (1789), Darwin's plants freely expressed



REGNUM VEGETABILE. 837

*CLAVIS SYSTEMATIS SEXUALIS.*

NUPTIÆ PLANTARUM.

Actus generationis incolarum Regni vegetabilis.

*Florescentia.*

PUBLICÆ.

Nuptiæ, omnibus manifestæ, apertis celebrantur.

*Flores unicuique visibiles.*

MONOCLINIA.

Mariti & uxores uno eodemque thalamo gaudent.

*Flores omnes hermaphroditi sunt, & stamina cum pistillis in eodem flore.*

DEFINITAS.

Mariti inter se non cognati.

*Stamina nulla sua parte connata inter se sunt.*

INDIFFERENTISMUS.

Mariti nullam subordinationem inter se invicem servant.

*Stamina nullam determinatam proportionem longitudinis inter se invicem habent.*

| | |
|---|---|
| 1. MONANDRIA. | 7. HEPTANDRIA. |
| 2. DIANDRIA. | 8. OCTANDRIA. |
| 3. TRIANDRIA. | 9. ENNEANDRIA. |
| 4. TETRANDRIA. | 10. DECANDRIA. |
| 5. PENTANDRIA. | 11. DODECANDRIA. |
| 6. HEXANDRIA. | 12. ICOSANDRIA. |
| | 13. POLYANDRIA. |

SUBORDINATIO.

Mariti certi reliquis præferuntur.

*Stamina duo semper reliquis breviora sunt.*

14. DIDYNAMIA. | 15. TETRADYNAMIA.

AFFINITAS.

Mariti propinqui & cognati sunt.

*Stamina cohærent inter se invicem aliqua sua parte vel cum pistillo.*

| | |
|---|---|
| 16. MONADELPHIA. | 19. SYNGENESIA. |
| 17. DIADELPHIA. | 20. GYNANDRIA. |
| 18. POLYADELPHIA. | |

DICLINIA (a δὶς bis & κλίνω thalamus s. duplex thalamus.)

Mariti & Feminæ diffinctis thalamis gaudent.

*Flores masculi & femines in eodem specie.*

| | |
|---|---|
| 21. MONOECIA. | 23. POLYGAMIA. |
| 22. DIOECIA. | |

CLANDESTINÆ.

Nuptiæ clam inditantur.

*Flores oculis nostris nudis vix conspiciuntur.*

24. CRYPTOGAMIA.

BETH PHILLIPS *Courtesy of American Museum of Natural History*

VEGETABLE KINGDOM

*KEY OF THE SEXUAL SYSTEM*

MARRIAGES OF PLANTS.
*Florescence.*

- PUBLIC MARRIAGES.
  *Flowers visible to every one.*
  - IN ONE BED.
    Husband and wife have the same bed.
    *All the flowers hermaphrodite: stamens and pistils in the same flower.*
    - WITHOUT AFFINITY.
      Husbands not related to each other.
      *Stamens not joined together in any part.*
      - WITH EQUALITY.
        All the males of equal rank.
        *Stamens have no determinate proportion of length.*

| | |
|---|---|
| 1. ONE MALE. | 7. SEVEN MALES. |
| 2. TWO MALES. | 8. EIGHT MALES. |
| 3. THREE MALES. | 9. NINE MALES. |
| 4. FOUR MALES. | 10. TEN MALES. |
| 5. FIVE MALES. | 11. TWELVE MALES. |
| 6. SIX MALES. | 12. TWENTY MALES. |
| | 13. MANY MALES. |

      - WITH SUBORDINATION.
        Some males above others.
        *Two stamens are always lower than the others.*

| | |
|---|---|
| 14. TWO POWERS. | 15. FOUR POWERS. |

    - WITH AFFINITY.
      Husbands related to each other.
      *Stamens cohere with each other, or with the pistil.*

| | |
|---|---|
| 16. ONE BROTHERHOOD. | 19. CONFEDERATE MALES. |
| 17. TWO BROTHERHOODS. | 20. FEMININE MALES. |
| 18. MANY BROTHERHOODS. | |

  - IN TWO BEDS.
    Husband and wife have separate beds.
    *Male flowers and female flowers in the same species.*

| | |
|---|---|
| 21. ONE HOUSE. | 23. POLYGAMIES. |
| 22. TWO HOUSES. | |

- CLANDESTINE MARRIAGES.
  *Flowers scarce visible to the naked eye.*
  24. CLANDESTINE MARRIAGES.

**LINNAEUS'S KEY TO THE SEXUAL SYSTEM is elaborated in this chart (*left*) from his *Systema Naturae* (1759 edition). The translation (*right*), from *A General System of Nature, Through the Three Grand Kingdoms,* by William Turton (1802), reveals that plants are classified according to the types of marriages they contract. Plants with visible flowers have public marriages; others have clandestine marriages. If male and female flowers occur on the same plant, then "husband" and "wife" share the same house, but if male and female parts occur in the same flower, the couples share the same bed. The number of husbands in the bed defines the classes of plants. The orders, based on the number of wives, are not shown.**

FROM *OVER DE VOORTTELING EN WONDERBAERLYKE VERANDERINGEN DER SURINAMSCHE INSECTEN* (1730); P. DE BAY *Courtesy of the Trustees of the Victoria and Albert Museum*

COURTESY OF ÖFFENTLICHE KUNSTSAMMLUNG, KUPFERSTICHKABINETT, BASEL

**MARIA SIBYLLA MERIAN (*above*) traveled to Dutch Surinam in 1699, at the age of 52, to draw its flora and fauna. This hand-colored engraving (*left*) depicts the passionflower and its fruit. Although Merian was an accomplished observer of plant and insect life, she did not venture into the male province of classification.**

every imaginable form of heterosexual union. The fair *Collinsonia,* sighing with sweet concern, satisfied the love of two brothers by turns. The *Meadia*—an ordinary cowslip—bowed with "wanton air," rolled her dark eyes and waved her golden hair as she gratified each of her five beaux. Darwin portrayed the tragic outcome of an ordinary *Gloriosa superba* repulsing the incestuous advances of her son:

> Fierce on the fair he fix'd his ardent gaze;
> Dropped on one knee, his frantic arms outspread,
> And stole a guilty glance toward the bed;
> Then breath'd from quivering lips a whisper'd vow
> And bent on heaven his pale unrepentant brow;
> "Thus, thus!" he cried, and plung'd the furious dart,
> And life and love gush'd mingled from his heart.

Darwin may well have been using the cover of botany to propagandize for the free love he practiced after the death of his first wife. There is no evidence, however, that Linnaeus intended his sexual vision of botany to undermine social custom. Raised in an upright, thrifty, Protestant family, he was conservative in his religious views (all of nature celebrated the glory of its Creator) and in his attitudes toward women. He would not allow his four daughters to learn French for fear that along with the language they would adopt the liberties of French custom. When his wife placed their daughter Sophia in school, Linnaeus immediately took her out again, stopping what he considered "nonsensical" education. (He did, however, allow his eldest daughter to develop a mild interest in botany: Elisabeth Christina contributed a paper entitled "Remarks on a Luminous Appearance of the Indian Cresses" to the *Transactions* of the Royal Academy of Science.)

It seems unlikely, then, that Linnaeus introduced his sexual imagery as an affront to middle-class sensibilities. Befitting the social order of the day, he simply saw anything female as a wife. He called Madeleine Basseporte, the celebrated botanical illustrator who worked at the Jardin du Roi in Paris, his second wife. He considered "Dame Nature" his other wife and true helpmate. Linnaeus called his own wife "my monandrian lily," the lily signifying virginity.

Linnaeus's system of classification was but one among many proposed systems. By 1799, 20 years after Linnaeus's death, when the English naturalist Robert Thornton published his popular version of the Linnaean system, he counted 52 different systems of bot-

any; the "system-madness," one authority complained, was truly "epidemical." A century earlier in England, John Ray had devised a means of establishing genera based on the flower, calyx, seed and seed coat. In France, Joseph Pitton de Tournefort defined genera principally on the characteristics of the corolla and fruit. And in 18th-century Switzerland, Albrecht von Haller argued that geography was crucial to an understanding of plant life and that development as well as appearance should be represented in a system of classification.

Despite the number and variety of systems, Linnaeus's, being simple and convenient, was widely adopted across Europe, and especially in Britain, after 1740. But bloody and protracted battles erupted almost immediately over the scientific and moral implications of Linnaeus's classification. "Anti-sexualists"—those opposing the system—attacked his work, beginning a controversy that spilled over into the next century. What man, fumed Johann Siegesbeck, a professor in St. Petersburg, could believe that God Almighty would introduce such "loathsome harlotry" into the plant kingdom? In 1790 Smellie blasted the "alluring seductions" of the analogies on which the sexualist hypothesis was founded, and he maintained that it did not hold up to facts of experience. Many animals (he mentioned polyps and millipedes) reproduced without sexual embraces, and if these were destitute of "all the endearments of love," what, he asked, should induce us to fancy that the oak or mushroom enjoyed these distinguished privileges?

In addition to his ontological qualms, Smellie denounced Linnaeus for taking his analogy "far beyond all decent limits," claiming that Linnaeus's metaphors were so indelicate as to exceed the most "obscene romance-writer." Smellie's sentiments were shared by others. In 1808 the Reverend Samuel Goodenough, later bishop of Carlisle, wrote to the Linnaean Society that "a literal translation of the first principles of Linnaean botany is enough to shock female modesty." In the face of such opposition, the authors who popularized Linnaeus's system in England made little use of his sexual imagery, with the audacious exception of Erasmus Darwin.

In the uproar that surrounded Linnaeus's ardent sexing of plants, no one noticed that his taxonomy, built as it was on sexual difference, imported into botany traditional beliefs about European sexual hierarchy. Linnaeus was among the first to highlight the biological importance of sexual reproduction in plants. But the success of his system did not rest on the fact that it was "natural," capturing God's order in nature—Linnaeus's desirable but still unattainable goal. Indeed, he readily acknowledged that his system was highly artificial. He focused on purely morphological features such as the number of sexual partners. But in fact, the number of stamens and pistils can vary among different flowers of the same plant. Linnaeus did not supply resolutions to such conflicts; he sometimes placed plants with different numbers of stamens in the same class, thus making nonsense of his numerical system.

### Taxonomic Sexism

Furthermore, Linnaeus devised his classification system in such a way that the number of a plant's stamens determined the class to which it was assigned, whereas the number of its pistils determined its order. In the taxonomic tree, class stands above order. In other words, Linnaeus gave male parts priority in determining the status of the organism in the plant kingdom. There is no empirical justification for this outcome; rather Linnaeus brought traditional tenets of gender hierarchy into science. Although today his classification of groups above the rank of genus has been abandoned, many of his genus and species labels have survived.

Why did the study of plant sexuality become a priority for many 18th-century botanists? There are, after all, many different ways of knowing nature. One important factor drawing naturalists' attention to plant and animal sexuality was a keen interest in gender differences among humans. This era of democratic awakening brought with it the "women question"—the question of women's social rights and privileges. Sexual difference weighed heavily on the minds of many, as the enlightened of Europe issued the challenge that "all men are by nature equal."

If women were not to be given equal rights in the new democratic states (and they were not), natural causes had to justify their exclusion. That Linnaeus supposedly found European sexual hierarchies reconfirmed within the plant kingdom indicated to thinkers of the time the "naturalness" of women's continued subordination to their fathers and husbands. Rousseau spoke for many when he wrote that natural philosophers were to read in the great book of nature "everything which suits the constitution of her [woman's] species and her sex in order to fulfill her place in the physical and moral order."

Had women been among 18th-century taxonomists, would the story have been different? It is difficult or even impossible to say. The sex of the scientist should not influence the results of science. But in the modern division of labor that crystallized during the Enlightenment, science was part of the terrain that fell to the male sex. Researchers read nature through the lens of social relations in such a way that the new language of botany, among other natural sciences, incorporated fundamental aspects of the social world as much as those of the natural world.

During the past few decades, the feminist critique of science combined with the process of more and more women becoming engaged as makers of knowledge has had a tremendous impact in the humanities, social sciences and many of the sciences. We are just beginning to unravel how deeply gender has been worked into nature's body. Historical exposé, of course, is not enough, for what we unravel by night is often rewoven by day in the institutions of science. Scientists need to become aware of not only how culture shapes science but also how what is studied (or neglected) grows out of who is doing the studying and toward what ends.


### *The Author*

LONDA SCHIEBINGER is professor of history and women's studies and founding director of the Women in the Sciences and Engineering Institute at Pennsylvania State University. She received her doctorate from Harvard University in 1984. Schiebinger is active in the movement to increase the number of women and minorities in the sciences and is currently writing a book on women in contemporary scientific culture to be published next year by Harvard University Press.


### *Further Reading*


FEMINISM AND SCIENCE. Edited by Nancy Tuana. Indiana University Press, 1989.

THE MIND HAS NO SEX? WOMEN IN THE ORIGINS OF MODERN SCIENCE. Londa Schiebinger. Harvard University Press, 1989.

MYTHS OF GENDER: BIOLOGICAL THEORIES ABOUT WOMEN AND MEN. Anne Fausto-Sterling. Basic Books, 1992.

SECRETS OF LIFE, SECRETS OF DEATH: ESSAYS ON LANGUAGE, GENDER AND SCIENCE. Evelyn Fox Keller. Routledge Press, 1992.

NATURE'S BODY: GENDER IN THE MAKING OF MODERN SCIENCE. Londa Schiebinger. Beacon Press, 1993.

# Quarks by Computer

*Yearlong computations have helped to confirm the fundamental theory behind quarks—and, using its principles, even to identify a new particle*

by Donald H. Weingarten

Running a supercomputer continuously for several years to generate a dozen or so results might raise a few impatient eyebrows. But to my collaborators and me, it was time well spent. The tasks we had given the machine are so complicated that they simply cannot be carried out with paper and pencil, but they address key problems in the theory of elementary particles. Calculations lasting a few years seemed reasonable.

The problems with which we were concerned arise from quantum chromodynamics, the theory governing the behavior of quarks. Formulated in the 1970s, QCD—as it is often called—describes how quarks combine in pairs or triplets to form hadrons, the class of particles that is subject to the strong nuclear force. The familiar proton and neutron belong to this family, as do more exotic kinds of particles that briefly appear in high-energy collisions in particle accelerators or in cosmic-ray showers. Experiments that scattered electrons off hadrons, which were designed to probe the hadron's internal structure, provided the initial data showing that QCD was very likely correct. The theory now represents one of the linchpins of the Standard Model of particle physics.

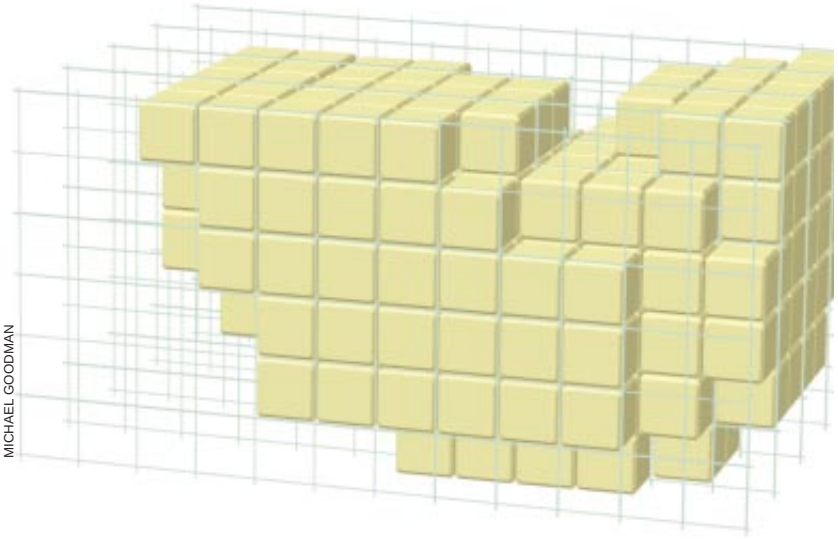
An important piece of evidence, however, was lacking. As the fundamental theory of quarks, QCD should enable one to calculate the mass of the proton or the neutron—in fact, of any hadron. (Of course, the heft of many hadrons had already been measured in experiments; the masses of the proton and of the neutron, for instance, were uncovered early this century.) But solving the mathematics needed to derive such masses from theory was beyond the power of hand calculations. Even the fastest computer of the 1970s would have required more than 100 years of continuous operation.

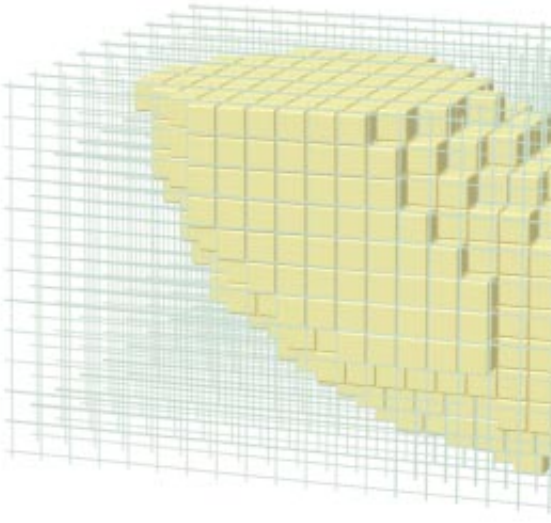
For this reason, in 1983 my colleagues and I at the IBM Thomas J. Watson Research Center began to design a parallel-processing computer dedicated solely to QCD calculations. The machine would perform about 11 billion arithmetic operations a second—and would thus be several hundred times more powerful than the fastest machines of the day. (Progress has made that speed less impressive: it is comparable to running about 200 desktop computers using the most advanced Pentium chips.)

In 1991, after eight years of designing, financing and constructing, we were able to begin QCD evaluations. After about a year of continuous number crunching, the computer, called GF11, produced its first results, including values for the masses of the proton and seven other hadrons. The differences between our predictions and the experimental results were below 6 percent and arose from the statistical algorithm we used, rather than reflecting uncertainties in QCD itself.

In November 1995 we completed another set of calculations, this time having run the computer continuously for two years. The machine found the mass and decay rate of an elusive subnuclear particle belonging to a subfamily of hadrons called glueballs. Using these data and searching back through a tabulation of experimental results, my colleagues and I found that this glueball had indeed appeared in past laboratory experiments but had gone unrecognized. The QCD calculation hence represents the first instance in which massive computation identified a new particle.

As a result, we believe our work on hadron masses and on the glueball mass and decay rate serves as a significant confirmation of QCD. Our work may





MICHAEL GOODMAN

also provide an example of how problems in fundamental physics will be tackled in the future. As the cost of particle accelerators and the complexity of theory conspire to limit our exploration of nature, large-scale computation can solve problems that can be approached in no other manner.

### Quarks and the Color Field

What makes quark behavior so complex? A small part of the reason is the variety of ways quarks can combine. They come in six "flavors": up, down, strange, charmed, top and bottom. In addition, they have one of three "colors," called color charges, that are often labeled red, green or blue. Hence, 18 different combinations of quark characteristics are possible. In addition, antimatter versions of quarks exist; they have their own corresponding set of 18 combinations of antiflavors and anticolors. QCD forbids individual quarks or antiquarks from roaming freely, so physicists deduce their properties by observing interactions between pairs of hadrons and between hadrons and other elementary particles, such as electrons and photons.

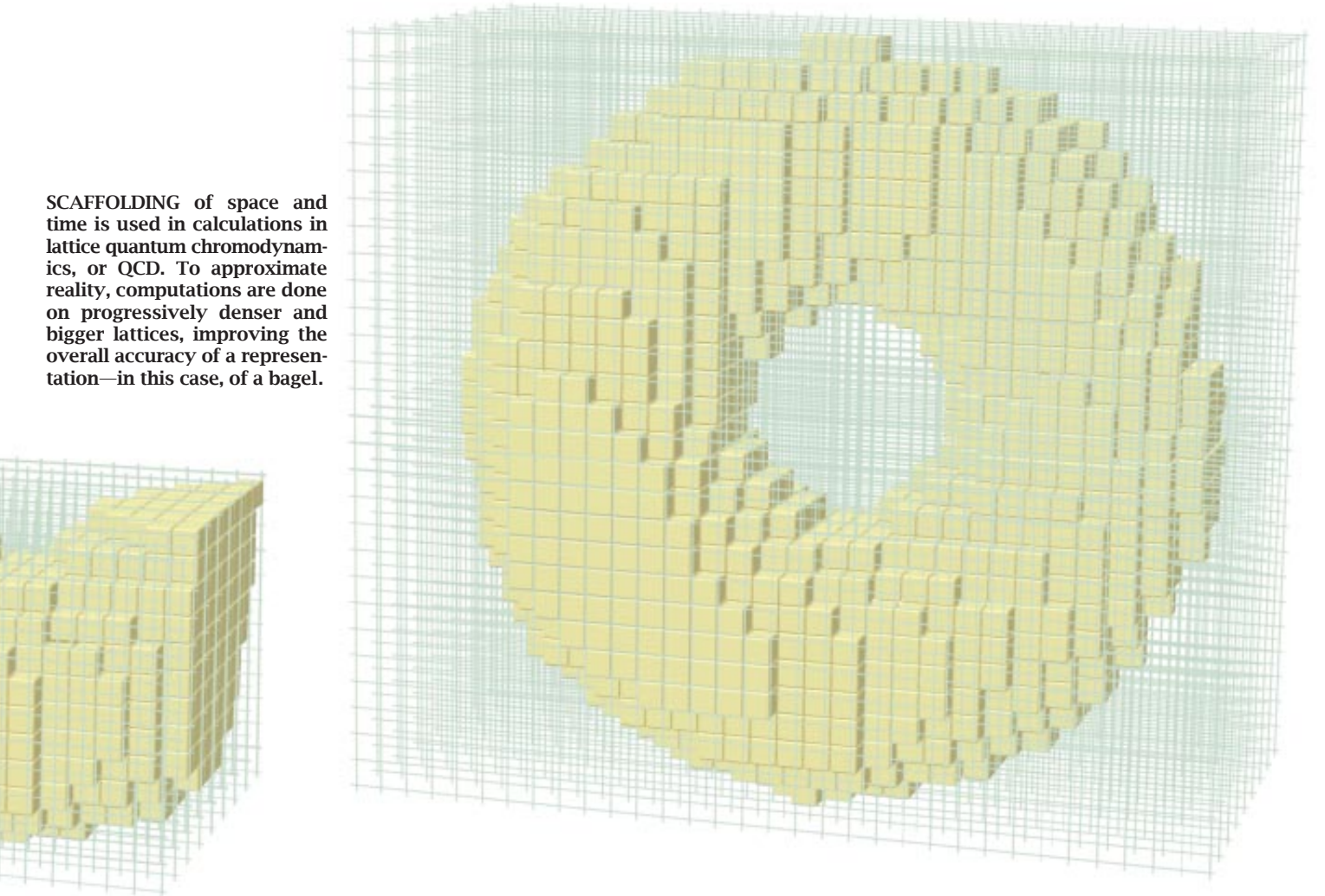
The primary source of the complexity of QCD, however, is that the color charges give rise to an energy pocket. This pocket is referred to as a chromoelectric field. In this fashion, quark color charges are analogous to ordinary electrical charges, which create an electromagnetic field around themselves. And just as an electromagnetic field connects the negatively charged electrons to the positively charged nucleus of an atom, the chromoelectric field binds together the quarks and antiquarks within hadrons. (Matter and antimatter versions can coexist nearby for limited periods before annihilating one another.) QCD thus accounts for a range of possible hadrons as the 36 different types of quarks and antiquarks combine and arrange themselves in various ways.

The analogy between an ordinary electromagnetic field and a chromoelectric field goes only so far. Unlike the electromagnetic field, the chromoelectric field can interact strongly with itself to create different kinds of condensed lumps. Within a hadron, the field condenses into a thin, intense, stringlike shape. The string carries a significant fraction of the hadron's total mass and is the main reason one cannot derive the mass of a hadron simply by adding up the masses of the quarks it contains. This strong self-interaction of the chromoelectric field is, in the end, the main difficulty in evaluating hadron masses from QCD.

The stringlike pieces of the chromoelectric field can also form a closed loop—thereby, according to QCD, creating a particle without quarks and antiquarks. Because the chromoelectric field "glues" together hadrons made of quarks, particles composed solely of the field are called glueballs.

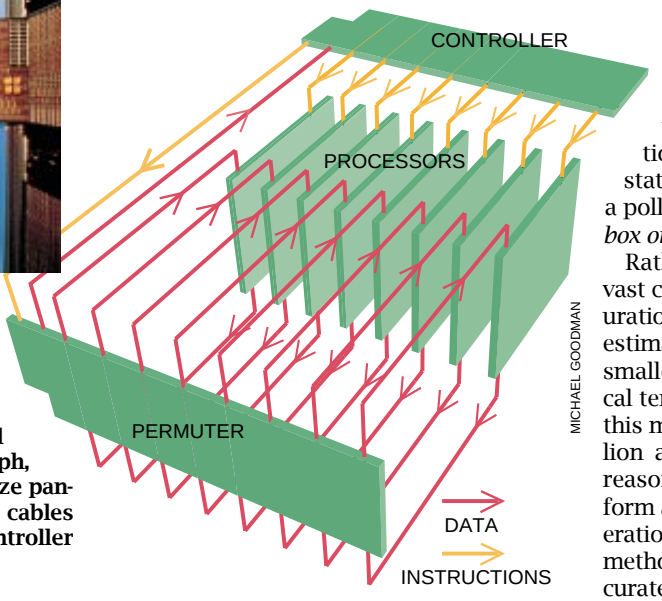
### Turning Space into a Lattice

Making predictions from these general ideas about QCD demands a more specific mathematical interpretation of the behavior of quarks and the chromoelectric field. The numerical methods (computer algorithms) that have been applied to QCD calculations



**SCAFFOLDING of space and time is used in calculations in lattice quantum chromodynamics, or QCD. To approximate reality, computations are done on progressively denser and bigger lattices, improving the overall accuracy of a representation—in this case, of a bagel.**

IBM THOMAS J. WATSON RESEARCH CENTER




MICHAEL GOODMAN

**GF11 SUPERCOMPUTER consists of 566 parallel processors. The "permuter" collects and feeds data to the processors, and the controller orchestrates all the activity. In the photograph, the processors are the bronze panels that slide into racks; the cables belong to the permuter-controller system.**

are all based on a formulation of the theory proposed in 1974 by Kenneth G. Wilson, now at Ohio State University.

The approach is rather technical. Briefly, Wilson's version of QCD can be taken as a description of how the configuration of quarks and the chromoelectric field in the world changes over time. For instance, a possible configuration at noon might consist of one quark heading for a collision with another quark at one mile an hour, with the chromoelectric field everywhere in space equal to zero. Another condition, observed perhaps at 2 P.M., might find the quarks heading in opposite directions 90 degrees from their original paths, with the chromoelectric field nonzero at some points in space. Wilson's form of QCD assigns a probability to this change over the period from noon to 2 P.M. Deriving quantitative predictions from QCD requires analyzing, with some mathematical legerdemain, such transition probabilities.

Wilson's formulation applies to a rather unusual world: the continuous, unbounded space and time of reality are approximated by a discrete lattice of points, somewhat like the vertices of a four-dimensional checkerboard (three dimensions for space and one for time). This scaffold is restricted to a finite volume, and as a result all of space and time are approximated by a finite set of points. A specific configuration of quarks and the chromoelectric field occurs at each of these points. Not surprisingly, Wilson's form of the theory is known as lattice QCD.

Results for the real world are supposed to be obtained by doing calculations on a sequence of lattices that have progressively smaller spacing between vertices and progressively larger volumes (in other words, shrinking the distance between lattice points to zero and expanding the volume of the network to infinity). The limit of this sequence of results represents lattice QCD's prediction for the real world.

The problem in computation is that Wilson's rule for determining the probability of a transition requires the summation of a vast collection of terms. The sum includes one contribution for each possible way of constructing a kind of two-column table. The table would list every point of the lattice in one column and the configuration of quarks and fields occurring at that point in the other column.

The number of different ways of filling in the configuration information in such a table is, in fact, astronomical. At each point on the lattice, a set of 32 real numbers represents the chromoelectric field. On a lattice with 10 sites in each dimension (10 by 10 by 10 by 10, about the smallest size that can provide even a rough approximation of reality), a table listing the field at each point specifies 320,000 real numbers—32 numbers for each of the 10,000 lattice sites. Even if we ignore the additional complication of keeping track of quarks and introduce the unrealistic simplification that each of the 320,000 numbers can be only zero or one, each table gives a sequence of 320,000 zeros and ones. The total count of such sequences—and therefore the total count of possible terms in Wilson's sum—becomes two to the power of 320,000, which is about one followed by 96,000 zeros.

### Monte Carlo and Lattice QCD

In 1979 Michael J. Creutz, Laurence A. Jacobs and Claudio Rebbi, all then at Brookhaven National Laboratory, introduced a practical numerical method to get around this extensive computation, working in part from a suggestion of Wilson's. Their method is a version of Monte Carlo integration, which approximates a solution to a complex problem through statistical sampling, much like taking a poll to predict an election winner [*see box on opposite page*].

Rather than explicitly summing the vast collection of possible lattice configurations, a sampling process is used to estimate the sum by examining a much smaller set of randomly selected, typical terms. Assessing a 10-unit lattice in this manner requires only about 20 billion arithmetic operations. In 1979 a reasonably fast computer could perform about half a million arithmetic operations in a second. So the Monte Carlo method could generate a sufficiently accurate approximation in less than a day.

Unfortunately, the calculation seemed reasonable because it applies to a simplified form of QCD: it included only the chromoelectric field and ignored the quarks. Performing a similar calculation with the full theory led to a series of technical problems. In 1981 several groups of investigators suggested versions of the Monte Carlo algorithm that could take into account the presence of quarks. These physicists included Donald N. Petcher and me, then working together at Indiana University, and Federico Fucito, Enzo Marinari and Giorgio Parisi, all at the University of Rome, working in collaboration with Rebbi. The revised algorithms included a measure of how hard it would be for quarks and antiquarks to travel from the initial configuration to the final one.

The cost of including quarks and antiquarks in the revised algorithms was a great increase in the number of arithmetic operations required to do meaningful calculations. Even the fastest computers of the time could not have crunched numbers on large enough lattices to obtain, in a reasonable amount of time, predictions for the real world. No one could afford to wait the 100-plus years it would have taken.

The origin of this stumbling block was an additional contribution to the energy of each configuration of quarks and chromoelectric field. This extra energy stems from the brief appearance and annihilation of quark-antiquark pairs, which occurs in any region of space carrying a chromoelectric field.

In 1981 I suggested a technique to compensate for this effect. Called the valence approximation, it consists of deleting the contribution of quark-antiquark pairs and at the same time dividing all color charges by a "chromodielectric" constant.

To evaluate the zero-spacing, infinite-volume limit of mass predictions

in the valence approximation, we eventually found we needed to work with lattices ranging in size up to 32 units on each side. The total cost of a full calculation becomes about $10^{17}$ arithmetic operations.

The fastest computers of the early 1980s could perform about 40 million arithmetic operations per second, or $10^{15}$ operations per year. So even with the valence approximation method, calculating the hadron masses might require 100 years. Hence, insufficient computer power prevented researchers from fully extracting real-world predictions from lattice QCD.

### Predictions with GF11

My IBM colleagues Monty Denneau and John Beetem and I began designing a computer expressly for QCD calculations in 1983. It would be capable of 11 billion arithmetic operations a second—specifically, 11 billion floating-point operations per second, or 11 gigaflops (hence the computer's name, GF11). This machine would be 250 times faster than the best computers of the time. Reaching this speed required us to incorporate 566 processors running in parallel. The machine cycle—how fast a processor can execute one addition or multiplication—was 50 nanoseconds. Each processor, capable of 20 megaflops, can feed one datum to other processors every 200 nanoseconds. (Several dozen supercomputers in the world are now faster than GF11, with the fastest computing perhaps 10 times the rate sustained by GF11.)

James C. Sexton and I wrote the operating system, a compiler and a set of hardware diagnostic programs. The two of us also brought the machine into operation, working in collaboration with a group headed by David George, including Michael Cassera, Molly Connors, Manoj Kumar, Edward Nowicki and Michael Tsao. Sexton, Chi-Chai Huang, Lan Wong and I completed the machine's final commissioning.

By late 1991 my colleagues Frank Butler, Hong Chen, Alessandro Vaccarino, Sexton and I were able to begin calculating hadron masses. We looked for the masses of 11 hadrons composed of combinations of only the three lightest flavors of quarks and antiquarks: up, down and strange. Up and down quarks are the constituents of neutrons (one up, two down) and protons (two up, one down). Strange quarks occur in a variety of unstable hadrons. We calculated each hadron mass on lattices of various sizes. For each choice of lattice volume and spacing between points, we calibrated certain parameters against the

## QCD by Monte Carlo

To circumvent the enormous amount of computation that lattice QCD would otherwise entail, researchers rely on a shortcut called the Monte Carlo method. The technique can find the "transition amplitudes"—the probability that quarks and the chromoelectric field go from one particular starting point to some final state. A down-to-earth analogy is the rolling of a die, whereby the transition amplitude is simply the odds that the die settles with a particular face up. The catch is that the die is loaded, and you do not know which sides are weighted. Moreover, the outcome depends on the orientation of the die before the throw and the length of time it is shaken before being thrown.

The Monte Carlo method consists of throwing the die and seeing what happens, as one would in actual gambling (hence the method's name). To find the die's transition amplitude between, say, the position with one dot up and the position with six dots up one minute later, you would hold the die initially with one dot up, shake it for a while, throw it so that it came to rest one minute later and then observe whether six dots are facing up. You then repeat this procedure several times. The fraction of trials in which the die lands with six dots up gives the desired transition amplitude.

Finding the transition amplitude for QCD is not much different. Initially, one specifies the desired field configuration at the initial time (that is, the hand holds the die with a certain orientation). Then the computer creates a random time sequence of fields stretching across the lattice's time dimension (it shakes the die). It stops on some field at the specified final time (it throws the die). One then examines the outcome to determine whether it happens to be the desired final configuration (you look to see if you rolled a six). The fraction of tries that yields the specified final configuration gives the desired transition amplitude. Once the transition amplitude is found, the actual masses of hadrons can be uncovered with a little arithmetic.

MICHAEL GOODMAN

**ROLLS OF A DIE contribute to the transition amplitude only if the die comes up showing a six (*first and third rows*).**

In a typical QCD calculation, each of the 566 parallel processors of the GF11 computer is given a chunk of the lattice to compute. The algorithms for generating random chromoelectric fields and for determining quark movements appear as programs that make each processor "walk" across its chunk of the lattice. At a lattice point, each processor stops to "shake" the chromoelectric field—the QCD analogue of shaking a die—in a manner that depends on the field at that site and at neighboring sites.

With GF11, we tried to determine how our calculated values would match up with the masses observed in the real world—that is, when the spacing between lattice points is zero and the lattice volume is infinite. We calculated masses on different lattice volumes and spacings.

First, we took the limit of large volumes. We chose $1.5 \times 10^{-14}$ centimeter as the spacing between lattice points, which had been known to give results that are not too far from mass values at the zero-spacing limit. Using this value of spacing, we computed masses on lattices of different spatial sizes. The results generated on a lattice 16 units on a side were within 5 percent of the infinite-volume limit, and on a 24-unit lattice the results were within 1 percent.

In the second step, we tackled the limit of the spacing between lattice points. We calculated masses on a 24-unit lattice with a spacing between points of $1.0 \times 10^{-14}$ centimeter and on a lattice 32 units a side with spacing of $0.75 \times 10^{-14}$ centimeter. Both lattices have the same volume as our original 16-unit lattice. These calculations enabled us to extrapolate our predictions to a lattice-point spacing of zero—in other words, to compare the numbers with the observed values of the masses. *—D.H.W.*

known masses of three hadrons. We were left with eight mass predictions for each type of lattice. To find the limiting values for a lattice with an infinite volume and zero spacing between points, we extrapolated from the predictions achieved when the lattice spacing was made smaller and the volume made larger. These extrapolated results mark QCD's predictions for the real world.

| HADRON | MASS (MILLIONS OF ELECTRON VOLTS) CALCULATED | OBSERVED |
|---|---|---|
| KAON (IN EXCITED STATE) | 898 ± 12 | 896 |
| PROTON | 936 ± 80 | 941 |
| PHI | 1,026 ± 25 | 1,022 |
| DELTA | 1,205 ± 94 | 1,235 |
| SIGMA (IN EXCITED STATE) | 1,391 ± 62 | 1,388 |
| XI + SIGMA − PROTON | 1,484 ± 57 | 1,576 |
| XI (IN EXCITED STATE) | 1,582 ± 50 | 1,537 |
| OMEGA | 1,768 ± 69 | 1,676 |
| LIGHTEST GLUEBALL | 1,740 ± 70 | 1,710 |

LISA BURNETT

**CALCULATED HADRON MASSES are compared with their observed values. Combining the xi-sigma-proton masses eased the calculation.**

The difference between each mass prediction and the corresponding value found in experiments was in all cases less than 6 percent. Viewed together, these uncertainties were of a size consistent with the errors expected from the Monte Carlo algorithm we had used. Although these answers were found in the valence approximation, we believe it is unlikely that the valence approximation could coincidentally give eight correct results if QCD itself were wrong. Thus, these results confirm both the mass predictions of QCD and the reliability of the valence approximation.

More recently, Sexton, Vaccarino and I explored the properties of the lightest glueball (the closed loop of the chromoelectric field mentioned above). Earlier calculations, using slower computers, contained significant uncertainties, with the result that no glueball had been unambiguously identified in experiments when we began our work. Running 448 of the GF11's processors continuously for two years, we calculated the mass of the lightest glueball and the rate at which it decays into other, more stable particles. The answers surprised us, because according to our numbers, this particle had appeared in several different experiments during the past 12 years. No one recognized it because there was no sufficiently detailed profile of its properties. The calculation is the first "discovery" of a subnuclear particle by computer.

Despite our success, a direct confirmation of QCD without any approximations remains to be done. Obtaining these results must await either a computer hundreds of times faster than GF11 or algorithms hundreds of times more efficient than the best ones today. One machine in this range is now planned by a group at Columbia University, another by a group at the University of Rome and a third by a collaboration of several Japanese institutions. A more difficult task than mass calculation, perhaps requiring entirely new algorithms, would be using QCD to explain the results of certain kinds of particle-scattering experiments.

### Experimental Theoretical Physics

Besides finding hadron and glueball masses, our computations are a sign of a fundamental shift in the direction some areas of elementary particle physics are headed. Even though numerical methods on the computer have been applied in other branches of the physical sciences, paper and pencil ruled theoretical analysis of particle behavior. Our work on QCD became necessary after analytic methods failed to make a dent in the theory's mathematics; the calculations became possible only after computers became powerful enough for the essential algorithms.

A qualitative difference between the paper-and-pencil method and the numerical approach arises in the degree of certainty that can be attached to the results. In theoretical work with paper and pencil, every step can (in principle) be verified to follow logically from the preceding step, and the final result can then be confirmed to be correct.

In contrast, both for experimental physics and in large-scale computation, it is almost always the case that results cannot be verified as true in the same mathematical sense. Instead plausible results are accepted if sufficiently many attempts to prove them false fail. We might accept the experimental result that all rocks fall down, not up, because no one has ever found a rock that falls up. The possibility remains, however, that someday someone might find a rock that falls up. In the strict mathematical sense, then, the claim that all rocks fall down remains unproved.

An analogous, unsuccessful search for a rock that falls up took place in our work. It remains a possibility that if we had calculated the proton mass on a lattice with spacing much smaller than that we used, we would have gotten, say, 2,000 million electron volts (or MeV, where 1 MeV equals about $1.78 \times 10^{-30}$ gram), contradicting our extrapolated zero-spacing lattice value of 936 MeV. Therefore, our claimed limit for the proton mass is tested by the sequence of numbers entering our extrapolation but is not truly proved.

The calculations we have done could be described with the seemingly contradictory phrase "experimental theoretical physics." Given the complexity of physical theories, the expense of huge projects and the continuing increase in computational power, experimental theoretical physics may become, for some segments of physics, the most practical probe of nature.

## *The Author*

DONALD H. WEINGARTEN is a research staff member at the IBM Thomas J. Watson Research Center in Yorktown Heights, N.Y. He received his Ph.D. from Columbia University and is a fellow of the American Physical Society. Besides pursuing studies in numerical quantum chromodynamics, he explores quantum theories of gravity and various kinds of computer architecture.

## *Further Reading*

GAUGE THEORY OF ELEMENTARY PARTICLE PHYSICS. Ta-Pei Cheng and Ling-Fong Li. Oxford University Press, 1984.

LATTICE QUANTUM CHROMODYNAMICS. D. Weingarten in *Particles, Strings and Supernovae.* Edited by A. Jevicki and C.-I. Tan. World Scientific, Singapore, 1989.

LATTICE GAUGE THEORIES. H. J. Rothe. World Scientific, Singapore, 1992.

HADRON MASS PREDICTIONS OF THE VALENCE APPROXIMATION TO LATTICE QCD. F. Butler, H. Chen, J. Sexton, A. Vaccarino and D. Weingarten in *Physical Review Letters,* Vol. 70, No. 19, pages 2849–2852; May 10, 1993.

QUARKS, GLUONS AND LATTICES. M. Creutz. Cambridge University Press, 1993.

QUANTUM FIELDS ON A LATTICE. I. Montvay and G. Munster. Cambridge University Press, 1994.



# Growing Seedlings at Less Than 1 G

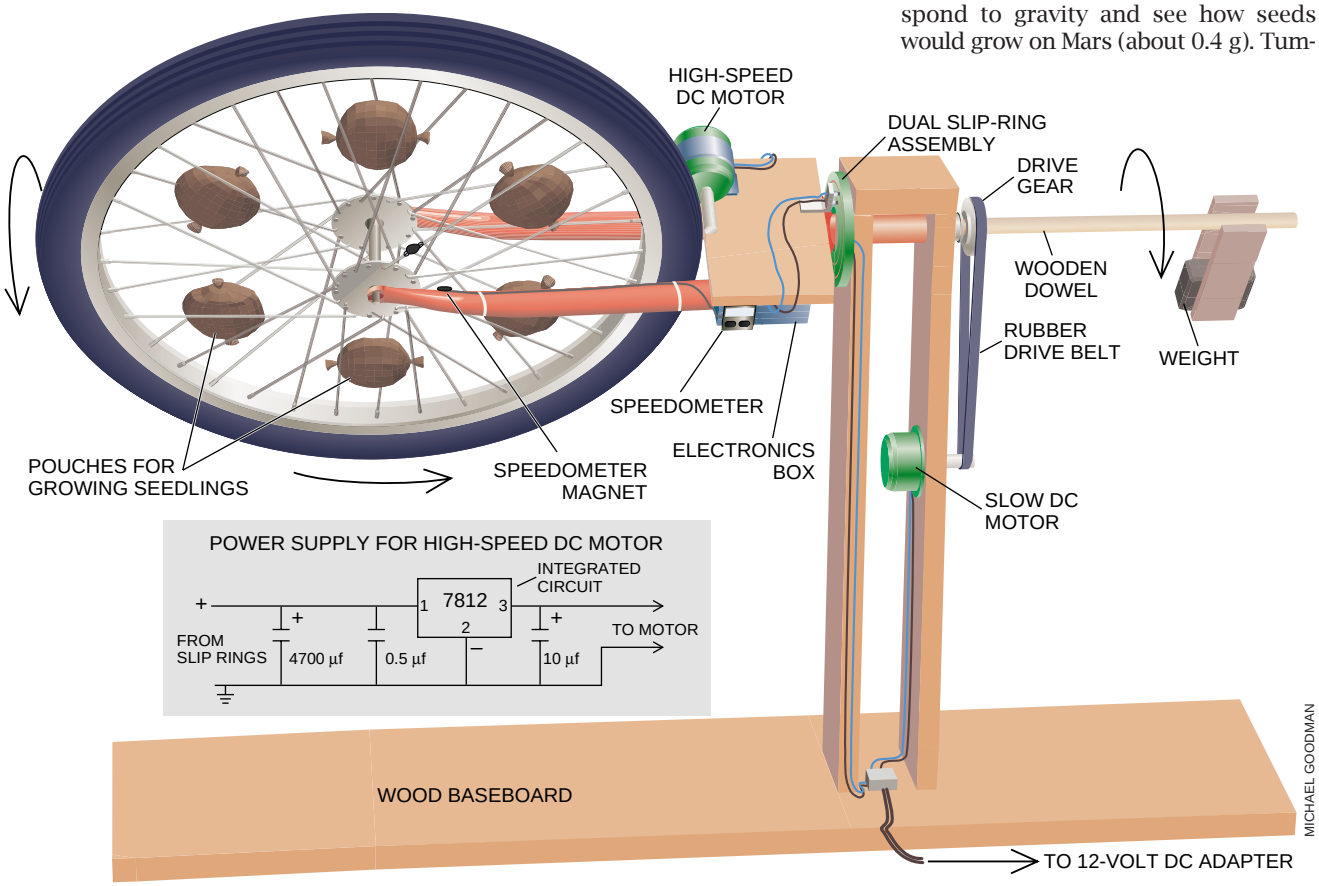
Something remarkable happens when you tip a plant. Special hormones, called auxins, begin to collect in the underside of its roots and stem. Auxins stimulate stem cells to grow and divide. The bottom of the stem then outgrows the top, causing the stem to bend skyward. Auxins in the root cells act differently: they retard growth. The auxin-poor cells near the top of the root then outgrow the auxin-rich cells near the bottom, and the root bends downward. In this way, a tipped plant makes internal adjustments to realign itself with the pull of gravity.

Botanists call a plant's response to gravity geotropism. In the early 1800s experimenters explored geotropism by growing plants on a rotating wheel, thereby exposing them to both the earth's gravity and centrifugal forces. The plants grew against the vector of the resultant force—that is, against the direction of the combined forces.

But scientists soon learned that plants respond to gravity only sluggishly. Most plants must be tipped for at least a minute before the auxins start to redistribute. By the turn of the century, scientists had invented a device, called a clinostat, that tricks plants into thinking they are growing in near-zero-gravity environments. Clinostats are still used today. By slowly rotating a plant vertically and horizontally, the clinostat prevents the plant from fixing on gravity. It then grows almost as if there were no gravity at all. Clinostats have fascinated amateur scientists for generations. Don Graham, my grandfather and an amateur scientist extraordinaire, published his explorations into geotropism in these pages 26 years ago [see "The Amateur Scientist," SCIENTIFIC AMERICAN, June 1970].

Oddly, professional biologists have paid little attention to what is perhaps the most interesting region to investigate—between 0 and 1 g (the acceleration caused by the earth's gravity, equal to an increase in speed every second of 9.8 meters per second). This is good news for the amateur, for a little dedication could reward you with original discoveries.

The rotating platform of the device described here is a bicycle wheel. By properly choosing the wheel's rotation speed and placing the seeds at different distances from the pivot, you can germinate seeds at any gravity. Observe the thresholds at which plants first respond to gravity and see how seeds would grow on Mars (about 0.4 g). Tum-




*BICYCLE WHEEL set tumbling and spinning can convince seeds, in pouches wedged between spokes, that they are sprouting on board a spaceship. When constructing the circuit for the high-speed motor's power supply, be sure to mount the 4,700-microfarad capacitor at least five centimeters from the type 7812 integrated-circuit chip.*

bling the wheel is also necessary. The earth's gravity then averages to zero, so that the seeds consistently experience only the centrifugal acceleration from the spinning.

Besides the wheel, you will also need to scavenge parts of the bicycle frame—in particular, the front mounting forks and its hollow shaft that slides up the head tube (the part through which the handlebar stem goes). I bought them all for $15 from a bike rental shop, which kept them around as spares.

Pieces of wood support the frame and wheel. Cut a hole through the upper ends of two slats of one-inch-thick pine shelving; attach the lower ends to a pine baseboard so that the slats stand upright [*see illustration on opposite page*]. Thread the mounting-fork shaft through the holes. For a drive gear, try the plumbing department of a hardware store. You can make an excellent one by cutting two inches off the end of a large-diameter plastic or rubber pipe. Shore up the inside of the circular ring with a wooden plug. Cut a hole in the center of the plug, then thread the assembly over the shaft and epoxy it in place.

Next, thread a dowel through the drive gear and into the shaft, then attach a weight, such as those scuba divers use on their belts, to the dowel. Bolt a three-by-seven-inch wood slat to the end of the dowel. Temporarily tie the weight to the slat. Position the wheel so that the bend of the mounting forks points upward. Slide the dowel into the shaft tube and adjust the weight's position, both horizontally and vertically, until the wheel is balanced against its closest support. Epoxy the dowel in place and secure the weight with bolts.

To negate the effects of gravity, the bicycle wheel must complete a tumble about once each minute. I used a slow motor (0.28 revolution per second, or rps) and connected the motor shaft to the drive gear with a flat rubber belt (check the power-tools section of your local hardware store). So attached, the drive gear turned at the right speed. If your motor spins at a rate much different from 0.28 rps, then you will have to fiddle with the size of your drive gear or add another gear to the motor shaft to produce the correct ratio.

A second, faster motor spins the bicycle wheel to create the artificial gravity. I used a 12-volt DC motor (Radio Shack No. 273-255), which spins at 192 rps. Make sure the tire's tread is smooth—tires from road bikes work well. The radius of the motor's shaft is about one millimeter. To figure out at what frequency to spin your wheel, consult the box on this page.

In my clinostat the radius of the wheel was 30 centimeters; at 192 rps, my motor spun the wheel at 0.64 rps. You can increase the motor's effectiveness by wrapping a few layers of cloth tape around the spool; the extra material will easily boost the rotation frequency to 1.5 rps. Power comes from a DC adapter and feeds the DC motor through two slip rings. The circuit shown in the schematic on the opposite page regulates the power.

Commercial bicycle speedometers let you easily monitor the acceleration. Get the kind that uses a small magnet placed on one of the spokes. Mounted near the axis, the magnet can measure the speed to the nearest 0.1 mile per hour or (better yet) 0.1 kilometer per hour.

So little work has been done in this area that you can grow just about anything and find something new. I've been focusing on corn. I let the seeds germinate for several days and then measured the total length of the sprouts and their "angularity"—the sum of the bend angles along the stock. At 1 g, the plants grow very straight; near 0 g, they become quite crooked. Growing seedlings at a number of locations along the radius of the wheel enables you to see the effects of gravity "turn on" inside the plant.

Put five seeds into a small handful of potting soil and place them inside the leg of an old nylon stocking. Cut the ends around the soil, then tie them off with a bit of twine to create a small pouch. These lightweight packets hold the seeds in place and make them easy to water. To get good statistics, you will need to sprout roughly 30 seedlings for each acceleration, so you should place six of these bundles at the same distance from the center. Arrange them symmetrically between the spokes of the wheel to keep the wheel balanced. You can make more bundles and insert them at different distances, so that you can experiment at different accelerations simultaneously.

Remove the seeds after they have germinated for three to seven days (or until they begin to poke out of the pouches). Cut the seedlings at their bends, then lay all the pieces end to end so that all the bends are in the same direction. The angularity is the angle between the first piece and the last piece. To measure the length of a seedling, lay a string along the piece. For each acceleration, divide the angularities by the seedling lengths and average the results. Plot this average versus acceleration, and you'll see how the sensors in your plants respond to different gravitational fields. You might also consider replanting the seedlings in normal conditions to see how they subsequently do—or try to figure out the best food plant that would grow on Mars.

*For more information about this project, consult the Society for Amateur Scientists' Web page at http://www.thesphere.com/SAS/ or* Scientific American*'s area on America Online, or send $5 and your address to the Society for Amateur Scientists, 4951 D Clairemont Square, Suite 179, San Diego, CA 92117. I am grateful to SAS member John Michaelson for perfecting the instrument's construction and for building the first prototype.*

*SHAWN CARLSON is an adjunct professor of physics at San Diego State University and executive director of the Society for Amateur Scientists.*

## Calculating G Forces

A seed on the spinning wheel experiences a centrifugal acceleration of $(2\pi f)^2 r$, where $f$ is the rotation frequency, and $r$ is the distance from the seedlings to the pivot. The formula to determine at what frequency to spin your bicycle wheel is $f = 1/2\pi\sqrt{(a/r)}$, where $a$ is the acceleration. For example, to produce 9.8 meters per second per second (1 g) at the rim of a wheel 0.3 meter in radius requires the wheel to spin 0.91 revolution per second (rps).

The frequency at which the Radio Shack motor will drive the wheel, $f_w$, is $f_d (r_m/r_w)$, where $f_d$ is the frequency of the drive motor (192 rps), $r_m$ is the radius of the motor shaft (about one millimeter), and $r_w$ is the radius of your wheel.

To find the acceleration at any distance from the wheel's center, use the formula

$$a = 19.98 \frac{v_{mph}^2}{r_w} \frac{r}{r_w} = 7.72 \frac{v_{kph}^2}{r_w} \frac{r}{r_w}$$

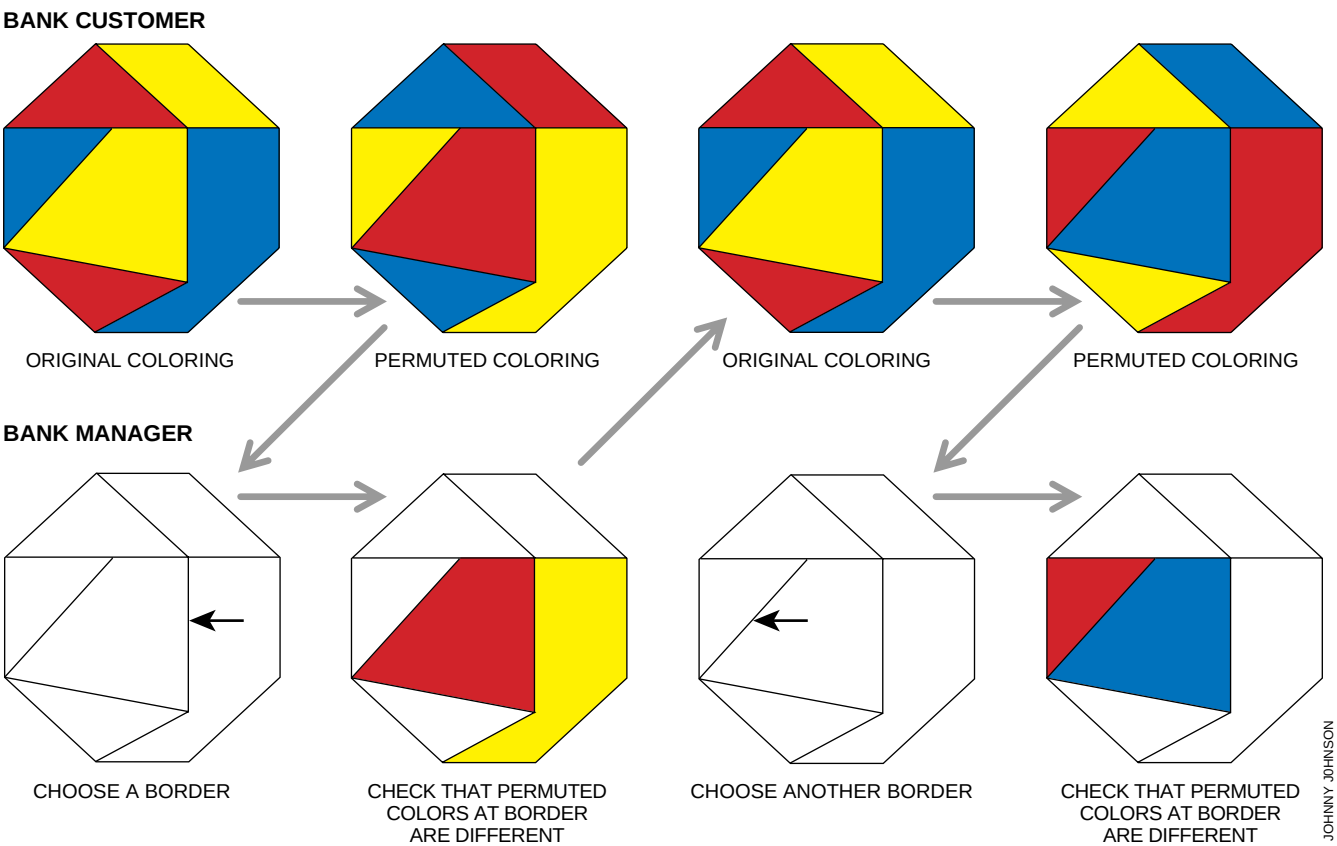
where $r$ and $r_w$ are expressed in centimeters, and $v$ is the speedometer reading, expressed in either miles per hour or kilometers per hour.

# Proof of Purchase on the Internet

Suppose you want to make a purchase on the Internet using your credit card. If you simply transmit your card number to a merchant, someone else can readily intercept your number and use it for himself or herself. So, too, your personal identification number, or PIN, can be purloined from the Net. To avoid such problems, most network systems employ encryption methods to encode, or scramble, private messages. Provided no one can crack the code, the encrypted message—say, a credit-card number—remains secure. But on the Internet, it has become increasingly clear that this isn't good enough. Although many secure encryption codes exist, users want proof that no one can access confidential information. The best way to keep a message private, of course, is to avoid sending it in the first place. And surprisingly enough, so-called zero-knowledge protocols let you do just that.

The method allows you to convince someone, such as a bank manager, that you possess some key item of information, such as a PIN, without actually revealing the information itself. The principle behind zero-knowledge protocols is best illustrated by problems called map colorings. In 1852 Francis Guthrie, a graduate student at University College London, first conjectured the Four Color Theorem, which states that every two-dimensional map can be colored using no more than four colors so that no two adjacent countries are the same color. More than a century later, in 1976, Kenneth Appel and Wolfgang Haken of the University of Illinois proved the theorem using a computer.

Given fewer than four colors, only some maps can be colored so that the adjacent countries are different hues. Now, suppose your bank manager sends you a complex map, and you wish to convince her, without directly showing her, that you can color it using three hues. Then, imagine that you can construct an electronic device linked to two touch-sensitive screens—one in the bank and one at your home [*see illustration below*]. First, you program into this machine your map coloring, say, by touching specific regions on your screen: one touch for red, two for blue and three for yellow. Next, the bank manager selects a border where two countries meet. The machine permutes your coloring scheme at random, perhaps systematically changing your red countries to blue, your blue countries to red and leaving your yellow countries unaltered. The manager's screen displays the new colors of the two countries sharing the selected border. If your original coloring was valid, then these two colors are different.




***ZERO-KNOWLEDGE PROTOCOL** shown above lets the bank customer convince the bank manager that he or she can color a three-color map. The bank manager need only check that the permuted colors differ at each border for proof.*

By repeating this operation on each border, the manager can determine whether you have a three-color map, as you claimed. If your original coloring failed and two adjacent countries had the same color, then, when the bank manager selects the border between them, she sees two permuted colors that are also identical. If, on the other hand, the permuted colors are different across every border, then your original map was valid. Because the machine chooses one of six possible color permutations at random, the manager cannot deduce your original coloring. She can merely confirm that it worked.

Experts on zero-knowledge protocols prefer a more rigorous argument, based on a process of "simulation." Imagine a fake system that chooses different colors at random rather than permuting your choices. This fake system might generate many different sequences of pairs of colors; one of them will be the sequence of responses based on your map. Suppose for a moment that the bank manager could determine your map from the responses the true system presented. Then she could also determine your map when, on rare occasion, the fake system produced those same responses. But for the fake system, there is no such thing as your map, and so such a deduction must be impossible. If the bank manager cannot deduce your coloring choices from the machine's responses, no one else can.

A more elaborate zero-knowledge protocol allows you to convince your bank manager that you know two prime factors, call them $p$ and $q$, of a particular number, $n$. Provided that $n$ is fairly large—say, 200 digits—there is no known algorithm that can find its factors within the lifetime of the universe. But algorithms do exist for testing whether the factors $p$ and $q$ are prime. Thus, your bank manager can cook up two primes, work out their product and treat the primes as a PIN, which you are told when you open an account. Over a suitable communications channel, you could convince her that you knew the PIN, without divulging the primes to her or anyone else.

By a suitable communications channel, I mean a so-called oblivious-transfer channel, which lets you send your bank manager two items of information in such a manner that she can read only one of them [*see box above*]. You don't know which item she can read. You can construct an oblivious-transfer channel using number theory. (For details, see *A Course in Number Theory and Cryptography,* by Neal Koblitz, published by Springer-Verlag in 1994.) The method requires that you use two 100-digit numbers instead of the usual four-digit PIN. Certainly, more practical strategies exist. They are, however, more difficult to describe.

Two centuries ago Carl Friedrich Gauss, one of the greatest mathematicians and a founder of number theory, described his brainchild as "the queen of mathematics." Queens are regal, but they are also largely decorative, and this nuance was not lost on Gauss. Only 20 years ago number theory, or the study of whole numbers and their properties, was seen as the most impractical branch of mathematics—pure thought from the ivory tower, beautiful but useless.

Today number theory is paying its way in the world of commerce, thanks to computers and modern communications. Digital information is numbers made real. So processes carried out via digital information, such as coding methods, are almost inevitably applications of number theory. Zero-knowledge proofs and oblivious-transfer protocols are just two of the curious ideas that have risen from number-theoretic depths. As an applied science, number theory has been a slow burner—but now the sparks are starting to fly.

## An Example of Oblivious Transfer

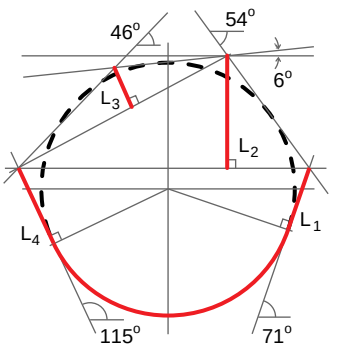
You and your bank manager both know two prime numbers, $p$ and $q$, and their product, $n$. A trusted, independent source gives both of you a sequence of binary bits, from which you can construct any random numbers required in the protocol. You can convince your bank manager that you know $p$ and $q$, without revealing them, as follows:

1. The independent source generates a random integer $x$ and sends you and your manager the remainder, $r$, left after dividing $x^2$ by $n$ (that is, $r = x^2 \bmod n$).
2. According to number theory, $r$ has exactly four different square roots in mod $n$. You use your knowledge of $p$ and $q$ to find them. One of them is $x$, and the other three are $n - x$, $y$ and $n - y$, for some $y$. (If you don't know $p$ and $q$, there is no efficient algorithm to find these square roots; from all four, however, you can easily deduce $p$ and $q$.)
3. You choose one of these four numbers at random: call it $z$.
4. You choose a random integer $k$ and send your manager the integer $s = k^2 \bmod n$. You then calculate two numbers, $a$ and $b$, so that $a = k \bmod n$, and $b = kz^z \bmod n$. Send $a$ and $b$ to your bank manager by oblivious transfer.
5. The manager can read exactly one of the two messages. She checks that its square mod $n$ is either $s$ (if she reads message $a$) or $rs$ (if she reads message $b$).
6. These steps are repeated $T$ times. At the end, your manager knows, with probability $1 - 2^{-T}$, that you know the factorization.

Notice that there is no communication from your bank manager back to you—that is, the protocol is not interactive.

## Feedback

John Day of Columbus, Ohio, sent me a series of letters about "The Great Drain Robbery" (September 1995). He first found a solution having two trenches and a total length of 481.89 yards—an improvement over the one reported in the December column "Feedback." His second letter described an even shorter solution that involved three trenches. And his third letter provided an even better three-trench solution, having a total length of 479.98 yards (*shown at right*). Does anyone have any four-trench solutions? —*I.S.*




JOHNNY JOHNSON

**LENGTH OF TRENCHES**

| | |
|---|---|
| $L_1$ = 52.05 YARDS | $L_4$ = 68.73 YARDS |
| $L_2$ = 87.88 YARDS | ARC = 237.36 YARDS |
| $L_3$ = 33.96 YARDS | TOTAL = 479.98 YARDS |

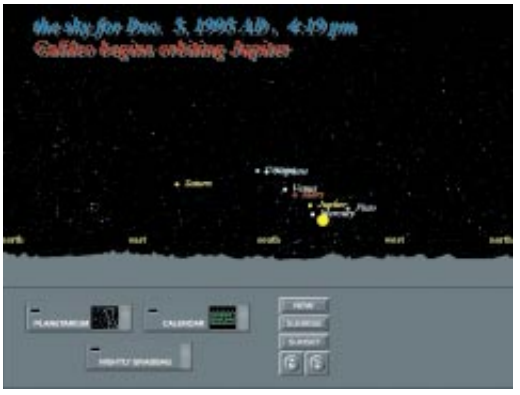


## Hands-on Astronomy

*Review by Timothy Ferris*

**DEEP SPACE: THE OBSERVER'S GUIDE TO THE NIGHT SKY,** Version 5. David Chandler Company, 1995 (CD-ROM for MS-DOS, $129). **DISTANT SUNS: FIRST LIGHT,** Version 1.0. Virtual Reality Laboratories, Inc., 1995 (CD-ROM for Windows, $99.95). **MEGASTAR: DEEP SKY ATLAS,** Version 2.11. E.L.B. Software, 1995 (CD-ROM for Windows, $99). **THESKY: ASTRONOMY SOFTWARE LEVEL III,** Version 2.0. Software Bisque, 1995 (CD-ROM for Windows, $199).

In an era when science seems to be growing ever more specialized and remote from the average citizen, it is reassuring to see that technology sometimes turns the tide, making academic-quality research accessible to hobbyists. Nowhere is that trend more evident than in astronomy. Amateurs who 20 years ago were largely limited to observing relatively nearby stars and planets can now probe deep into the wider universe of galaxies. Today many backyard stargazers have graduated from modest, spyglasslike refracting telescopes to mammoth yet inexpensive reflecting telescopes whose light-gathering mirrors are the size of serving platters. Equipped with sensitive light detectors known as charge-coupled devices (CCDs), these instruments can exceed the reach of the legendary five-meter Hale telescope at Mount Palomar in its pre-CCD days.



**DISTANT SUNS** *opens with a view of the current sky, showing both stars and planets.*

With better tools for exploration comes a demand for better ways to navigate the sky. The typical amateur of a decade or two ago could get by with a slim volume such as *Norton's 2000.0 Star Atlas,* useful as a guide to double stars and bright galaxies and nebulae but not much else. Now the same observer may be able to see 15,000 to 20,000 galaxies, with many thousands more accessible via CCD. A hard-copy atlas that presented everything visible to an ambitious amateur observing with one of the portable 30-inch reflectors now showing up at star parties would fill a bookshelf—and might cost more than the telescope.

Enter the CD-ROM, a single plastic disk capable of holding all the data in an armload of atlas and star catalogues yet much less expensive and more efficient to use. The four programs under review here are far more powerful than anything on the market five years ago. Three months of hands-on testing while observing with an 18-inch telescope show that all these disks have plenty to offer but that they differ considerably in strengths and weaknesses. Two of them—*Distant Suns* and *TheSky*—feature an ease of use and a range of features that should appeal to students and casual amateurs. The other two—*MegaStar* and *Deep Space*—are less flexible but provide imposing resources for the knowledgeable observer.

*Distant Suns* and *TheSky* are CD-ROM versions of programs that have already won a large following in their floppy-disk incarnations. Both are "planetarium programs," capable of displaying the entire sky or a piece of it, as seen at various times and places. They can be used to plan observing sessions and to print star charts; their extensive help files can supplement the education of astronomical novices.

*Distant Suns* sports some of the most pleasing graphics in the business, including an uncommonly evocative presentation of the night sky. Its animated features commend it to younger enthusiasts and their teachers. The animations include clips from the Apollo missions, the *Galileo* space probe's dramatic color view of the earth rotating in space and NASA's simulated flyovers of Mars and Venus—though these, alas, depict a ludicrously exaggerated vertical relief that sacrifices accuracy for the sake of a quick thrill.

Whereas *Distant Suns* scores points aesthetically, *TheSky* is oriented more toward the serious observer. It has four display options: black sky, a white-sky "chart mode" and two red-light displays that facilitate using the program at the telescope without dazzling the observer's dark-adapted eye (red is the least visually disruptive color). *TheSky* can interface with a properly equipped telescope, so that the display is always centered on the field of view seen through the eyepiece. Using a supplementary program, the observer can use *TheSky* to send operational commands through a modem to telescopes at distant locations, a capability that puts advanced amateurs pretty much neck-and-neck with the professionals. The program can produce star charts capable of satisfying the needs of casual users, if not always adequate for deep-sky explorers using larger telescopes.

Skygazers who suffer from "aperture fever"—the hunger for large telescopes that can carry them deep into intergalactic space—will want to boot up *Deep Space* or *MegaStar.*

*Deep Space* is intended for intensely committed observers like its author, David Chandler. "I know who you are!" Chandler writes, in the lively instruction manual. "You plan your vacations around the phase of the moon and you consider a sunny day with high cirrus to be lousy weather.... *Deep Space* was designed with you in mind, because I really wrote it for myself." Chandler, a

**MEGASTAR** *can generate detailed, customized maps of stars and galaxies (here centered on M81).*

science teacher, emphasizes the program's usefulness in education, describing among other things how to print unlabeled constellation charts to help students learn their way around the night sky. But the principal use of *Deep Space* is as a deep-space aid for the involved amateur. As a DOS program, it does without the friendly graphic design familiar to users of Windows or Macintosh operating systems; this austerity can make it a bit difficult to use at first. Nevertheless, *Deep Space* has emerged as a favorite among those who want a program that incorporates a big database and can interface with their mammoth telescopes.

*MegaStar* is quite limited in some ways. It cannot show a piece of sky larger than 20 degrees. It has never heard of the moon and is ignorant of every planet save Pluto. It offers nothing in the way of pretty pictures or video animations. But within these limitations it emerges as an impressive tool for the earnest student of the sky. It can link up to a computer-controlled telescope, and its functions are well organized for easy use by a bleary-eyed observer who has been awake most of the night. Its data filters make it simple to generate displays tailored to the observer's interests and to prevailing conditions—for example, one can eliminate all stars dimmer than the limits of one's equipment. An "eyepiece" function conveniently replicates the view through the telescope. *MegaStar* can be employed to print out star charts that simply blow away anything commercially available in hard-copy form. For a study of the rich Virgo cluster of galaxies, I printed and pasted together nine adjacent *MegaStar* charts. The result was a document any extragalactic observer would envy and one few could have readily obtained before.

All these programs have their glitches. Some—*Distant Suns,* for instance—look a lot better on-screen than they do on a printout. Each excels at specific tasks where others falter. Buffs on a budget would do well to consult with the local astronomy club and see the software packages in action before making a purchase. But stargazers who find one that fills their needs will soon wonder how they ever got along without it. These disks show what CD-ROM technology can do when used not as a flashy attempt to imitate Hollywood movies but as a down-to-earth means for bringing professional-caliber scientific tools to the general public.

*TIMOTHY FERRIS is a faculty member at the University of California, Berkeley.*

## Rethinking Green Thoughts

*Review by Thomas E. Lovejoy*

**A MOMENT ON THE EARTH: THE COMING AGE OF ENVIRONMENTAL OPTIMISM,** by Gregg Easterbrook. Viking Penguin, 1995 ($27.95). **NOAH'S CHOICE: THE FUTURE OF ENDANGERED SPECIES,** by Charles C. Mann and Mark L. Plummer. Alfred A. Knopf, 1995 ($24). **LOSING GROUND: AMERICAN ENVIRONMENTALISM AT THE CLOSE OF THE TWENTIETH CENTURY,** by Mark Dowie. MIT Press, 1995 ($25).

The modern environmental movement has achieved some remarkable successes. These days we almost take it for granted that automobiles should be as clean as possible, that some pesticides are too dangerous to use, that governments should act to protect the ozone layer. Now it seems that the moment has arrived when an evaluation of the strengths and weaknesses of the environmental agenda is due, perhaps even overdue. These three volumes are prominent among a considerable, recent spate of such reevaluations. Gregg Easterbrook examines environmentalism generally; Charles C. Mann and Mark L. Plummer focus on biological diversity in the U.S., most particularly on the Endangered Species Act; Mark Dowie looks at the history and political science of the American environmental movement. Taken together these books paint a fascinating portrait of the perceptions and realities of environmentalism, especially in the U.S., and of the challenges now facing the movement.

Many of my colleagues have been upset by *A Moment on the Earth.* In the book, Easterbrook advances what he calls the "Ecorealist Manifesto," a program of reduced governmental intervention based on the notion that many current environmental crises are not so dire as commonly claimed. I anticipated that I would disagree with some of Easterbrook's ideas, but I also expected to find important ones. In general, however, I was stunningly disappointed by the book's rambling prose and profusion of inconsistency and error.

Near the beginning of the book, for example, Easterbrook faults Rachel Carson for predicting that American robins would be seriously reduced by pesticides and claims that "nothing Carson forecast in *Silent Spring* came to pass." Yes, much of what Carson forecast did not come to pass. Why? Because she raised the alarm about the risks of chlorinated hydrocarbons early on, before scientists understood the mechanism by which those substances strangle the calcium metabolism of birds at the ends of long food chains (such as various birds of prey and the brown pelican). That warning inspired research that advanced knowledge of pesticide risks from correlation to causality.

Refined understanding in turn led to policy changes—most notably the banning of chlorinated hydrocarbons such as DDT—that enabled the affected species to recover. Easterbrook himself reports that the bald eagle population is now growing at 5.4 percent annually, while ignoring its previous downward trend, which was reversed only after the harmful pesticides were made illegal. He also extolls the recovery of the peregrine falcon and its reintroduction to New York City. Surely these qualify as environmental science and policy success stories.

It is hard to comprehend how a reporter like Easterbrook, who has followed environmental matters for years, could have failed to grasp such a fundamental issue as this one. He (as well as Mann and Plummer in *Noah's Choice*) misunderstands the essential purpose of bleak projections such as those publicized by Carson—namely, to highlight unfavorable trends so that the potential calamities will not come to pass. Sad to say, it is in fact quite rare for such projections to be totally wrong. To dismiss such efforts as doomsaying, and to portray the brave and prescient individuals who raise such warnings as biological Cassandras, does a disservice to society.

Easterbrook does make some significant points in *A Moment on the Earth.* He is correct that environmentalists have scored some major triumphs, such as achieving measurably cleaner air and water in the U.S. and in some other parts of the world. This country possesses an impressive set of environmental laws, although many of these are coming under fire. Far-reaching international agreements watch over the ozone layer, climate change and biodiversity. The numerous environmental successes give us just cause to be proud, but they do tend to get taken for granted. News, at least what people most vividly remember, is more about problems and moments of peril than about increments of progress. We all need some retrospection, caught up as we are in the daily hurly-burly, if we are to appreciate the overall improvements; Easterbrook's recounting of how far we have come serves a useful function.

Easterbrook touches on another noteworthy topic when he chastises environmentalists for overlooking problems of developing nations: drinkable water, clean air, food supply, disease control. Some of his criticisms are on target, al-

though a lot of the sustainable development guidelines that emerged from the United Nations Earth Summit in Rio de Janeiro do speak to these points, and multilateral development plans have increasingly taken them into consideration. He is wrong, however, to question the prominence of climate change as an environmental issue at Rio: this problem is one that can be addressed only at the international level.

Again and again, Easterbrook fails to grapple with the complexity of the topics he raises. He is rightly optimistic about clean technology, but his belief that "almost every pollution issue will be solved within the lifetimes of readers of this book" seems overly exuberant. The Delaney clause, a seemingly sensible guideline that prohibits carcinogens in food at any level of detectability, approaches never-never land because of ever improving detection technology; there is no simple solution to setting a "safe" level of carcinogens. The Ecorealist Manifesto gives high priority to conserving biodiversity because of the irreversibility of extinction—in itself, an admirable stance. Yet Easterbrook's book seems not to recognize a primary implication of his goal, that harder problems relating to the use of land and waters lie ahead.

The overall problem with *A Moment on the Earth* is that it consists of such a jumble of value judgments, anecdotes and errors that the constructive thoughts are often obscured. Easterbrook's most absurd assertion, building from a misunderstanding of evolutionary biologist Lynn Margulis's work, is that cooperation is dominant in nature. That notion completely ignores the existence of food chains, competition and disease. Yes, there are a lot of social insects, but vaguely mentioning that "for instance, deer greatly outnumber wolves" does not prove that "cooperative species are far more numerous than combative species and usually have larger populations." There are also problems of sheer sloppiness. It is hard to believe an editor overlooked the nonsense geography of the statement that "the North American population when Columbus landed may have been as high as 100 million, with most of this number living in Central and South America."

The innumerable errors and careless assertions are all the more frustrating because *A Moment on the Earth* could have been so much better. Easterbrook is obviously well versed in environmental matters, and he could have packed an enormous amount of information into the book's 698-page text. Yet he leaves potentially important topics undeveloped, such as the question of why environmentalists worry about the development of immunity in species targeted by pesticides but do not count on such immunity developing in nontarget species.

This book is so long and muddled that I suspect few people will plow through it in its entirety—and therein lies its danger. A casual reading could lead to the conclusion that all environmental problems have been solved. The novitiate and casual layperson who recognizes some of Easterbrook's numerous mistakes will find it hard to sort out the valid from the invalid. And it is unfortunate that the book excludes most names of people from the index.


SHAHN KERMANI *Gamma Liaison*

***PROTESTERS DENOUNCE** logging in old-growth forests in Oregon. Such controversies have promoted an unfortunate caricature of environmentalists as antibusiness extremists.*

In contrast with the broad but addled *A Moment on the Earth, Noah's Choice* is superbly written but disingenuously selective. The book deals, as its title implies, with biodiversity conservation in the context of the Endangered Species Act. The act is best known to the public from a handful of highly publicized instances (such as the efforts to protect the snail darter from the effects of the Tellico Dam) in which a vested interest came head-to-head with an endangered plant or animal having an esoteric, easily ridiculed name. From such stories, one might readily conclude that protecting endangered species comes down to a question of the future of the spotted owl versus jobs for loggers. For those unfamiliar with the true situation, the act seems severely tilted against people.

This kind of choice—protecting the ecosystem or safeguarding human interests—is the focus of Mann and Plummer's outstandingly readable book. *Noah's Choice* opens with an account of the discovery of the rare American burying beetle (which has a fascinating natural history) in Oklahoma and the difficulties that ensue as regulations dictated by the Endangered Species Act threaten to terminate the construction of a new highway. The authors follow with a couple of chapters that provide background information on biological diversity, the classification of organisms and estimation of extinctions.

Mann and Plummer, like Easterbrook, find it confusing to reconcile the relatively small number of officially acknowledged extinctions with projections—one of which I made in 1980—that major extinctions loom in the near future. Their book presents a lucid and reasonably fair discussion of species-versus-area curves, a common method for estimating biodiversity and projecting extinctions. Mann and Plummer reflect the debate over those estimates much better than Easterbrook does. But all three authors seem not to understand that the organizations that "officially" confer extinction status are enormously conservative; they often wait for decades before pronouncing a species extinct. Also, many biodiversity projections take into account the numerous extinctions that will occur once the remnant populations living in isolated fragments eventually wink out. Such losses are inevitable unless something is done to alter the situation.

Mann and Plummer's examples of endangered species and ecosystems—the American burying beetle, the Karner Blue butterfly, the Balcones Canyonlands development in Texas—are all difficult ones in which it is easy to empathize with local people caught up in the exercise. Although these instances are undeniable and represent one kind of experience, it seems unfair that the authors give little representation to the multitude of success stories, said to number in the tens of thousands, in which matters worked out reasonably.

The Endangered Species Act is portrayed as unsuccessful because the number of species being added to the list is so much larger than the number removed. That imbalance is more a matter of the act trying to catch up with reality than anything else. *Noah's Choice*

also contains whiffs of the old humans-versus-environment illusion, and the use of modifiers seems unnecessarily value-laden: "slimy things," "creepy crawlies," "shrieking predictions."

Mann and Plummer offer various suggestions about how to proceed in the future, but for the most part they seem to ignore the difference between flaws in the Endangered Species Act per se and problems with how it has been used. In 1978 Wildlife and America propheticall-ly noted that when the act is "used to combat all that is wrong about society's approach to the environment, the legislation stands in danger of losing public support and being weakened."

In reality, the biggest obstacle to saving endangered species has been the lack of intervention until a conservation situation grows so serious that it demands invoking the regulatory powers of the act. During the 1980s, for example, the conservation community warned about the status and unfavorable trends of old-growth forests in the Northwest, but the federal government did nothing. Instead the situation languished and deteriorated until the spotted owl qualified as the first endangered species of the ecosystem. The government could have designed a much better compromise of economic and conservation interests if it had acted earlier. The Northwest forest plan came just in the nick of time for a large number of species. The biodiversity of those forests now will survive only with a great deal of on-the-ground survey work and ongoing management.

Like Easterbrook, Mann and Plummer recognize the value of biodiversity conservation. They endorse the visionary effort, spearheaded by Secretary of the Interior Bruce Babbitt, to create a National Biological Survey. They also recognize that "Noah's Principle" (to save every species) is unattainable. Consequently, they offer "Noah's Choice," a way to make fundamentally impossible biological choices in a politically feasible manner. Their plan emphasizes prohibiting harm to individuals of endangered species over limiting the destruction of those species' habitat. The authors overlook the slippery slope this approach would invite: their scheme would automatically expand the list of endangered species (the individuals of which cannot be harmed) while offering no protection to the ecosystem in which those species live. The result, in all likelihood, would be a ballooning list of endangered species and an enforcement and management nightmare.

The "Noah's Choice" solution encourages procrastination, waiting until there is a problem rather than trying to avoid creating one in the first place. As a possible offset, the authors suggest protecting biodiversity on public land. That plan works fine for Nevada, 83 percent of which is federal land, but hardly at all for Texas, which has almost none. More promising is Mann and Plummer's proposal for the formation of a national biodiversity trust, which could provide incentives for private landowners. Part of the difficulty so far has certainly been that current laws impose costs on a few to generate benefits for many. Efforts to conserve biodiversity tend to come into conflict with our system of private property and especially with the plight of the small landowner. It is one of the most fundamental conundrums of environmental regulation.

The unflinching historical view in *Losing Ground* may have ruffled its own share of environmental feathers, but the book is well written, thought-provoking and plows the most new ground of the three. Mark Dowie, a former editor at *Mother Jones,* divides the history of environmentalism in the U.S. into four periods, which he terms waves. The first consisted of private individuals who were engaged in conservation (but certainly Theodore Roosevelt took it firmly into the public realm). In the second, it became an environmental political movement, starting in the mid-1960s and ending when it hit the wall of the Reagan administration.

Dowie focuses most of his attention on the third wave, from the 1980s to the present; he is harshly critical of much of what he sees. As it gained professional status, the environmental movement lost passion. Fund-raising gained ascendancy over ecological vision. The movement got too close to the corporate world and succumbed too easily to compromise. It remained an elitist group, paying little attention to poor neighborhoods, ethnic groups, environmental justice or the stark poverty of Third World nations. There is sufficient truth to all of the above to strike home, but environmentalism has never been monochromatic. Significant exceptions also exist to most of the above. That said, these attacks represent real challenges that must be dealt with in a clear-eyed manner, for they restrict the movement's effectiveness.

There are plenty of irritating errors in *Losing Ground,* but they are mostly ones of interpretation rather than of scientific fact. The collaboration between Conoco and the Natural Resources Defense Council in Ecuador is much more about learning to work with indigenous peoples than about giving in to industry or self-importance, as Dowie asserts. Market incentives and emissions trading, described in a section called "A Market for 'Bads,' " are yielding successes as well as valuable experiences about how to fine-tune such approaches. A discussion of Vice President Al Gore's potentially misleading remark linking environment and health issues does not accurately reflect his evident, deep understanding of the linkage. The description of the Northwest forest plan is relatively one-sided, giving a false impression that many environmentalists and the Clinton administration caved in to logging interests.

Dowie ends the book with his view of an emerging fourth wave. Its main characteristic is a swing, already under way, back toward the grass roots. Dowie broadly anticipates what the new environmentalism will encompass, including some very pragmatic elements and some fundamental value shifts. My main criticism of his vision is that it will not work by itself: it will need the assistance of the third-wave types who work on national policy and government standards; no effort will succeed without the committed involvement of the private sector as well.

All three of these books will make a lot of environmentalists unhappy. *Noah's Choice* and especially *A Moment on the Earth* display a surprising lack of understanding about how science works: perpetually self-testing, advancing by hypothesis and critical examination, inconveniently nonlinear but wonderfully up front about uncertainty. By and large, humanity acts as if the scientific method is disconnected from our daily lives, and yet a wider awareness of that method would help greatly in framing the current environmental debates.

I hope many environmental professionals will read these books and winnow out some crucial messages we should be thinking about. These kinds of reevaluations have mostly appeared in the U.S., perhaps reflecting the progress of American environmentalism or our own particular national style. In any case, these books show that we have not been listening enough. We have become portrayed as extremists saying, "Don't. Stop. You can't do this," when our real motivation is conservation of opportunity for health, wealth and a better future. We must get that message across if the movement is to persevere.

---

*THOMAS E. LOVEJOY is counselor to the secretary of the Smithsonian Institution for biodiversity and environmental affairs and senior adviser to the executive director of the United Nations Environment Program.*

# WONDERS

by Phylis and Philip Morrison

## *The Physics of Binary Numbers*

The label we give the millennial turn of the calendar looming ahead is a string of four digits, written as 2000 in the everyday convention of number base 10 (of course, the true new millennium begins a year later on January 1, 2001—fine for purists, but a less aesthetically pleasing number). Count leftward from the right-hand end: at the zero power of 10, enter no 1's (that is, place a 0 there); none either at the first power of 10; the second power, again none. But for the third power—$10^3$, or 1,000—you want to count two of them. Total: two times $10^3$ and no more, thus 2,000. You need to distinguish just 10 digits in this scheme: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. String them out as far as you like; each place has meaning.

There are simpler number bases that one could choose. Try the choice closest to primordial, just a set of identical strokes, |||||||||||||||||||... and so on, using only one form of mark, base 1. Now you need a string of 2,000 strokes, dozens of lines of type, nearly uncountable and unusable. Would you prefer a recipe *alla romana?* That is easy: 2,000 becomes just MM. The ancient Romans also used repeated strings, as in MDCCCCLXXXXVI, supplemented with a few more symbols. The mix is quite practical. What they did recalls the assembly of chemical elements that make up the natural world: interminable arrays of atoms spell out all the varieties of matter as repetitive piles composed of fewer than 100 different symbols, rather in the Roman way but without arithmetic.

Try a modern practicality, hexadecimal numbers, base 16. (Consult the calculator held in your computer software; you will find hexadecimal math lurking there.) The number we call 2,000, when written out in conventional hex, is 7D0 = $7(16^2) + 13(16^1) + 0(16^0)$, where the 10 decimal digits are supplemented by six letters, A through F, to make a total of 16 alphanumeric symbols.

Finally, we have the binary numbers, base 2, needing only two digits, 1 and 0 (or on-off, true-false and many another dichotomy). To write the year name in binary, pure machine discourse, first note that it factors into $250 \times 8$ and that multiplication is very easy in base 2. Thus, $250 = 128 + 64 + 32 + 16 + 8 + 2$, which is $2^7 + 2^6 + 2^5 + 2^4 + 2^3 + 0(2^2) + 2^1 + 0(2^0)$, or 11111010 in base 2 notation. Multiply that by 1000 (the way you would write eight, $2^3$, in base 2), to find 11111010000 for 2,000—a simple enough process, if longish.

The choice of base is plainly open to meet our needs. The tally of line strokes—an unwieldy string all of one single symbol—uses the minimal list of digits, just one. It has the longest strings but of the simplest sort. At the other extreme lies a maximal scheme whose list of distinct "digits" is unending: every number becomes its own one-place symbol. Placeholding vanishes; all numbers have only one digit, the shortest possible, but there is an interminable symbol list to learn.

*Matter, gravity and the simplest symmetry of life lead directly to the binary numbers.*

The fertile metaphysical imagination of the Argentine writer Jorge Luis Borges conjured up a character whose mind was able to fathom that logical limit. In his short story "Funes, the Memorious," Borges introduces us to a poor young Uruguayan, a hapless, friendless prodigy by the name of Ireneo Funes. His blessing and curse was that he could forget nothing. Where we might "perceive at a glance three wineglasses on the table, Funes saw all the shoots, clusters, and grapes of the vine." He could easily envision in memory every visible star in the sky. His arithmetic had no placeholders, only arbitrary names corresponding to every number. To count to 7,013, he said, "*Máximo Perez,*" and for 7,014, he said, "*The Train.*" Funes the Memorious employed an unending number base—every number a new symbol, unrelated to any other, without computational structure. Ordinary people cannot learn an unending list of such complicated symbols; Funes's system was just as pure and just as impractical as the opposite limit, base 1.

The minimal useful number base, base 2, forms the underpinning of binary logic and digital computing. Logic is a splendid path to mathematics, but in fact the oldest known use of binary notation was derived not from logic nor from abstraction but arose, surprisingly, from a strictly physical use in the Indus River cities of the Harappān culture, more than four millennia back.

Between the world wars, the Boston Museum of Fine Arts dug in the big, brick-built ruins of Mohenjo-Daro, once a great city, on the Indus. Among the other finds, an odd assortment that the museum displayed in a small glass case caught our attention years ago. At first glance, it seemed only half a dozen pebbles, such as any child might collect. Not quite: it was a display of weights used in the beadmakers' quarter of that old city, long a center for the export of beads, often the mineral carnelian, an ancient specialty of the region.

The case contained a nice set of carefully made stone cubes and a few other examples of formal weights. But what was that handful of unworked pebbles? Although they were in no way modified by the hand of their ancient collectors, these stones were nonetheless a set of weights, having closely correct weight ratios that doubled along the set in the pattern, 1, 1, 2, 4, 8, 16.... The ancient weights were much like those still in use today; in fact, the Harappan's fundamental unit was just a little lighter than our modern ounce. Here is an artifact of pure information from the mind of the long-vanished collector. It seems likely that our old European standard of 16 ounces to the pound is a relic of the same idea. We cannot say whether it was shared across the millennia or reinvented by independent craftsmen, located far apart, solving a parallel problem of measurement.

Phylis visited the glacier-borne pebbles that litter the South Beach of Martha's Vineyard, part of the great terminal moraine left behind by the retreating ice sheets from the last Ice Age. There she undertook to replicate those artifacts so cleverly assembled by the Indus people long ago. She first made the instrument we know lies behind the pebble weights: the equal-arm balance, examples of which are found among the Indus ruins and which remain in

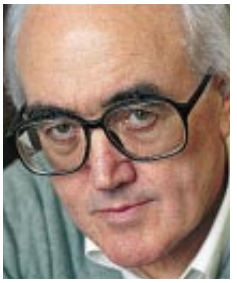
use there today. Such a device consists only of two light pans held by fibers to the ends of a uniform stick, itself suspended, balanced at the middle and adjusted to ride level, then checked by an exchange of pan loads. That balance has looked about the same for at least 6,000 years; you can see it represented still in your local courthouse.

To use it, choose any pebble that suits your needs for a minimum weight. Put the stone in one pan and seek patiently by trial and error for another pebble that balances the first. Then you have the unit twice, 1 and 1. Put them both into one pan and find another to balance their combined weight. That next balance pebble represents 2 (or 1 and 1). Put those three pebbles in one pan, and you can seek 4, then 8. Now you can reach any integer weight below 16 units by combination—just as you can write every integer number up to 16 using five sequential slots marked with a 0 or 1 at each place.

The two of us still have our set from Martha's Vineyard, natural weights, made true by human selection alone, our own artifact of pure information. The binary scale used by the Harappān people and re-created by us is realized through an unworked but by no means unstudied set of small weights. The intrinsic twofold, bilateral symmetry of the equal-arm balance expresses itself in the visible identities $1+1=2$, $1+1+2=4$, and so on. It is addition without mathematics. The only counting is that of the real weights that rest on each pan. Placeholding happens not by convention but by the physical weight sensed by hand and eye.

The amazing additivity of gravity gives the same result at any human scale. That property may seem self-evident, even inevitable, but imagine how different our world would be if it were not so. Matter, gravity and the simplest of geometric symmetries—the bilateral one shared by many organisms, ourselves included—lead directly to the binary numbers.

There are other ancient users of binary patterns, not from weighing but from counting a heap of similar discrete objects to build a dichotomy. That task is often done by using a random handful that is then paired out into an even or an odd number. The Yorubans divined by grabbing palm nuts; the Chinese, with the *I Ching,* by casting batches of yarrow stalks. But long before either of these innovations, the dwellers along the Indus had developed a real craftsman's path to this same mathematical idea, with its simplicity, profundity and breadth of use. Viva the ancient symmetry of the balance!



# CONNECTIONS

by James Burke

## *Show Time!*

Even though I've spent most of my adult life working for the small screen (television, that is), I can't resist that magic moment in the cinema theater when the lights go down, the pictures come up in total-everything-scope, and I am enveloped in Surround-sound. Even when the actual product is kid stuff, as it was recently when I was the adult-in-charge at a visit to a Disney movie. I found myself thinking, as I watched, about what a pity is it that Hollywood so thoroughly sanitizes the plotlines; these days you hardly ever get to see the real thing.

I'm referring to the tales by the Brothers Grimm, from which Disney took so much, but which were not quite the saccharine stories we've grown accustomed to. In the originals, Cinderella's sisters had their eyes pecked out, Rapunzel got pregnant, and the Big Bad Wolf ate both Red and her granny. Even the Grimms' own editor toned it all down for the second edition.

Of course, the two brothers were dealing with pretty raw stuff in the Germanic folktales they had collected as material for their stories. Around 1806 they began copying down the oral fables passed around by shepherds, wagoners, tinkers, gypsies and peasants. Over the years the Grimms listened to what they thought of as the ancient voice of the Fatherland: legends laced with violence, cruelty, racism, contempt for foreigners, authoritarianism. Stuff that would warm the hearts of Nazis more than a century later.

The Grimms were part of the great back-to-prehistory craze sweeping Romantic Europe in the early 19th century. This folklore movement drew its inspiration from the new science of linguistics and from the work of a Welsh jurist in Calcutta. Sir William Jones's job was to apply British law to Hindus in British India, so he studied the local culture. During the course of his investigations, he came across the amazing ancient Indian language of Sanskrit and promptly spread the word about its words. These were so similar to comparable words in Greek and Latin (and Germanic and Persian and Celtic and even Armenian and Albanian) that Jones reckoned Sanskrit had to be nearly everybody's ancient mother tongue.

Germans went particularly ape about Sanskrit because its existence suggested that the Teutonic races might have a cultural heritage that extended back every bit as far as that of the French (who, in the Napoleonic Wars, were clobbering Germany at the time).

The Indo-European mania kicked off by the discovery of Sanskrit affected even the most respectable types. One of these was a mathematician called Carl Gauss, who was at the University of Göttingen with the Grimms. Gauss's musings carried a lot of weight because of his reputation as something of a genius. He had earned his position at Göttingen in the pluck of youth: in 1794, at age 17, he developed a trick for calculating a planetary orbit even if all you had were three quick glimpses.

Which more or less describes the discovery of the first asteroid, Ceres. On New Year's Day, 1801, the Italian astronomer Giuseppe Piazzi spotted the new heavenly body but watched it through only nine degrees of its orbit before he fell sick. When Piazzi's health had improved, the weather had not. By the time the Palermo sky was finally clear enough for him to look again, there was Ceres: gone. Now, the discovery of a new planet had caused quite a Euro-stir, and the loss of it, even more so. But all was not gloom and doom. Thanks to young Gauss's fancy math (officially, his "method of least squares"), bingo, a year later, Ceres turned up exactly where Gauss's figures said it would be. That feat made Gauss an instant celebrity, with astronomical job offers from all over, of which he chose Göttingen.

One of Gauss's fans was so impressed by the whole affair that he couldn't bring himself to give his own name to the new element he'd just discovered. The chap in question was a stout, hypochondriac Swedish gourmet and ladies' man who gave chemistry lessons to royalty, who in return gave him a title. Baron Jöns Jakob Berzelius went everywhere, wrote travelogues and knew everybody worth knowing. And the next time you're enjoying a refreshing draught of $H_20$, you might want to recall that it was Berze-

lius who came up with modern chemical symbology.

Big B. was also the hottest thing in Europe on the blowpipe, a device resembling a set of bellows, which he used to raise flame temperatures to 1,500 degrees Celsius. Under such ferocious heat, substances become incandescent and glow in a characteristic way that reveals their constituent elements and compounds. In this way, Berzelius went about analyzing all kinds of mineral specimens for such hip-hop types as Goethe (Berzelius also inspected meteorites, ancient Egyptian mortar and a Canadian trapper's gastric juice). Which was why, in 1803, he was well able to examine a strange stone, identify the new element I mentioned and, in honor of Gauss, name it "cerium" instead of "berzelium."

One of the things I love about exploring the web of change is just the kind of connection we've stumbled across. Here we are in northern Scandinavian nowhere, watching this dapper fat man who was about to make yet *another* notable contribution to the modern world, thanks to the activity of somebody thousands of miles away who probably never even heard of him. I'm talking about that great inventor and even greater self-publicist, Thomas Alva Edison, and his two dazzling contributions to our story: the creation of the incandescent lightbulb (deliberate) and the effect it had (accidental).

In 1882 the opening of Edison's power station on Pearl Street in New York City must have made it clear to even the dimmest investors that they should get out of gaslight company stocks, which were heading for the floor. But wait. Suddenly, another aristocrat-to-be appears on the scene, armed with a brilliant solution for the gas companies, thanks to Berzelius and his mystery stone.

Carl Auer von Welsbach of Vienna discovered that if you impregnated a gauzy wrapping (or "mantle") made of Sea Island cotton with the right minerals, 1 percent of which was good old cerium nitrate, and then stuck this gauze around a gas flame, the cerium glowed enough to increase by seven times the light ordinarily given off by the flame. That 1885 invention, known as the Welsbach mantle, netted Welsbach an Austro-Hungarian baronetcy and the right to choose a (stupendously expectable) family motto: "More Light."

And the new mantle did keep the gaslight business in business up until World War I (you can still see it around campfires today). Some time around 1900, in a seedy Chicago rooming house, a fellow with the improbably Hollywood name of Lee De Forest was looking at a Welsbach light and got the idea for a flame detector. He envisioned a gizmo that would pick up variations in the charge from an (electrified) gauze mantle, as the current in the mantle was affected by the presence of a flame. De Forest's problem was how to make the imperceptible variations in current big enough to perceive. So he co-opted a strange phenomenon that Edison had accidentally stumbled across when he had noticed his hot lightbulb filament giving off particles that dirtied the metal baseplate of the bulb.

The phenomenon—which Edison named the Edison Effect, patented, filed and forgot—was eventually revealed to result from the steady stream of electrons that boils off the filament and rushes headlong for the baseplate. De Forest placed a metal gauze between filament and baseplate, then used the powerful flow of electrons (the suddenly handy Edison Effect) to magnify tiny charge fluctuations in the gauze. He called this all-purpose electrical signal booster an audion. Great for flame detectors. Even better for telephone amplifiers. In 1914 the first New York–San Francisco telephone line opened, audion-amplified all the way.

Now comes a real bit of irony. One of the many clever things Baron Berzelius had done 100 years earlier with his blowpipe was to isolate another element, which he called selenium and which turned out to be useful in electric circuitry. One day in 1873 an operator at the transatlantic telegraph cable terminal on the Irish island of May noticed that the current in his selenium resistors varied strangely: high in sunlight, low in dark. Turned out that the selenium was giving off electricity in response to light.

From 1900 a Professor Joseph T. Tykociner at the University of Illinois worked on developing a selenium-based voice recorder. He shone a light (whose intensity varied according to the strength of the vibrations of a microphone membrane) through film stock, to make prints of varying exposures. These in turn would allow varying amounts of light to penetrate the film and to hit a selenium cell, which in response gave off varying current that then re-created the original sound on a loudspeaker. This was great, except for the fact that the signal was too weak to hear.

Then, voilà, in 1923, the new Movietone system from Western Electric, incorporating the amplifying power of the audion. And suddenly movies became talkies. Which is why I heard every word of the Grimm fairy tale in that Disney show. That's all, folks!

ESSAY by James Randi

# Investigating Miracles, Italian-Style

It appears that Italians are not asking as many questions as they used to. At an earlier time, such curious souls as Leonardo, Galileo and Giordano Bruno dared to bring scientific skepticism to bear on the popular fallacies of their day. Today the Italian media and public swoon over weeping plaster casts of religious figures, charismatic faith healers and anything that bleeds, sweats or moves in an unexpected way. But I take it as a cautionary sign that rampant disregard for rational thought is sweeping a nation that has made so many contributions to modern science.

Fortunately, Italy still has people like Luigi Garlaschelli, an energetic and observant scientist in the department of organic chemistry at the University of Pavia. When not occupied with his bread-and-butter research, he and his colleague Franco Ramaccini work with the Italian Committee for the Investigation of the Paranormal (CICAP), providing rational explanations for supposed miracles.

The Italian marvel of the moment consists of a nationwide epidemic of painted plaster or ceramic Madonna figures that cry tears of water or blood. The owners of these religious objects call on devout observers, who are kept at a respectful distance, to testify to the wonder. In one recent case, no less than a chief of police affirmed that he had witnessed the phenomenon. With that sort of validation, the case for a supernatural origin seems—to the media and thus to much of the public—confirmed beyond a reasonable doubt.

Even the slightest scientific prying shows that something suspicious is going on. The bloody tears of one well-promoted Madonna, for instance, tested to be of masculine origin; undaunted, the local priest declared that the Madonna would of course shed male blood—the blood of her Son. A magistrate investigating the case demanded that the owner of the wondrous statue submit to DNA tests of his own blood, but the parishioner piously objected on the grounds that such miracles should not ever be questioned, let alone tested.

It is not hard to understand why. Garlaschelli has demonstrated that almost any ceramic or plaster figure can be prepared to exude mock tears or blood by drilling a tiny hole in the top of the head, injecting a liquid and then scratching away a thin line of glaze below the eyes. More simply, a hollow plaster figure can be filled with a liquid, then drained. The porous material retains some fluid, and tearlike drops will begin to flow from the eyes while a pool of liquid gathers about the base (an attendant condition that is often described). The process leaves no obvious clues. The reluctance of owners to let skeptics such as Garlaschelli examine their Madonnas has made it difficult to assess his explanation, however.

JAMES RANDI

From time to time, another variety of sanguinary miracle is trumpeted in the Italian press. Back in 1264, Pope Urban IV established the Feast of Corpus Christi in honor of a mass at Bolsena, a town near Rome, when the communion host allegedly dripped blood. Since then, bread, polenta, cooked potatoes and other foodstuffs have regularly been reported to bleed.

The bright crimson patches that spontaneously appear on the food certainly resemble blood, but chemical analysis has revealed no trace of hemoglobin, believed to be a requirement even of supernatural blood. In contrast, the fungus *Serratia marcescens*—a harmless microorganism known to thrive on most starchy, nonacidic foods in warm, moist environments—has been very evident at the scene of the miracle. *S. marcescens,* it should be noted, has a startling, blood-red color. Garlaschelli has cultured the fungus on ordinary white bread and finds that it exactly duplicates the acclaimed miracle.

Johanna Cullen, a pre-med student at George Mason University, has researched the history of the Miracle of Bolsena. She finds that miraculous appearances occur from May to September (with a dramatic peak in July), when conditions are most likely to be warm and humid; the red spots appear on a variety of foods, including chicken, that satisfy the requirements for the growth of *S. marcescens.* Cullen concludes that the most celebrated miracle of the 13th century "may be more microbiological than metaphysical."

The most enduring religious curiosity that has attracted Garlaschelli's sleuthing is a heavily publicized, annual rite that takes place at the Cathedral of Naples. The story goes that when the third-century Christian known as Gennaro was martyred by decapitation, a bystander bottled some of the victim's blood and saved it. Every year an archbishop at the Cathedral of Naples trots out two vials said to contain the martyr's congealed blood. When inverted before a packed congregation, the substance in the vials liquefies and changes from reddish brown to bright red, a transformation taken as a sign that all is well with Naples.

Garlaschelli and Ramaccini have whipped up a mixture that replicates the transformation: it liquefies and changes color with a simple shake, even though neither of the investigators is an archbishop. They made it from materials found in the vicinity of the relic, using techniques available to medieval tinkerers. As for the "blood" in the vials, the church has steadfastly refused to allow samples to be taken, so no chemical analysis has been performed.

I do not object to faith in wonders so long as it does not insist on being taken as fact. But when blind belief refuses scientific inquiry, I bristle. We have fought long and hard to escape from medieval superstition. I for one do not wish to go back.

---

*JAMES RANDI is a professional magician, author and lecturer.*