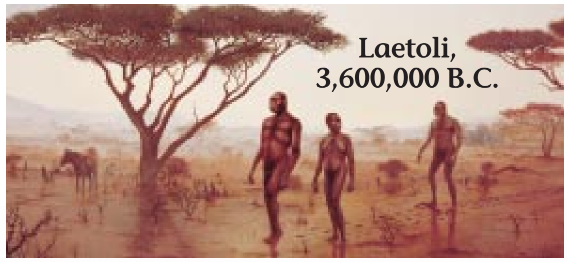
WEIGHTLESSNESS AND HEALTH • NEW ELEMENTS • ATTENTION-DEFICIT DISORDER

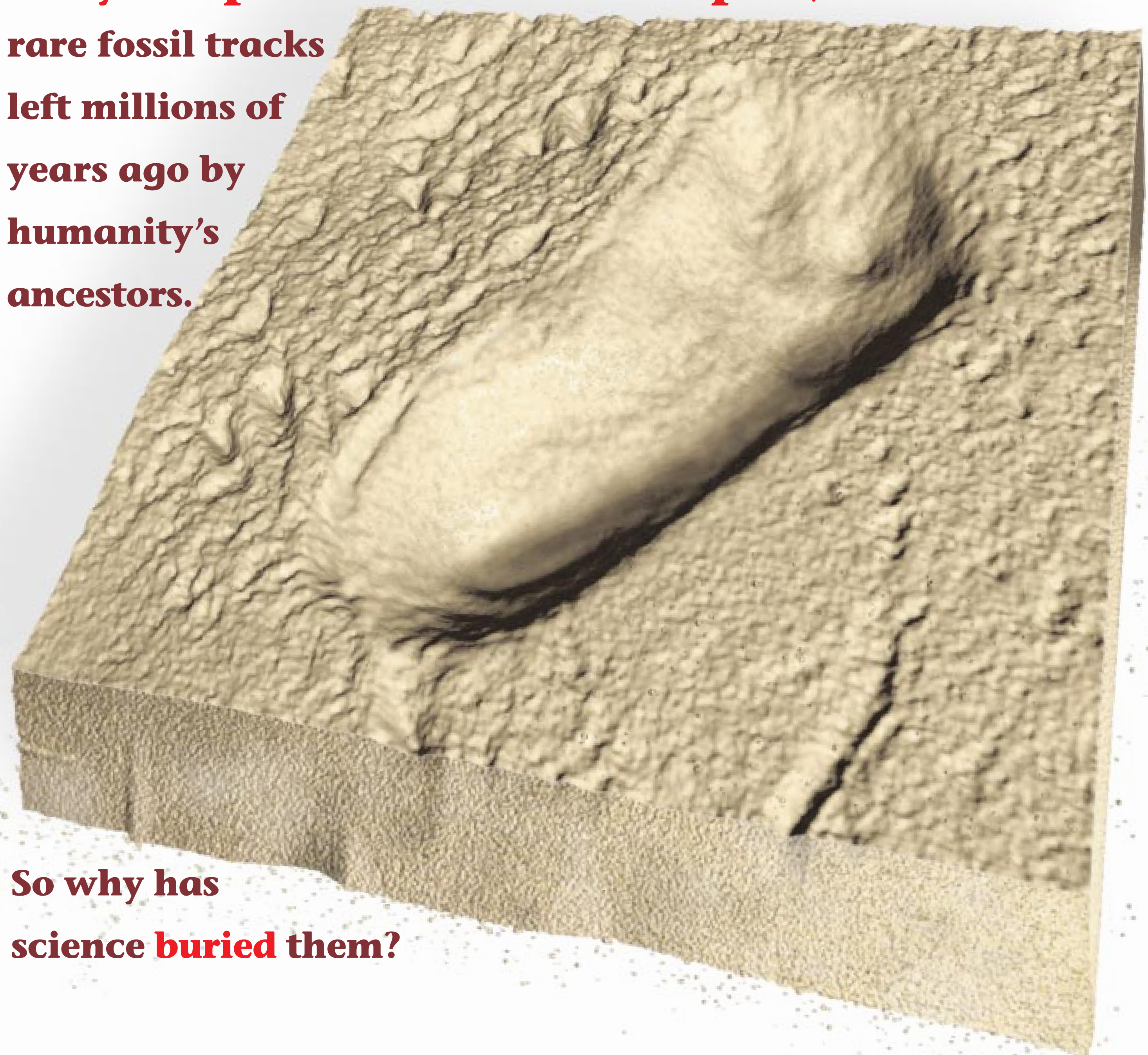
SCIENTIFIC AMERICAN

SEPTEMBER 1998 $4.95



They are precious clues to the past,

rare fossil tracks left millions of years ago by humanity's ancestors.



So why has science buried them?

A Last Look at Laetoli

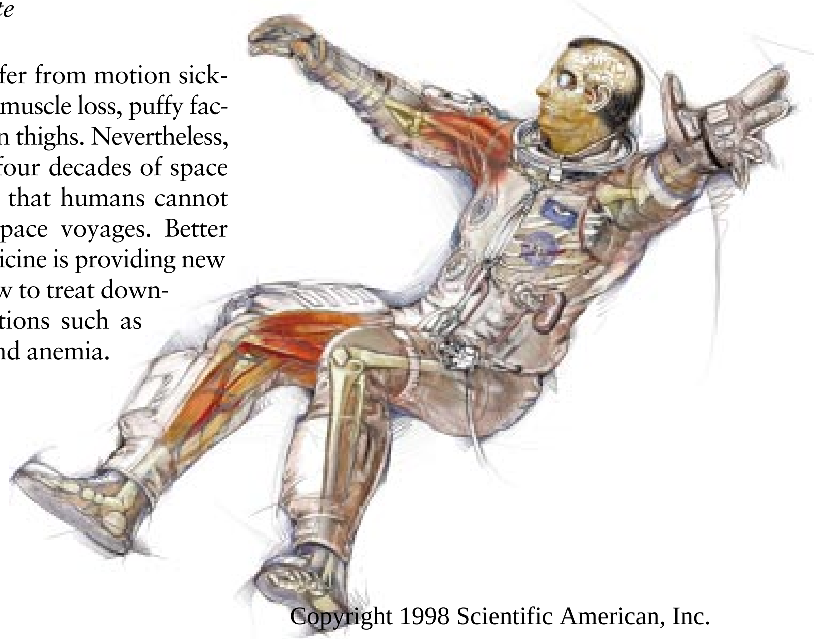
SCIENTIFIC AMERICAN

September 1998 Volume 279 Number 3

A rare dugong
(page 20)

The 3,600,000-year-old footprints found 20 years ago in the Laetoli area of northern Tanzania vividly evoked how early human ancestors may have lived. To protect those tracks, scientists have now painstakingly reburied them. The authors, who led the conservation project, explain why and how it was done. In addition, with anthropologist Ian Tattersall and artist Jay H. Matternes, they describe how views of the footprint makers have changed.

Astronauts suffer from motion sickness, bone and muscle loss, puffy faces and shrunken thighs. Nevertheless, no ailment in four decades of space travel suggests that humans cannot survive long space voyages. Better still, space medicine is providing new clues about how to treat down-to-earth conditions such as osteoporosis and anemia.



About the Cover
Footprint made by an ancestor of *Homo sapiens* proves that early hominids were fully bipedal—long before the invention of stone tools or the expansion in brain size. Image by Slim Films.

THE SCIENTIFIC AMERICAN WEB SITE

Voyage to an undersea volcano on board the submersible *Alvin:* http://www.sciam.com/explorations/1998/070698atlantis/index.html
And check out enhanced versions of this month's feature articles and departments, linked to other science resources on the World Wide Web.

www.sciam.com

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 1998 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian BN No. 127387652RT; QST No. Q1015332537. Subscription rates: one year $34.97 (outside U.S. $49). Institutional price: one year $39.95 (outside U.S. $50.95). 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 sacust@sciam.com **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.**

FROM THE EDITORS

Go Ahead, Walk in the Mud

Percy Bysshe Shelley's poem *Ozymandias* describes the cracked and toppled statue of an ancient potentate: "My name is Ozymandias, king of kings: Look on my works, ye Mighty, and despair!" The destruction of his once great empire might be mistaken as punishment for hubris. The grimmer reality is that Ozymandias could have been the soul of modesty and nature would have ground his works to powder just the same.

Three and a half million years ago a trio of furry bipeds walked across an African savanna caked with damp volcanic ash. Maybe it was a happy family stroll on a Saturday afternoon; maybe they crept fearfully through a predator's hunting ground. We will never know (but we can present the best, most recent guess; see page 44 in "Preserving the Laetoli Footprints," by Neville Agnew and Martha Demas). Odds are that for those creatures, it was just another ordinary walk on another ordinary day. The hidden struggles of their lives, the glimmerings of hope and pride they nurtured are all gone forever. Meanwhile their muddy footprints have lasted 700 times longer than recorded history. Where is the poetic justice?

J. PAUL GETTY TRUST

AN EARLY STEP in human evolution

We moderns can expect no better treatment. Wood rots, paper burns, stone splits, plastic corrodes, glass shatters, metal rusts. If humans disappeared tomorrow and no one was left to mow the lawns, paint the walls and fix the pipes, even the sturdiest of our concrete and steel structures would be mossy rubble in roughly 10,000 years. The irony is that already ancient stone monoliths like the Egyptian pyramids might be among the last artifacts to vanish from view.

To put it another way, imagine watching the events of the next million years on that uninhabited Earth, all compressed into a 100-minute feature film. Don't be late finding a seat in the theater: within the first 60 seconds nearly every large trace of civilization will have melted into the terrain. Nothing to do then but watch the forests grow (talk about a slow second act). Our world would survive as a stratum of buried junk.

So future anthropologists may not be assessing the heights of our accomplishments from the Mona Lisa, or Shakespeare, or the Golden Gate Bridge, or a space shuttle. They may be measuring the tooth marks on our chewed pencils; checking the metallurgy of old screwdrivers; deducing the economy from phone books in landfills. Perhaps the act for which you will be longest remembered was something you wrote in a wet cement sidewalk when you were six years old: I WUZ HERE.

JOHN RENNIE, *Editor in Chief*
editors@sciam.com

SCIENTIFIC AMERICAN, INC.
415 Madison Avenue
New York, NY 10017-1111
(212) 754-0550
PRINTED IN U.S.A.

LETTERS TO THE EDITORS

Letters about the article by Robert Plomin and John C. DeFries on the genetics of intelligence began pouring in as soon as the May issue hit subscribers' mailboxes. Some readers looked forward to a day when such research could, for instance, help teachers make "a preemptive strike against reading problems," as suggested by Jonathan Bontke of St. Louis. But many more people expressed concern about potential misuse of these findings. Henry D. Schlinger, Jr., a professor of psychology at Western New England College, even questioned the rationale for seeking an intelligence gene: citing the authors' assertion that biology is not destiny, Schlinger wondered, "Then why should we care about what the heritability of a particular trait is?"

BRAIN TEASER

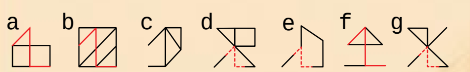
I was intrigued by the interesting article "The Genetics of Cognitive Abilities and Disabilities," by Robert Plomin and John C. DeFries [May]. I was puzzled, however, by a sample test that appeared in the article (*right*). According to the answer, the figure specified appears in only a, b and f. But it also appears in d, e and g. Although the test's instructions state explicitly that the figure must always be in the position shown, not upside down or on its side, nothing is said about the figure needing to be the same size as illustrated. This example demonstrates a problem facing anyone who has taken a cognitive ability or related test—having to gauge the intelligence and perceptions of the people who wrote the test. Are they even aware of all possible answers? In this case, the answer is no, and the child pays the penalty.

JIM BAUGHMAN
West Hollywood, Calif.

Plomin and DeFries's article left me skeptical as to whether the authors' approach is likely to be scientifically fruitful. For instance, their reliance on tests of so-called cognitive ability (what used to be called IQ) as measures of general categories of intellectual functioning is questionable. Psychologists' claims about what such tests measure rely entirely on assumed correlations between test performance and cognitive ability. In the physical and medical sciences, inferring causes from correlations alone is typically considered an error. Imagine Galileo and Newton proceeding similarly: on noticing a high (but imperfect) correlation between the speed at which falling objects hit the ground and the height from which they fall, the scientists simply regard falling objects as a new type of measuring instrument for estimating height, forgoing a scientific understanding of the reason for the correlation (gravity), as well as its imperfect character (air friction).

J. M. CRONKHITE
Department of Physics
Georgia Institute of Technology

4. HIDDEN PATTERNS: Circle each pattern below in which the figure appears. The figure must always be in this position, not upside down or on its side.



JENNIFER C. CHRISTIANSEN

OFFICIAL ANSWERS (red) *to this test of cognitive ability do not include the alternative answers* (dashed red lines) *suggested by several readers.*

Until tests are devised to measure the full array of human cognitive abilities, including teamwork and leadership skills, rhythm, curiosity, attention, self-confidence, imagination and so on, we will have little luck in teasing apart the various genetic and environmental mechanisms of "intelligence." It would be more profitable to explore human cognitive abilities from a different angle. How can we help all children take maximum advantage of their unique genetic endowments? How close has anyone come to reaching his or her genetic potential, and how was this achieved?

BOB KOHLENBERGER
Burlingame, Calif.

Plomin and DeFries reply:

When tests of specific cognitive abilities are actually administered, practice items and examples clearly illustrate the types of responses that are considered to be correct, so the problem encountered by Baughman should not be an issue. In response to the comment by Cronkhite, the old shibboleth that correlations do not prove causation is not relevant. Tests of statistical significance, including analysis of variance and covariance, can be incorporated as special cases of multiple regression and correlation analysis.

A more relevant issue is the experimental power of the design. Although twin and adoption studies are quasiexperimental in that people are not randomly assigned to be members of a set of twins or to be adopted, the studies do provide considerable power to address the questions of nature and nurture in relation to cognitive abilities and disabilities.

We agree that there is much more to life than cognitive abilities, including teamwork, leadership skills and so on. But these traits are not highly correlated with cognitive abilities and thus were not discussed in our article. And we are interested in trying to help children maximize their genetic endowments: our review concerned the extent to which such endowments for cognitive abilities and disabilities are important.

X-RAY VISION

As an otolaryngologist, I was interested in W. Wayt Gibbs's article on radiation therapy in the May issue ["Taking Aim at Tumors," News and Analysis]. I noticed, however, that the caption for the picture on page 20 is incorrect. Most standard x-rays and CT scans show high-density areas, such as bone, to be light, and they show low-density areas, such as air, to be dark. The black area in the picture, listed as the esophagus, is actually air inside the larynx. The walls of the esophagus are generally collapsed on each other and don't contain any air.

JACK ALAND, JR.
Birmingham, Ala.

REMOTE CONTROL

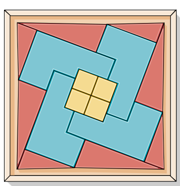
As a retired communications engineer, I was naturally very intrigued with your articles on the upcoming improvement in the technology of television ["The New Shape of Television," May]. Now, if we could only see corresponding improvement in the quality of the programming, we would have something worth watching.

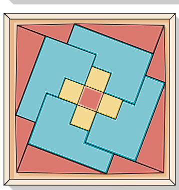
EUGENE V. KOSSO
Gualala, Calif.

Letters to the editors should be sent by e-mail to editors@sciam.com or by post to Scientific American, 415 Madison Ave., New York, NY 10017. Letters may be edited for length and clarity. Because of the considerable volume of mail received, we cannot answer all correspondence.

Solution to the Martin Gardner Puzzle

In "A Quarter-Century of Recreational Mathematics," by Martin Gardner [August], the author presented his Vanishing Area Paradox, illustrated by the two figures below. Each pattern is made with the same 16 pieces, but the lower pattern has a square hole in its center. Where did this extra bit of area come from? The key to the paradox is that the large and small right triangles are not similar—their acute angles are slightly different. Because of this difference, the upper pattern is concave: the angles at the corners are slightly less than 90 degrees, so the sides of the figure buckle inward. In the lower pattern, the corner angles are slightly more than 90 degrees, so the sides bulge outward. The difference in area between the two figures is equal to the area of the square hole in the lower pattern.





IAN WORPOLE

SCIENTIFIC AMERICAN
Advertising and Marketing Contacts

NEW YORK
Kate Dobson, PUBLISHER
tel: 212-451-8522, kdobson@sciam.com
415 Madison Avenue
New York, NY 10017
fax: 212-754-1138

Thomas Potratz, EASTERN SALES DIRECTOR
tel: 212-451-8561, tpotratz@sciam.com

Kevin Gentzel
tel: 212-451-8820, kgentzel@sciam.com
Randy James
tel: 212-451-8528, rjames@sciam.com
Stuart M. Keating
tel: 212-451-8525, skeating@sciam.com
Wanda R. Knox
tel: 212-451-8530, wknox@sciam.com

Laura Salant, MARKETING DIRECTOR
tel: 212-451-8590, lsalant@sciam.com
Diane Schube, PROMOTION MANAGER
tel: 212-451-8592, dschube@sciam.com
Susan Spirakis, RESEARCH MANAGER
tel: 212-451-8529, sspirakis@sciam.com
Nancy Mongelli, PROMOTION DESIGN MANAGER
tel: 212-451-8532, nmongelli@sciam.com

ASSISTANTS: May Jung, Beth O'Keeffe

DETROIT
Edward A. Bartley, MIDWEST MANAGER
3000 Town Center, Suite 1435
Southfield, MI 48075
tel: 248-353-4411, fax: 248-353-4360
ebartley@sciam.com
OFFICE MANAGER: Kathy McDonald

CHICAGO
Randy James, CHICAGO REGIONAL MANAGER
tel: 312-236-1090, fax: 312-236-0893
rjames@sciam.com

LOS ANGELES
Lisa K. Carden, WEST COAST MANAGER
1554 South Sepulveda Blvd., Suite 212
Los Angeles, CA 90025
tel: 310-477-9299, fax: 310-477-9179
lcarden@sciam.com
ASSISTANT: Stacy Slossy

SAN FRANCISCO
Debra Silver, SAN FRANCISCO MANAGER
225 Bush Street, Suite 1453
San Francisco, CA 94104
tel: 415-403-9030, fax: 415-403-9033
dsilver@sciam.com
ASSISTANT: Rosemary Nocera

DALLAS
The Griffith Group
16990 Dallas Parkway, Suite 201
Dallas, TX 75248
tel: 972-931-9001, fax: 972-931-9074
lowcpm@onramp.net

International Advertising Contacts

CANADA
Fenn Company, Inc.
2130 King Road, Box 1060
King City, Ontario
L7B 1B1 Canada
tel: 905-833-6200, fax: 905-833-2116
dfenn@canadads.com

EUROPE
Roy Edwards, INTERNATIONAL ADVERTISING DIRECTOR
Thavies Inn House, 3/4, Holborn Circus
London EC1N 2HB, England
tel: +44 171 842-4343, fax: +44 171 583-6221
redwards@sciam.com

BENELUX
Reginald Hoe Europa S.A.
Rue des Confédérés 29
1040 Bruxelles, Belgium
tel: +32-2/735-2150, fax: +32-2/735-7310

MIDDLE EAST
Peter Smith Media & Marketing
Moor Orchard, Payhembury, Honiton
Devon EX14 OJU, England
tel: +44 140 484-1321, fax: +44 140 484-1320

JAPAN
Tsuneo Kai
Nikkei International Ltd.
CRC Kita Otemachi Building, 1-4-13 Uchikanda
Chiyoda-Ku, Tokyo 101, Japan
tel: +813-3293-2796, fax: +813-3293-2759

KOREA
Jo, Young Sang
Biscom, Inc.
Kwangwhamun, P.O. Box 1916
Seoul, Korea
tel: +822 739-7840, fax: +822 732-3662

HONG KONG
Stephen Hutton
Hutton Media Limited
Suite 2102, Fook Lee
Commercial Centre Town Place
33 Lockhart Road, Wanchai, Hong Kong
tel: +852 2528 9135, fax: +852 2528 9281

OTHER EDITIONS OF SCIENTIFIC AMERICAN

Le Scienze
Piazza della Repubblica, 8
20121 Milano, ITALY
tel: +39-2-655-4335
redazione@lescienze.it

Spektrum der Wissenschaft
Verlagsgesellschaft mbH
Vangerowstrasse 20
69115 Heidelberg, GERMANY
tel: +49-6221-50460
redaktion@spektrum.com

Investigacion y Ciencia
Prensa Científica, S.A.
Muntaner, 339 pral. 1.ª
08021 Barcelona, SPAIN
tel: +34-93-4143344
precisa@abaforum.es

Pour la Science
Éditions Belin
8, rue Férou
75006 Paris, FRANCE
tel: +33-1-55-42-84-00

Majallat Al-Oloom
Kuwait Foundation for the Advancement of Sciences
P.O. Box 20856
Safat 13069, KUWAIT
tel: +965-2428186

Swiat Nauki
Proszynski i Ska S.A.
ul. Garazowa 7
02-651 Warszawa, POLAND
tel: +48-022-607-76-40
swiatnauki@proszynski.com.pl

Nikkei Science, Inc.
1-9-5 Otemachi, Chiyoda-Ku
Tokyo 100-8066, JAPAN
tel: +813-5255-2821

Svit Nauky
Lviv State Medical University
69 Pekarska Street
290010, Lviv, UKRAINE
tel: +380-322-755856
zavadka@meduniv.lviv.ua

Ke Xue
Institute of Scientific and Technical Information of China
P.O. Box 2104
Chongqing, Sichuan
PEOPLE'S REPUBLIC OF CHINA
tel: +86-236-3863170

50, 100 AND 150 YEARS AGO

SCIENTIFIC AMERICAN

SEPTEMBER 1948

THE TRANSISTOR APPEARS—"Within the past few months a group of physicists at the Bell Telephone Laboratories has made a profound and simple finding. In essence, it is a method of controlling electrons in a solid crystal instead of in a vacuum. This discovery has yielded a device called the transistor (so named because it transfers an electrical signal across a resistor). Not only is the transistor tiny, but it needs so little power, and uses it so efficiently (as a radio amplifier its efficiency is 25 per cent, against a vacuum tube's 10 per cent) that the size of batteries needed to operate portable devices can be reduced. In combination with printed circuits it may open up entirely new applications for electronics."

PRIMARY CARE—"Primitive medicine men learned long, long ago what modern medicine is just rediscovering—that distinctions between the mind and the body are artificial. The primitive doctor understands well the nature of psychogenic illness. Among pre-literate peoples, as among those in more civilized societies, these emotional discomforts are easily translated into neurotic symptoms. This illustration shows a sand painting made by a Navaho medicine man, designed to treat mind and body in a curing ceremony. The painting is made on the floor of a hut, the patient is laid upon it and paint is rubbed over him."

Navaho sand painting for a curing ceremony

SEPTEMBER 1898

ON EVOLUTION—"At the Cambridge Congress of Zoology Prof. Ernst Haeckel read a fascinating paper on the descent of man. He does not hesitate to say that science has now definitely established the certainty that man has descended through various stages of evolution from the lowest form of animal life, during a period of a thousand million years. 'The most important fact is that man is a primate, and that all primates—lemurs, monkeys, anthropoid apes, and man—descended from one common stem. Looking forward to the twentieth century, I am convinced it will universally accept our theory of descent.'"

NO TRANSISTOR NEEDED—"Mr. José Bach describes in *L'Illustration* an instrument by means of which the Brazilian Indians communicate with each other at a distance. In each *malocca*, or dwelling, there is a *cambarisa*, a sort of wooden drum buried for half of its height in sand. When this drum is struck with a wooden mallet, the sound is distinctly heard in the other drums situated in the neighboring *maloccas*. The blows struck are scarcely audible outside of the houses in which the instrument is placed, so it is certain that the transmission of the sound takes place through the earth, the drums doubtless resting upon the same stratum of rock."

SEPTEMBER 1848

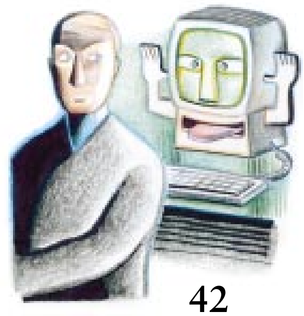
GOLD!—"News has reached us from California of the discovery of an immense bed of gold of *one hundred* miles in extent, near Monterey. It is got by washing out river sand in a vessel, from a tea saucer to a warming pan. A single person can gather an ounce or two in a day, and some even a hundred dollars' worth. Two thousand whites and as many Indians are on the ground. All the American settlements are deserted, and farming nearly suspended. The women only remain in the settlements. Sailors and captains desert the ships to go to the gold region." [*Editors' note: The Sutter's Mill find led to the 1848 California gold rush.*]

FOSSIL THEORY—"The fossiliferous rocks in the sedimentary strata present us with the different objects of bygone periods, and it is astonishing what minute and delicate objects have been transmitted to us: the traces of footsteps on wet sand; undigested food; even the ink bag of the sepia [cuttlefish] has been found so perfect that the same material which the animal employed centuries, nay, thousands of years ago, to preserve itself from its enemies, has served for color to paint its likeness with!—Alexander Humbolt."

FLOATING TUNNEL—"One of the most extraordinary plans submitted to the French Academy of Sciences is that of M. Ferdinand, engineer, who proposes a floating tunnel from Calais to Dover, for the wires of the electric telegraph, and large enough to be traversed by small locomotives, for the conveyance of passengers. A tunnel for the wires of the electric telegraph we believe to be perfectly practicable and requires no great genius to conceive or construct, but a floating tunnel for locomotives is as preposterous as it is useless." [*Editors' note: See News and Analysis, "Tunnel Visions," July 1997, for an update on useful floating tunnels now being planned.*]

NEWS AND ANALYSIS

18
SCIENCE AND THE CITIZEN



IN FOCUS

FORESTALLING VIOLENCE

American youths are suffering an epidemic of violence, both in and out of the classroom. Designing effective prevention programs is proving difficult

MARK PETERSON SABA

***CONFLICT RESOLUTION** training courses are required in 61 percent of U.S. school districts. Despite the popularity of such programs, many have not been evaluated for effectiveness.*

Heading back to school brings to mind shiny new notebooks, multicolored pens, the latest clothes and some free time for parents. This fall, however, parents, teachers and students have an additional concern: school shootings. Although only 1 percent of all homicides—and suicides—of school-age children in the U.S. occur on school grounds, this statistic represents a dramatic increase. According to a survey by the National School Safety Center (NSSC), the number of violent deaths in schools rose 60 percent last year to a total of 41, nearly half of which were multiple shootings. Experts worry that an epidemic of school violence is under way. As Ronald D. Stephens, executive director of the NSSC, describes it, there have been attempted cases of "copycat killings," particularly after the shootings in March at a Jonesboro, Ark., middle school that killed four students and a teacher.

Anxious to stop this trend, teachers and administrators around the country have embraced a variety of preventive techniques—everything from metal detectors to daily classes in controlling anger. But in many instances, these programs have not been graded for efficacy. Even more troubling is the fact that, according to recent studies, certain popular methods simply do not work.

Preventing violence depends in large part on understanding what causes it. The school shootings are not isolated but are clearly part of a larger problem. During the past decade, homicides and suicides among young people have more than doubled; the rate of death as a result of firearms among American children 15 years and younger is 12 times higher than it is in 25 other developed countries combined. Although the causes for these developments are myriad, studies have documented that the standard complaints—ready access to guns as well as exposure to brutality, both at home and on-screen—do have an effect on kids. Initial results from the National Longitudinal Study on Adolescent Health (an ongoing survey of 12,000 adolescents) showed that children who are able to get ahold of guns at home were more likely to behave violently. The study also indicated that good parental and family relationships correlated somewhat with reductions in violent behavior.

For its part, the correlation with television mayhem is longstanding. As far back as the 1960s, psychologist Leonard Eron and his colleagues at the University of Illinois demonstrated that the more violence children watched on television, the more aggressive their behavior at school. The final report of the National Television Violence Study, conducted by the National Cable Television Association (NCTA) and released this past spring, "confirms that TV portrays violence in a way that increases the risk of learning aggressive attitudes," says John C. Nelson of the American Medical Association, one of the organizations that was part of the NCTA advisory council.

Although experts have been able to make headway in understanding some of the roots of violence, their efforts to forestall it have been less successful. Most violence-prevention programs are run locally, often through the school system. Because policy at each of the some 100,000 U.S. schools is typically set by local school boards, there is considerable diversity in approach. Yet "many of the programs being implemented [in schools] have not been rigorously evaluated" by researchers, according to Linda L. Dahlberg of the National Center for Injury Prevention and Control at the Centers for Disease Control and Prevention (CDC). In 1992, to remedy this problem, the CDC began a large-scale effort to review violence-prevention initiatives around the country.


DANIEL SHEEHAN *Gamma Liaison Network*

STUDENTS MOURN
victims of a shooting in Springfield, Ore., in May of this year.

Preliminary results from the CDC and other studies are beginning to come in, Dahlberg says, and they are "a mixed bag." For instance, intervention programs that start very early—some in kindergarten—can actually introduce children to ideas about violence that might not have occurred to them otherwise. Young children in such programs have described more violent, aggressive thoughts and fantasies than researchers anticipated. "We want to intervene early," Dahlberg notes. "But when? And what should we do?" She suggests that early-intervention programs should focus not just on the child but on the family and community.

At a conference earlier this year in Charleston, S.C., Delbert S. Elliott of the University of Colorado's Center for the Study and Prevention of Violence reported that "the evidence for programs that focus on family relationships and functioning, particularly on family management and parenting practices, is quite strong and consistent." His findings were based on a study of more than 450 prevention programs. Elliott also described conflict resolution training, peer counseling and peer mediation as ineffective when implemented alone: only when used as part of a more comprehensive prevention approach did they did show positive results.

More extensive programs, however, require more resources—money, people and time. Dahlberg points out that some of the less effective techniques were used in what she calls "schools in crisis," where teachers and administrators were preoccupied with other problems, such as overcrowding or deteriorating buildings, or were not supportive of the program.

Quick fixes such as metal detectors do not seem to do much good either. Researchers point out that such sensors are often expensive and will keep only some of the weapons out. At the same time, because most violence occurs outside of school, they do little to address the general problem of youth violence.

In the aftermath of recent school shootings, experts emphasized the importance of watching for warning signs of violence, but again, such monitoring is not foolproof. In June the NSSC released a list of 20 potential indicators for violent behavior, including having a history of bringing weapons to class or having been bullied in school. Even so, NSSC executive director Stephens says, "for all the high-tech strategies we have, there is not a scanner around that can predict how and when a child might explode" in anger and violence.

Some researchers are even concerned that this analytical approach could wind up harming kids. Edward Taylor of the University of Illinois, who is developing a study for identifying predictors of violence in children, offers words of caution: "We certainly don't want a school system that every time a child throws a temper tantrum, every time a child says something aggressively, that they are immediately suspect of becoming mentally ill and violent."

Notwithstanding the debates about prevention and the various attempts to reduce youth violence, many experts worry that the broader context is being forgotten: until programs consider youth violence against a societal backdrop of violence, they may have only limited success at best. Mike Males of the University of California at Irvine, whose book *Framing Youth: Ten Myths about the New Generation* will be published in October, argues that "the youth culture of violence is the adult culture of violence." Nearly 10 times as many children die at the hands of their parents as die at school. The tradition of learning by example has rarely had such tragic consequences. —*Sasha Nemecek*

PHARMACOLOGY

HERB REMEDY

Exploring ways to administer marijuana as a medicine

The exact location is a secret. But somewhere between London and Brighton a compound ringed by high fences and razor wire will house the world's only pot farm primarily devoted to commercial drug development. In June the British Home Office gave a startup pharmaceutical company a license to grow 20,000 marijuana plants of varied strains.

Geoffrey W. Guy, chairman of GW Pharmaceuticals, intends to proceed to clinical trials with a smokeless, whole-plant extract, while also supplying marijuana to other investigators interested in medical research and pharmaceutical development. The 43-year-old entrepreneur-physician wants to capitalize on what he sees as the unexploited opportunity to legitimize marijuana as medicine. "Cannabis has been much maligned," Guy says. "There are over 10,000 research articles written on the plant, and there's something well worth investigating here."

The idea of giving this alternative medicine a place alongside antibiotics and aspirin in the physician's standard pharmacopoeia is by no means a new one. Marijuana and its chemical constituents have aroused interest as a treatment for conditions ranging from the nausea induced by cancer drugs to the fragility of brain cells harmed by stroke. In the U.S., oral doses of delta-9-tetrahydrocannabinol (THC)—a synthetic version of the chemical in marijuana that both relieves nausea and gets a person high—have been available on the market since 1986.

But the makers of Marinol (the trade name for the THC synthetic) have had trouble competing with dealers on the street. A swallowed pill takes too long to relieve nausea. "The maximum levels of THC and the active metabolites you see after you swallow a capsule occur at anywhere from two to four hours," says Robert E. Dudley, senior vice president of Unimed Pharmaceuticals in Buffalo Grove, Ill., Marinol's manufacturer. "That's contrasted with a marijuana cigarette, where the peak levels might occur from five to 10 minutes."

Unimed and other companies are in various stages of developing nasal sprays, sublingual lozenges, vaporizers, rectal suppositories or skin patches that will deliver THC into the bloodstream quickly. But new interest in marijuana as pharmaceutical goes beyond just substitutes for smoking. Guy's motivation for establishing GW borrows a page from the herbal medicine literature. He hypothesizes that the plant's 400 chemicals, including dozens of cannabinoids such as THC, may interact with one another to produce therapeutic effects. A few studies have shown that one cannabinoid, called cannabidiol, may dampen some of THC's mind-altering effects. And synthetic THC users sometimes report feeling more anxious than smokers of the drug, perhaps because of the absence of cannabinoids other than THC.

GW Pharmaceuticals wants to test whole-plant extracts for a series of medical conditions. A Dutch company, HortaPharm, will provide seeds to GW for plants that contain mainly one cannabinoid. Different single cannabinoid plant extracts can be blended to provide the desired chemical composition.

ROBERT C. CLARKE IHA

GOOD BREEDING allows the Dutch firm HortaPharm to grow medical marijuana that contains predominantly one cannabinoid.

Interest in whole-plant medicinal marijuana has even stirred in the U.S., where research on the drug has been stymied for 20 years. That bias may be shifting, as witnessed by a 1997 National Institutes of Health advisory panel that recommended more research on the subject. Robert W. Gorter, a professor at the University of California at San Francisco, has received approval from the Food and Drug Administration to perform a clinical trial on an orally administered whole-plant extract—and he is also organizing a separate investigation with patients in Germany and the Netherlands. "Various cannabinoids in the plant appear to work in a little symphony," Gorter observes.

Pushing whole marijuana as medicine is not a task for the fainthearted. Financing pharmaceutical development for a controlled substance may not come easy. "I need the right type of people as backers," Guy says. "I don't want people from Colombia turning up with suitcases full of dollar bills."

In addition, some scientists observe that evidence for cannabinoid synergies is relatively slim. "There has never been an effect of marijuana that has not been reproduced with pure delta-9-THC," says John P. Morgan, a professor of pharmacology at the City University of New York. "Herbal medicine advocates think that plants are better because there's a mix of natural substances. There's not much basis for most of these claims."

Ultimately, advocates of marijuana as natural medicine may find their work superseded by developments stemming from discoveries of cannabinoid receptors in the human body—and of molecules that bind to them. Some research groups are seeking analogues to the binding molecules naturally present in the body that might provide therapeutic benefits superior to those of plant-based cannabinoids.

Receptor research is also shedding light on the role played by the cannabinoids found in marijuana. NIH investigators reported in the *Proceedings of the National Academy of Sciences* in early July that THC and cannabidiol serve as powerful antioxidants. In laboratory rat nerve cells, the compounds can prevent the toxic effects of excess glutamate, which can kill brain cells after stroke. (After reading this report, legalization advocates reveled at the notion that marijuana may actually protect brain cells.)

To proponents of legalization of the smokable herb, arguments about alternatives remain academic. "Because patients are receiving full relief right now from smoking the whole plant, we shouldn't let them suffer while science plods along trying to come up with synthetic analogues that may not have the same beneficial effect," says Allen F. St. Pierre, executive director of the National Organization for the Reform of Marijuana Laws Foundation.

Some medical users would rather fight than switch from joints or brownies. Elvy Musikka, a glaucoma patient in Hollywood, Fla., is one of eight people enrolled in a federal program that supplies the drug for medical reasons. She maintains that if her legal supply is cut off she will move to a country where she can grow her own. "I think for the pharmaceutical companies to think they produce a better product than God is totally presumptuous," she says. Pharmaceutical makers may find that Musikka's attitude—shared by thousands—becomes the biggest impediment to successful drug development. —*Gary Stix*

STALKING THE WILD DUGONG

An undersea elephant remains elusive

From the porch where I am slumped, exhausted by the heat, I stare in astonishment at a man walking up the forest trail from the beach, snorkel dangling from one hand. I have just arrived at Dugong Creek, a remote corner of Little Andaman Island in the Bay of Bengal, to meet the Onges, a group of hunter-gatherers believed to be descended from Asia's first humans. I hadn't expected to find other visitors.

"You know there are crocodiles," I say, indicating his snorkel.

"A hazard of the trade," he grins.

Himansu S. Das of the Salim Ali Center for Ornithology and Natural History in Coimbatore, India, is a sea-grass ecologist. Because dugongs, Old World relatives of the manatee, feed on underwater greenery, he had guessed that Dugong Creek would have beds of sea grass nearby. The animals themselves, though, were likely to be long gone. Once seen in the hundreds or even thousands along the tropical coasts of Africa and Asia, these sea elephants are all but extinct in most of their range and occur in reasonable numbers only in Australia. In five years of exploration, Das has gathered evidence of at most 40 dugongs throughout the Andaman and Nicobar archipelago. To his surprise, he has just learned from the Onges that a family of four still lives in Dugong Creek, down one since their hunt of two weeks ago.

The grass beds nourish not only these rare mammals but also marine turtles and a variety of fish and shellfish. With the help of a grant from UNESCO, Das is estimating the impact of humans on the ecology. In fact, it is the local peoples who point him to the beds, more predictably than do the satellite images on which he initially relied.

The next afternoon, under a blistering sun, we set out for an Onge camp a kilometer or so along the shore. At one point we have to ford a creek. Halfway across, in chest-deep water and with my sandals held aloft in one hand, it strikes me.

KELVIN AITKEN *Peter Arnold, Inc.*

DUGONG, A VEGETARIAN MAMMAL,
needs fields of marine greens where it can graze in peace.

"Crocodiles?" I ask.

"Just keep walking," he replies.

I do. Saltwater crocodiles are the most ferocious of them all, and I've seen their tracks on the beach.

The Onges, we discover, have not seen a dugong but have harpooned two turtles. One is being cooked, and the other is secured at the end of a long rope stretched into the sea. When Koira, an Onge man, pulls on the leash, a head sticks anxiously out of the water as the animal looks to see where it is being drawn. It is an endangered green sea turtle, small, about 15 kilograms.

Neither of us begrudges the Onges their meal. They have lived on Little Andaman for millennia with no harm to its biodiversity and now, because of pressure from recent settlers, will probably vanish long before the turtles. The main threat to the sea-grass beds and to the creatures that depend on them is the silt that muddies the water as the dense tropical forest is cut down: the marine plants die of darkness. Overexploitation of fish, shellfish and other marine species by immigrants from mainland India and by fishers from as far away as Thailand is another pressing problem.

As the grass patches shrink, the dugongs become confined to ever smaller regions that are also the local fishing grounds. Some fishers set their nets around the beds to catch predators, such as sharks, that come to feed on smaller fish, but the nets entangle turtles as well as an occasional dugong. Das will be recommending to the Indian authorities that some sea-grass beds be protected as sanctuaries. But as we return—the tide has gone out, mercifully leaving the creek just knee-deep—I realize with sadness that it's already too late for the Andaman dugong. *—Madhusree Mukerjee in the Andaman Islands*

IN BRIEF

Alexander's Fate

An ancient conspiracy theory held that rivals poisoned Alexander the Great, who died unexpectedly at the age of 32 in 323 B.C. But a new analysis, published in the *New England Journal of Medicine* on June 11, finds otherwise: Alexander probably fell victim to typhoid fever. The authors—including infectious-disease expert David W. Oldach of the University of Maryland and historian Eugene N. Borza of Pennsylvania State University—were puzzled by historical accounts stating that Alexander's body did not begin to decay for days after his death. They believe he most likely succumbed to ascending paralysis, a complication of typhoid fever that can slow down a person's breathing and make them look dead.

NIMATALLAH/ART RESOURCE

Science Knowledge

Interest in science is at an all-time high, according to a survey of 2,000 U.S. adults that was presented to Congress in July. But basic knowledge remains poor. Jon D. Miller, director of the International Center for the Advancement of Scientific Literacy, conducted the survey for the National Science Foundation last year. Although 70 percent of the subjects said they were curious about science and technology, only 11 percent could define "molecule," half believed that humans and dinosaurs had at one time coexisted, and only 48 percent knew that the earth orbits the sun once every *year.*

Hello, SOHO?

Engineers have tried to reach the Solar and Heliospheric Observatory (SOHO) since losing contact with the craft on June 24 during routine maintenance operations. For the past two and a half years, SOHO has provided researchers from NASA and the European Space Agency with a wealth of information about the sun. Now the observatory is apparently spinning in such a way that its solar panels do not receive enough light. In case SOHO comes out of the dark, the team is issuing frequent signals to activate its transmitters.

More "In Brief" on page 24

COSMOLOGY

THE FLIP SIDE OF THE UNIVERSE

New cosmological observations confirm inflation

Late into the night astronomers Angelica de Oliveira-Costa and Max Tegmark worked to analyze their observations of the cosmic microwave background radiation. The next morning the young wife-and-husband team were due to present what their data revealed about the single most important unknown fact in cosmology: the shape of the universe. Their previous results, from a telescope in Saskatoon, Canada, between 1993 and 1995, had suggested that the universe is flat—the first observations to substantiate a long-held belief among cosmologists. But intrinsic uncertainties in the measurements made it impossible to be sure.

So in 1996 the QMAP team (de Oliveira-Costa, Tegmark and five colleagues from the Institute for Advanced Study in Princeton, N.J., and the University of Pennsylvania) flew instruments on a balloon 100,000 feet (30 kilometers) above Texas and New Mexico. When they finally processed the data—the night before their announcement at the Fermi National Accelerator Laboratory this past May—the situation looked grim. The Saskatoon and the balloon results were completely different.

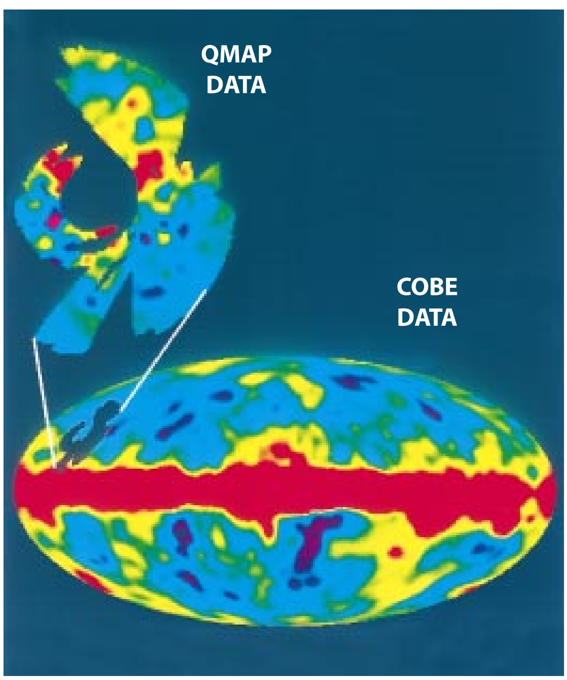
Suddenly, however, de Oliveira-Costa realized that Tegmark had accidentally plotted the map upside down. When righted, it matched the Saskatoon data exactly. "That was my most exciting moment as a scientist, when I realized we'd flipped that map," Tegmark says. "It was then I realized, yes, Saskatoon was right. The universe is flat."

The QMAP balloon discerned much finer details in the radiation than the Cosmic Microwave Background Explorer (COBE) satellite did eight years ago. In some areas this radiation is slightly dimmer (*blue, in illustration below*); in others, brighter (*red*). The red stripe down the middle represents the Milky Way galaxy, whose own microwave emission overpowers the cosmic signal; to avoid it, QMAP focused on a clear patch of sky around the North Star.

When the brightness fluctuations are exaggerated 100,000 times, blobs become clear. They correspond to clumps of matter that existed 300,000 years or so after the big bang. Their apparent size depends on the geometry of the universe and, in turn, on the cosmic density of matter and energy.

Combined with other observations, including those of distant supernova, the QMAP results corroborate the prevailing theory of inflation—with the twist that the universe is only one third matter (both ordinary and dark) and two thirds "quintessence," a bizarre form of energy, possibly inherent in empty space. Despite Tegmark's enthusiasm, however, this conclusion is not definitive. Astronomers are still waiting for results from two upcoming satellites, the Microwave Anisotropy Probe and Planck; meanwhile other groups are flying balloons or taking ground-based measurements. They all hope to hold up or shoot down inflationary theory. "It's like an Indiana Jones movie," says Paul Steinhardt of Penn. "Everyone sees that holy grail."

—*George Musser*



MAX TEGMARK

***COSMIC BACKGROUND RADIATION** was remeasured with high resolution, revealing finer details.*

In Brief, continued from page 22

Perfecting Microwaved Foods

For the sake of convenience, most people put up with microwaves—even though the french fries get soggy, the potatoes stay frozen and the eggs (don't try it yourself) explode. But that could soon change, thanks to new mathematical and computer models developed by Ashim K. Datta and graduate student Hua Zhang of Cornell University. The models predict how edibles of various shapes and consistencies will heat, and Datta hopes they will help manufacturers devise more successful products. Ninety percent of new microwavable foods introduced to the marketplace every year fail.

R. PLEASANT *FPG*

Treating Tuberculosis

Deadly multidrug-resistant strains of TB have emerged in large part because the standard therapy is so intensive. Patients must follow two months of daily doses, followed by four months of twice-weekly treatments. Many simply do not finish. But now the Food and Drug Administration has approved a new medication, rifapentine, that should make adherence easier. TB sufferers need only take rifapentine once a week during the last four months of recovery. The drug, which will be marketed under the name Priftin, is the first anti-TB agent approved in 25 years.

Triggering Tourette's Syndrome

The ailment—which produces involuntary movements or vocalizations called tics—tends to run in families, but it appears in only a small percentage of those children who inherit one copy of the responsible gene. Now scientists think they have discovered what can push these kids over the edge into illness: a streptococcus infection. Harvey Singer and his colleagues at Johns Hopkins University looked for antineuronal antibodies—which the body can make in response to bacterial infections and which attack brain tissue—in 41 Tourette's patients and 39 controls. They found that those in the former group had higher levels of the antibodies in a region of the brain that helps to control movement. The finding could someday lead to new means of prevention and treatment.

More "In Brief" on page 26

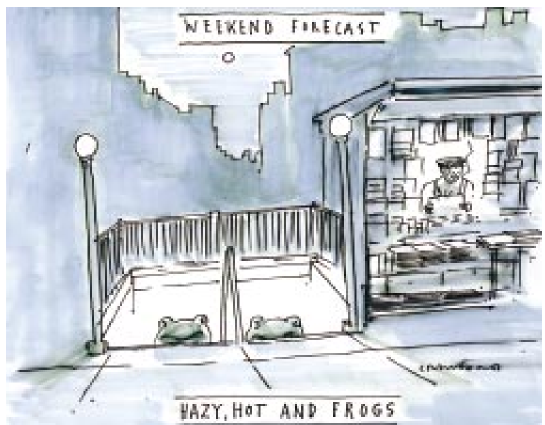
ANTI GRAVITY

Tomorrow, Partly Froggy

Television evangelist and sometime presidential candidate Pat Robertson recently shocked the world by revealing that a science existed that he knew even less about than paleontology. Delving into the latter discipline, Robertson once contended that "there is no case where we have remains or fossils of an animal that died during the evolutionary process." In fact, every fossil ever found is of an organism that died during the "evolutionary process." But I digress. In June, Robertson added meteorology to the list of sciences about which he has theories—and few facts.

The self-appointed forecaster took to the airwaves to warn residents of Orlando, Fla., about wicked weather possibly headed in their direction. Orlando stood in the way of some righteous wrath, he maintained, as a result of the city's decision to allow gay organizations to fly rainbow flags in a local celebration called Gay Days. "This is not a message of hate; this is a message of redemption," he insisted. "But if a condition like this will bring about the destruction of your nation, if it will bring about terrorist bombs, if it will bring about earthquakes, tornadoes and possibly a meteor, it isn't necessarily something we ought to open our arms to. And I would warn Orlando that you're right in the way of some serious hurricanes, and I don't think I'd be waving those flags in God's face if I were you."

Robertson the scientist might be expected to be an expert on hot air masses. But he's not going out on a limb long enough to snap off in a 100-mile-per-hour wind by prophesying hurricanes in Florida. This just in: Buffalo gets snow in January! (And Florida weather is even easier than most. I lived in Miami one summer, and I can give you a fairly decent forecast for any day in July or August without looking at a single satellite map: temperature in the 90s with high humidity, good chance of an afternoon thunderstorm. Rinse. Repeat.)



The issue of whether God has a face in which to wave a flag is an open one for some people, but assuming He or She does, ears are probably attached, and assuming Robertson has God's ear, maybe he could ask Him or Her for some more detailed meteorological information. Which he could then include as a regular feature on his 700 Club broadcasts, launching into something like this:

"Odious lectures by Stephen Jay Gould on evolution at the American Museum of Natural History will lead to cherub-size hail in the early evening in New York City. Continued homosexual activity in the large cities of the East Coast and, of course, San Francisco, will initiate a severe low-pressure front with associated torrentially heavy rains late in the day. Look for heavy flooding, especially in theater districts. Because of some isolated pockets of freethinking, the Midwest will see a 90 percent chance of frogs this afternoon with vermin, especially in low-lying regions, so please drive with the low beams on. Frogs diminishing toward dusk, followed by scattered murrain. Putting the map in motion now, we see that general sinfulness across the country will bring darkness after sundown."

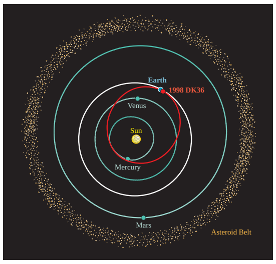
In the interest of an even-handed attack on wacky ideas, I'd like to point out the obvious fact that no religious group has a monopoly on them. Some of the more vocal members of some groups, however, do have their own broadcast outlets, so it's easier to notice when they say something outrageous. The good news is that such spokespeople serve the larger purpose of reminding everyone of the simple and profound words written by Gould: "The enemy of knowledge and science is irrationalism, not religion." Amen. *—Steve Mirsky*

MICHAEL CRAWFORD

In Brief, *continued from page 24*

Asteroids on the Inside

Scientists have long looked far afield for asteroids. Now they have found one circling the sun inside the earth's orbit.



DAVID J. THOLEN

David J. Tholen and graduate student Robert Whiteley of the University of Hawaii spotted the object, named 1998 DK36, using a specialized camera on the 2.24-meter telescope atop Mauna Kea in February. Preliminary calculations show that nearly 1.3 million kilometers always separate the earth from the asteroid as it passes through the daytime sky—good news, given that DK36 appears to be 40 meters in diameter. The asteroid that devastated the Tunguska region of Siberia in 1908 was about the same size.

Phone Home

Still no sign of extraterrestrial life, according to SERENDIP III, the most sensitive sky survey to date. The search, led by Stuart C. Bowyer of the University of California at Berkeley, used a detector mounted on the world's largest radio telescope in Arecibo, Puerto Rico. Starting in 1992, the instrument analyzed 500 trillion signals, looking in a radio band centered on a wavelength of 70 centimeters—a region typically reserved for communications. No luck. But the team hasn't given up hope. SERENDIP IV, now in the works, should be 40 times more sensitive than its predecessor; it will simultaneously examine 168 million frequency channels every 1.7 seconds.

Just Add Water

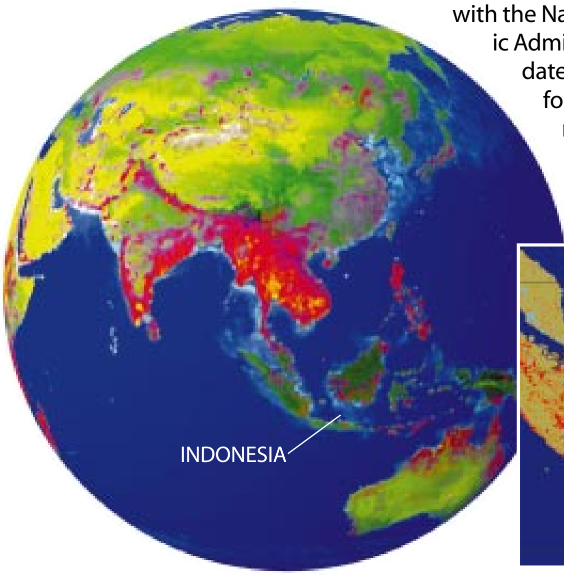
The wonders of modern science never cease. Ryuzo Yanagimachi and colleagues at the University of Hawaii have produced live mice from dead sperm. The workers added water back to freeze-dried sperm and injected it into mouse eggs using a procedure called intracytoplasmic sperm injection (ICSI). They found that freeze-drying had preserved the genetic information in the sperm well enough to regenerate healthy mice. The tactic is expected to be an improvement on previous methods of storing genetic information from mice used in research. *—Kristin Leutwyler*

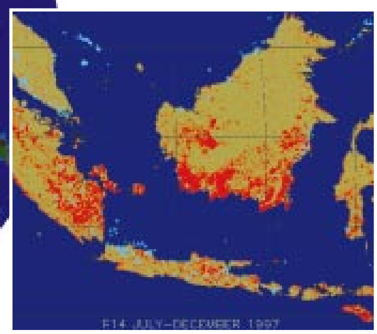
SA

ECOLOGY

A World Aflame

Every year fire scorches some 71 million hectares (175 million acres) of forest and grassland. In 1997 drought brought on by El Niño exacerbated fires, many of which were deliberately set, the world over. In Indonesia, for instance, the devastation was particularly extreme because of the worst drought the country had seen in 50 years. According to the World Wildlife Fund's 1997 report *The Year the World Caught Fire,* Indonesia lost two million hectares to flame. A satellite image from last year shows the extent of the damage (*red in image at right*). A composite of data from 1992 to 1995 (*at left*) shows fires in the region in red and purple. Fires this year in the Amazon, Mexico, Florida and elsewhere promise to make 1998 another record year. The National Aeronautics and Space Administration has teamed up with the National Oceanic and Atmospheric Administration to provide weekly updates on fires around the world. Information can be found at http://modarch.gsfc.nasa.gov/fire_atlas/fires.html on the World Wide Web. *—Sasha Nemecek*





SIMMON/SUTTON/STOCKLI NASA *Goddard EOS AM-1 Visualization Team (globe)*; DEFENSE METEOROLOGICAL SATELLITE PROGRAM *(map)*

BIOLOGY

RIVER OF VITRIOL

The Rio Tinto in Spain abounds in acid—and unexpected organisms

Against the dark stand of pine trees, the waters of the Rio Tinto appear even more vividly red than usual. Here, near its headwaters in southwest Spain, the strong smell of sulfur overwhelms even the fragrance of the dense forest. The crimson river—infamous for its pH of two, about that of sulfuric acid, and for its high concentration of heavy metals—seems dead, a polluted wasteland and a reminder of the ecological devastation mining can entail.

Yet the remarkable Rio Tinto is hardly lifeless, as scientists have discovered in the past several years. Even in parts of the river where the pH falls below two—and the water is painful to touch—green patches of algae and masses of filamentous fungi abound. "Each time we go there we find something new," says Ricardo Amils, director of the laboratory of applied microbiology at the Center for Molecular Biology at the Autonomous University in Madrid, who discovered the river's wild ecosystem in 1990. "We have now collected about 1,300 forms of life living here, including bacteria, yeast, fungi, algae and protists. But the real number is surely much higher."

Before Amils and his colleagues studied the 93-kilometer (58-mile) river, it was assumed that the acidic waters were purely the result of the Rio Tinto copper mine, one of the world's largest and oldest. The microbiologist now believes that industry—in particular, the sulfuric acid associated with copper mining and the discarded metal tailings—is not entirely responsible for the condition of the water. He has found that historical records refer to the river's long-standing acidity. Amils postulates that the river's strange chemistry led its first miners—the Tartessians in 3000 B.C.—to investigate the banks for deposits. Soon after, the Romans, who extracted great quantities of gold and silver, called the river *Urbero,* Phoenician for "river of fire." And the Arab name for it was "river of sulfuric acid."

Amils's argument is bolstered by other observations. The low pH and high concentration of metals—including iron, arsenic, copper, cadmium and nickel—is consistent throughout the entire river, becoming less acidic where the Rio Tinto meets the Atlantic Ocean. Typically, waterways that receive mining waste have acidic concentrations only near the source of pollution. Strong rains also cannot seem to reduce the acidity of the river.

To explain how this amazing condition came about—and is perpetuated—the researchers point to what they have learned about the species found there. They are convinced that the extreme conditions are produced by bacteria. *Thiobacillus ferrooxidans,* for example, is abundant in the river. This microorganism is capable of oxidizing sulfur and iron—thereby giving the Rio Tinto its red hue and name. Amils and his colleagues recently documented the presence of another bacterium, *Leptospirillum ferrooxidans,* which feeds exclusively on iron and is even more abundant than *T. ferrooxidans*. Evidence of its corrosive capability sits on the banks, where abandoned railroad cars were flooded with river water. In service just 15 years ago, these cars now appear as skeletons, devoured by the Rio Tinto microbes. "These bacteria are a kind of metallic piranha," Amils says.

Other bacteria—not as well understood for the time being—appear to feed on the immense deposits of metal sulfides, creating the sulfuric acid that, together with oxidized iron, produces the very conditions that lead to heavy metals in solution.

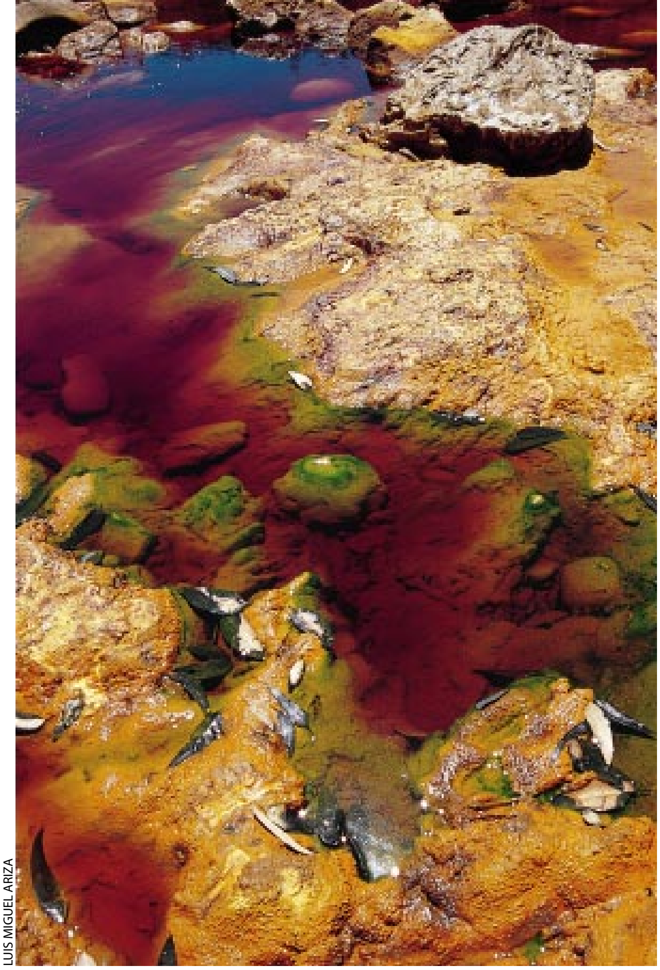
The most abundant organisms in the river appear to be algae, which produce oxygen in champagnelike bubbles that the researchers watch float to the surface. How algae work in this acid inferno, however, has the scientists mystified. "We need more research to explain how algae can collect light and produce organic matter and oxygen in such conditions," Amils notes.

One possibility is that some of the Rio Tinto algae and fungi have established certain symbiotic associations. "Through evolution we can see that symbionts can thrive successfully in a habitat that otherwise would be inhospitable," explains Lynn Margulis, a biologist at the University of Massachusetts at Amherst. "Long-term association can create new species through symbiogenesis."

Understanding this symbiosis—if it is present in the Rio Tinto—could help Margulis, Amils and others understand the development of early life on the earth. "In my view, the river is a better model for the life that flourished in the Proterozoic, with an abundance of oxygen and algae, than it is for the anoxic Archean eon," Margulis says. During the Proterozoic—between 2.5 billion and 600 million years ago—anaerobic and aerobic organisms survived in extreme conditions, perhaps assisting one another.

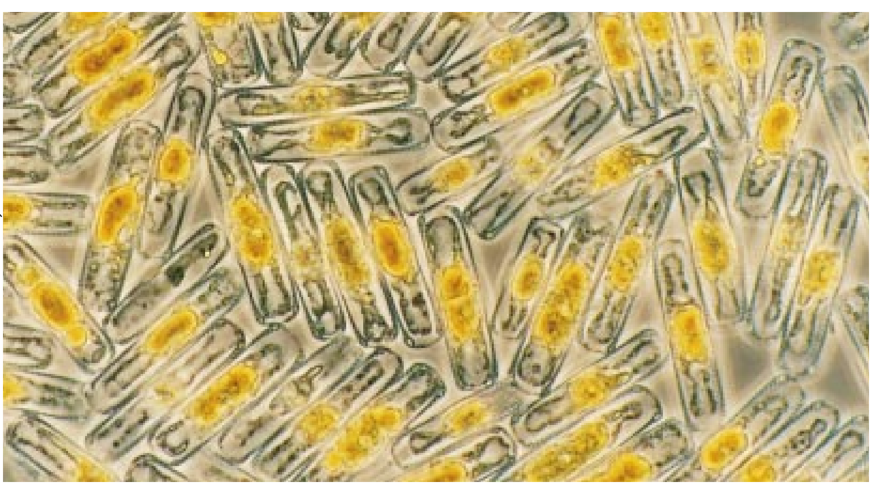
The Rio Tinto could also offer further insights. Perhaps Amils and his crew are seeing the kind of life that thrived on Mars millions of years ago. The versatility of these bacteria—particularly those that work in anoxic conditions on mineral substrates such as iron sulfides—make them good candidates for a model of extraterrestrial life. "I cannot say that Martians were like this, but Mars would be a perfect bite for many bacteria living here, that is for sure," Amils says.

—Luis Miguel Ariza in Madrid



LUIS MIGUEL ARIZA

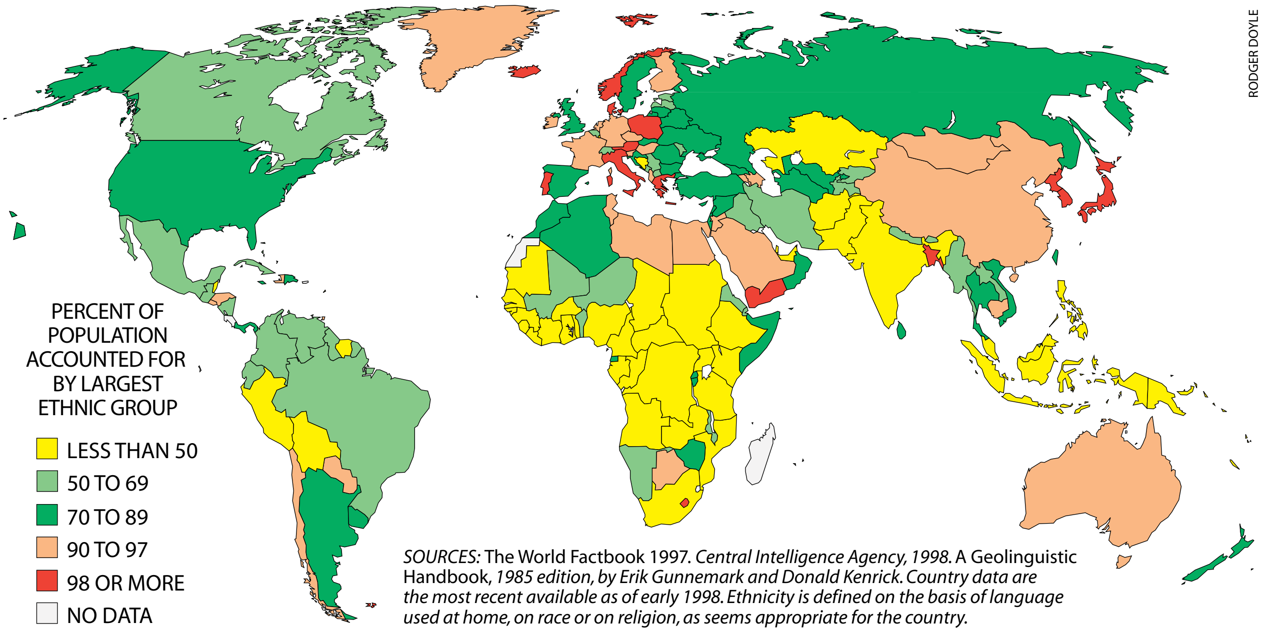
RED WATERS
and high acidity of the Rio Tinto do not deter a wealth of microscopic life, including blooms of algae.



ANABEL LÓPEZ-ARCHILLA *Autonomous University of Madrid*

DIATOMS
from the Rio Tinto are among the 1,300 species found to live there.

Ethnic Groups in the World



SOURCES: The World Factbook 1997. *Central Intelligence Agency, 1998.* A Geolinguistic Handbook, *1985 edition, by Erik Gunnemark and Donald Kenrick. Country data are the most recent available as of early 1998. Ethnicity is defined on the basis of language used at home, on race or on religion, as seems appropriate for the country.*

Many of the world's problems stem from the fact that it has 5,000 ethnic groups but only 190 countries. This situation is illustrated on the map, which shows that few states are ethnically homogeneous and that many, particularly in Africa, have no majority ethnic group. Since 1945 some 15 million people have been killed in conflicts involving ethnic violence, although ethnic tensions have not necessarily been the catalyst. Among the worst incidents were the 1994 civil war in Rwanda, which resulted in more than a million dead and three million refugees, and the 1947 communal riots in India, which left several hundred thousand dead and 12 million refugees.

Why are some multiethnic countries plagued by violent, persistent ethnic conflict and others not? There are no completely satisfactory answers, but it is evident that several factors affect the outcome. One of these is the presence or absence of political institutions that give minorities protection against the tyranny of majority rule. Federal systems, such as the one instituted after 1947 in India, can help dampen ethnic tensions by giving minorities regional autonomy. Intermarriage—between Thais and Chinese in Thailand, say, or Taiwanese and Mainlanders in Taiwan—erodes ethnic differences. And free-market forces tend to mitigate ethnic tensions. For instance, Russia has not adopted an irredentist policy—there are nearly 25 million Russians in neighboring republics—arguably because it would interfere with the goal of achieving a Western-style market economy.

The region with perhaps the most intransigent ethnic rivalries is sub-Saharan Africa, which has about 1,300 language groups in 42 countries, the boundaries of which were imposed by the colonial powers with little regard for ethnicity. In addition to language differences, religious divisions exist—most prominently between Muslims and Christians. These widespread ethnic and religious divides have contributed heavily to instability in countries such as Nigeria, where Hausa, Fulani, Yoruba and Ibo tribes contend for political power. Nigeria has suffered six military coups and two civil wars since gaining independence from Britain in 1960. Following decolonization, about three fourths of the sub-Saharan African countries have undergone coups or civil wars.

India has had a generally successful record in dealing with ethnic tensions since independence despite its 300 languages, thousands of castes and major religious fault lines. One explanation may be its extreme diversity: a country with so many divisions may be at less risk of violence than one in which just a few groups contend, because no single group can dominate. At its inception, India was blessed with a large, well-educated, democratically inclined elite that used its prestige to build a multiethnic political machine—the Congress Party—that was a potent force in mitigating tensions. But now there is uneasiness about India's future because of the rise to power of the nationalist Bharatiya Janata Party, whose more extreme supporters shout slogans such as "For Muslims there are only two places, Pakistan or the grave."

There is a widely held belief that ethnic violence in the former Yugoslavia arose from ancient ethnic hatreds. But this view ignores a history of peaceful coexistence and extensive intermarriage among ethnic groups going back generations. It is unlikely that the recent conflict in the region would have progressed to genocide had it not been for political leaders such as Slobodan Milosevic, who distorted history to create a myth of a Serbia wronged by ancient enemies and now again threatened by these same enemies. The notion of "ancient ethnic hatreds," at least in the Balkans, India and Africa, seems to have limited basis in fact.

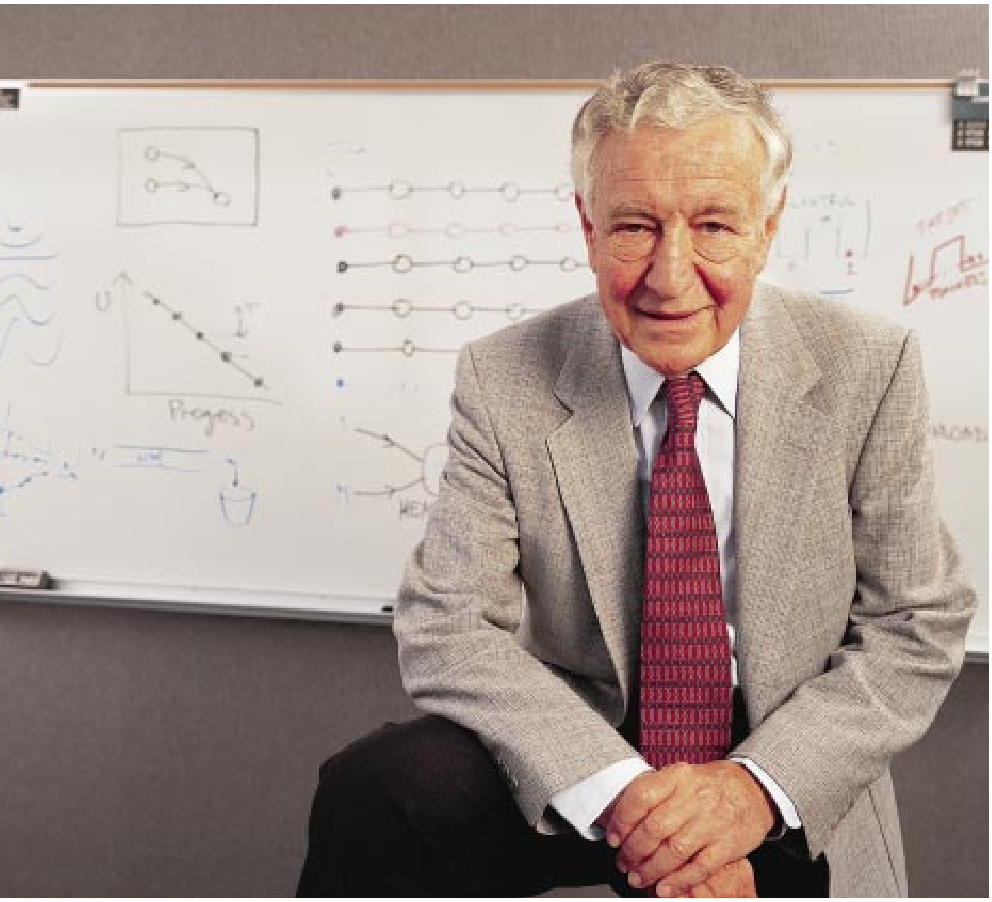
—Rodger Doyle (rdoyle2@aol.com)

PROFILE

Riding the Back of Electrons

Theoretician Rolf Landauer remains a defining figure in the physics of information

In 1952 a young physicist visited an aging industrial building in Poughkeepsie, N.Y., that its owners called the "pickle works." The trim 25-year-old was looking for a new position after becoming disenchanted with his two-year-old job at the National Advisory Committee for Aeronautics (predecessor of the National Aeronautics and Space Administration), where he had some involvement with a project to build a nuclear-powered jet aircraft.



NAJLAH FEANNY SABA

***LIMITS OF COMPUTATION** and the kinetics of small structures are ideas that physicist Rolf Landauer prefers to communicate with pictures, not numbers.*

IBM, the company he was visiting, had just made the onerous transition from computing machines that used electromechanical relays to those that incorporated vacuum tubes. Even then, it was looking warily ahead. The company needed physicists as part of a small semiconductor research team, established to guard against the unlikely possibility that transistor technology would ever amount to anything. "The future of IBM is in semiconductors, and they don't even know it," confided the manager who conducted the interview.

Rolf Landauer, the erstwhile job applicant, recalls those words more than 45 years later in his tidy office in the sweeping Eero Saarinen–designed glass edifice that houses IBM's primary research facility in northern Westchester County outside of New York City. "I was very lucky to have my career coincide with the period of high adventure that followed," he observes.

Moving from nuclear aircraft to microcircuitry closely follows the technological trajectory of the latter half of the 20th century, a period in which fascination with moon shots and nuclear power has given way to a preoccupation with the movement of electrons in small spaces. Landauer has helped define at the most fundamental level how much useful work a computing machine can perform in these lilliputian confines. "He really started the field of the physics of computation," says Seth Lloyd, a leading theorist on using quantum-mechanical principles for computing and a professor of mechanical engineering at the Massachusetts Institute of Technology.

The basic tenet of Landauer's worldview is that information is not some mathematical abstraction. Instead it is a physical entity—whether it be stored on an abacus, on a punched card or in a neuron. Early in his research Landauer began to wonder how much energy is required for each computational step. Identifying a minimum would establish a basic limit akin to the laws of thermodynamics, which calculated the efficiency of 19th-century steam engines, or to Claude E. Shannon's information theory, which figures how many bits can be shipped over a wire.

The seminal insight of Landauer's career came in 1961, when he challenged the prevailing idea—put forth by mathematician John von Neumann and others—that each step in a computer's binary computation required a minimum expenditure of energy, roughly that of the thermal motion of an air molecule. Landauer's paper in the *IBM Journal of Research and Development* argued that it was not computation itself but the erasing of information that releases a small amount of heat. Known as Landauer's principle, the idea that throwing away bits, not processing them, requires an expenditure of energy was criticized or ignored for years. More recently, this cornerstone of the physics of information has become the underpinning for advanced experimental computers.

One person who did take notice of this early work was an unorthodox postdoctoral student at Argonne National Laboratory. At a conference in Chicago in 1971, Charles H. Bennett explained to Landauer how a computer might be designed that would circumvent Landauer's principle by not discarding information and therefore dissipating virtually no energy. Bennett expanded on Landauer's work by showing that each step in the computation can be carried out in a way that allows the input to be deduced from the output—in essence, the machine can be run backward. Such a machine can first save the answer and then put itself into reverse until each step is undone. It avoids the energy losses stipulated by Landauer's principle; no information is erased, and accordingly no energy is lost. Other researchers have borrowed these ideas for reversible logic circuits that may help avoid the potentially fatal amounts of heat generated by very small circuits in future computers.

Landauer's relationship with Bennett bears a resemblance to the circularity of reversible computation. At first, Landauer served as a mentor, convincing the younger man to take a job at IBM. As Bennett refined and extended Landauer's original work, Bennett became a mentor to Landauer. The evidence of this inversion can be seen in a 1996 paper published by Landauer in the journal *Science*. The reversibility of an information-processing operation—which Bennett suggested in response to Landauer's work—became a central idea in the paper, which dealt with the limits of communication. The paper showed that no minimal energy expenditure is required to ship a bit of information. In contrast, ordinary communications links, where the signal energy is thrown away at the receiver, undergo a distinctly irreversible operation. Landauer evokes the image of a ski lift in which each chair has two seats, each of which represents a 0 or a 1. Skiers—the bits of a message—sit in one of the two chairs for the trip up the mountain. There the bits are all switched to the 0 chair so that they can be hauled down the mountain for reuse in another message.

Landauer has always understood the gap between thought and practice. As a theorist, he has explored the bounds of computation. Yet he has also served as a leading critic of the practicality of technologies—such as quantum computation—that have pushed limits. He has even voiced doubts about the fate of reversible computers, a field related to his own work. "All technology proposals come with a penalty," he notes. "A reversible machine would require more complex and slower circuitry."

M.I.T.'s Lloyd recounts that it is not unusual for him to receive a letter from Landauer after publication of one of his papers on quantum computation. (A quantum computer is a type of reversible machine that performs many calculations simultaneously based on quantum-mechanical principles that allow a single bit to coexist in many states at once.) Landauer invariably suggests that a disclaimer should be affixed to the publication: "Warning: quantum computers are unlikely to work in the real world."

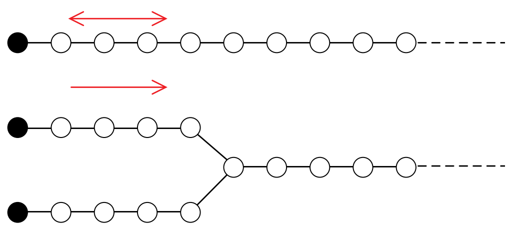
The contrast between his adventurousness as a theorist and his conservative pragmatism stems from his tenure during the 1960s as director of IBM's physical sciences department, where he had to make judgments about funding one technology or another. He managed the beginnings of the company's programs for developing the semiconductor laser and integrated circuits, even coining the term "large-scale integration."

His realist's bent may also be rooted in his experience as a childhood refugee from Hitler's Germany. Born in Stuttgart in 1927 to a well-to-do Jewish family, Landauer recalls the patriotism of his architect father, who died in 1935, his life shortened by a wound sustained while fighting during World War I in the German army: "Like a lot of Jews who were good Germans, he always thought that this craziness has got to stop. I'm alive today because he died in 1935. We would never have left Germany if he had continued to be head of the family."

The family settled in New York City, where Landauer made his way through the public school system, eventually graduating from the renowned Stuyvesant High School. After his acceptance to Harvard University, his uncle urged him to major in electrical engineering, not physics. "He felt that physics was a hard way to make a living, particularly for someone who was Jewish at a time when the universities had official or unofficial quotas," Landauer says in his still noticeable German accent. Landauer eventually pursued courses that mixed both physics and electronics. But when given a choice later in life, he never concealed his preference for theory over management. Fortunately, his career coincided with the golden era of industrial research, when musings on galaxy formation or limits to computation did not have to be tied to a specific technology development, as they do today.

Besides his other work, Landauer became known for basic theories related to the physics of small structures. He can describe each of his theories in a crisp, methodical manner, intelligible to those who are not initiates in the subtleties of quantum mechanics. Folders rest on one corner of his otherwise clean desk, each containing papers or charts he has set aside for a visitor. One folder corresponds to electron transport—that is, the movement of electrons through small spaces—and another to statistical mechanics, a contemplation of the tiniest device that can hold a 0 or 1 bit without being knocked out of its state by noise in the environment. Landauer eschews heavy use of mathematics in favor of a "lowbrow, intuitive style" that relies on word pictures: "It's like I'm sitting on the back of an electron and watching the world go by." One major contribution, known as Landauer's formula, calculates electrical conductance (how much current can be achieved for a given voltage) from the probability that an electron entering a small structure will make it to the other side rather than bouncing out of its entry point.

Although the physics of information is now well established, Landauer remains preoccupied with a number of questions that he knows he may never be able to answer. How large, for instance, can one make a computer memory in a finite universe? How precisely can the world be described? "The vision of a totally precise computer doesn't exist in the real world any more than one can hypothesize detecting seven angels on the head of a pin," Landauer declares. One of his legacies may be to underline questions that will define the boundary between real computation and mere angel watching. *—Gary Stix*



***REVERSIBLE COMPUTATION** allows operations to be run backward: the input can be deduced from the output* (horizontal chain). *Therefore, no information is lost, no energy is dissipated, and Landauer's principle is preserved. Irreversibility occurs when one cannot track back to the input from the output* (two paths merging).

JOHNNY JOHNSON

TECHNOLOGY AND BUSINESS

ELECTRONICS

THE REINVENTION OF PAPER

Cheap, lightweight, low-power electronic displays have been made in the lab

Twenty years ago Nicholas K. Sheridon got his big idea, the kind that scientists—if they are talented and fortunate—get just once or twice in a career. Sheridon hit on a way to draw images electronically that would be far more portable than heavy cathode-ray tubes, far cheaper than liquid-crystal panels. In theory, his invention could bring to digital displays many of the advantages of paper. They would be thin and flexible yet durable. They would consume only tiny amounts of power yet would hold images indefinitely. They could be used for writing as well as reading, and they could be reused millions of times. Yet they would be as cheap as fine stationery. Sheridon named his idea Gyricon, and he applied for and received a patent on it.

But there his good fortune failed him. Twenty years ago Sheridon's managers at the Xerox Palo Alto Research Center were sitting on many of the inventions that would eventually propel the personal computing revolution: the windows and mouse interface, the laser printer, Ethernet. Like those innovations, Gyricon drew only yawns from Xerox's blinkered managers. "The boss said, 'Xerox really isn't interested in displays. Why don't you work on printing technologies?' So I did," Sheridon recalls.

Fifteen years later the soft-spoken scientist returned to his inspiration, and today the incarnation of his idea sits flashing the PARC logo in black-and-white on his desk. Although it is an early prototype, the 15-by-15-centimeter device is quite legible and thin. More impressive is the fact that the device is powered entirely by a pinky-size solar cell. When Sheridon removes the power altogether, the logo stops changing shade, but it does not fade.

Gently peeling a sheet off its plastic backing, Sheridon shows me what Gyricon is made of. The material is no thicker than a latex glove, and it feels about as rubbery. That is no coincidence: the substance is made by mixing tiny plastic balls, each just 0.03 to 0.1 millimeter in diameter, into molten, transparent silicone rubber. Every ball is white on one side, black on the other. Cooled on slabs and cut into sheets, the rubber is next soaked in oil, which it sucks up like a mop. As it does, the sheets expand and oil-filled pockets form around each ball, which can then float and rotate freely.

Through a chemical process that Xerox is holding as a trade secret, "each ball is given an electric charge, with more on one side than on the other," Sheridon explains. So when an electric field is applied to the surface of the sheet, the balls are lifted in their oil-filled cells, rotated like the needles of tiny compasses to point either their black or their white hemispheres eyeward, and then slammed against the far wall of the cell. There they stick, holding the image, until they are dislodged by another field. At high voltages, the balls stick before completing their rotation, thus producing various shades of gray. Sheridon's group has also produced red-and-white displays and is working on combining balls of various hues to produce full-color ones.

In his early work on Gyricon, Sheridon had figured out everything except a cheap way to make billions of plastic balls, all colored on one side only and all the size of a pinpoint. "This is the secret," Sheridon says, holding up a steel disk slightly smaller than a CD. The disk is spun on a spindle at 2,700 rpm. White plastic is pumped onto the top of the spinning disk, black onto the bottom. The plastic streams skitter off into jets that join at the edge and break up into precisely bicolored, spherical droplets.

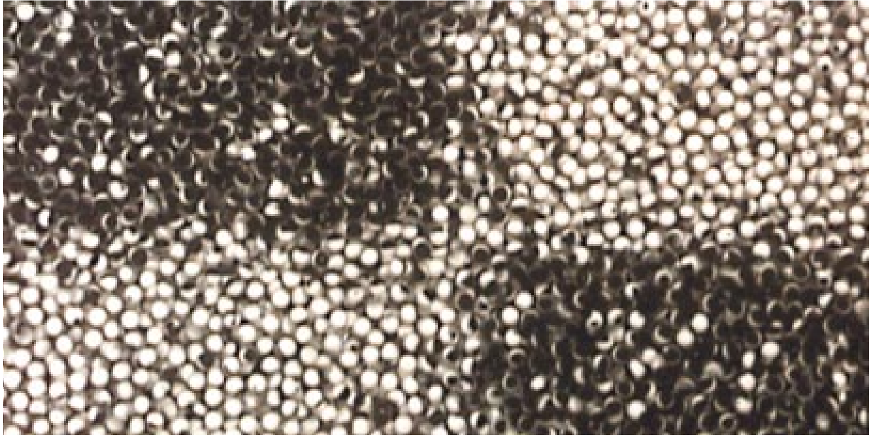
Sheridon has demonstrated that the Gyricon material remains stable after more than two years and three million erasures. His group has built displays at resolutions of up to 220 dots per inch (200 percent finer than most LCDs) and sizes up to a foot square. "We would like to get higher resolution and better whiteness," Sheridon admits. "But we know how to do that: make the balls smaller and pack them more closely."

BRIAN TRAMONTANA *Xerox PARC*

PROTOTYPE OF XEROX'S GYRICON DISPLAY,
as thin as seven sheets of paper, will hold its image for months without power.

For certain applications, such as large commercial signs, the technology appears to be only a few years from market. Gyricon displays might find their way into laptop and handheld computers soon after that. "It would probably allow you to run a laptop for six months on a few AA batteries," Sheridon says, because the device requires neither a backlight nor constant refreshing, as LCDs do.

But the real goal, Sheridon says, is also the most distant: an electronic surrogate for paper. Engineer Matt Howard hands me a wooden pencil that is plugged into



BRIAN TRAMONTANA *Xerox PARC*

TINY, BICOLORED PLASTIC BALLS,
seen here under a microscope, flip to create black, white or gray-scale images.

a weak power supply. As I write on the sheet, the tiny electric field conducted through the pencil's graphite core darkens the screen wherever the tip touches. Howard is working on a handheld wand that will receive text and images from a computer and scan them onto a Gyricon page, which would then be annotated, photocopied, erased—but not discarded.

The effect of such an invention on business—especially Xerox's business—is hard to overestimate. Had PARC been more farsighted, or Sheridon more ambitious, would electronic paper have become commonplace a decade ago? Quite possibly. As it is, Gyricon now must compete with liquid crystals and electrophoretic displays (which use charged particles of one color suspended in liquid dye of another) being developed by the Massachusetts Institute of Technology and E Ink in Cambridge, Mass. Sheridon grimaces briefly as he concedes, "It's a horse race."

—W. Wayt Gibbs in Palo Alto, Calif.

MEDICINE

HEALING CANCER

Vaccines that prod the body to cure itself are finally being readied for market

If all goes as expected, the first vaccine against cancer will be approved for sale before the end of September. The vaccine will neither prevent cancer nor cure it, and it would first be sold in Canada, not the U.S. It will be a significant event nonetheless, because it will demonstrate that a long-held dream—of attacking cancer by guile from within, rather than assaulting the body by brute force from without—is beginning to come true.

More than half a dozen large-scale tests of cancer vaccines are under way in clinics around the world. Most aim at the same malignancy as this first drug: melanoma, a fast-spreading skin cancer that strikes about one person in 100. Ribi ImmunoChem Research, a biotech firm in Hamilton, Mont., was simply the first to file for market approval. European regulators are also evaluating the company's clinical results, and Ribi plans to put its new medicine, Melacine, before U.S. Food and Drug Administration reviewers later this year.

What those experts will see is evidence that nearly all the 70 terminal patients injected with this cocktail of ripped-up tumor cells and bits of the bacteria that cause tuberculosis felt significantly fewer ill effects than the 70 given standard chemotherapy. The vaccine made patients' lives easier but not longer. At least not on average; the lucky few who responded well to the vaccine did survive longer than those who responded well to conventional drugs.

There is good reason to hope that other, more sophisticated vaccines still in clinical trials will improve on those modest gains, in two ways. They may contain more potent adjuvants, additives such as the bacterial fragments in Melacine that awaken the body's immune system to the fact that the cancer doesn't belong there. And they may use more effective antigens, fragments of tumor cells that train antibodies and killer *T* cells to recognize cancer when they see it.

In March, for example, Steven A. Rosenberg and his colleagues at the National Cancer Institute reported good news about a vaccine they have made from a particular protein fragment and interleukin-2, a chemical secreted by *T* cells when they stumble on foreign bodies. Of the 31 patients with widespread melanoma who were immunized with the new medicine, 13 saw their tumors shrink by more than half.

Rosenberg's group went to great effort to identify just the right section of protein to use in its vaccine, but a shotgun technique may also work against some cancers. Michael G. Hanna, chairman of Intracel in Rockville, Md., announced in July that a decade-long test of its OncoVaxCL vaccine for colon cancer had succeeded. Intracel borrowed parts of tumors removed from patients' colons, digested them with enzymes and then injected each patient with his or her own tumor cells, along with a bacterial adjuvant. In people suffering from stage II colon cancer, the vaccine appeared to cut the rate at which the disease resurged by 61 percent over about five years, when compared with patients treated by surgery alone. Intracel is planning to file for FDA approval of its drug later this year.

Other large-scale vaccine trials are just getting started. Progenics Pharmaceuticals in Tarrytown, N.Y., had by July enrolled more than half the 800 American skin cancer patients it wants for its study. They will test a concoction of a carbohydrate antigen and an adjuvant derived from the bark of the South American soap tree, says Robert J. Israel, the company's chief scientist.

ImClone Systems in New York City is gearing up for a trial of similar size to see whether its vaccine will prevent the recurrence of small-cell lung cancer. "Virtually all patients with this disease relapse after their initial treatment," says Harlan W. Waksal, ImClone's chief operating officer. "The disease usually comes back within a year—and with a vengeance. We hope to stop that." ImClone's antigen might seem like an unlikely champion, constructed as it was by making an antibody to an antibody of a sugar-fat compound on cancer cells. But in small-scale trials, Waksal reports, about 40 percent of people given the vaccine survived five years, despite odds predicting that fewer than 5 percent of them would hang in that long.

The John Wayne Cancer Institute in Santa Monica, Calif., is coordinating an even larger and longer study, to span five years and eight nations and to include 1,100 people whose melanoma has spread into their lymphatic system. The subjects will be given either irradiated melanoma tumor cells or interferon alfa-2b, a drug that forces tumors to display the antigens that make them susceptible to attack. (The two drugs need not be exclusive; Ribi is running a clinical trial to see whether they work well together.)

The largest trial of a cancer vaccine so far, however, is unfortunately an unscientific one. In the past year, reportedly upward of 50,000 people in China have been injected with *kang lai te,* an extract from seeds of the herb Job's tears (*Coix lacryma-jobi*) that the government has endorsed as a treatment for cancers of the lung, liver and stomach.

It is too early to say whether science can coax and coach the human body to defend itself successfully against itself. But at the very least, medicine now seems poised to offer a more palatable exit strategy than poison, radiation or the blade.

—W. Wayt Gibbs in San Francisco

DATA STORAGE

THE DOPE ON HOLOGRAPHY

A new technique could fulfill holography's promise for capturing information

The burgeoning demand for storing colossal amounts of digital data has spurred academic and industrial researchers to seek new memory technologies. One promising idea, dating from the 1970s, is to use holograms: "frozen" interference patterns created by lasers. A thumb-tip-size block of the right material can potentially store holograms representing thousands of billions of data bits, thereby offering a much greater density of information than do today's data storage devices. Moreover, holograms can be read very quickly. So far, though, holographic memory has not become commercially viable. But recent efforts by a group at the California Institute of Technology may soon change all that.

Until now, the technology has been dogged by a key disadvantage: holograms tend to be "volatile." In other words, reading them quickly degrades their content. A hologram is created when two laser beams—one of which encodes data—interfere with each other. The interference pattern created by the beams is captured as electric fields in a susceptible material—for example, lithium niobate that has been doped with a tiny amount of some other metal. To read the hologram, a single laser beam is shone at it; the hologram then diffracts the light in a pattern that holds the stored data. But the laser light also "washes out" the hologram as it illuminates it.

Although engineers have experimented with various ways to make holograms more durable, all have had serious limitations. One technique requires that the storage medium be heated, and another method needs very high power lasers.

The Caltech group has discovered what it claims is a far more practical way to make nonvolatile holograms. Karsten Buse, Ali Adibi and Demetri Psaltis reported in *Nature* in June that they used some special, thin crystals of lithium niobate that incorporated trace amounts of iron and manganese atoms. When excited by different kinds of light, these doping atoms liberate electrons that can be taken up by either nearby iron or nearby manganese atoms to capture the electric fields of a hologram. Iron gives up its electrons in response to either ultraviolet light or red light, whereas manganese must have ultraviolet light.

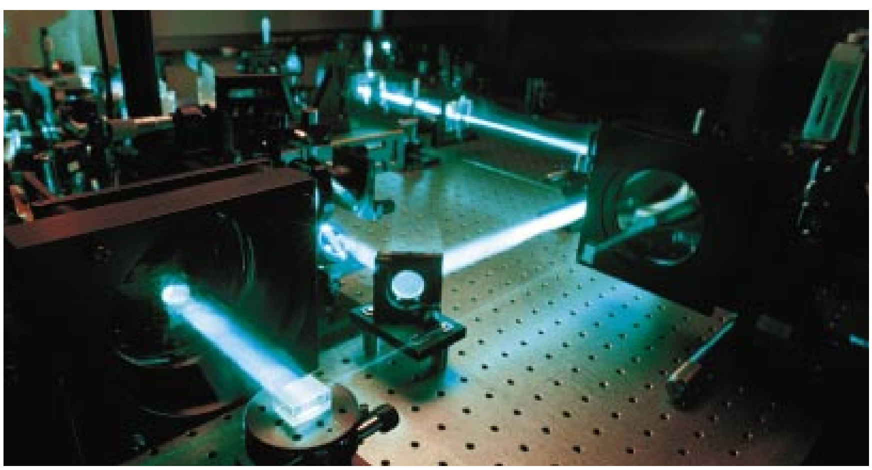
The researchers found that they could record data durably in their crystals by illuminating them with ultraviolet light (not from a laser) at the same time that they made a hologram with two red laser beams. The UV light stimulated both the manganese and the iron atoms to liberate electrons. Doing this ensured that the hologram created by the red laser beams was stored by both iron and manganese atoms—despite the insensitivity of manganese to red light.

After recording, the UV light was turned off. The resulting hologram could be read by illuminating it with red laser light alone. The red light did not excite the UV-triggered manganese atoms, so they retained the imprinted data without loss. (The light signal from the iron atoms did diminish during reading, as expected, but there were more than enough manganese atoms to preserve a strong hologram for very long periods.) Turning the ultraviolet light back on allowed a new durable hologram to be written.

The work "is a step toward a practical holographic storage device," according to Hans Coufal of the IBM Almaden Research Center, who works in the field. But Coufal notes that materials sensitive to dimmer light will be needed for mass-market devices. And Lambertus Hesselink of Stanford University says he doubts whether the Psaltis group's technique will still look promising when the experiments are repeated with crystals thick enough to hold useful amounts of data.

Psaltis, however, says the technique—which is patent pending—is in important ways adequate for a commercial read-write memory, although he hopes to improve it. The double-doped crystals the Caltech researchers used were made in Europe 20 years ago. Theory suggests that up to a 100-fold improvement in performance should be possible with better crystals, states Psaltis, who is fabricating new versions. Double-doped crystals incorporating cerium instead of iron look particularly intriguing, he adds. If Psaltis is right, the new age might be dominated by crystals after all.

—Tim Beardsley in Washington, D.C.



CHARLES O'REAR

DOPED WITH DIFFERENT METALS, THE LITHIUM NIOBATE CRYSTAL **(translucent block at bottom)** ***can improve holographic memory.***

Opaque Transparency

If the frustrations of all the world's computer users were brought together, the resulting explosion would make the big bang look like a Roman candle. This is true even though computers have come a long way in the decade since the industry pronounced "usability" a necessity. Nevertheless, we still have a plethora of frustrating functions: inconsistent commands (drag a file in Windows, for instance, and you could end up moving it, copying it or perhaps even creating a link back to it from another directory—who knows?), programs that rename files according to their tastes instead of yours, "help" screens that explain options but not what they mean or what their consequences are, and inscrutable error messages. Home systems—such as Windows 95 or 98—hide their inner workings from users, whereas professional systems—such as Windows NT—keep those inner workings accessible but design them only for experts.

Nowhere, it seems, is there a system designed for people who know how to use a computer but aren't techies. I fall in that category, and the upshot is that I spend some part of every day in an absolute rage at the bozos who designed the computer I live with, which is refusing to let me do one or another simple thing.

The underlying presumption of software experts is that there are only two kinds of people: those who already know all about computers and those who don't want to know anything about them but just want to use them to complete tasks. The computer industry, with its usual fine grasp of language, calls this last notion "transparency"—as in, "the computer should be transparent to the user." A good example of this way of thinking is the Windows help system, which lets you click on a button to go directly from the topic you looked up to whatever nested bit of the program you need. Fine, but then the system never tells you where that bit was or how to go there directly so that next time you don't have to go through the help system. That's like welding the training wheels to your bicycle: you can never actually learn anything.

The notion that users should not have to worry or care about how the innards worked was first implemented in a mainstream computer in 1984, when the Apple Macintosh changed people's notions of what a computer could be. At the time, liberating users from the petty bureaucracy of command lines meant people were free to do all kinds of work they either could not have done before or could not have afforded to do before—desktop publishing is just one example.

From the mid-1980s to the early 1990s, as graphic interfaces took hold

DAVID SUTER

and computer companies set up usability labs, the idea that users should not have to think about the computer lying between them and their task was an enormous step forward. "The point cannot be overstressed: make the computer system invisible," wrote Donald Norman in his 1988 design classic *The Psychology of Everyday Things*. Norman went on to imagine the perfect appointment calendar, which looked like a paper calendar but which could send messages and reminders to remote systems.

The usability efforts, however, addressed only one of the two problems every computer poses: how to accomplish tasks and how to manage the computer itself. The emphasis for the past 15 years has been on tasks, and rightly so. Most people do not buy computers because they think it will be fun to rearrange files and directories. Computers, though, do still have to be managed—just as cars have to be serviced—and the accessibility trend has made this job far more difficult.

Some recent changes—such as Windows 95/98/NT's registry, apparently designed to be read and edited in the original Martian by people who like walking over a 1,000-foot canyon on a tightrope with no safety net—seem to me purely lethal. Here is a single database that most people don't know how to back up and whose corruption or loss wipes out all your customization and configuration settings and makes your computer forget it has any software installed. It would have been perfectly possible to design the registry to be forgiving and to track all changes, so that you could go back and undo the last change you made or the changes the program you just installed made that knocked out your network. Instead changes are made on the fly, and there's no way back.

The next leap in computing is to embed computers in objects all around us. The new CrossPad is a crude example: you write on a pad of ordinary paper with a special pen, and the device stores the image and indexes the pages by the keywords you circle. If only you could use a fountain pen and the pad came in jacket-pocket size.... Even so, I want one.

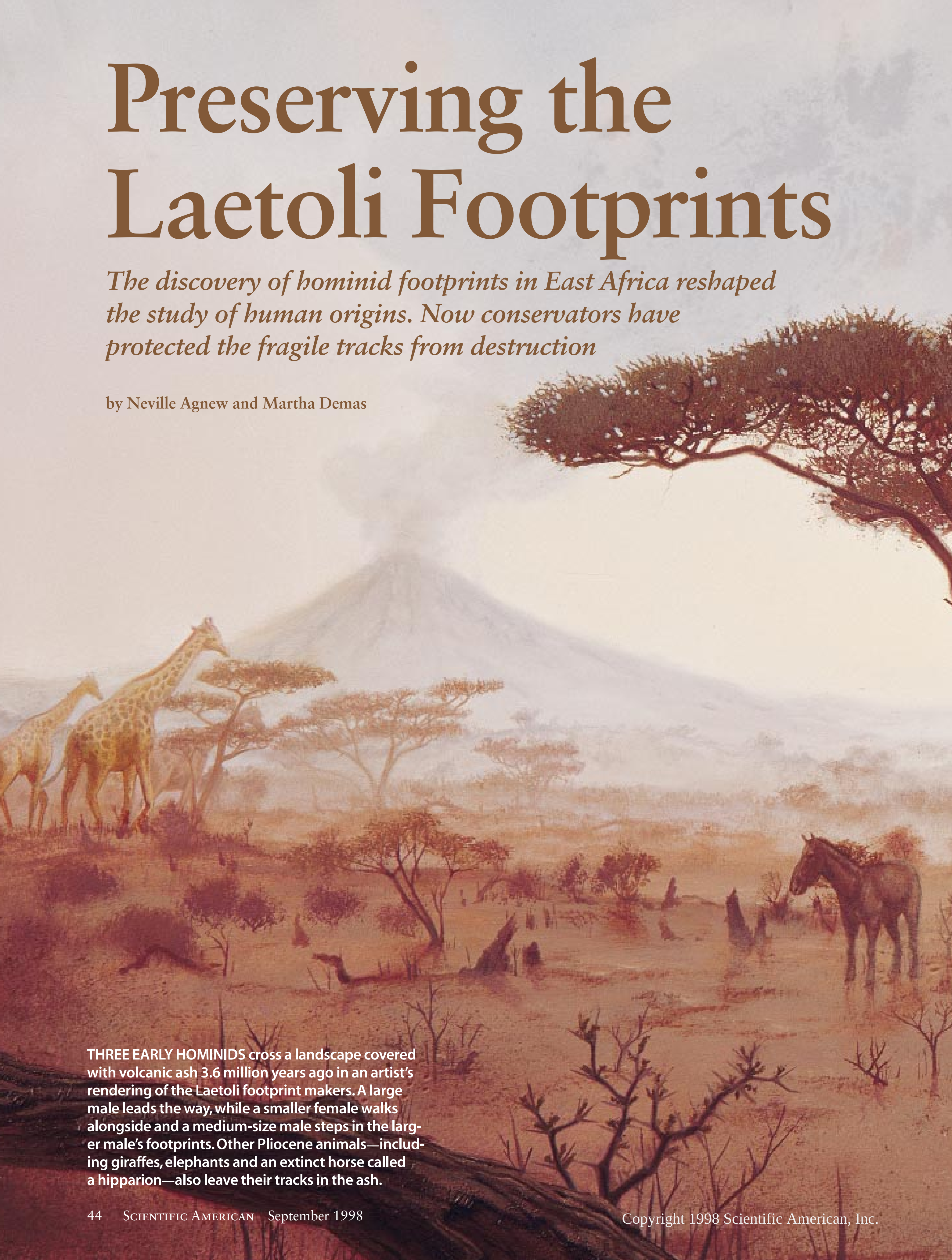
But the whole mess still has to be uploaded to a PC before you run out of storage space, and that introduces the same old conundrum: how to get "inside" your computer so as to manage it. As we increasingly talk about "smart" toasters and doorknobs and refrigerators (imagine: they could tell your computer to order milk before you run out), we are not just talking about enhanced capabilities for these objects, we are also talking about putting in place underlying systems to manage all these things. I just spent four days trying to network three PCs that are supposed to network "right out of the box." Will installing your new toaster in a smart world be any easier than trying to network those PCs? As the comic actor Edmund Kean might have said, complexity is easy; simplicity is hard.

—Wendy M. Grossman in London

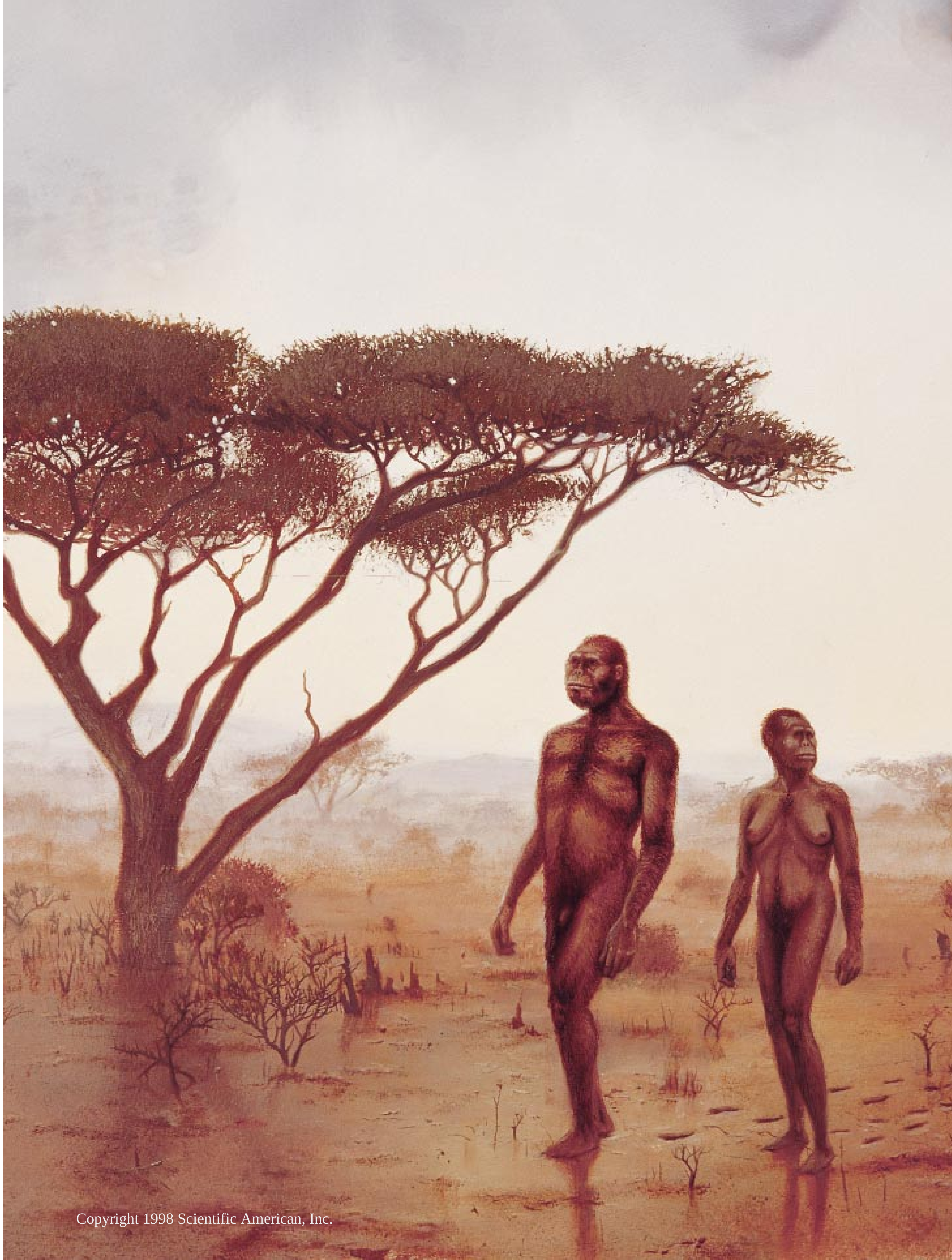
Preserving the Laetoli Footprints

The discovery of hominid footprints in East Africa reshaped the study of human origins. Now conservators have protected the fragile tracks from destruction

by Neville Agnew and Martha Demas

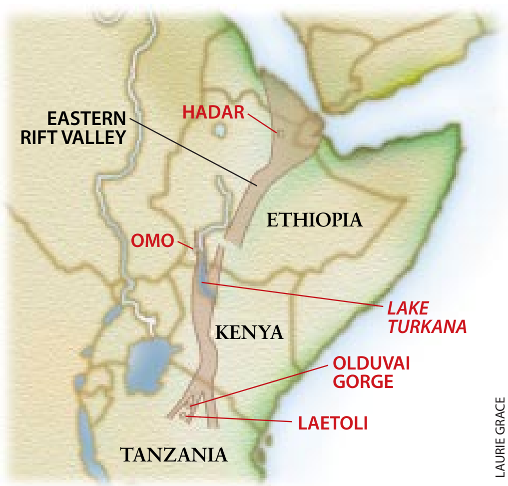


THREE EARLY HOMINIDS cross a landscape covered with volcanic ash 3.6 million years ago in an artist's rendering of the Laetoli footprint makers. A large male leads the way, while a smaller female walks alongside and a medium-size male steps in the larger male's footprints. Other Pliocene animals—including giraffes, elephants and an extinct horse called a hipparion—also leave their tracks in the ash.



One of the most remarkable events in the annals of anthropology occurred 20 years ago in an area of northern Tanzania called Laetoli. A team led by famed archaeologist Mary D. Leakey was searching for fossils of the early hominids that ranged through East Africa millions of years ago. In the summer of 1976, after a long day in the field, three visitors to Leakey's camp engaged in some horseplay, tossing chunks of dried elephant dung at one another. When paleontologist Andrew Hill dropped to the ground to avoid getting hit, he noticed what seemed to be animal tracks in a layer of exposed tuff—a sedimentary rock created by deposits of volcanic ash. On closer inspection of the area, the scientists found thousands of fossilized tracks, including the footprints of elephants, giraffes, rhinoceroses and several extinct mammal species. But the most extraordinary find came two years later, when Paul I. Abell, a geochemist who had joined Leakey's team, found what appeared to be a human footprint at the edge of a gully eroded by the Ngarusi River.

ALFRED T. KAMAJIAN



LAURIE GRACE

LAETOLI AREA in northern Tanzania lies in the eastern branch of the Great Rift Valley, where many hominid fossils have been found. Other well-known hominid sites include Hadar and Omo in Ethiopia, Lake Turkana in Kenya and Olduvai Gorge in Tanzania.

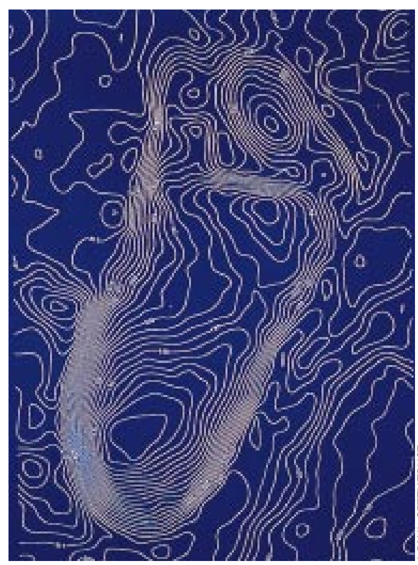
Excavations of the Footprint Tuff—as it came to be known—in 1978 and 1979 revealed two parallel trails of hominid footprints extending some 27 meters (89 feet). The volcanic sediments were dated radiometrically to be between 3.4 million and 3.8 million years old. The discovery settled a long-standing scientific debate: the Laetoli footprints proved that early hominids were fully bipedal—they had an erect posture and walked on two feet—long before the advent of stone toolmaking or the expansion in size of the human brain. What is more, the trackway provided information about the soft tissue of the hominids' feet and the length of their strides—information that cannot be ascertained from fossil bones. For these reasons, the Laetoli footprints attracted a huge amount of attention from scientists and the general public. Leakey, who died in 1996, regarded the discovery as the crowning achievement of her six decades of work in East Africa.

That the footprints have scientific value is obvious: they have answered fundamental questions about humanity's past. But they also have a profound cultural symbolism. In a powerfully evocative way, the tracks of those early hominids represent the long evolutionary history of humankind. The footprints bear witness to a defining moment in the development of our species and speak to us directly across thousands of millennia.

For the past six years, the Getty Conservation Institute—a Los Angeles–based organization concerned with the preservation of cultural heritage—has worked with Tanzanian authorities to ensure that the Laetoli footprints stay intact for years to come. A team of conservators and scientists recently completed a project to protect the footprints from erosion, plant growth and other causes of deterioration that have threatened the trackway since its discovery.

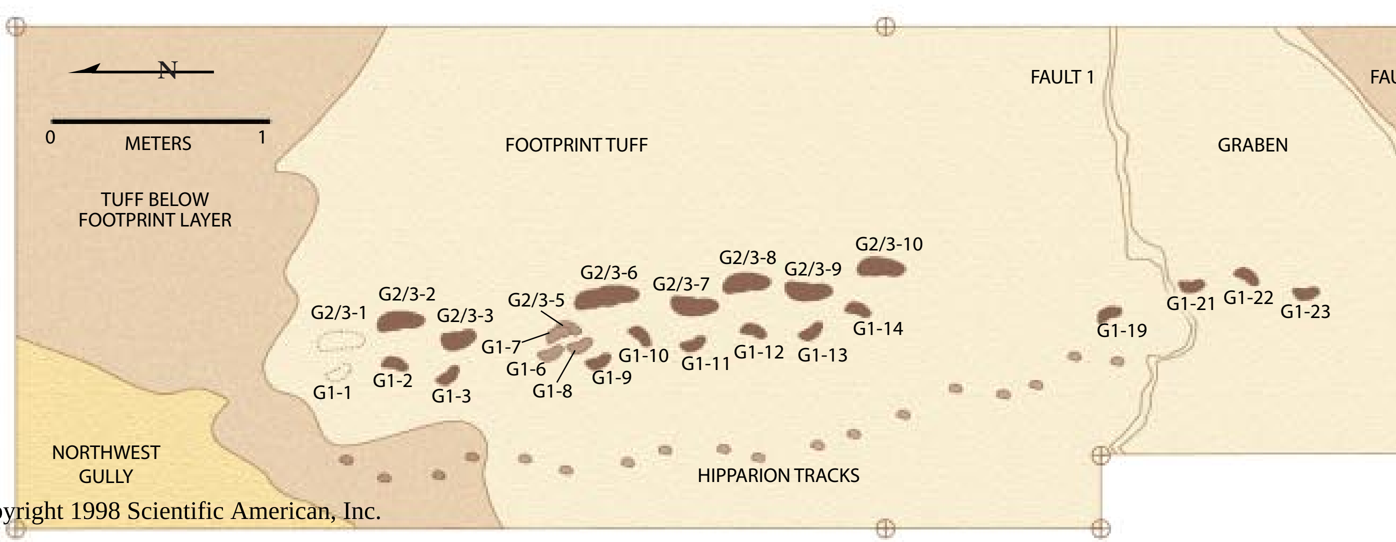
A Pliocene Eruption

Skeletal remains stand a better chance of survival in the fossil record than impressions in mud or volcanic ashfall. Yet traces of many animals dating back to the Paleozoic era, some as old as 500 million years, are known throughout the world. Because an animal leaves many tracks during its lifetime but only one set of bones when it dies, statistically it is not so surprising that some of the tracks survive as fossil imprints. The number and variety of tracks preserved in the Laetoli exposures is nonetheless unusual. At the largest of the 16 sites at Laetoli where tracks have been found, there are an estimated 18,000 prints,



J. PAUL GETTY TRUST

CONTOUR MAP of hominid footprint G1-36 (*right*) was created by taking two overlapping photographs of the print with a high-resolution camera. The deep impression at the bottom of the print indicates that the hominid walked like a modern human, placing its full weight on its heel. The length of the footprint is about 20 centimeters (eight inches). On the next page, two views of footprint G1-25 show that it suffered little damage between its discovery in 1979 and its reexcavation in 1995. The reexcavated print (*far right*) is shown next to a photograph of the print taken in 1979 by a member of Mary Leakey's team.



representing 17 families of animals, in an area of about 800 square meters.

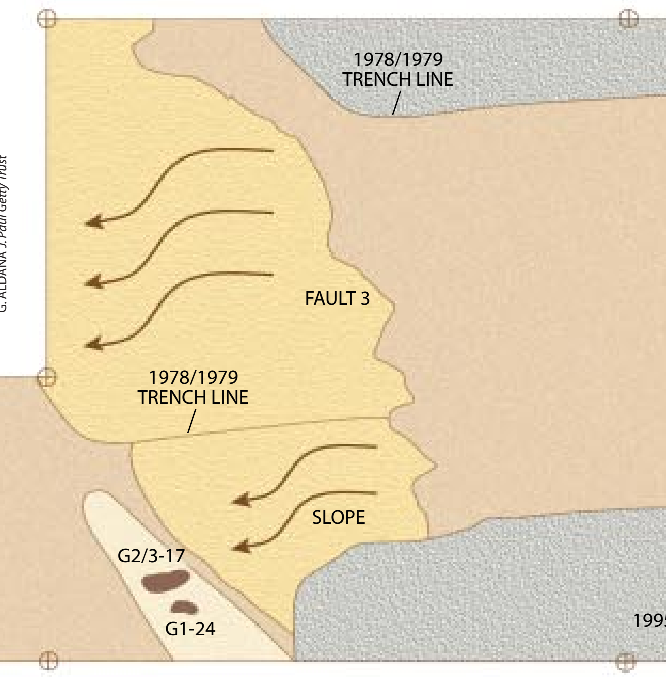
Laetoli lies in the eastern branch of the Great Rift Valley, a tectonically active area. About 3.6 million years ago, during the Pliocene epoch, the Sadiman volcano—located 20 kilometers (12 miles) east of Laetoli—began belching clouds of ash, which settled in layers on the surrounding savanna. At one point in the volcano's active phase, a series of eruptions coincided with the end of an African dry season. After a light rainfall, the animals that lived in the area left their tracks in the moist ash. The material ejected from Sadiman was rich in the mineral carbonatite, which acts like cement when wet. The ash layers hardened, preserving the thousands of animal footprints that covered the area. Shortly afterward Sadiman erupted again, depositing additional layers of ash that buried the footprints and fossilized them. Finally, erosion over millions of years reexposed the Footprint Tuff.

The two parallel trails contained a total of 54 footprints that could be clearly identified as hominid tracks. The soil covering varied from a few centimeters at the northern end of the trackway—the area where the footprints had first been discovered—to 27 centimeters (11 inches) at the southern end. To the north, the footprints ended at the wide, deep gully cut by the Ngarusi River; to the south, faulting and erosion precluded any chance of picking up the trail. The trackway itself shows faulting, too, with a graben—a section that had dropped 20 to 40 centimeters because of tectonic activity—near the midpoint. Part of the trackway is also heavily weathered: in this section the tuff had changed to dried mud and the footprints were poorly preserved. But in the less weathered part of the trackway the preservation was good, allowing clear recognition of soft-tissue anatomical features such as heel, arch and big toe.

As so often happens in the field of paleoanthropology, disagreement soon broke out regarding the interpretation of the evidence. One point in dispute was the species of the hominids that made the footprints. Leakey's team had found fossilized hominid bones in the Laetoli area that were the same age as the trackway. Most scientists believe these hominids belonged to the species *Australopithecus afarensis,* which lived in East Africa between 3.0 million and 3.9 million years ago. In fact, one of the Laetoli hominid remains—a mandible with nine teeth in place—became the type specimen, or defining fossil, for *A. afarensis.* (The famous hominid skeleton known as "Lucy," discovered in 1974 in Ethiopia, is another representative of this species.) But Leakey did not accept that the Laetoli hominids were specimens of *A. afarensis;* she resisted assigning them to any species. (Leakey was cautious about interpreting her discoveries.) She did believe, however, that the makers of the Laetoli footprints stood in the direct line of human ancestry.

Another dispute concerned the number of hominids that made the two parallel trails. In one trail, the footprints were small and well defined,

G. ALDANA J. Paul Getty Trust



HOMINID TRACKWAY consists of 54 footprints running north in two parallel trails. In the G1 trail the prints are small and well defined. In the G2/3 trail the prints are larger and poorly defined, indicating that the trail may have been

but in the other the prints were larger and less clear. Some scientists speculated that the trails were made by two hominids—possibly a female and a male—walking abreast or close to each other. [For artistic representations of this interpretation, see “The Footprint Makers: An Early View,” by Jay H. Matternes, on page 52, and “The Laetoli Diorama,” by Ian Tattersall, on page 53.] Other scientists believed the trails were made by three hominids. In this view—which most paleoanthropologists now share—the trail of larger footprints was made by two individuals, with the second hominid purposely stepping in the tracks of the first [see “A New Look at Laetoli,” at right].

The footprints prompted other intriguing questions: Where were the hominids going? What caused them to break stride—which is indicated by the position of four footprints in the northern section of the trackway—as though to look back on where they had come from? Were they a family group? Were they carrying anything? And how did they communicate? These tantalizing questions will never be answered, but scientists can use the evidence gleaned from the Laetoli site to attempt to re-create the moment when the hominid tracks were made.

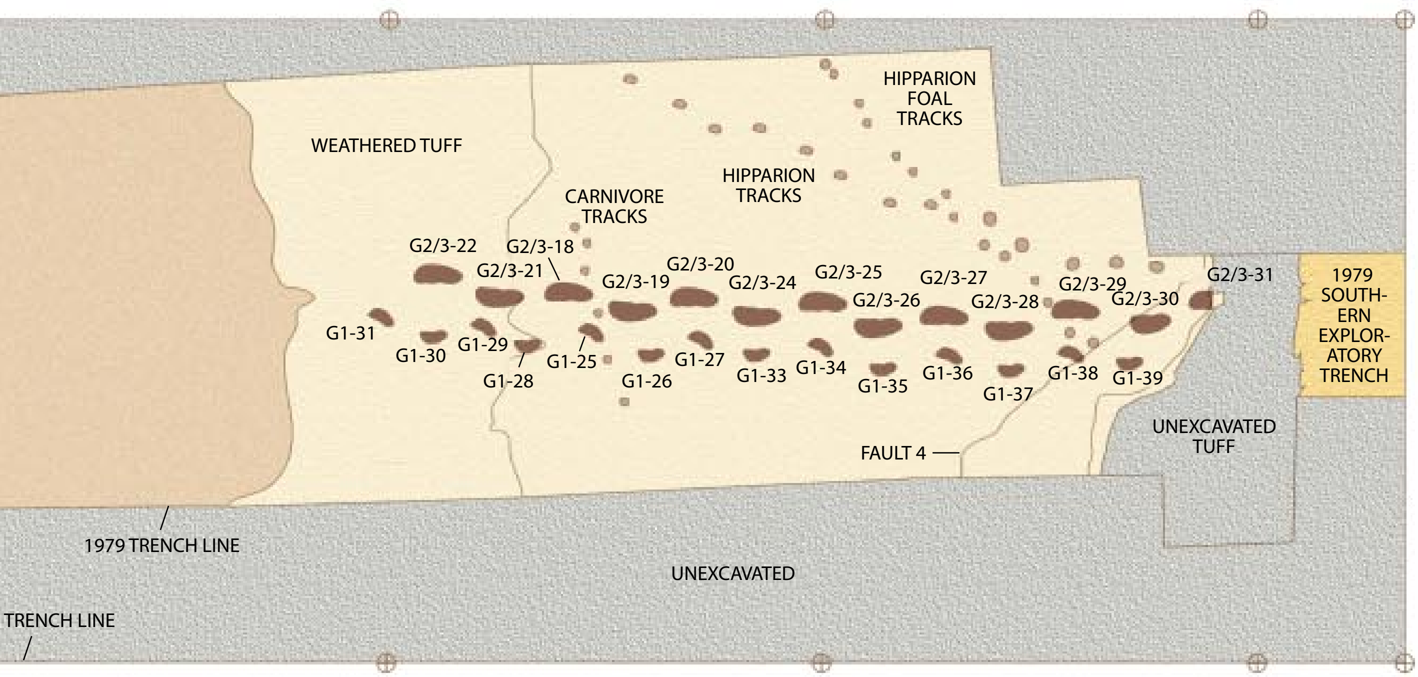
A New Look at Laetoli

The artist’s rendering of the Laetoli footprint makers on pages 44 through 46 reflects the widely accepted interpretation that the trackway was made by three hominids. Many of the larger tracks at the site have features indicating that they may be double footprints. The evidence suggests that a relatively large hominid—about five feet tall, based on the size of its footprints—walked first, and a hominid four and a half feet tall deliberately stepped in the leader’s footsteps, perhaps to make it easier to cross the slick, ash-covered ground. A smaller hominid—about four feet tall—apparently made the parallel trail of well-defined footprints. The trackway indicates that this hominid adjusted its stride to keep up with one or both of the other hominids.

The illustration shows the two larger hominids as males and the smaller individual as a female, but this was not necessarily the case: the smallest member of the trio could have been a child. The female is shown walking slightly behind the lead male because the two could not have walked abreast without jostling each other.

—The Editors

Much of the controversy over the footprints arose because few scientists had the opportunity to study the prints firsthand. At the end of each field season, Leakey’s team reburied the trackway for its protection. But the team members made casts of the best-preserved sections of the trails and documented the site fully. Researchers created three-dimensional contour maps of some of the footprints by photographing them from two perspectives—a process called photogrammetry. Leakey later published her work with several co-authors in a monumental monograph that dealt not only with the hominid prints but also with the many animal tracks and the geology of the Laetoli area. The evidence collected by Leakey’s group—which also included fossilized pollen and impressions of vegetation—provides an unparalleled record of the African



LAURIE GRACE

made by two hominids walking in tandem. The two northernmost tracks (*far left*) were destroyed by erosion between their discovery in 1978 and reexcavation in 1996. Four other tracks in the northern section—G1-6, G1-7, G1-8 and G2/3-5—mark the point where the hominids apparently broke stride. Also present are the tracks of a hipparion.

savanna during Pliocene times and a context in which to understand better the hominid trackway.

The Root Problem

Fieldwork on the Laetoli footprints ended with the 1979 season, and Leakey's team used local river sand to rebury the site. Because the tuff is soft and easily damaged, the mound of sand was covered with volcanic boulders to armor it against erosion and the animals that sometimes roam across the site—particularly elephants and the cattle of the Masai people living in the area. We now know that seeds of *Acacia seyal,* a large, vigorously growing tree species, were inadvertently introduced with the reburial fill. The loose fill and the physical protection and moisture retention provided by the boulders created a microenvironment conducive to germination and rapid plant growth. Over the following decade, the acacias and other trees grew to heights of over two meters. Scientists who occasionally visited the Laetoli site began to voice concern that the roots from these trees would penetrate and eventually destroy the hominid footprints.

In 1992 the Antiquities Department of the Tanzanian government approached the Getty Conservation Institute, which has extensive experience in preserving archaeological sites, to consider how the trackway might be saved. The following year a joint team from the institute and the Antiquities Department excavated a sample trench in the reburial mound to assess the condition of the hominid footprints. The assessment revealed that tree roots had indeed penetrated some of the tracks. But in the areas where no root damage had occurred, the preservation of the prints was excellent. Leakey's intuitive decision to rebury the site had been the right one. With hindsight we can now say that perhaps greater care should have been taken in how the site was buried. Also, periodic monitoring and maintenance—including the removal of tree seedlings before they became established—would have avoided the need for a long and costly conservation effort.

The Getty Conservation Institute and the Tanzanian government agreed to collaborate on the project, but before fieldwork could begin, various options had to be considered. Fossil bones are routinely brought into the laboratory for study and permanent safekeeping.


T. MOON *J. Paul Getty Trust*


N. AGNEW *J. Paul Getty Trust*


A. BASS *J. Paul Getty Trust*


N. AGNEW *J. Paul Getty Trust*


M. DEMAS *J. Paul Getty Trust*

REEXCAVATION began in 1995 with the southern section of the trackway (*top right*). Conservators extracted the acacia tree roots that had penetrated the Footprint Tuff (*middle right*), then removed the fill from the footprints (*top left*). The reexcavated trackway (*bottom left*) was photographed with a Polaroid camera (*bottom right*) to record conditions.

Indeed, to leave them in the field would be irresponsible: they would certainly be lost or damaged. But could the entire hominid trackway be lifted and moved to a museum in Tanzania? Was it technically possible to do this without damaging the footprints? Some scientists were vehement in their belief that this was the only way to save the tracks.

Removal would have been very risky, however, because the techniques for cutting out, lifting and transporting such a large trackway had not been proved. The Footprint Tuff is far from being a homogeneous stratum. It consists of many thin layers of volcanic ash, each with different weathering, hardness and cohesion. Without strengthening the tuff with resin—an intervention with unknown long-term consequences—fracturing would probably occur during removal. What is more, removing the trackway or the individual footprints would separate them from the many animal tracks that had been made at the same time. Part of the significance of the hominid trails—their setting in the savanna landscape of East Africa together with the tracks of other Pliocene species—would be lost.

An alternative proposal was to shelter the trackway, erecting a protective building over it. The site could then be opened to the public, and the footprints could be studied by visiting scholars. The Laetoli area, however, is remote. There is no road to the site and no water or power lines nearby. Experience in Tanzania has shown that without proper financing, trained personnel and an adequate infrastructure, sheltering the site could prove disastrous: it could result in the deterioration of the trackway rather than its preservation. Even in countries where resources are plentiful, archaeological sites have been damaged when planning has been inadequate or when climate-controlled enclosures have not performed as expected. Moreover, no shelter could fully protect the trackway from weathering: moisture from the ground below would rise to the surface seasonally through capillary action. Soluble salts in the water would crystallize on the surface, causing stress that would eventually rupture the trackway. During the dry season, dust accumulation in the prints would require frequent cleaning, which would inevitably lead to damage.

The third option was to reexcavate the trackway, remove the vegetation that had damaged it and then rebury the site more carefully, taking steps to prevent root growth that might harm the footprints. Reburial is a proved preservation method. The trackway survived underground for thousands of millennia; if reburied, it would be protected from erosion, physical damage and rapid fluctuations of moisture. Reburial is also readily reversible: the tuff can be uncovered in the future if the other options become more feasible. For these reasons, the Getty Conservation Institute recommended reburial. In 1993 Tanzania's Antiquities Department decided to proceed with this recommendation, and a committee was set up to assist the implementation of the plan. Participating in the discussions were Leakey and other eminent paleoanthropologists, Tanzanian officials and a regional representative from the United Nations Educational, Scientific and Cultural Organization.


A. BASS J. Paul Getty Trust


F. LONG J. Paul Getty Trust

LEAKEY'S CAST OF THE TRACKWAY was used to guide the final stages of the reexcavation of the footprints (*top*). Once the tracks were exposed and photographed, conservators recorded the condition of each print, noting any damage caused by root growth or erosion (*bottom*).

Saving the Footprints

The conservation project began in 1994. During that year's field season, the trees and shrubs growing on and near the reburial mound were cut down. To prevent regrowth, the conservation team applied the biodegradable herbicide Roundup to the tree stumps. In all, 150 trees and shrubs were killed, 69 of them directly on the reburial mound.

Reexcavation of the trackway took place during the 1995 and 1996 field seasons, beginning with the southern section. This section was where the densest revegetation had occurred and, coincidentally, where the best preserved footprints had been found in 1979. Archaeologists and conservators used Leakey's photographs of the trackway to find the exact positions of the hominid footprints. Also useful was the original cast of the trackway, which was replicated, cut into conveniently short sections and used as a guide for the final stages of reexcavation. A temporary shelter erected over the excavated area protected it from direct sunlight and shaded those who were working on the trackway.

In the southern section of the trackway the trees had fortunately developed shallow, adventitious roots rather than deep taproots because of the hardness of the tuff. As a consequence, there was far less damage than had been feared, and most of the footprints were generally in good condition. In areas where the tuff was weathered, however, roots had penetrated the prints. Here the conservation team surgically removed stumps and roots after strengthening adjacent areas of disrupted tuff with a water-based acrylic dispersion. Team members used miniature rotary saws to trim the roots and routers to extract the parts that had penetrated the surface of the trackway. The holes created by root removal were filled with a paste of acrylic and fumed silica to stabilize them against crumbling.

Recording the condition of a site is

The Footprint Makers: An Early View

by Jay H. Matternes

I worked on my painting of the Laetoli footprint makers during the early fall of 1978, shortly after the discovery of the hominid trackway. As part of my research, I flew to Africa to confer with Mary Leakey and her associates at their base camp in Tanzania's Olduvai Gorge. When I boarded the plane, the only information I had on the project consisted of a few photographs of the footprints and the surrounding area, along with a report on the geology of the Laetoli site and a list of the animal tracks found there.

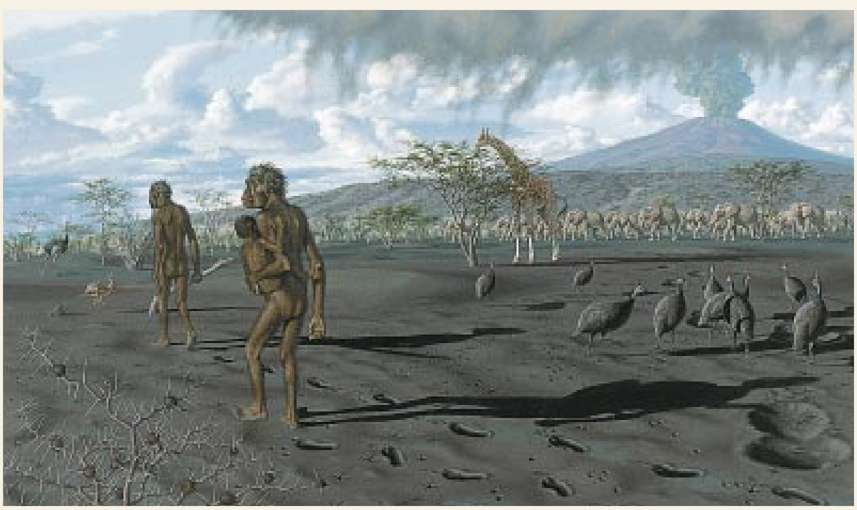
While at the base camp, I consulted with Leakey and made a number of drawings of proposed layouts. She drove me to Laetoli so I could familiarize myself with the main features of the terrain. The analysis of the Laetoli sediments indicated that there had been several types of volcanic ashfalls in the area—some settling undisturbed on the ground, some redeposited by wind—but all the ash had come from the Sadiman volcano. Geologists believe the color of this ash was light gray, not very different from the color of the hardened tuff in which the footprints were discovered.

I based my reconstruction of the two walking figures on the descriptions of *Australopithecus afarensis*. Fossil specimens of this species had been found at Laetoli and the Afar Triangle of Ethiopia; the bone fragments and dental evidence indicated that the two hominid populations looked roughly the same and lived at the time the footprints were made. I inferred the limb proportions of the adults from the skeleton of "Lucy," the female *Australopithecus* whose fossil remains had been found in Ethiopia in 1974. I assumed these hominids would have been lean, energetic bipeds, capable of exploiting a variety of habitats. For this reason, they would have probably had relatively little body hair, to ensure rapid heat loss. They would have also developed a dark skin to counteract the injurious effects of ultraviolet radiation.

At the time I worked on the painting, only a few fragments of *A. afarensis* skulls had been found. I had to base the facial features of the female figure on those of *A. africanus,* a species I had earlier reconstructed. Leakey wanted me to emphasize the small stature of these hominids, so I painted several guinea fowl near the figures. The male figure carries a digging stick, presumably the only tool of this species (the earliest stone tools did not appear until much later). The female carries her toddler on her hip, probably the most convenient position for a habitual biped. The theory that the trails had been made by three hominids was not put forth until after I finished the painting.

The final depiction (*below*) accorded with the few facts of the Laetoli site that were then known. The painting first appeared in the April 1979 issue of *National Geographic* magazine to illustrate an article by Leakey about the trackway.

JAY H. MATTERNES is an artist who specializes in the depiction of hominids and extinct mammals. His work has appeared in museums worldwide.



JAY H. MATTERNES

HOMINID FAMILY members leave their tracks in the ash from the Sadiman volcano.

one of the most important and challenging conservation activities. The team conducted a full survey of the exposed trackway to provide the baseline data that will allow future investigators to assess changes. Using a Polaroid camera, team members made eight-by-10-inch color photographs of the footprints. They then laid acetate sheets over the photographs and noted the places where there were fractures, loss of tuff and intrusive root growth, as well as any other salient information.

During the reexcavation, the conservators noted dark stains in and around each hominid footprint. This darkening was the result of the application of Bedacryl, an acrylic consolidant that Leakey's team had used to strengthen the footprints before making molds of them. (Silicone rubber was applied to the trackway to create molds, which were then peeled off and used to make fiberglass casts.) The staining was an unforeseen side effect: although the Bedacryl did not damage the footprints, it impaired their legibility and thus their scientific value. The Bedacryl could be removed by gently poulticing the footprints with acetone and tissue paper, but because there was a risk of damage to the prints where the underlying tuff was fragile, only two prints were cleaned.

In consideration of the fact that few researchers had ever seen the exposed footprints—most of the scientific literature is based on casts and photographs—Tanzania's Antiquities Department invited a group of scientists to reexamine the trackway while the conservation and recording work was going on. Bruce Latimer, curator of physical anthropology at the Cleveland Museum of Natural History, Craig S. Feibel, a geologist at Rutgers University, and Peter Schmid, curator of the anthropology museum at the University of Zurich, were nominated by specialists in the field of paleoanthropology to come to Laetoli. Their studies included a formal description of the footprints, stature and gait of the hominids and an examination of the thin layers of the Footprint Tuff.

Once the footprints were uncovered and the root damage repaired, a team of photogrammetrists recorded the trackway to make new contour maps of the prints. The new maps are accurate to within half a millimeter, which is far better than the maps made by Leakey's team in 1979. The Laetoli trackway may now be one of the most thoroughly documented paleontological sites. The new

The Laetoli Diorama

by Ian Tattersall

Only very rarely does the fossil record provide evidence of an actual event in human prehistory. So in the late 1980s, when we were considering subjects for presentation in diorama form in the American Museum of Natural History's Hall of Human Biology and Evolution, the making of the Laetoli footprints seemed an obvious choice. Constructing lifelike sculptures of extinct humans involves many tricky decisions [see "Evolution Comes to Life," by Ian Tattersall; SCIENTIFIC AMERICAN, August 1992]. The decisions involving the Laetoli hominids were particularly difficult because the 3.6-million-year-old creatures are so remote from modern-day humans. Our Laetoli diorama posed an additional problem: it was designed to represent a specific event—the journey of the hominids across a plain of volcanic ash—but the evidence from that event is a little ambiguous.

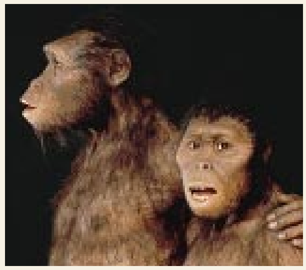
Willard Whitson, the museum hall's designer, and I visited the Laetoli site in Tanzania and discussed our plans for the diorama with Peter Jones, an archaeologist who was part of Mary Leakey's team when the trackway was discovered in 1978. We also consulted paleoanthropologist Ron Clarke, who excavated many of the footprints. Nobody disputes that the two parallel trails were made by beings who were walking bipedally (although they may have been tree climbers as well). The footprints in the westernmost trail were much smaller and more clearly defined than the prints in the eastern trail, but Jones pointed out that the stride lengths were the same. Clearly, the hominids were walking in step and accommodating each other's stride—which meant that the two trails were made at the same time. What is more, the trails are so close together that the hominids must have been in some kind of physical contact when they made them.

Some anthropologists concluded that the trails had been made by a group of three hominids. The western trail, they claimed, was made by a relatively small individual, whereas the eastern trail was made by two larger hominids walking in tandem, with one individual deliberately stepping into the tracks of the other. But Clarke disagreed with this view. He claimed that because the footprints in the eastern trail had so many consistent features, they must have been made by a single large hominid. The larger footprints were more poorly defined than the smaller prints, Clarke argued, because the feet of the large hominid had slid more in the rain-slickened ash.

These facts and theories were our starting point. The rest had to be conjecture. Individuals of different body sizes could have meant a number of things: male and female, parent and child, older and younger siblings. And although we suspected that the two hominids were in physical contact, we had no idea how they were supporting each other. Were they holding hands? Walking arm in arm? Carrying something between them?

The scene as we finally rendered it (*above*) shows two *Australopithecus afarensis,* a large male and a smaller female, walking side by side through a sparsely vegetated landscape. We opted for a male and a female partly to maximize visual interest but also to show the large sex difference in body size that is believed to have existed in *A. afarensis,* the presumed maker of the trails. The male's arm is draped over the female's shoulder. In the explanatory text we emphasize that this scenario is consistent with the few facts we have but is not the only one possible.

Feminists have excoriated us for the "paternalistic" nature of the scene, but in fact we decided to show the figures joined this way because it seemed to carry the fewest unwanted implications. Indeed, a look at the faces of these creatures, brilliantly sculpted by English artist John Holmes, shows that both are worried, the male as much as the female. Here are two small, slow and rather defenseless individuals moving through open country that almost certainly teemed with predators. These early hominids were clearly bipeds when they were on the African savanna, but this dangerous and difficult environment was probably not their preferred milieu. Plausibly, they were crossing this hostile territory to get from one more congenial region to another. Their tracks were headed almost directly toward the well-watered Olduvai basin, where the lakeside forest and its fringes would have felt much more like home.



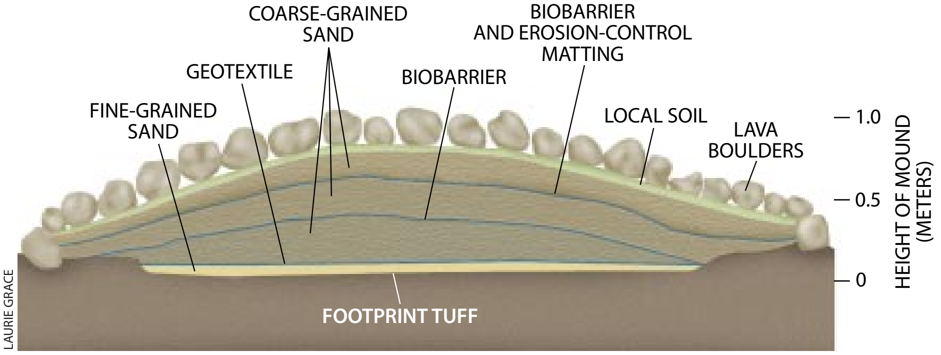
LIFELIKE SCULPTURES of the Laetoli hominids bear worried expressions (*right*). The diorama's background shows the stark landscape (*below*).

IAN TATTERSALL is a curator in the department of anthropology at the American Museum of Natural History in New York City.

PHOTOGRAPHS COURTESY OF AMERICAN MUSEUM OF NATURAL HISTORY

photography, mapping and detailed condition survey have added an enormous archive of data to the base record compiled during the Leakey field seasons. This material is being integrated into an electronic database developed in collaboration with the department of geomatics at the University of Cape Town.

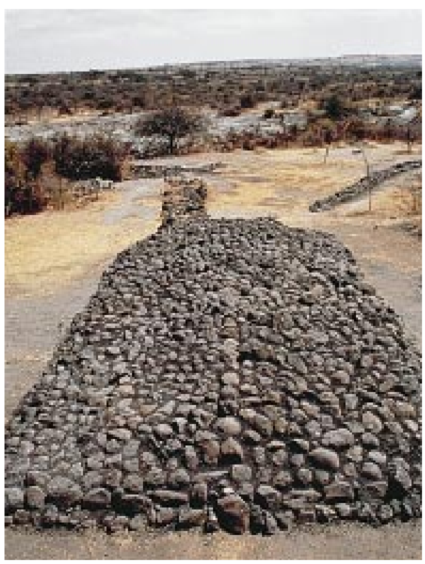
When conservation and documentation were complete, the trackway was reburied under multiple layers of sand and soil from the surrounding area and from the nearby Ngarusi and Kakesio rivers. The fill was sieved to remove coarse material and acacia seeds. The conservation team poured fine-grained sand on the footprint surface, then placed sheets of geotextile—a water-permeable polypropylene material—about five centimeters above the surface to serve as a marker. Then the team mem-



REBURIAL MOUND over the hominid trackway includes five layers of sand and soil (*diagram*). The conservation team poured fine-grained sand directly on the Footprint Tuff (*top photograph*). The reburial layers are separated by polypropylene geotextiles and erosion-control matting (*middle*). The mound is capped with lava boulders to protect the trackway from cattle and other animals (*bottom*).

T. MOON *J. Paul Getty Trust*

N. AGNEW *J. Paul Getty Trust*



M. DEMAS *J. Paul Getty Trust*

bers poured a layer of coarse-grained sand and covered it with a special kind of geotextile called Biobarrier, which is designed to block root intrusion into the burial fill.

Biobarrier is studded with nodules that slowly release the root inhibitor trifluralin, a low-toxicity, biodegradable herbicide. Trifluralin is not soluble in water, so it is nonleaching and nonmigrating: it inhibits root growth but does not kill the plants whose roots contact the nodules. The effective life of Biobarrier depends on the temperature of the soil and the depth of burial. Based on the manufacturer's data, the material will have an effective life of about 20 years at the Laetoli site. Above the Biobarrier, the conservators added another layer of coarse-grained sand, then laid down a second covering of Biobarrier and a synthetic erosion-control matting.

The conservation team topped the mound with a layer of local soil and a bed of lava boulders to provide a physical armor for the reburial fill. The mound, which is one meter high at its apex, will be allowed to revegetate with grasses; because they are shallow-rooted, they will stabilize the reburial soil without posing any danger to the trackway surface. But the staff of the Antiquities Department will regularly monitor the site and remove any tree seedlings that take root. The geotextiles are a second line of defense should the maintenance lapse. The shape of the mound, which has a slope of about 14 degrees on each side, will facilitate the runoff of surface water.

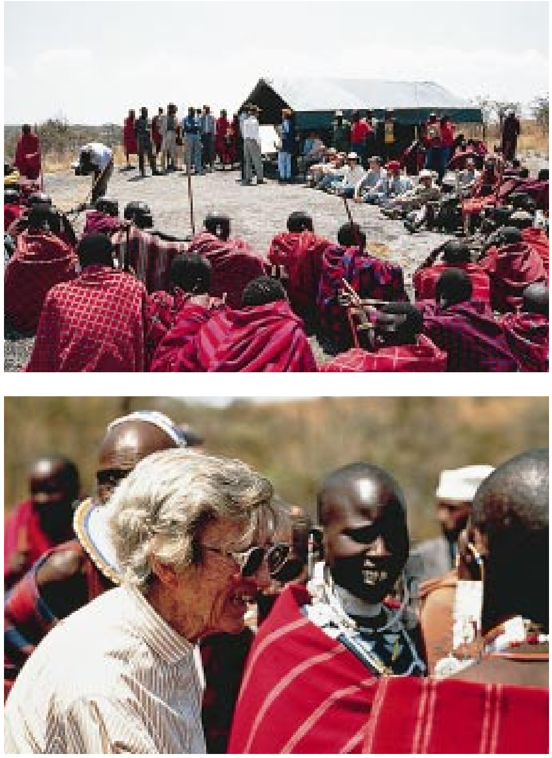
The entire process was repeated for the northern section of the hominid trackway during the 1996 field season. This section had suffered the most erosion because surface water from the surrounding area drains into the Ngarusi River across the northern end of the trackway. It was this drainage that exposed the first hominid footprint found by Abell in 1978; unfortunately, the same drainage resulted in the loss of this print and an adjacent one in the 18 years between the burial of the trackway and its reexcavation. To prevent further erosion, simple berms were constructed from lava boulders around the trackway to divert runoff from nearby areas. Two gullies that were threatening the northern end of the trackway were also stabilized by placing lava boulders and erosion-control matting on their slopes.

Near the trackway, the team members dug a monitoring trench, 2.5 meters square, which was reburied according to the same method used on the trackway. Parts of this trench will be periodically reexcavated to assess the subterranean conditions and the continued effectiveness of the Biobarrier. Acacia trees have been permitted to survive around the monitoring trench to see how well the Biobarrier can block the tree roots. Although polypropylene materials may be expected to last for many years underground, their use in tropical environments such as Laetoli where large numbers of termites live has not been properly evaluated. The monitoring trench will allow the Antiquities Department staff to check the performance of the geotextiles without disturbing the trackway itself.

A Sacred Ceremony

Experience has shown that successful preservation of remote sites requires the cooperation of local people. If they feel excluded, there are frequently adverse results, from neglect to deliberate harm. Most of the people in the Laetoli area are Masai. They have maintained to a large degree their traditional way of life, which centers on their herds of cattle. Cattle grazing on and around the trackway site would cause erosion of the reburial mound and the destruction of the system of berms and drains for diverting the surface runoff. While tending the cattle, herders with time on their hands might also be tempted to interfere with the reburial mound. Everyone in the region knows of the intensive activity at the site in recent years, and some local people have been curious about the Biobarrier and other materials used in the reburial.

Laetoli lies within the Ngorongoro Conservation Area, a vast tract set aside by the government to preserve both the natural environment and the Masai community's way of life. This extraordinary undertaking, perhaps unique in Africa, has a good chance of succeeding under capable management. We frequently consulted the conservation area's regional coordinator—who became a member of the advisory committee for the Laetoli project—and the chairmen of the two closest villages, Endulen and Esere. On their recommendation, a meeting at the site was called by the *Loboini* of the region, the traditional religious leader and healer.



PHOTOGRAPHS BY J. C. LEWIS *J. Paul Getty Trust*

CEREMONIAL BLESSING OF THE TRACKWAY took place in August 1996, when men and women from the Masai community gathered at the Laetoli site (*top*). Leakey attended this event and reacquainted herself with the local people (*bottom*). The great archaeologist died just four months later.

In a daylong meeting attended by about 100 people, including men and women of all ages, the *Loboini* emphasized the significance of the trackway and explained the need for its protection. A sheep was sacrificed and a sacred ceremony held to include the site among the places revered by the Masai people. In 1996, after the northern section of the trackway had been reexcavated, the ceremony was repeated. Leakey herself attended this meeting and was greeted by some of the older people who recalled her work in the Laetoli area in the 1970s.

Ultimately, the survival of the site will depend on the vigilance of the Tanzanian authorities and the international community. The Antiquities Department has appointed two Masai men from the area as full-time guards and instituted a detailed monitoring and maintenance plan. The plan calls for regular photography from specified perspectives around the site, periodic removal of all seedlings—especially acacias—and repair to the berms and drainage system.

Because the Laetoli site is not open to visitors, we have installed a permanent display at the Olduvai Museum, which overlooks the gorge where Leakey and her husband, Louis S. B. Leakey, made so many of their famous discoveries. The museum is a short distance off the dirt road that runs from the Ngorongoro caldera to the Serengeti Plain; it is accessible to both local people and international visitors. The room devoted to Laetoli contains the cast of the southern section of the trackway, along with text and photographs that explain why the site was reburied and how it is being protected. In the past, the Olduvai Museum primarily served international tourists en route to the Serengeti Plain. But the text of the Laetoli exhibit is in Swahili as well as English, so it is hoped the local people—particularly Tanzanian schoolchildren—will come to the museum to learn more about the Laetoli footprints and will be inspired to care for the site.

Footprints are evocative. When astronaut Neil Armstrong trod on the surface of the moon, images of his footprints were instantly recognized as symbols of humankind's first steps into the cosmos. Between the Laetoli footprints and those on the moon lies a 3.6-million-year-long evolutionary journey. Looking at the myriad animal tracks at Laetoli, one has the sense that hominids were not frequently encountered on that landscape—their tracks are too few in number compared with those of the other fauna. These creatures must have belonged to an insignificant species that somehow escaped the inevitable extinctions in the harsh environment. The wistful trail of three small figures carefully making their way across the recently fallen ash from Sadiman is both humbling and stirring. These fragile traces of humankind's beginnings on the plains of Africa deserve to be given every care and protection for their future survival. SA

The Authors

NEVILLE AGNEW and MARTHA DEMAS led the Getty Conservation Institute's project at Laetoli in Tanzania. Agnew received his Ph.D. in chemistry from the University of Natal in Durban, South Africa. He headed the conservation section of the Queensland Museum in Brisbane, Australia, before joining the Getty institute in 1988. He has undertaken conservation projects in China, Ecuador and the U.S. and is now the institute's group director for information and communications. Demas earned a doctorate in Aegean archaeology from the University of Cincinnati and a master's in historic preservation from Cornell University. She joined the Getty in 1990 and is currently involved in developing and managing conservation projects in the Mediterranean region and China.

Further Reading

The Fossil Footprints of Laetoli. Richard L. Hay and Mary D. Leakey in *Scientific American,* Vol. 246, No. 2, pages 50–57; February 1982.

Disclosing the Past. Mary D. Leakey. Doubleday, 1984.

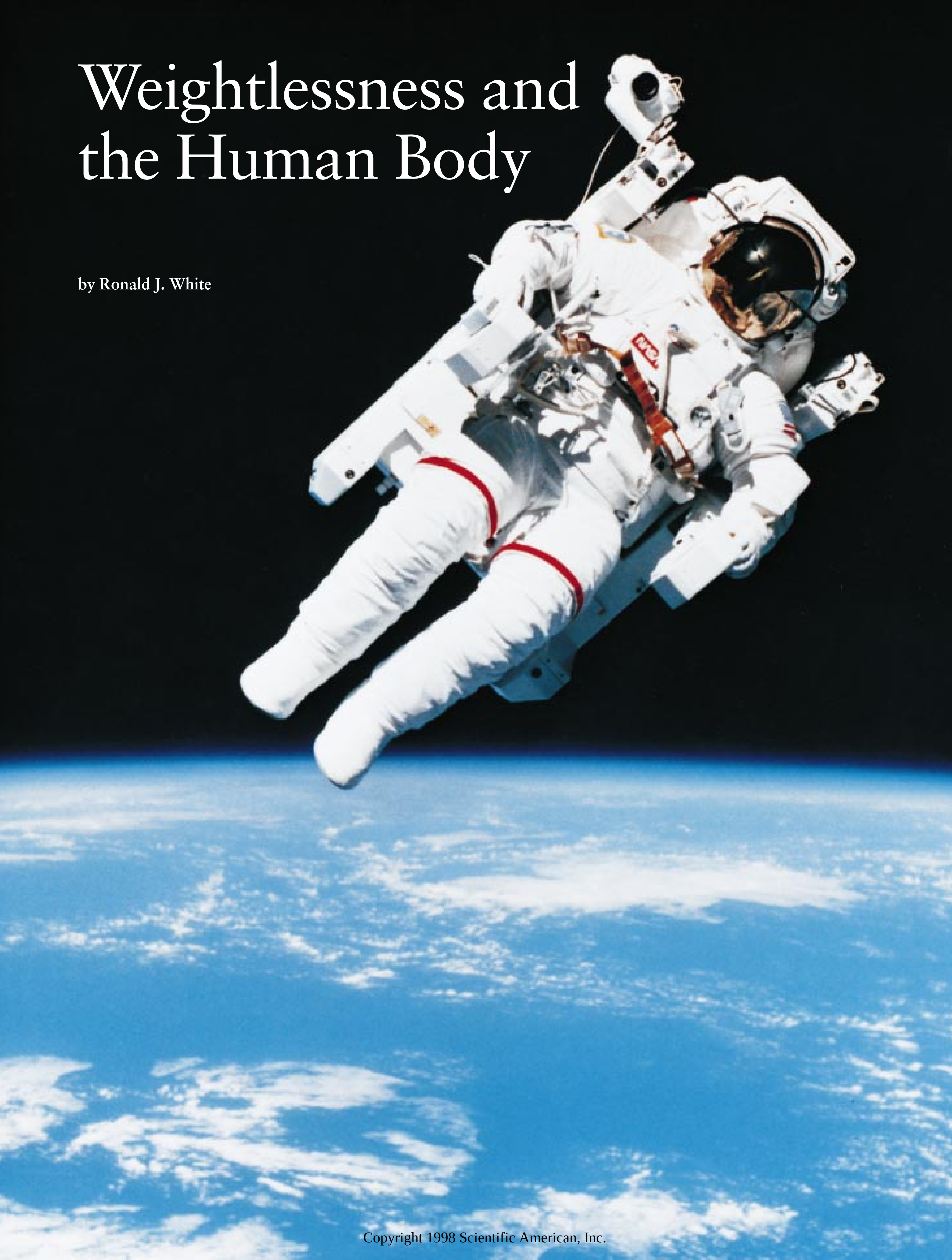
Hominid Footprints at Laetoli: Facts and Interpretations. Tim D. White and Gen Suwa in *American Journal of Physical Anthropology,* Vol. 72, No. 4, pages 485–514; 1987.

Laetoli: A Pliocene Site in Northern Tanzania. Edited by M. D. Leakey and J. M. Harris. Clarendon Press, 1987.

Missing Links: The Hunt for Earliest Man. John Reader. Penguin Books, 1988.

Weightlessness and the Human Body

by Ronald J. White



The effects of space travel on the body resemble some of the conditions of aging. Studying astronauts' health may improve medical care both in orbit and on the ground

When a healthy Valeri Polyakov climbed out of his *Soyuz* capsule on March 22, 1995, after a world-record 438 days on the Mir space station, he had demonstrated that humans can live and work in space for months at a time. It was not always clear that this would be the case.

In 1951, more than 10 years before Yuri Gagarin's first short flight (108 minutes), SCIENTIFIC AMERICAN published an article by Heinz Haber of the U.S. Air Force School of Aviation Medicine that anticipated many of the medical effects of space travel and, in particular, of weightlessness [see "The Human Body in Space," January 1951]. Some of his predictions, such as the occurrence of space motion sickness at the beginning of a flight, have been borne out. Others, such as the notion that space travelers would feel as if they were being jerked back and forth or that they would suddenly start to spin around during normal motion in space, have not.

As most doctors can attest, it is difficult to predict what will happen when a brand-new challenge is presented to the human body. Time and again, space travel has revealed its marvelous and sometimes subtle adaptive ability. But only in the past few years have scientists begun to understand the body's responses to weightlessness, as the data—the cumulative experience of nearly 700 people spending a total of 58 person-years in space—have grown in quantity and quality. Pursuit of this knowledge is improving health care not only for those who journey into space but also for those of us stuck on the ground. The unexpected outcome of space medicine has been an enhanced understanding of how the human body works right here on Earth.

Feeling Gravity's Pull

Although many factors affect human health during spaceflight, weightlessness is the dominant and single most important one. The direct and indirect effects of weightlessness precipitate a cascade of interrelated responses that begin in three different types of tissue: gravity receptors, fluids and weight-bearing structures. Ultimately, the whole body, from bones to brain, reacts.

When space travelers grasp the wall of their spacecraft and pull and push their bodies back and forth, they say it feels as though they are stationary and the spacecraft is moving. The reason is embedded in our dependence on gravity for perceptual information.

The continuous and pervasive nature of gravity removes it from our daily consciousness. But even though we are only reminded of gravity's invisible hand from time to time by, say, varicose veins or an occasional lightheadedness on standing up, our bodies never forget. Whether we realize it or not, we have evolved a large number of silent, automatic reactions to cope with the constant stress of living in a downward-pulling world. Only when we decrease or increase the effective force of gravity on our bodies do we consciously perceive it. Otherwise our perception is indirect.

Our senses provide accurate information about the location of our center of mass and the relative positions of our body parts. This capability integrates signals from our eyes and ears with other information from the vestibular organs in our inner ear, from our muscles and joints, and from our senses of touch and pressure. Many of these signals are dependent on the size and direction of the constant terrestrial gravitational force.

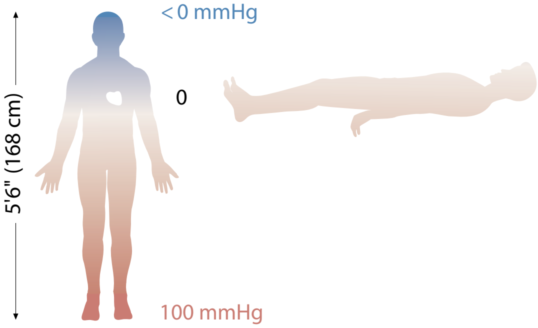
The vestibular apparatus in the inner ear has two distinct components: the semicircular canals (three mutually perpendicular, fluid-filled tubes that contain hair cells connected to nerve fibers), which are sensitive to angular acceleration of the head; and the otolith organs (two sacs filled with calcium carbonate crystals embedded in a gel), which respond to linear acceleration. Because movement of the crystals in the otoliths generates the signal of acceleration to the brain and because the laws of physics relate that acceleration to a net force, gravity is always implicit in the signal. Thus, the otoliths have been referred to as gravity receptors. They are not the only ones. Mechanical receptors in the muscles, tendons and joints—as well as pressure receptors in the skin, particularly on the bottom of the feet—respond to the weight of limb segments and other body parts.

Removing gravity transforms these signals. The otoliths no longer perceive a downward bias to head movements. The limbs no longer have weight, so muscles are no longer required to contract and relax in the usual way to maintain posture and bring about movement. Touch and pressure receptors in the feet and ankles no longer signal the direction of down. These and other changes contribute to visual-orientation illusions and feelings of self-inversion, such as the feeling that the body or the spacecraft spontaneously reorients. In 1961 cosmonaut Gherman Titov reported vivid sensations of being upside down early in a spaceflight of only one day. Last year shuttle payload specialist Byron K. Lichtenberg, commenting on his earlier flight experiences, said, "When the main engines cut off, I immediately felt as though we had flipped 180 degrees." Such illusions can recur even after some time in space.

The lack of other critical sensory cues also confuses the brain. Although orbital flight is a perpetual free fall—the only difference from skydiving is that the spacecraft's forward velocity carries it around the curve of the planet—space travelers say they do not feel as if they are falling. The perception of falling probably depends on visual and airflow cues along with information from the direct gravity receptors. This contradicts a prediction made in 1950 by Haber

NASA JOHNSON SPACE CENTER

FREEDOM FROM WEIGHT, alluring though it is, has unfortunate side effects, such as motion sickness, head congestion and bone loss. On a space walk in 1984, Bruce McCandless II tried out NASA's new rocket pack—becoming the first astronaut to venture outside a spacecraft without a tether.



DMITRY KRASNY

HYDROSTATIC PRESSURE in blood vessels changes dramatically when a person stands up. Pressure increases with depth below the heart, up to 100 millimeters of mercury (mmHg) in a person of average height; above the heart, pressure decreases. As a result, fluid settles into the lower body and blood flow diminishes. Conversely, in a prone position (or in weightlessness), pressure equalizes and fluid sloshes into the upper body.

and his colleague Otto H. Gauer: "In the absence of gravity there would necessarily be a sensation of falling in free space. It is expected that one would gradually get accustomed to this state."

The aggregate of signal changes produces, in half or more of space travelers, a motion sickness that features many of the symptoms of terrestrial motion sickness: headache, impaired concentration, loss of appetite, stomach awareness, vomiting. Space motion sickness usually does not last beyond the first three days or so of weightlessness, but something similar has been reported by cosmonauts at the end of long flights.

At one time, scientists attributed space motion sickness to the unusual pattern of vestibular activity, which conflicts with the brain's expectations. Now it is clear that this explanation was simplistic. The sickness results from the convergence of a variety of factors, including the alteration of the patterns and levels of motor activity necessary to control the head itself. A similar motion sickness can also be elicited by computer systems designed to create virtual environments, through which one can navigate without the forces and sensory patterns present during real motion [see News and Analysis, "Virtual Reality Check," by W. Wayt Gibbs; Scientific American, December 1994].

Over time, the brain adapts to the new signals, and for some space travelers, "down" becomes simply where the feet are. The adaptation probably involves physiological changes in both receptors and nerve-cell patterns. Similar changes occur on the ground during our growth and maturation and during periods of major body-weight changes.

The way we control our balance and avoid falls is an important and poorly understood part of physiology. Because otherwise healthy people returning from space initially have difficulty maintaining their balance but recover this sense rapidly, postflight studies may allow doctors to help those non–space travelers who suffer a loss of balance on Earth.

Bernard Cohen of the Mount Sinai School of Medicine and Gilles Clément of the National Center for Scientific Research in Paris undertook just such a study after the Neurolab shuttle mission, which ended on May 3. To connect this work with patients suffering from balance disorders, Barry W. Peterson of Northwestern University and a team of researchers, supported by the National Aeronautics and Space Administration and the National Institutes of Health, are creating the first whole-body computer model of human posture and balance control.

Space Sniffles

The second set of weightlessness effects involves body fluids. Within minutes of arriving in a weightless environment, a traveler's neck veins begin to bulge, and the face begins to fill out and become puffy. As fluid migrates to the chest and head, sinus and nasal congestion results. This stuffiness, which is much like a cold on Earth, continues for the entire flight, except during heavy exercise, when the changing fluid pressures in the body relieve the congestion temporarily. Even the senses of taste and smell are altered; spicy food retains its appeal best. In the early days of spaceflight, doctors feared that the chest congestion might be dangerous, much as pulmonary edema is to cardiac patients; fortunately, this has not been the case.

All these events occur because the fluids in the body no longer have weight. On average, about 60 percent of a person's weight is water, contained in the cells of the body (intracellular fluid), in the arteries and veins (blood plasma) and in the spaces between the blood vessels and cells (interstitial fluid). On Earth, when a person stands up, the weight of this water exerts forces throughout the body. In the vascular system, where the fluid columns are directly connected, blood pressure increases hydrostatically, just as pressure increases with depth in water. For a quietly standing individual, this hydrostatic effect can be quite large. In the feet, both arterial and venous pressures can increase by approximately 100 millimeters of mercury—double the normal arterial pressure and many times the normal venous pressure. At locations above the feet but below the heart, the pressure increases by zero to 100 millimeters of mercury. Above the heart, arterial and venous pressure fall below atmospheric pressure [*see illustration at left*].

The hydrostatic effect has only a small influence on blood flow through tissue because both arterial and venous pressures increase by the same amount. But it does affect the distribution of fluid within the body by increasing the amount of blood that leaks from capillaries into the interstitial space. Going from a prone to a standing position moves fluid into the lower part of the body and reduces the flow of blood back to the heart. If unchecked, quiet standing can lead to fainting; soldiers sometimes swoon when standing at attention. Two other hydrostatic effects are varicose veins, which have become permanently distorted by the extra fluid, and swollen feet after long periods of quiet sitting (such as an airplane flight).

In space, the hydrostatic pressure disappears, causing fluids to redistribute naturally from the lower to the upper body. Direct measurements of leg volumes have shown that each leg loses about one liter of fluid—about a tenth of its volume—within the first day. The legs then stay smaller for the whole time in space. (Actually, fluids begin to shift toward the head while space travelers are still on the launch pad, because they sit for several hours in couches that elevate their feet above their heads.) As fluid moves around, the body adapts by further redistributing water among its compartments. Plasma volume decreases rapidly (by nearly 20 percent) and stays low.

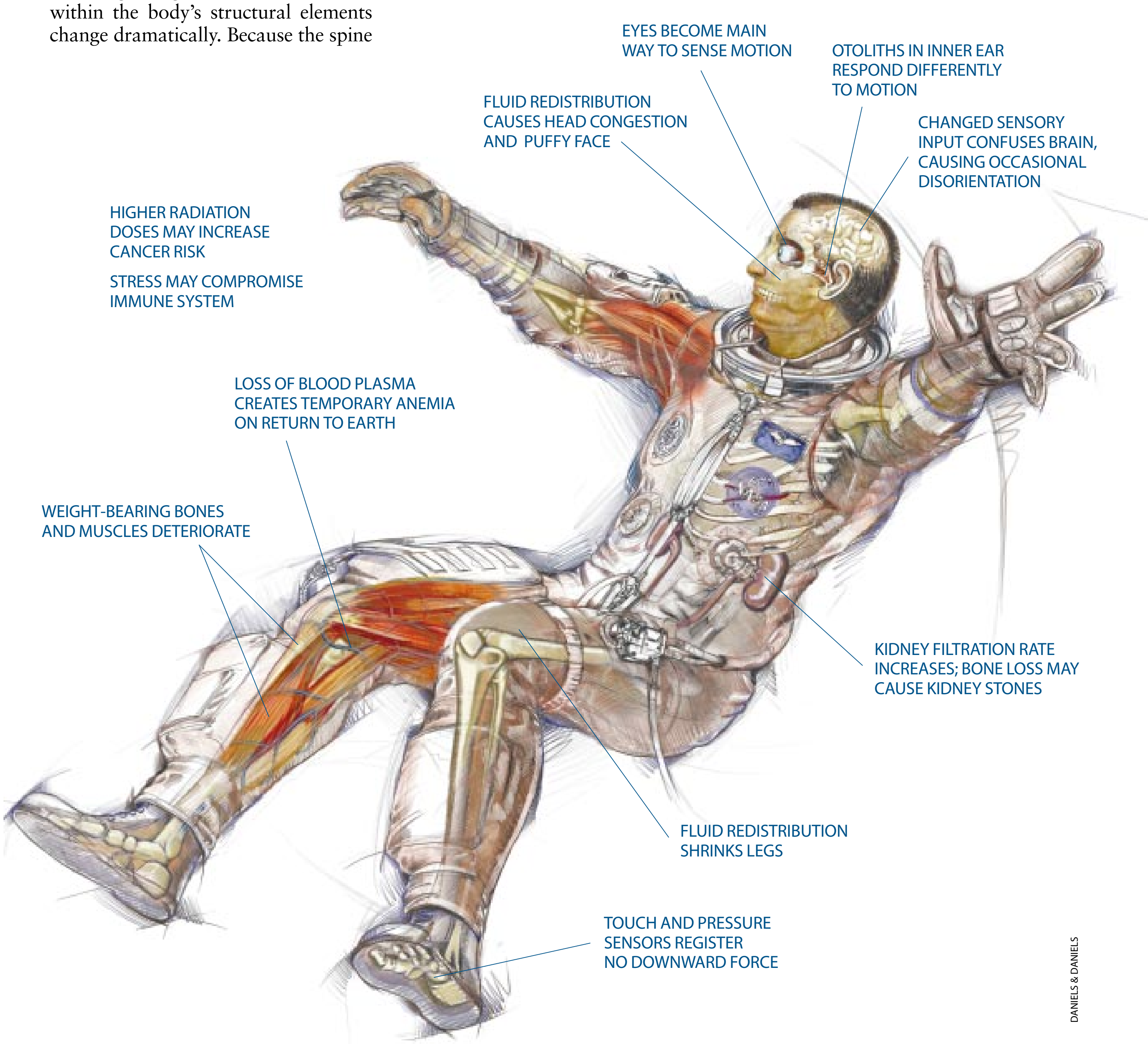
These fluid shifts in turn initiate a cascade of interacting renal, hormonal and mechanical mechanisms that regulate fluid and electrolyte levels. For example, the kidney filtration rate, normally stable, increases by nearly 20 percent and remains at that level for the first week in space. In addition, returning space travelers have a special form of anemia, even after flights as short as a few days. Over the past few years, Clarence Alfrey of the Baylor College of Medicine has shown that the loss of plasma and the concomitant decrease in vascular space lead to an overabundance of red blood cells. The body responds by stopping production of and destroying new red blood cells—using a mechanism that was not fully appreciated by hematologists before Alfrey's research on space travelers.

A third set of effects caused by weightlessness relates to muscle and bone. People who travel in space for any length of time come home with less of both. Is this a cause for concern?

During weightlessness, the forces within the body's structural elements change dramatically. Because the spine is no longer compressed, people grow taller (two inches or so). The lungs, heart and other organs within the chest have no weight, and as a result, the rib cage and chest relax and expand. Similarly, the weights of the liver, spleen, kidneys, stomach and intestines disappear. As F. Andrew Gaffney said after his 1991 flight: "You feel your guts floating up. I found myself tightening my abdomen, sort of pushing things back."

Meanwhile muscles and bones come to be used in different ways. Skeletal muscle, the largest tissue in the body, evolved to support our upright posture and to move body parts. But in space, muscles used for antigravity support on the ground are no longer needed for that purpose; moreover, the muscles used for movement around a capsule differ from those used for walking down a hall. Consequently, some muscles atrophy rapidly. At the same time, the nature of the muscle itself alters, changing from certain slow-twitch fibers that are useful for support against gravity to

The Effects of Space Travel on the Body



DANIELS & DANIELS

LYING IN BED mimics weightlessness in its effect on the human body. At the NASA Ames Research Center, volunteers lie head-down at a six-degree angle—which, over a period of weeks, is not as comfortable as it might look. Fluids drain from the legs to the upper body, muscles atrophy and bones weaken. These subjects try out various restorative exercises, diets and drugs. Seated on the right is astronaut Charles Brady, who carried out medical tests during a 1996 Spacelab mission.

faster contractile fibers, useful for rapid response. None of these changes presents a problem to space travelers as long as they perform only light work. But preventing the atrophy of muscles required for heavy work during space walks and preserving muscle for safe return to Earth are the subject of much current experimentation.

Bone metabolism, too, changes substantially. One of the strongest known biological materials, bone is a dynamic tissue. Certain cells, the osteoblasts, have the job of producing it, whereas others, the osteoclasts, destroy it. Both usually work harmoniously to rebuild bones throughout life. These cellular systems are sensitive to various hormones and vitamins in the blood and to mechanical stress on the bone.

Bone contains both organic materials, which contribute strength and stability, and inorganic materials, which contribute stiffness and serve as a mineral reservoir within the body. For example, 99 percent of the calcium in the body is in the skeleton. Stabilized calcium levels in the body's fluids are necessary for all types of cells to function normally.

Joint Russian-American studies have shown that cosmonauts have lost bone mass from the lower vertebrae, hips and upper femur at a rate of about 1 percent per month for the entire duration of their missions. Some sites, such as the heel, lose calcium faster than others. Studies of animals subjected to spaceflight suggest that bone formation also declines.

Needless to say, these data are indeed cause for concern. During spaceflight, the loss of bone elevates calcium levels in the body, potentially causing kidney stones and calcification in soft tissues. Back on the ground, the bone calcium loss stops within one month, but scientists do not yet know whether the bone recovers completely: too few people have flown in space for long periods. Some bone loss may be irreversible, in which case ex-astronauts will always be more prone to broken bones. A 1996 Spacelab mission was partly devoted to these questions; a team of scientists from Italy, Sweden, Switzerland and the U.S. carried out eight investigations related to muscle and bone changes.

These uncertainties mirror those in our understanding of how the body works here on Earth. For example, after menopause women are prone to a loss of bone mass—osteoporosis. Scientists understand that many different factors (activity, nutrition, vitamins, hormones) can be involved in this loss, but they do not yet know how the factors act and interact. This complexity makes it difficult to develop an appropriate response. So it is with bone loss in space, where the right prescription still awaits discovery. Up to now, various types of exercise have been tried [see "Six Months on Mir," by Shannon W. Lucid; Scientific American, May] with little verified success.

VIGOROUS EXERCISE, typically lasting several hours, is a daily routine for astronauts. Terence T. Henricks works out on the shuttle *Atlantis* in 1991, while Mario Runco, Jr., wired with medical sensors, waits on deck. Although such exercise may slow atrophy of muscles, its effectiveness is still not clear.

Heavy Breathing

Disorientation, fluid redistribution, and muscle and bone loss are not the only consequence of weightlessness. Other body systems are affected directly and indirectly. One example is the lung.

John B. West and his group at the University of California at San Diego, along with Manuel Paiva of the Free University of Brussels, have studied the lung in space and learned much they could not have learned in earthbound laboratories. On the ground the top and bottom parts of the lung have different patterns of airflow and blood flow. But are these patterns the result only of gravity or also of the nature of the lung itself? Only recently have studies in space provided unequivocal evidence for the latter. Even in the absence of gravity, different parts of the lung have different levels of airflow and blood flow.

Not everything that affects the body during spaceflight is related solely to weightlessness. Also affected, for example, are the immune system (the various physical and psychological stresses of spaceflight probably play roles in the immunodeficiency experienced by astronauts) and the multiple systems responsible for the amount and quality of sleep (lighting levels and work schedules disrupt the body's normal rhythms). Looking out the spacecraft window just before retiring (an action difficult to resist, considering the view) can let enough bright light into the eye to trigger just the wrong physiological response, leading to poor sleep. As time goes on, the sleep debt accumulates.

For long space voyages, travelers must also face confinement in a tight volume, unable to escape, isolated from the normal life of Earth, living with a small, fixed group of companions who often come from different cultures. These challenges can lead to anxiety, insomnia, depression, crew tension and other interpersonal issues, which affect astronauts just as much as weightlessness—perhaps even more. Because these factors operate at the same time the body is adapting to other environmental changes, it may not be clear which physiological changes result from which factors. Much work remains to be done.

Finally, spaceflight involves high levels of radiation. An astronaut spending one year in a moderately inclined, low-Earth orbit would receive a radiation dose 10 times greater than the average dose received on the ground. A year's stay on the moon would result in a dose seven times higher still, whereas a flight to Mars would be even worse. Sudden outflows of particles from the sun, of the type that occurred in August 1972, can deliver a dose more than 1,000 times the annual ground dose in less than a day. Fortunately, such events are rare, and spacecraft designers can guard against them by providing special shielded rooms to which astronauts can temporarily retreat.

Obviously, the radiation hazard to long-duration space travelers—and the consequent cancer risk—is disconcerting. The problems of space radiation are difficult to study because it is nearly impossible to duplicate on Earth the radiation environment of space, with its low but steady flux of high-energy cosmic rays. Even so, researchers generally believe that with proper shielding and protective drugs, the risks can be brought within acceptable limits.

Down to Earth

When space travelers return to the world of weight, complementary changes occur. If the effects of weightlessness are completely reversible, everything should return to its normal condition at some time after the flight. We now know that most systems in the body do work reversibly, at least over the intervals for which we have data. We do not yet know whether this is a general rule.

Space travelers certainly feel gravitationally challenged during and just after their descent. As one person said after nine days in space: "It's quite a shock. The first time I pushed myself up, I felt like I was lifting three times my weight." Returning space travelers report experiencing a variety of illusions—for example, during head motion it is their surroundings that seem to be moving—and they wobble while trying to stand straight, whether their eyes are open or closed.

Most of the body's systems return to normal within a few days or weeks of landing, with the possible exception of the musculoskeletal system. So far nothing indicates that humans cannot live and work in space for long periods and return to Earth to lead normal lives. This is clearly good news for denizens of the upcoming International Space Station and for any future interplanetary missions. In fact, the station, assembly of which should begin late this year or early next year, will provide researchers with a new opportunity to investigate the effects of space travel on humans. On its completion in five years, the station will have 46,000 cubic feet of work space (nearly five times more than the Mir or Skylab stations) and will include sophisticated laboratory equipment for the next generation of medical studies. Recognizing the need for a comprehensive attack on all the potential human risks of long-duration space travel, NASA has selected and funded a special research body, the National Space Biomedical Research Institute, to assist in defining and responding to those risks.

Many of the "normal" changes that take place in healthy people during or just after spaceflight are outwardly similar to "abnormal" events occurring in ill people on Earth. For example, most space travelers cannot stand quietly for 10 minutes just after landing without feeling faint. This so-called orthostatic intolerance is also experienced by patients who have stayed in bed for a long time and by some elderly people. Prolonged bed rest also results in muscle and bone wasting. The parallel is so close that bed rest is used to simulate the effects of spaceflight [*see top illustration on opposite page*].

Other age-related changes in function also seem to match changes caused by spaceflight. The wobbliness following flight resembles the tendency of the elderly to fall; the loss of bone during flight is analogous to age-related osteoporosis; immunodeficiency, poor sleep quality and loss of motor coordination afflict both astronauts and the elderly. Although parallels in symptoms do not imply parallels in cause, the data are so striking that NASA and the National Institute on Aging began an investigative collaboration in 1989. The flight in October of Senator John Glenn of Ohio, the oldest person ever scheduled to fly in space, should draw more attention to the ongoing research in this area. SA

NASA JOHNSON SPACE CENTER

JOHN GLENN, astronaut turned senator turned astronaut again, emerges from a shuttle-training mock-up in 1989. Glenn is due to fly on board the shuttle *Discovery* in late October. Some of the medical effects of space travel resemble the symptoms of aging, such as poor sleep quality; Glenn, 77, will participate in sleep experiments.

The Author

RONALD J. WHITE is associate director of the National Space Biomedical Research Institute, a consortium of universities led by the Baylor College of Medicine in Houston. He has been a professor at Baylor since 1996, before which he served as the chief scientist of the Life Sciences Division at the headquarters of the National Aeronautics and Space Administration. A specialist in space life sciences and biomedical research, White served as NASA program scientist for Spacelab missions in June 1991 and January 1992. He is the author of various scientific papers and currently serves on the board of trustees of the International Academy of Astronautics.

Further Reading

PROCEEDINGS OF A CONFERENCE ON CORRELATIONS OF AGING AND SPACE EFFECTS ON BIOSYSTEMS. Edited by R. L. Sprott and C. A. Combs in *Experimental Gerontology,* Vol. 26, Nos. 2–3, pages 121–309; 1991.

ORIENTATION AND MOVEMENT IN UNUSUAL FORCE ENVIRONMENTS. James R. Lackner in *Psychological Science,* Vol. 4, No. 3, pages 134–142; May 1993.

SPACE PHYSIOLOGY AND MEDICINE. Third edition. A. E. Nicogossian, C. L. Huntoon and S. L. Pool. Lea & Febiger, 1993.

APPLIED PHYSIOLOGY IN SPACE. Special issue of the *Journal of Applied Physiology,* Vol. 81, No. 1, pages 3–207; July 1996.

Attention-Deficit Hyperactivity Disorder

A new theory suggests the disorder results from a failure in self-control. ADHD may arise when key brain circuits do not develop properly, perhaps because of an altered gene or genes

by Russell A. Barkley

YAN NASCIMBENE

As I watched five-year-old Keith in the waiting room of my office, I could see why his parents said he was having such a tough time in kindergarten. He hopped from chair to chair, swinging his arms and legs restlessly, and then began to fiddle with the light switches, turning the lights on and off again to everyone's annoyance—all the while talking nonstop. When his mother encouraged him to join a group of other children busy in the playroom, Keith butted into a game that was already in progress and took over, causing the other children to complain of his bossiness and drift away to other activities. Even when Keith had the toys to himself, he fidgeted aimlessly with them and seemed unable to entertain himself quietly. Once I examined him more fully, my initial suspicions were confirmed: Keith had attention-deficit hyperactivity disorder (ADHD).

Since the 1940s, psychiatrists have applied various labels to children who are hyperactive and inordinately inattentive and impulsive. Such youngsters have been considered to have "minimal brain dysfunction," "brain-injured child syndrome," "hyperkinetic reaction of childhood," "hyperactive child syndrome" and, most recently, "attention-deficit disorder." The frequent name changes reflect how uncertain researchers have been about the underlying causes of, and even the precise diagnostic criteria for, the disorder.

Within the past several years, however, those of us who study ADHD have begun to clarify its symptoms and causes and have found that it may have a genetic underpinning. Today's view of the basis of the condition is strikingly different from that of just a few years ago. We are finding that ADHD is not a disorder of attention per se, as had long been assumed. Rather it arises as a developmental failure in the brain circuitry that underlies inhibition and self-control. This loss of self-control in turn impairs other important brain functions crucial for maintaining attention, including the ability to defer immediate rewards for later, greater gain.

ADHD involves two sets of symptoms: inattention and a combination of hyperactive and impulsive behaviors [*see table on next page*]. Most children are more active, distractible and impulsive than adults. And they are more inconsistent, affected by momentary events and dominated by objects in their immediate environment. The younger the children, the less able they are to be aware of time or to give priority to future events over more immediate wants. Such behaviors are signs of a problem, however, when children display them significantly more than their peers do.

Boys are at least three times as likely as girls to develop the disorder; indeed, some studies have found that boys with ADHD outnumber girls with the condition by nine to one, possibly because boys are genetically more prone to disorders of the nervous system. The behavior patterns that typify ADHD usually arise between the ages of three and five. Even so, the age of onset can vary widely: some children do not develop symptoms until late childhood or even early adolescence. Why their symptoms are delayed remains unclear.

Huge numbers of people are affected. Many studies estimate that between 2 and 9.5 percent of all school-age children worldwide have ADHD; researchers have identified it in every nation and culture they have studied. What is more, the condition, which was once thought to ease with age, can persist into adulthood. For example, roughly two thirds of 158 children with ADHD my colleagues and I evaluated in the 1970s still had the disorder in their twenties. And many of those who no longer fit the clinical description of ADHD were still having significant adjustment problems at work, in school or in other social settings.

To help children (and adults) with ADHD, psychiatrists and psychologists must better understand the causes of the disorder. Because researchers have traditionally viewed ADHD as a problem in the realm of attention, some have suggested that it stems from an inability of the brain to filter competing sensory inputs, such as sights and sounds. But recently scientists led by Joseph A. Sergeant of the University of Amsterdam have shown that children with ADHD do not have difficulty in that area; instead they cannot inhibit their impulsive motor responses to such input. Other researchers have found that children with ADHD are less capable of preparing motor responses in anticipation of events and are insensitive to feedback about errors made in those responses. For example, in a commonly used test of reaction time, children with ADHD are less able than other children to ready themselves to press one of several keys when they see a warning light. They also do not slow down after making mistakes in such tests in order to improve their accuracy.

The Search for a Cause

No one knows the direct and immediate causes of the difficulties experienced by children with ADHD, although advances in neurological imaging techniques and genetics promise to clarify this issue over the next five years. Already they have yielded clues, albeit ones that do not yet fit together into a coherent picture.

Imaging studies over the past decade have indicated which brain regions might malfunction in patients with ADHD and thus account for the symptoms of the condition. That work suggests the involvement of the prefrontal cortex, part of the cerebellum, and at least two of the clusters of nerve cells deep in the brain that are collectively known as the basal ganglia [*see illustration on page 69*]. In a 1996 study F. Xavier Castellanos, Judith L. Rapoport and their colleagues at the National Institute of Mental Health found that the right prefrontal cortex and two basal ganglia called the caudate nucleus and the globus pallidus are significantly smaller than normal in children with ADHD. Earlier this year Castellanos's group found that the vermis region of the cerebellum is also smaller in ADHD children.

The imaging findings make sense because the brain areas that are reduced in size in children with ADHD are the very ones that regulate attention. The right prefrontal cortex, for example, is involved in "editing" one's behavior, resisting distractions and developing an awareness of self and time. The caudate nucleus and the globus pallidus help to switch off automatic responses to allow more careful deliberation by the cortex and to coordinate neurological input among various regions of the cortex. The exact role of the cerebellar vermis is unclear, but early studies suggest it may play a role in regulating motivation.

What causes these structures to shrink in the brains of those with ADHD? No one knows, but many studies have sug-

CHILDREN WITH ADHD cannot control their responses to their environment. This lack of control makes them hyperactive, inattentive and impulsive.

Diagnosing ADHD

Psychiatrists diagnose attention-deficit hyperactivity disorder (ADHD) if the individual displays six or more of the following symptoms of inattention or six or more symptoms of hyperactivity and impulsivity. The signs must occur often and be present for at least six months to a degree that is maladaptive and inconsistent with the person's developmental level. In addition, some of the symptoms must have caused impairment before the age of seven and must now be causing impairment in two or more settings. Some must also be leading to significant impairment in social, academic or occupational functioning; none should occur exclusively as part of another disorder. (Adapted with permission from the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders.* ©1994 American Psychiatric Association.)

INATTENTION

- Fails to give close attention to details or makes careless mistakes in schoolwork, work or other activities
- Has difficulty sustaining attention in tasks or play activities
- Does not seem to listen when spoken to directly
- Does not follow through on instructions and fails to finish schoolwork, chores or duties in the workplace
- Has difficulty organizing tasks and activities
- Avoids, dislikes or is reluctant to engage in tasks that require sustained mental effort (such as schoolwork)
- Loses things necessary for tasks or activities (such as toys, school assignments, pencils, books or tools)
- Is easily distracted by extraneous stimuli
- Is forgetful in daily activities

HYPERACTIVITY AND IMPULSIVITY

- Fidgets with hands or feet or squirms in seat
- Leaves seat in classroom or in other situations in which remaining seated is expected
- Runs about or climbs excessively in situations in which it is inappropriate (in adolescents or adults, subjective feelings of restlessness)
- Has difficulty playing or engaging in leisure activities quietly
- Is "on the go" or acts as if "driven by a motor"
- Talks excessively
- Blurts out answers before questions have been completed
- Has difficulty awaiting turns
- Interrupts or intrudes on others

LISA BURNETT

gested that mutations in several genes that are normally very active in the prefrontal cortex and basal ganglia might play a role. Most researchers now believe that ADHD is a polygenic disorder—that is, that more than one gene contributes to it.

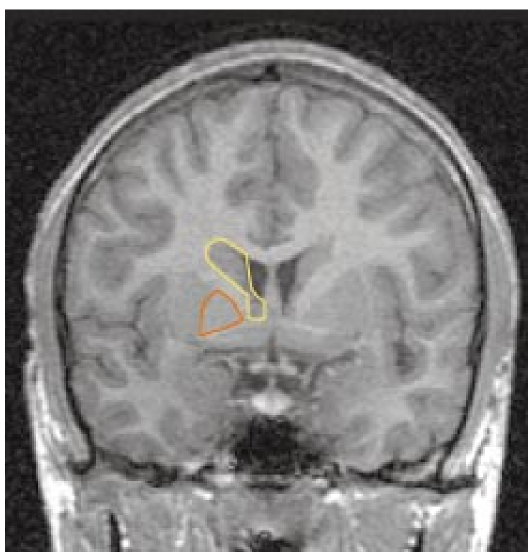
Early tips that faulty genetics underlie ADHD came from studies of the relatives of children with the disorder. For instance, the siblings of children with ADHD are between five and seven times more likely to develop the syndrome than children from unaffected families. And the children of a parent who has ADHD have up to a 50 percent chance of experiencing the same difficulties.

The most conclusive evidence that genetics can contribute to ADHD, however, comes from studies of twins. Jacquelyn J. Gillis, then at the University of Colorado, and her colleagues reported in 1992 that the ADHD risk of a child whose identical twin has the disorder is between 11 and 18 times greater than that of a nontwin sibling of a child with ADHD; between 55 and 92 percent of the identical twins of children with ADHD eventually develop the condition.

One of the largest twin studies of ADHD was conducted by Helene Gjone and Jon M. Sundet of the University of Oslo with Jim Stevenson of the University of Southampton in England. It involved 526 identical twins, who inherit exactly the same genes, and 389 fraternal twins, who are no more alike genetically than siblings born years apart. The team found that ADHD has a heritability approaching 80 percent, meaning that up to 80 percent of the differences in attention, hyperactivity and impulsivity between people with ADHD and those without the disorder can be explained by genetic factors.

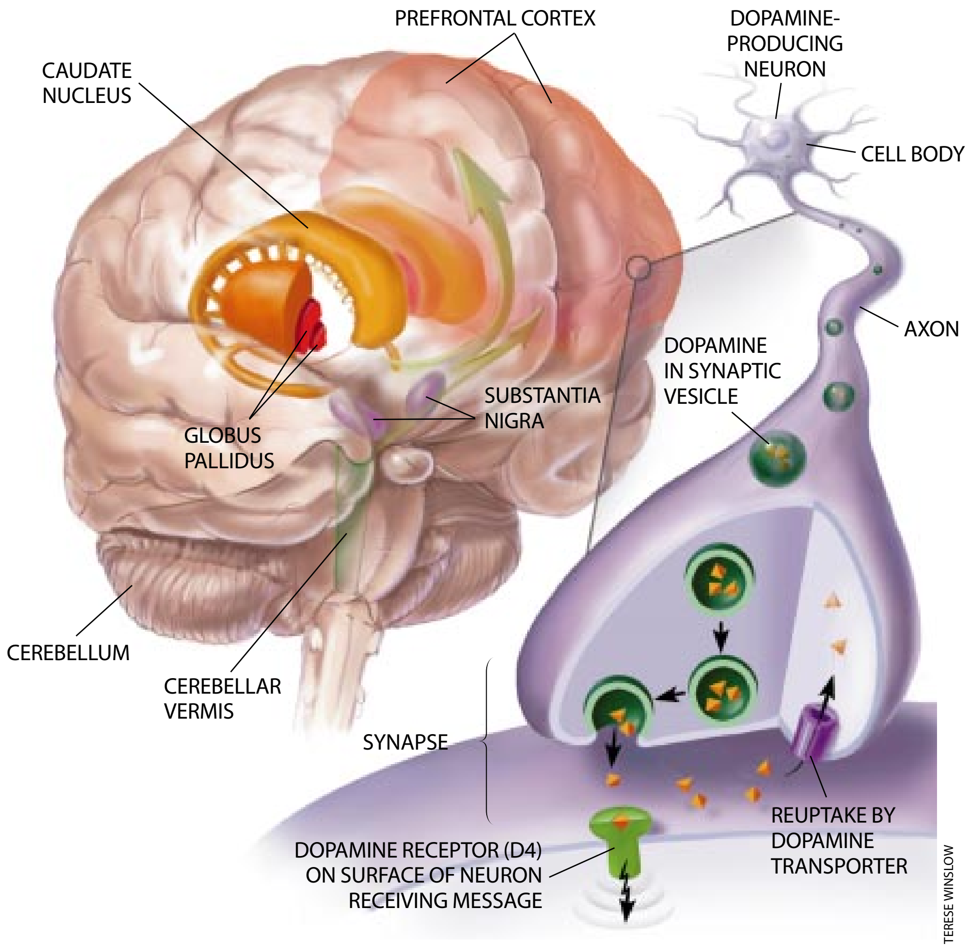
Nongenetic factors that have been linked to ADHD include premature birth, maternal alcohol and tobacco use, exposure to high levels of lead in early childhood and brain injuries—especially those that involve the prefrontal cortex. But even together, these factors can account for only between 20 and 30 percent of ADHD cases among boys; among girls, they account for an even smaller percentage. (Contrary to popular belief, neither dietary factors, such as the amount of sugar a child consumes, nor poor child-rearing methods have been consistently shown to contribute to ADHD.)

Which genes are defective? Perhaps those that dictate the way in which the brain uses dopamine, one of the chemicals known as neurotransmitters that convey messages from one nerve cell, or neuron, to another. Dopamine is secreted by neurons in specific parts of the brain to inhibit or modulate the activity of other neurons, particularly those involved in emotion and movement. The movement disorders of Parkinson's disease, for example, are caused by the death of dopamine-secreting neurons in



NORMAL BRAIN image outlines the right caudate nucleus (*yellow*) and the globus pallidus (*red*), brain structures that regulate attention and that are reduced in size in children with ADHD.

F. XAVIER CASTELLANOS AND JAY N. GIEDD *National Institute of Mental Health*



BRAIN STRUCTURES affected in ADHD use dopamine to communicate with one another (*green arrows*). Genetic studies suggest that people with ADHD might have alterations in genes encoding either the D4 dopamine receptor, which receives incoming signals, or the dopamine transporter, which scavenges released dopamine for reuse. The substantia nigra, where the death of dopamine-producing neurons causes Parkinson's disease, is not affected in ADHD.

a region of the brain underneath the basal ganglia called the substantia nigra.

Some impressive studies specifically implicate genes that encode, or serve as the blueprint for, dopamine receptors and transporters; these genes are very active in the prefrontal cortex and basal ganglia. Dopamine receptors sit on the surface of certain neurons. Dopamine delivers its message to those neurons by binding to the receptors. Dopamine transporters protrude from neurons that secrete the neurotransmitter; they take up unused dopamine so that it can be used again. Mutations in the dopamine receptor gene can render receptors less sensitive to dopamine. Conversely, mutations in the dopamine transporter gene can yield overly effective transporters that scavenge secreted dopamine before it has a chance to bind to dopamine receptors on a neighboring neuron.

In 1995 Edwin H. Cook and his colleagues at the University of Chicago reported that children with ADHD were more likely than others to have a particular variation in the dopamine transporter gene *DAT1*. Similarly, in 1996 Gerald J. LaHoste of the University of California at Irvine and his co-workers found that a variant of the dopamine receptor gene *D4* is more common among children with ADHD. But each of these studies involved 40 or 50 children—a relatively small number—so their findings are now being confirmed in larger studies.

From Genes to Behavior

How do the brain-structure and genetic defects observed in children with ADHD lead to the characteristic behaviors of the disorder? Ultimately, they might be found to underlie impaired behavioral inhibition and self-control, which I have concluded are the central deficits in ADHD.

Self-control—or the capacity to inhibit or delay one's initial motor (and perhaps emotional) responses to an event—is a critical foundation for the performance of any task. As most children grow up, they gain the ability to engage in mental activities, known as executive functions, that help them deflect distractions, recall goals and take the steps needed to reach them. To achieve a goal in work or play, for instance, people need to be able to remember their aim (use hindsight), prompt themselves about what they need to do to reach that goal (use forethought), keep their emotions reined in and motivate themselves. Unless a person can inhibit interfering thoughts and impulses, none of these functions can be carried out successfully.

In the early years, the executive functions are performed externally: children might talk out loud to themselves while remembering a task or puzzling out a problem. As children mature, they internalize, or make private, such executive functions, which prevents others from knowing their thoughts. Children with ADHD, in contrast, seem to lack the restraint needed to inhibit the public performance of these executive functions.

The executive functions can be grouped into four mental activities. One is the operation of working memory—holding information in the mind while working on a task, even if the original stimulus that provided the information is gone. Such remembering is crucial to timeliness and goal-directed behavior: it provides the means for hindsight, forethought, preparation and the ability to imitate the complex, novel behavior of others—all of which are impaired in people with ADHD.

The internalization of self-directed speech is another executive function. Before the age of six, most children speak out loud to themselves frequently, reminding themselves how to perform a particular task or trying to cope with a problem, for example. ("Where did I put that book? Oh, I left it under the desk.") In elementary school, such private speech evolves into inaudible muttering; it usually disappears by age 10 [see "Why Children Talk to Themselves," by Laura E. Berk; SCIENTIFIC AMERICAN, November 1994]. Internalized, self-directed speech allows one to

A Psychological Model of ADHD

A loss of behavioral inhibition and self-control leads to the following disruptions in brain functioning:

IMPAIRED FUNCTION	CONSEQUENCE	EXAMPLE
Nonverbal working memory	Diminished sense of time Inability to hold events in mind Defective hindsight Defective forethought	Nine-year-old Jeff routinely forgets important responsibilities, such as deadlines for book reports or an after-school appointment with the principal
Internalization of self-directed speech	Deficient rule-governed behavior Poor self-guidance and self-questioning	Five-year-old Audrey talks too much and cannot give herself useful directions silently on how to perform a task
Self-regulation of mood, motivation and level of arousal	Displays all emotions publicly; cannot censor them Diminished self-regulation of drive and motivation	Eight-year-old Adam cannot maintain the persistent effort required to read a story appropriate for his age level and is quick to display his anger when frustrated by assigned schoolwork
Reconstitution (ability to break down observed behaviors into component parts that can be recombined into new behaviors in pursuit of a goal)	Limited ability to analyze behaviors and synthesize new behaviors Inability to solve problems	Fourteen-year-old Ben stops doing a homework assignment when he realizes that he has only two of the five assigned questions; he does not think of a way to solve the problem, such as calling a friend to get the other three questions

LISA BURNETT

reflect to oneself, to follow rules and instructions, to use self-questioning as a form of problem solving and to construct "meta-rules," the basis for understanding the rules for using rules—all quickly and without tipping one's hand to others. Laura E. Berk and her colleagues at Illinois State University reported in 1991 that the internalization of self-directed speech is delayed in boys with ADHD.

A third executive mental function consists of controlling emotions, motivation and state of arousal. Such control helps individuals achieve goals by enabling them to delay or alter potentially distracting emotional reactions to a particular event and to generate private emotions and motivation. Those who rein in their immediate passions can also behave in more socially acceptable ways.

The final executive function, reconstitution, actually encompasses two separate processes: breaking down observed behaviors and combining the parts into new actions not previously learned from experience. The capacity for reconstitution gives humans a great degree of fluency, flexibility and creativity; it allows individuals to propel themselves toward a goal without having to learn all the needed steps by rote. It permits children as they mature to direct their behavior across increasingly longer intervals by combining behaviors into ever longer chains to attain a goal. Initial studies imply that children with ADHD are less capable of reconstitution than are other children.

I suggest that like self-directed speech, the other three executive functions become internalized during typical neural development in early childhood. Such privatization is essential for creating visual imagery and verbal thought. As children grow up, they develop the capacity to behave covertly, to mask some of their behaviors or feelings from others. Perhaps because of faulty genetics or embryonic development, children with ADHD have not attained this ability and therefore display too much public behavior and speech. It is my assertion that the inattention, hyperactivity and impulsivity of children with ADHD are caused by their failure to be guided by internal instructions and by their inability to curb their own inappropriate behaviors.

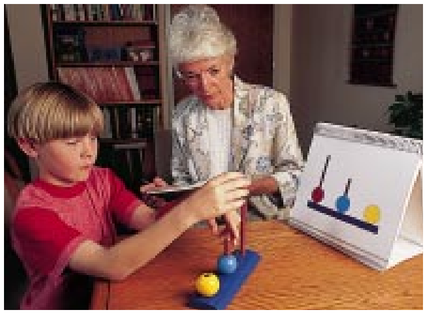
Prescribing Self-Control

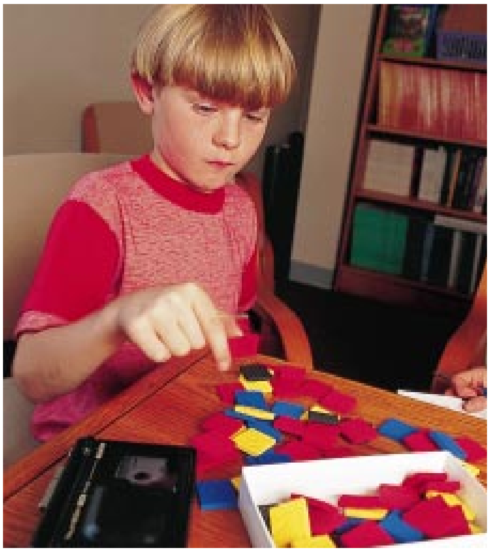
If, as I have outlined, ADHD is a failure of behavioral inhibition that delays the ability to privatize and execute the four executive mental functions I have described, the finding supports the theory that children with ADHD might be helped by a more structured environment. Greater structure can be an important complement to any drug therapy the children might receive. Currently children (and adults) with ADHD often receive drugs such as Ritalin that boost their capacity to inhibit and regulate impulsive behaviors. These drugs act by inhibiting the dopamine transporter, increasing the time that dopamine has to bind to its receptors on other neurons.

Such compounds (which, despite their inhibitory effects, are known as psychostimulants) have been found to improve the behavior of between 70 and 90 percent of children with ADHD older than five years. Children with ADHD who take such medication not only are less impulsive, restless and distractible but are also better able to hold important information in mind, to be more productive academically, and to have more internalized speech and better self-control. As a result, they tend to be liked better by other children and to experience less punishment for their actions, which improves their self-image.

My model suggests that in addition to psychostimulants—and perhaps antidepressants, for some children—treat-

PSYCHOLOGICAL TESTS used in ADHD research include the four depicted here. The tower-building test (*upper left*), in which the subject is asked to assemble balls into a tower to mimic an illustration, measures forethought, planning and persistence. The math test (*upper right*) assesses working memory and problem-solving ability. In the auditory attention test (*below*), the subject must select the appropriate colored tile according to taped instructions, despite distracting words. The time estimation test (*lower right*) measures visual attention and subjective sense of time intervals. The subject is asked to hold down a key to illuminate a lightbulb on a computer screen for the same length of time that another bulb was illuminated previously.





PHOTOGRAPHS BY STEPHEN ROSE *Gamma Liaison*

ment for ADHD should include training parents and teachers in specific and more effective methods for managing the behavioral problems of children with the disorder. Such methods involve making the consequences of a child's actions more frequent and immediate and increasing the external use of prompts and cues about rules and time intervals. Parents and teachers must aid children with ADHD by anticipating events for them, breaking future tasks down into smaller and more immediate steps, and using artificial immediate rewards. All these steps serve to externalize time, rules and consequences as a replacement for the weak internal forms of information, rules and motivation of children with ADHD.

In some instances, the problems of ADHD children may be severe enough to warrant their placement in special education programs. Although such programs are not intended as a cure for the child's difficulties, they typically do provide a smaller, less competitive and more supportive environment in which the child can receive individual instruction. The hope is that once children learn techniques to overcome their deficits in self-control, they will be able to function outside such programs.

There is no cure for ADHD, but much more is now known about effectively coping with and managing this persistent and troubling developmental disorder. The day is not far off when genetic testing for ADHD may become available and more specialized medications may be designed to counter the specific genetic deficits of the children who suffer from it. SA

The Author

RUSSELL A. BARKLEY is director of psychology and professor of psychiatry and neurology at the University of Massachusetts Medical Center in Worcester. He received his B.A. from the University of North Carolina at Chapel Hill and his M.A. and Ph.D. from Bowling Green State University. He has studied ADHD for nearly 25 years and has written many scientific papers, book chapters and books on the subject, including *ADHD and the Nature of Self-Control* (Guilford Press, 1997) and *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment* (Guilford Press, 1998).

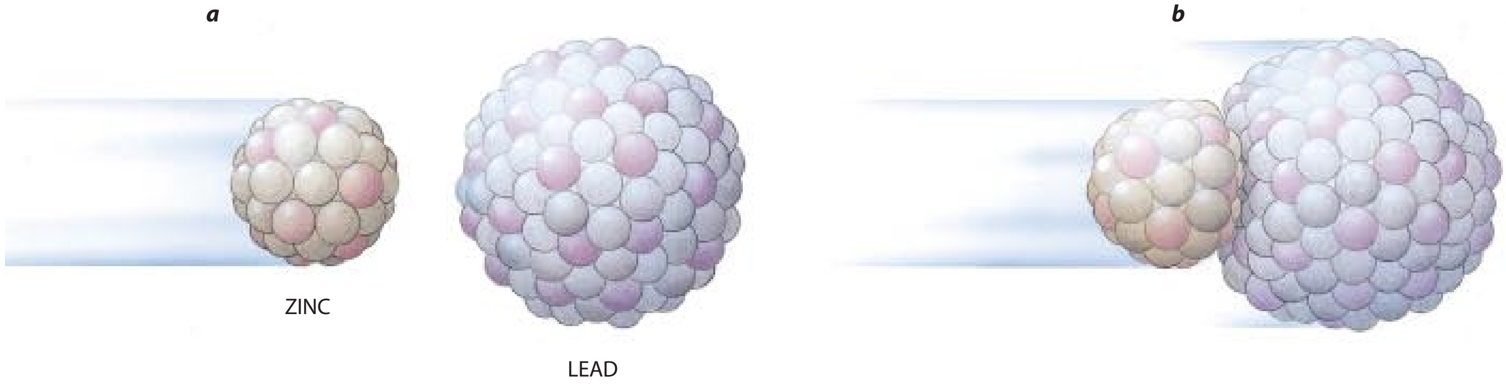
Further Reading

The Epidemiology of Attention-Deficit Hyperactivity Disorder. Peter Szatmari in *Child and Adolescent Psychiatric Clinics of North America*, Vol. 1. Edited by G. Weiss. W. B. Saunders, 1992.

Hyperactive Children Grown Up. Gabrielle Weiss and Lily Trokenberg Hechtman. Guilford Press, 1993.

Taking Charge of ADHD: The Complete, Authoritative Guide for Parents. R. A. Barkley. Guilford Press, 1995.

Dopamine *D4* Receptor Gene Polymorphism Is Associated with Attention Deficit Hyperactivity Disorder. G. J. LaHoste et al. in *Molecular Psychiatry*, Vol. 1, No. 2, pages 121–124; May 1996.



Making New Elements

Three new elements—110, 111 and 112—have been produced over the past several years. Scientists are now struggling to create 113 and 114. How many elements can they add to the periodic table?

by Peter Armbruster and Fritz Peter Hessberger

You can call it a gamble: there's just a narrow path to reach our goal—or to miss it. In this spirit, we start an experiment to make new, superheavy elements. These elements have to be produced in a lengthy, complicated procedure in which we smash atomic nuclei into one another at very high speeds and hope they undergo fusion. The resulting products will be extremely fragile, and most will break apart immediately.

Only under very extraordinary conditions will a new element have a chance to survive the production process and land in a stable configuration—what we call the ground state. But even when these conditions are met, the production rate of a new element is tiny. To make element 112, the heaviest artificial element produced to date, we conducted an around-the-clock experiment for 24 days and created only two atoms of 112, which lasted for only microseconds.

When we began hunting for element 113 this past spring, we expected the production rate to be a factor of two or three times lower. So we attempted to make 113 in an experimental run that lasted 42 days. We found nothing. We're still asking ourselves, Why didn't we see anything? Did we choose the wrong energy setting? Is the production rate smaller than we expected? Is there some unusual property of element 113 that makes it difficult to detect with our current equipment?

Despite the difficulties of making new elements that last for such a short time, these lingering questions make it impossible for us to give up the search. Over the past six decades, researchers have made 20 artificial elements. The question is, How many more can we create?

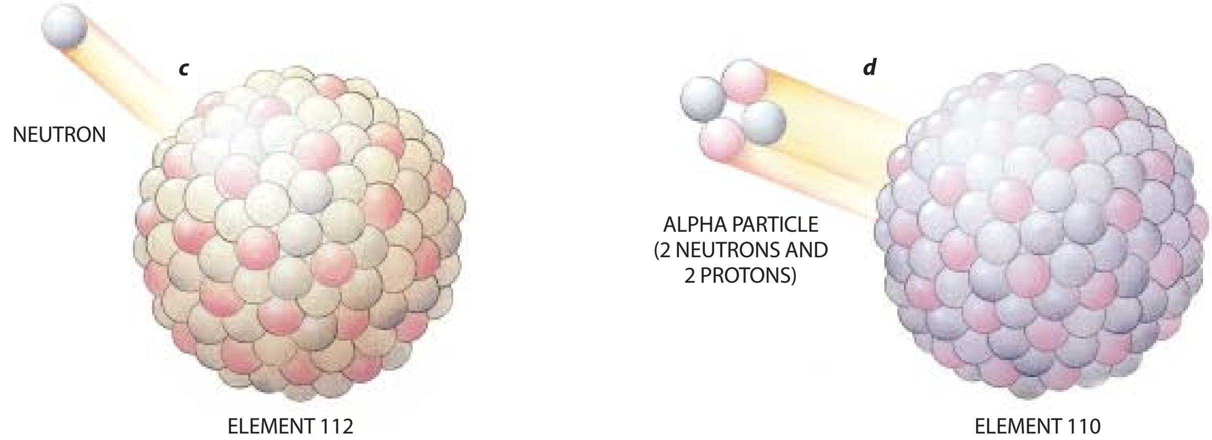
Artificial Elements

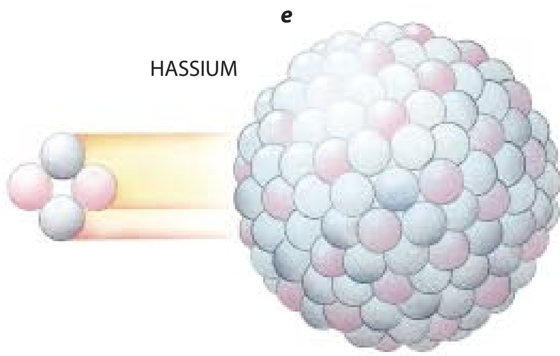
In 1936 physicist Emilio G. Segrè was working at the cyclotron in Berkeley, Calif., in the laboratory of his friend Ernest O. Lawrence. Segrè irradiated a sample of molybdenum with particles called deuterons; afterward, he took the sample back to the University of Palermo. There Segrè discovered the first man-made element, technetium, element number 43 (this so-called atomic number indicates an element's position in the periodic table as well as the number of positively charged particles, or protons, found in an atom's nucleus).

The inspiration for making technetium came from experiments carried out by another Italian physicist, Enrico Fermi. In 1934 Fermi, then at the University of Rome, proposed that new elements could be made by bombarding an atom's central core, or nucleus, with uncharged particles known as neutrons. Typically neutrons reside within atomic nuclei, but a lone neutron can penetrate the nucleus, where it may be captured. The resulting nucleus may be stable, or it may be radioactive. In the latter case, the neutron will, through a process known as beta-decay, change into a proton, an electron and an antineutrino (a chargeless, virtually massless, subatomic particle often released during nuclear decay). As a result of this process of neutron capture and beta-decay, the number of protons in the nucleus increases, the atomic number goes up with each proton and higher elements are formed.

This idea of irradiating the nuclei of various elements with neutrons was picked up by other research groups. In 1940 Edwin M. McMillan and Philip H. Abelson, both at the University of

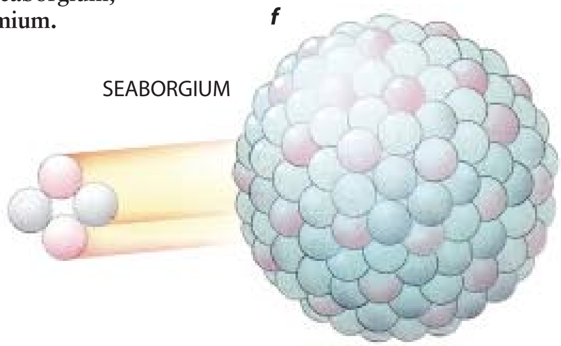
FERMIUM

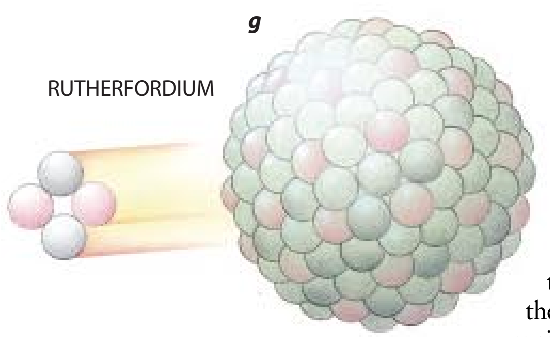


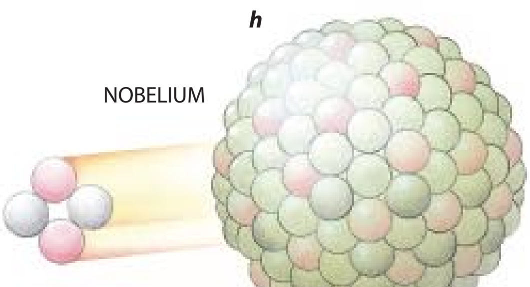


TOMO NARASHIMA

ELEMENT 112 forms when a zinc nucleus smashes into a lead nucleus with enough force to overcome the natural repulsion of the two positively charged nuclei (*a*). The two nuclei fuse together (*b*) to form a compound nucleus that is very unstable. The nucleus immediately releases a neutron to produce isotope 277 of element 112 (*c*). The new element lasts several hundred microseconds before undergoing radioactive decay to produce a chain of six "daughter" elements (*d–i*): element 110, hassium, seaborgium, rutherfordium, nobelium and fermium.







California at Berkeley, jumped ahead to the next unknown element and, by neutron irradiation at the cyclotron there, synthesized element 93—the first element beyond uranium, the heaviest naturally occurring element known at the time. (Appropriately, element 93 was eventually named neptunium, because Neptune is the first planet beyond Uranus.) During the 1940s and 1950s, the Americans—working primarily with Glenn T. Seaborg of Lawrence Berkeley National Laboratory and continuing the scientific work that began during wartime research into nuclear weapons—discovered a string of new elements: plutonium (element 94), americium (95), curium (96), berkelium (97), californium (98), einsteinium (99) and fermium (100).

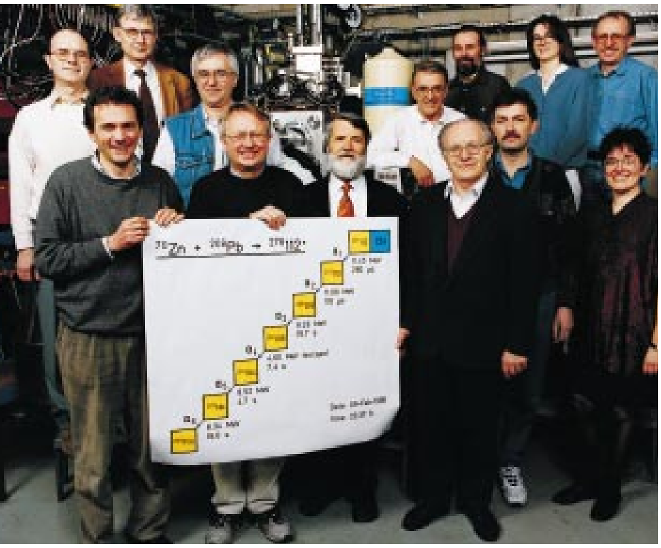
All eight of these elements can be produced in weighable quantities according to Fermi's suggested combination of neutron capture and beta-decay. But the quantity of each that can be synthesized varies widely. Today the world's reserve of plutonium is more than 1,000 tons (10^{30} atoms); the reserve of fermium, however, was never more than some trillionths of a gram (10^{10} atoms). Unfortunately, Fermi's recipe ends at fermium—beyond this point, beta-decay does not take place, so new elements cannot be produced by the technique. A new approach was needed.

International Rivalries

The option at hand was to bring together at very high speeds light elements such as carbon (element 6), nitrogen (7) or oxygen (8) and transuranic elements—such as plutonium (94) through einsteinium (99)—with the expectation that the energy of the collision would prompt the fusion of the two nuclei, creating even heavier elements. But the enormous practical disadvantages that arise when trying to slam two nuclei into each other posed major challenges in the laboratory.

To carry out these experiments, scientists improved on existing technology (namely, the cyclotron) and, for the first time, built linear accelerators. These devices allow researchers to accelerate high-intensity beams of ions at well-defined energies. The use of transuranic elements meant that this technique could be carried out only where there was access to the large nuclear reactors found in countries with nuclear weapons. As a result, during the cold war, the two research sites involved—Lawrence Berkeley National Laboratory in the U.S. and the Joint Institute for Nuclear Research in Dubna, Russia—were competing not only scientifically but also politically.

By 1955 the Berkeley group had produced element 101, mendelevium, by



GSI

TEAM EFFORT is needed when creating elements. The authors are pictured with the team that produced element 112, the heaviest artificial element synthesized to date. Peter Armbruster is standing in the front row, third from the left; Fritz Peter Hessberger is standing in the back row at the far left.

the fusion of helium (element 2) and einsteinium (99). Between 1958 and 1974 the two groups created the elements nobelium (102), lawrencium (103), rutherfordium (104), dubnium (105) and seaborgium (106). Those days were so full of tension that the U.S. and Russia still argue over who was first to discover the elements as well as what their names should be. The names mentioned here reflect the ones recognized by the International Union of Pure and Applied Chemistry [*see table on page 76*].

After making element 106, scientists hit another roadblock: the standard fusion technique did not succeed in producing additional new elements. It was at this time that Germany entered the race. With the founding in December 1969 of the Gesellschaft für Schwerionenforschung (GSI, or the Institute for Heavy-Ion Research) in Darmstadt, German chemists and physicists had begun to participate in this field that had been exclusively the domain of American and Soviet researchers. Several years later, in 1975, the heavy-ion accelerator UNILAC (Universal Linear Accelerator), an idea conceived by Christoph Schmelzer of the University of Heidelberg (the first director of the institute), was put into operation at GSI.

This novel machine was the first that could accelerate all types of ions, including uranium, at continuous, adjustable energies, offering great flexibility in attempts to fuse two nuclei. One of our main objectives for UNILAC was to make elements 107 through at least 114—what we call the superheavy elements. Why aim for 114? According to theoretical calculations, element 114 should be particularly stable because its nucleus should exist in what physicists call a closed shell.

In 1948 Otto Haxel, J. Hans D. Jensen and Hans E. Suess of the University of Heidelberg, as well as Maria Goeppert-Mayer of Argonne National Laboratory, observed interesting regularities in the number of neutrons and protons found in atomic nuclei: certain combinations of these two subatomic particles produced nuclei that were quite stable compared with their neighbors. Similar patterns had been recognized in the number of electrons found in atoms as well—specific configurations of electrons produced chemically stable elements. Scientists determined that such patterns reflected the way electrons fill successive energy shells surrounding atomic nuclei. Some shells hold only two electrons, whereas others can carry up to 14. Researchers found that in the most stable elements—the chemically inactive noble gases—the outermost occupied shell was completely full, or what they termed "closed."

Drawing on the idea of the filling of electron shells, Goeppert-Mayer and, independently, Jensen developed the shell model of the atomic nucleus and explained how many protons and neutrons together produced closed shells. The two shared the 1963 Nobel Prize in Physics for the discovery.

For every atom of a given element, the number of protons in the nucleus remains the same. But various forms, or isotopes, of a particular element can exist, each with a different number of neutrons in the nucleus. Isotopes are distinguished by what is known as atomic weight, a figure equal to the sum of the number of protons and neutrons. Stable isotopes with extremely stable nuclei include calcium 40 (20 protons and 20 neutrons), calcium 48 (20 protons and 28 neutrons) and lead 208 (82 protons and 126 neutrons). In addition, one isotope of element 114, with 114 protons and 184 neutrons, should also have a closed shell. Physicists described elements neighboring these closed shells as existing on an "island of stability" in a sea of otherwise unstable nuclei.

So once UNILAC was up and running, we were anxious to see if we could reach that island of stability. In the beginning, everything seemed promising. Theoreticians initially told us that calculations indicated elements close to 114 would have long half-lives, on the order of billions of years, comparable to the half-lives of certain lighter elements such as uranium and thorium. (An isotope's half-life is the time required for half of a given number of atoms to undergo radioactive decay.) We expected to be able to produce considerable quantities of the superheavy elements—possibly leading to novel materials for chemists to investigate and new atoms for nuclear and atomic physicists to study.

But in the early 1980s it became clear that the production of superheavy elements around element 114 was not going to be easy. All attempts to synthesize them using various combinations of projectiles and targets in nuclear reactions failed. Even further efforts to find these elements in nature yielded nothing.

Cold Fusion

After learning about an important 1974 discovery by Yuri Oganessian and his research partner Alexander Demin of the Dubna facility, we at GSI took up their new strategy for producing the superheavy elements. Oganessian and Demin bombarded a target made of lead (element 82) with ions of argon (18) and produced element 100, fermium. Oganessian recognized that during this process, the heating of the newly formed nuclei was much lower than that in the higher-impact collisions needed when irradiating the heaviest transuranic isotopes with very light ions, as had been common practice until then. As a result of the lower excitation energy, Oganessian concluded, far more nuclei could survive the fusion process and would not subsequently undergo fission and break apart.

Our GSI team became interested in this technique, which we termed “cold” fusion because of the lower excitation energy—and hence, reduced heat—of the nucleus during the procedure. (Of course, this method has nothing to do with the discussion some years ago of purported room-temperature fusion of deuterium nuclei in a test tube.) Researchers at Berkeley regarded cold fusion as a curiosity and did not take it seriously, but for us it was our only chance to enter the field. For one, the starting materials, lead and bismuth, were readily available in nature and did not need to be produced by nuclear reactors. In addition, the configuration of UNILAC allowed us to use all kinds of ions as projectiles and to vary their energies in small steps that could be easily reproduced at any value.

We also had the right device for isolating and detecting products from a fusion reaction. As UNILAC was being built at GSI, Gottfried Münzenberg led a team of scientists from both GSI and the Second Physics Institute of the University of Giessen in Germany that built a special filter, called the separator for heavy-ion reaction products, or SHIP, for this purpose. The filter removes projectiles and unwanted side products of the fusion reaction and efficiently focuses desired products onto a detector that enables us to identify them by their radioactive decay products.

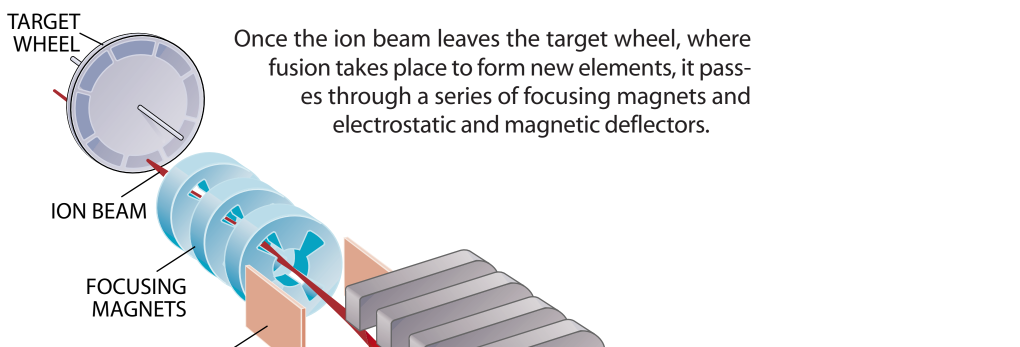
Elements 107 and Beyond

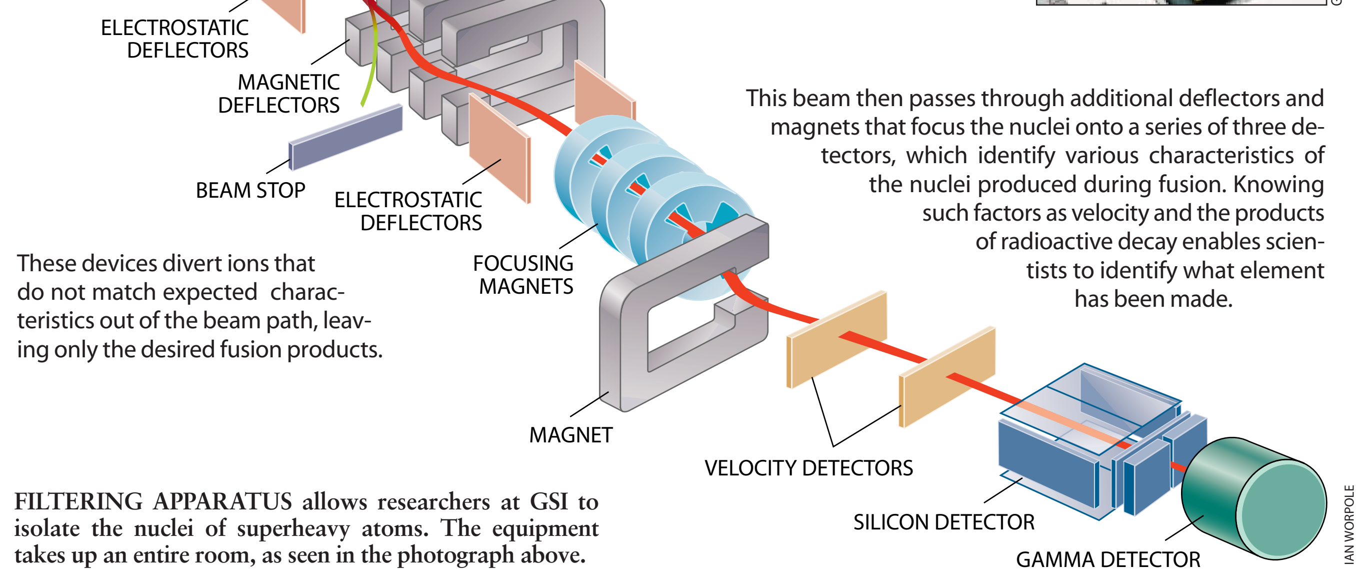
In the early 1980s our group was able to prove convincingly that cold fusion works: we identified bohrium (element 107), hassium (108) and meitnerium (109). Soon after, we described our technique in these pages [see “Creating Superheavy Elements,” by Peter Armbruster and Gottfried Münzenberg; Scientific American, May 1989].

Following this success, however, we recognized that taking the next step to element 110 would require improvements in our experimental techniques. To produce a single atom of meitnerium, element 109, for instance, we had to operate the accelerator for two weeks. Furthermore, our ability to produce these superheavy elements falls off by a factor of three with each addition to the periodic table—making them increasingly difficult to detect.

Despite these potential problems, we continued to pursue the method that had worked for us in the past. Teams at Berkeley and Dubna had by then returned to the previous method of “hot” fusion, so at least we had no competition at that time. Through a variety of structural changes to UNILAC, we were able to triple the intensity of its ion beam. We also enhanced by a factor of three the sensitivity of SHIP, the filtering device. Finally, our colleague Sigurd Hofmann built a new detector system with increased sensitivity.

Our team had changed somewhat as well. In addition to the old group (the two of us, Hofmann, Münzenberg, Helmut Folger, Matti E. Leino, Victor Ninov and Hans-Joachim Schött), a new crew joined us: Andrey G. Popeko, Alexandr V. Yeremin and Andrey N. Andreyev of the Flerov Laboratory of Nuclear Reactions in Dubna, as well as Stefan Saro and Rudolf Janik of Comenius University in Bratislava, Slovakia. We had stayed in close contact with the Dubna team since 1973, but changing politics now allowed these scientists to collaborate with us directly.





FILTERING APPARATUS allows researchers at GSI to isolate the nuclei of superheavy atoms. The equipment takes up an entire room, as seen in the photograph above.

The Politics of Naming

Disputes over who created elements 102 through 109 resulted in an international controversy over their names. In 1994 the International Union of Pure and Applied Chemistry (IUPAC) appointed the Commission on Nomenclature of Inorganic Chemistry (CNIC) to determine official names for these elements (many had been informally named by their discoverers). The list proposed by the CNIC, however, did not settle the controversy—it simply inflamed the debate. Last year the IUPAC solicited comments from chemists around the world concerning the naming of these elements and in August 1997 settled on final names, as shown below. The names for elements 110 through 112 have yet to be officially determined. *—P.A. and F.P.H.*

ELEMENT	DISCOVERER(S)	SUGGESTED NAMES (proposed by)	OFFICIAL NAME (chemical symbol)
102	Initial claim: Nobel Institute in Stockholm, Sweden Additional claims: University of California, Berkeley; Joint Institute for Nuclear Research, Dubna, Russia	Joliotium (Dubna) Nobelium (Nobel Institute, CNIC)	Nobelium (No); in honor of Alfred Nobel, Swedish inventor and founder of the Nobel Prizes
103	Disputed; both Dubna and Berkeley teams claim priority	Lawrencium (Berkeley, CNIC)	Lawrencium (Lr); in honor of Ernest O. Lawrence, inventor of the cyclotron
104	Disputed; both Dubna and Berkeley teams claim priority	Dubnium (CNIC) Kurchatorium (Dubna) Rutherfordium (Berkeley)	Rutherfordium (Rf); in honor of New Zealand–born physicist Ernest Rutherford, whose work was crucial to the early understanding of the nucleus
105	Disputed; both Dubna and Berkeley teams claim priority	Hahnium (Berkeley) Joliotium (CNIC) Nielsbohrium (Dubna)	Dubnium (Db); in honor of the Dubna laboratory
106	Berkeley (undisputed)	Rutherfordium (CNIC) Seaborgium (Berkeley)	Seaborgium (Sg); in honor of Glenn T. Seaborg, American chemist who was co-discoverer of 11 artificial elements
107	GSI, Darmstadt, Germany (undisputed)	Bohrium (CNIC) Nielsbohrium (GSI)	Bohrium (Bh); in honor of Danish physicist Niels Bohr, whose research contributed significantly to the modern understanding of the atom
108	GSI (undisputed)	Hahnium (CNIC) Hassium (GSI)	Hassium (Hs); named for the German state Hesse, where Darmstadt is located
109	GSI (undisputed)	Meitnerium (GSI, CNIC)	Meitnerium (Mt); in honor of Lise Meitner, the Austrian physicist who first conceived the idea of nuclear fission

When we resumed operation of UNILAC in 1993, we conducted a series of test runs with beams of the isotopes argon 40 (element 18) and titanium 50 (element 22), which had been used to produce various isotopes of both mendelevium and rutherfordium. We also conducted other runs to determine the right energy setting for making element 110. In all these trials, the success of our structural improvements was apparent: both the beam intensity and detector sensitivity were definitely better.

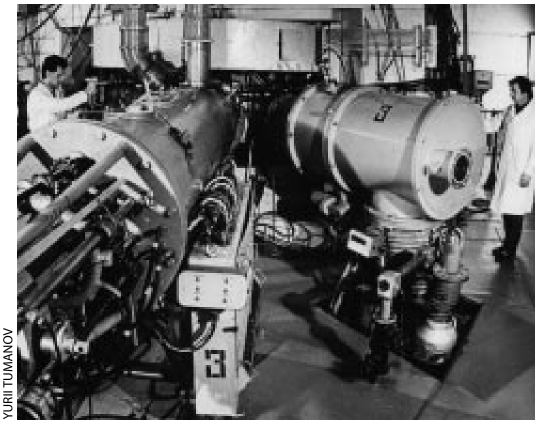
On November 9, 1994—after a break of more than 10 years—we were finally able to identify the decay products of yet another previously unknown element: during the fusion of lead 208 (element 82) with nickel 62 (element 28), a new compound nucleus was formed with 110 protons and an atomic weight of 270. This nucleus promptly discharged one neutron, resulting in isotope 269 of element 110. This new isotope had a half-life of 170 microseconds; we identified it by monitoring how the nucleus decayed into so-called daughter products. In this case, the nucleus discharged a series of four alpha particles, each a helium nucleus with two neutrons and two protons, to generate the daughter product isotope 257 of rutherfordium (element 104). In subsequent experiments, we also created isotope 271 of element 110. This isotope of the element was somewhat easier to make: its rate of production was about a factor of four higher than that of the previous isotope.

Just over one month later, on December 17, 1994, the 25th anniversary of GSI, we discovered element 111 after irradiating bismuth 209 with nickel 64. This experiment provided some important confirmation of the theory of nuclear shells. Two of the decay products of element 111 (isotope 268 of meitnerium and isotope 264 of bohrium) proved to be more stable than previously observed lighter isotopes of these elements, just as the theory predicted.

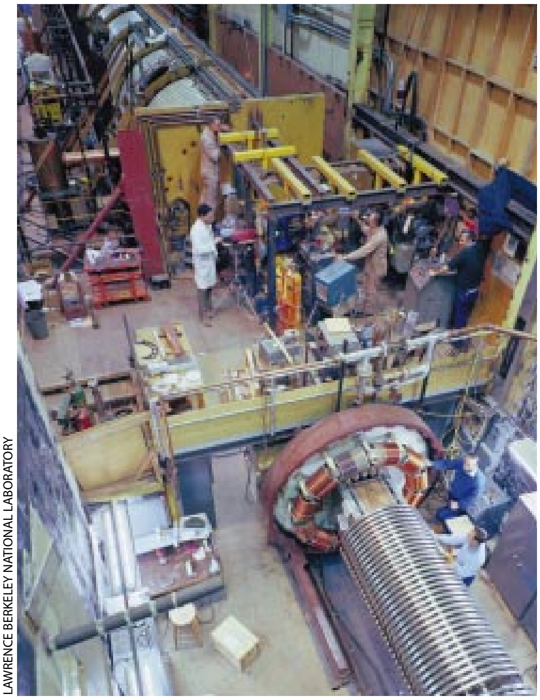
In our next set of experiments, we planned to use the neutron-rich zinc 70 as a projectile, hoping to generate elements 112 and 113. And in February 1996 we succeeded in producing element 112, the heaviest element ever made in the laboratory.

Our yield was much lower than expected; as mentioned earlier, we made two atoms in 24 days. The half-life of this isotope of element 112 was 240 microseconds. We were able to identify isotope 277 of element 112 by observing its radioactive decay and following its successive emission of six alpha particles to yield the daughter product fermium 253 [*see illustration on pages 72 and 73*]. In the process, we identified a new isotope of element 110 and a new isotope of hassium. The isotope of 112 was the first nucleus we made with over 162 neutrons, a quantity that constitutes a deformed closed nuclear shell. This factor contributed to the increased stability of certain products of the element's decay. Specifically, we found that hassium 269, produced during the decay of this isotope of 112, had a half-life of 9.3 seconds. In contrast, hassium 265, which had been observed in earlier experiments, had a half-life of only 1.7 milliseconds. Such findings further confirmed theoreticians' predictions that there would be a closed shell at 162 neutrons.

The creation of superheavy elements is the result of careful long-term planning. Success has come intermittently, and recent developments are no exception. The next step for us to take—making elements 113 and 114—is proving quite difficult. As mentioned earlier, labs here in Germany, as well as in the U.S. and Russia, have not, as of this writing, been successful in creating 113 or 114, despite lengthy and repeated efforts. Nevertheless, we have come a long way since the 1940s, when Niels Bohr predicted that fermium, element 100, would be the last element of the periodic table. We can now identify new elements with half-lives of less than 10 microseconds. If, out of 10 billion trials, two nuclei fuse together once to form one superheavy atom, we will find it.



YURI TUMANOV



LAWRENCE BERKELEY NATIONAL LABORATORY

HIGH-TECH LABS at the Joint Institute for Nuclear Research in Dubna, Russia (*top*), and at Lawrence Berkeley National Laboratory in the U.S. (*bottom*) have been home to many new elements.

Yet all evidence suggests that each new element in the periodic table will be harder to make than the previous one—a trend that shows no sign of changing. Continued improvements in our experimental techniques might bring us to the elusive element 114 using the successful cold fusion method. But the main obstacle to reaching higher elements reflects a fundamental law of physics: positively charged nuclei naturally repel one another, and these repulsive forces increase as the nuclei become larger. Although we have managed to circumvent these forces so far, we will not be able to continue indefinitely.

Today all laboratories searching for superheavy elements collaborate closely, and political competition between our countries is no longer a major driving force. But we hope that even without the political pressures, the countries involved will continue to support research on superheavy elements. We know that these efforts probably cannot be converted into practical applications, but the intellectual and technological achievements gained from this work have certainly justified our efforts. We know that the number of elements in the period table is finite. The question to be answered is, How far can we go? SA

The Authors

PETER ARMBRUSTER and FRITZ PETER HESSBERGER work together at Gesellschaft für Schwerionenforschung (GSI, or the Institute for Heavy-Ion Research) in Darmstadt, Germany. Armbruster has been at GSI since 1971. As senior scientist of the nuclear chemistry department, Armbruster was head of the research project to synthesize elements 107 through 112. His current interests also include developing methods for incinerating nuclear waste. Hessberger arrived at GSI in 1979 after studying physics at Technische Hochschule in Darmstadt. He worked with Armbruster on the synthesis of elements 107 through 112 and received his doctorate in 1985. Hessberger is also studying nuclear reactions and nuclear spectroscopy.

Further Reading

On the Production of Heavy Elements by Cold Fusion: The Elements 106 to 109. Peter Armbruster in *Annual Review of Nuclear and Particle Science,* Vol. 35, pages 135–194; 1985.

Recent Advances in the Discovery of Transuranium Elements. Gottfried Münzenberg in *Report on Progress in Physics,* Vol. 51, No. 1, pages 57–104; January 1988.

An Element of Stability. Richard Stone in *Science,* Vol. 278, pages 571–572; October 24, 1997.

Meine 40 Jahre Mit Rückstossspektrometern. Peter Armbruster in *Physics Blätter,* Vol. 53, pages 661–668; 1997.

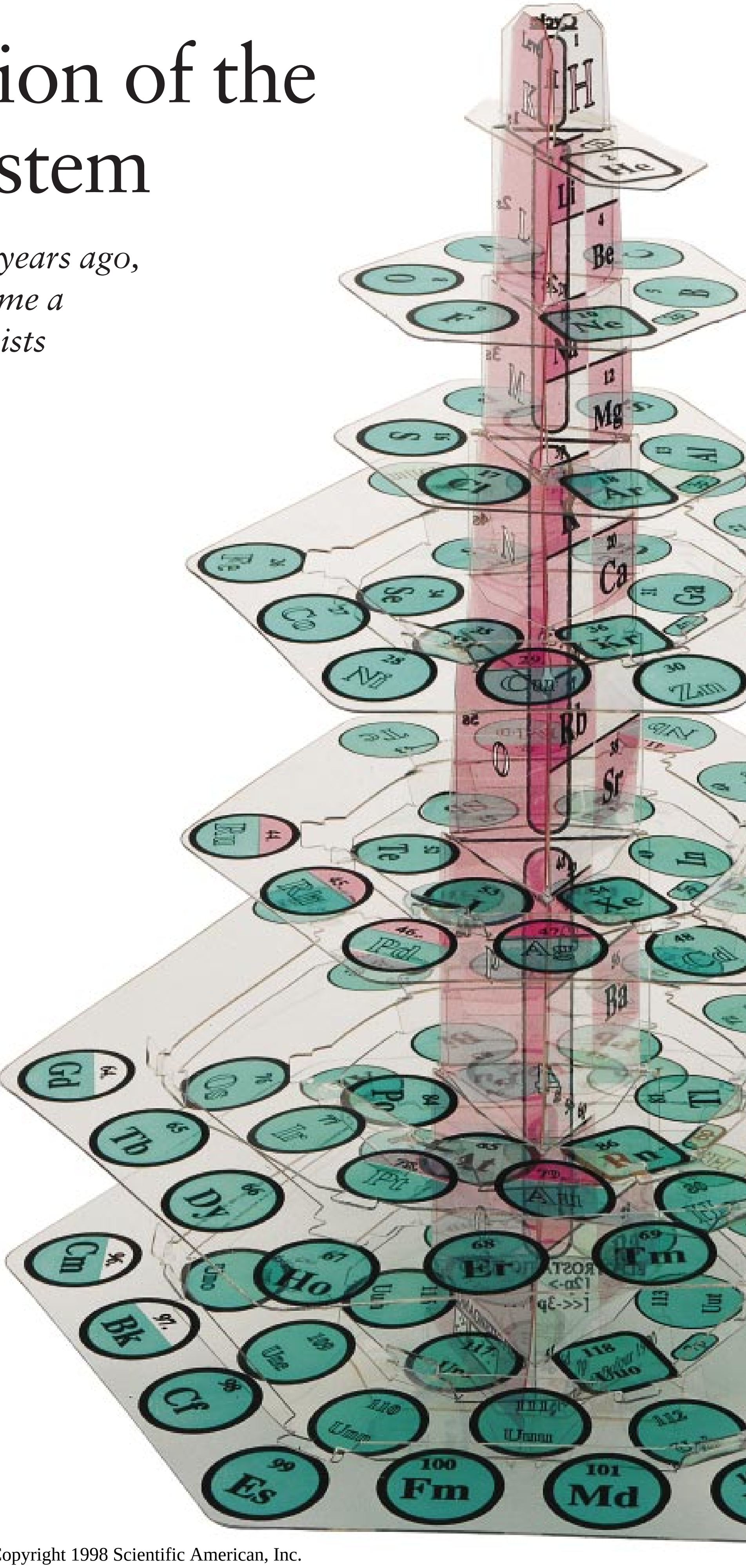
New Elements: Approaching Z = 114. Sigurd Hofmann in *Reports on Progress in Physics,* Vol. 61, No. 6, pages 639–689; June 1998.

GSI site is available at http://www.gsi.de/gsi.html (also available in German at http://www.gsi.de/gsi.welcome.deutsch.html) on the World Wide Web.

The Evolution of the Periodic System

From its origins some 200 years ago, the periodic table has become a vital tool for modern chemists

by Eric R. Scerri



The periodic table of the elements is one of the most powerful icons in science: a single document that consolidates much of our knowledge of chemistry. A version hangs on the wall of nearly every chemical laboratory and lecture hall in the world. Indeed, nothing quite like it exists in the other disciplines of science.

The story of the periodic system for classifying the elements can be traced back over 200 years. Throughout its long history, the periodic table has been disputed, altered and improved as science has progressed and as new elements have been discovered [see "Making New Elements," by Peter Armbruster and Fritz Peter Hessberger, on page 72].

But despite the dramatic changes that have taken place in science over the past century—namely, the development of the theories of relativity and quantum mechanics—there has been no revolution in the basic nature of the periodic system. In some instances, new findings initially appeared to call into question the theoretical foundations of the periodic table, but each time scientists eventually managed to incorporate the results while preserving the table's fundamental structure. Remarkably, the periodic table is thus notable both for its historical roots and for its modern relevance.

The term "periodic" reflects the fact that the elements show patterns in their chemical properties in certain regular intervals. Were it not for the simplification provided by this chart, students of chemistry would need to learn the properties of all 112 known elements. Fortunately, the periodic table allows chemists to function by mastering the properties of a handful of typical elements;

all the others fall into so-called groups or families with similar chemical properties. (In the modern periodic table, a group or family corresponds to one vertical column.)

The discovery of the periodic system for classifying the elements represents the culmination of a number of scientific developments, rather than a sudden brainstorm on the part of one individual. Yet historians typically consider one event as marking the formal birth of the modern periodic table: on February 17, 1869, a Russian professor of chemistry, Dimitri Ivanovich Mendeleev, completed the first of his numerous periodic charts. It included 63 known elements arranged according to increasing atomic weight; Mendeleev also left spaces for as yet undiscovered elements for which he predicted atomic weights.

Prior to Mendeleev's discovery, however, other scientists had been actively developing some kind of organizing system to describe the elements. In 1787, for example, French chemist Antoine Lavoisier, working with Antoine Fourcroy, Louis-Bernard Guyton de Morveau and Claude-Louis Berthollet, devised a list of the 33 elements known at the time. Yet such lists are simply one-dimensional representations. The power of the modern table lies in its two- or even three-dimensional display of all the known elements (and even the ones yet to be discovered) in a logical system of precisely ordered rows and columns.

In an early attempt to organize the elements into a meaningful array, German chemist Johann Döbereiner pointed out in 1817 that many of the known elements could be arranged by their similarities into groups of three, which he called triads. Döbereiner singled out triads of the elements lithium, sodium and potassium as well as chlorine, bromine and iodine. He noticed that if the three members of a triad were ordered according to their atomic weights, the properties of the middle element fell in between those of the first and third elements. For example, lithium, sodium and potassium all react vigorously with water. But lithium, the lightest of the triad, reacts more mildly than the other two, whereas the heaviest of the three, potassium, explodes violently. In addition, Döbereiner showed that the atomic weight of the middle element is close to the average of the weights for the first and third members of the triad.

Döbereiner's work encouraged others to search for correlations between the chemical properties of the elements and their atomic weights. One of those who pursued the triad approach further during the 19th century was Peter Kremers of Cologne, who suggested that certain elements could belong to two triads placed perpendicularly. Kremers thus broke new ground by comparing elements in two directions, a feature that later proved to be an essential aspect of Mendeleev's system.

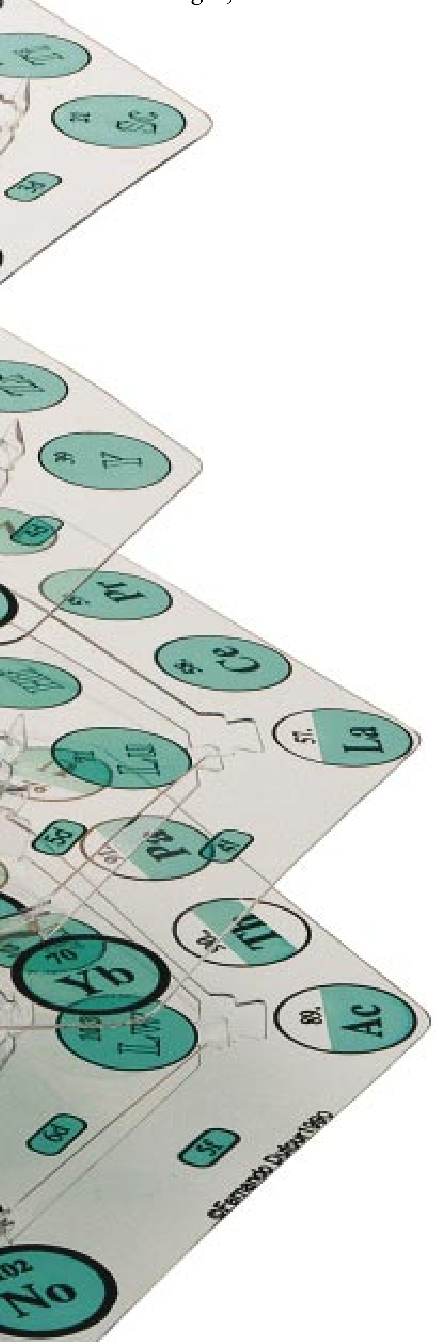
In 1857 French chemist Jean-Baptiste-André Dumas turned away from the idea of triads and focused instead on devising a set of mathematical equations that could account for the increase in atomic weight among several groups of chemically similar elements. But as chemists now recognize, any attempt to establish an organizing pattern based on an element's atomic weight will not succeed, because atomic weight is not the fundamental property that characterizes each of the elements.

Periodic Properties

The crucial characteristic of Mendeleev's system was that it illustrated a periodicity, or repetition, in the properties of the elements at certain regular intervals. This feature had been observed previously in an arrangement of elements by atomic weight devised in 1862 by French geologist Alexandre-Emile Béguyer de Chancourtois. The system relied on a fairly intricate geometric configuration: de Chancourtois positioned the elements according to increasing atomic weight along a spiral inscribed on the surface of a cylinder and inclined at 45 degrees from the base [*see illustration on next page*].

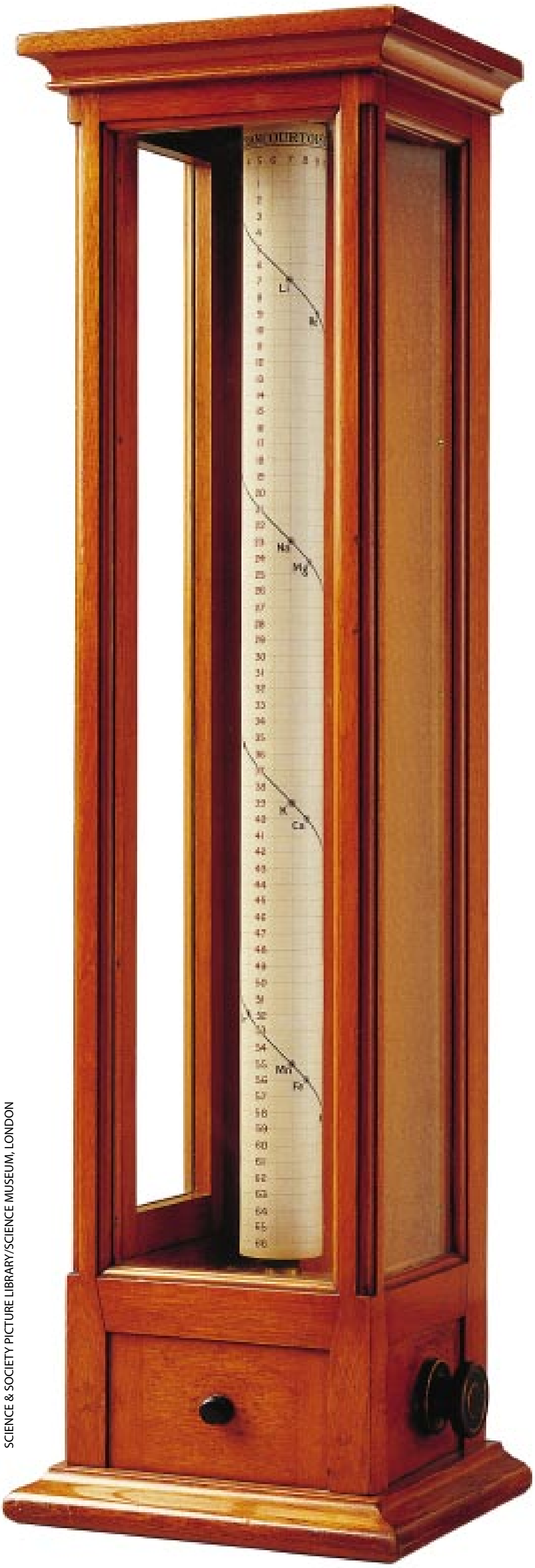
The first full turn of the spiral coincided with the element oxygen, and the second full turn occurred at sulfur. Elements that lined up vertically on the surface of the cylinder tended to have similar properties, so this arrangement succeeded in capturing some of the patterns that would later become central to Mendeleev's system. Yet for a number of reasons, de Chancourtois's system did not have much effect on scientists of the time: his original article failed to include a diagram of the table, the system was rather complicated, and the chemical similarities among elements were not displayed very convincingly.

Several other researchers put forward



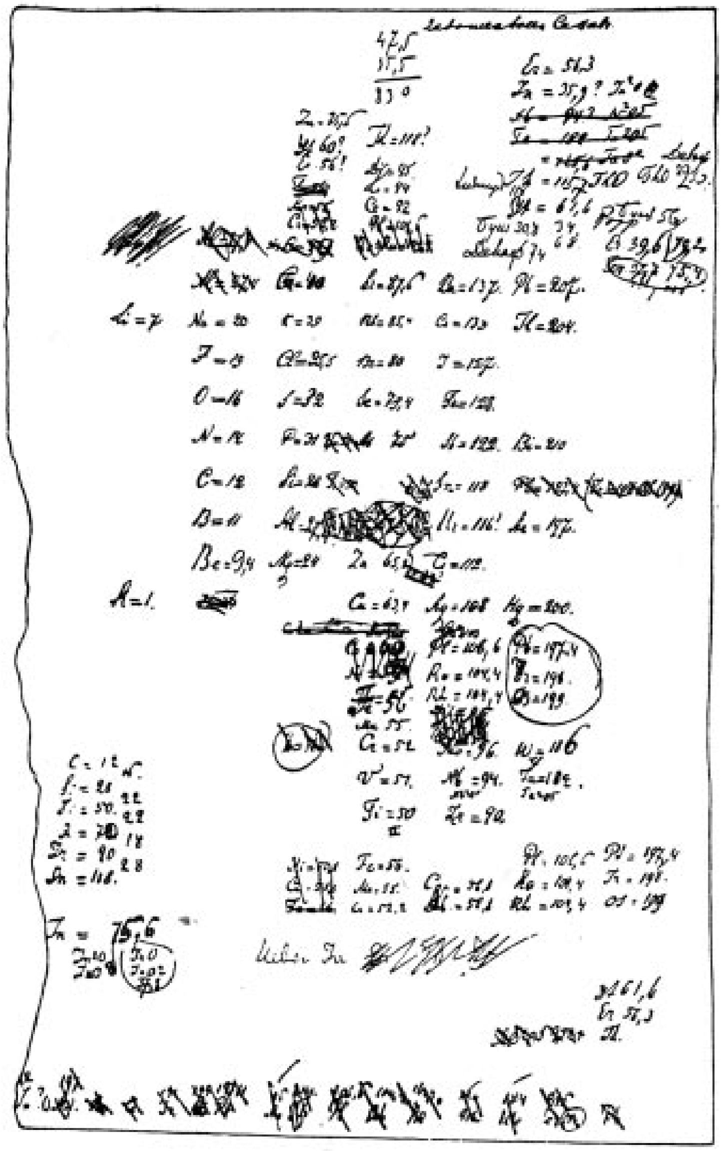
DAN WAGNER

THREE-DIMENSIONAL TABLE transforms the traditional periodic chart into a multilayered structure. The traditional vertical columns, which correspond to a group or family of elements, can be seen running down the central core of this structure (for example, H, Li, Na and so on) as well as through the layers. Elements that are positioned on top of one another in layers, such as He, Ne, Ar and so on, belong in the same group and thus have similar chemical properties. The horizontal rows, or periods, of the traditional table correspond to the multiple layers of the three-dimensional table. This version highlights the symmetrical and regularly increasing size of periods, a fundamental chemical feature not yet fully explained by quantum mechanics.



SCIENCE & SOCIETY PICTURE LIBRARY/SCIENCE MUSEUM, LONDON

EARLY VERSION of an organizing system for the known elements was designed in 1862 by French geologist Alexandre-Emile Béguyer de Chancourtois. Known as the telluric screw, it was the earliest discovery of chemical periodicity.



OESPER COLLECTION, UNIVERSITY OF CINCINNATI

FIRST PERIODIC TABLE was developed by Russian chemist Dimitri Ivanovich Mendeleev in February 1869. This draft shows groups of elements arranged horizontally rather than in the more familiar vertical columns. Mendeleev produced many tables of both kinds.

their own versions of a periodic table during the 1860s. Using newly standardized values for atomic weights, English chemist John Newlands suggested in 1864 that when the elements were arranged in order of atomic weight, any one of the elements showed properties similar to those of the elements eight places ahead and eight places behind in the list—a feature that Newlands called "the law of octaves."

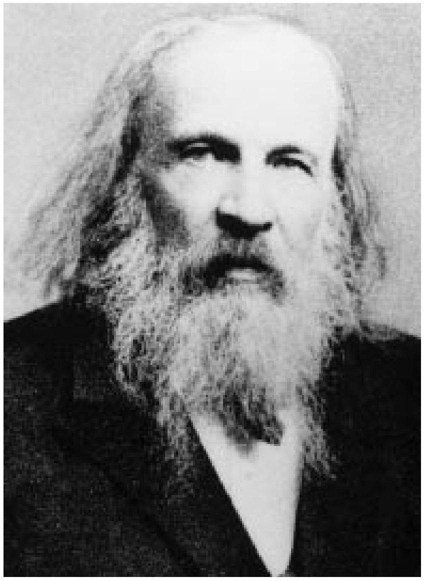
In his original table, Newlands left empty spaces for missing elements, but his more publicized version of 1866 did not include these open slots. Other chemists immediately raised objections to the table because it would not be able to accommodate any new elements that might be discovered. In fact, some investigators openly ridiculed Newlands's ideas. At a meeting of the Chemical Society in London in 1866, George Carey Foster of University College London asked Newlands whether he had considered ordering the elements alphabetically, because any kind of arrangement would present occasional coincidences. As a result of the meeting, the Chemical Society refused to publish Newlands's paper.

Despite its poor reception, however,

Newlands's work does represent the first time anyone used a sequence of ordinal numbers (in this case, one based on the sequence of atomic weights) to organize the elements. In this respect, Newlands anticipated the modern organization of the periodic table, which is based on the sequence of so-called atomic numbers. (The concept of atomic number, which indicates the number of protons present within an atom's nucleus, was not established until the early 20th century.)

The Modern Periodic Table

Chemist Julius Lothar Meyer of Breslau University in Germany, while in the process of revising his chemistry textbook in 1868, produced a periodic table that turned out to be remarkably similar to Mendeleev's famous 1869 version—although Lothar Meyer failed to classify all the elements correctly. But the table did not appear in print until 1870 because of a publisher's delay—a factor that contributed to an acrimonious dispute for priority that ensued between Lothar Meyer and Mendeleev.



CORBIS-BETTMANN (*left*); VAN PELT-DIETRICH LIBRARY, UNIVERSITY OF PENNSYLVANIA (*right*)

CHEMISTS Mendeleev (*left*) and Julius Lothar Meyer (*right*) developed the modern periodic chart almost simultaneously in the late 1860s. Mendeleev's table was published first, and he receives most of the credit for discovering the periodic system because he used it to make many successful predictions and vigorously defended its validity.

Around the same time, Mendeleev assembled his own periodic table while he, too, was writing a textbook of chemistry. Unlike his predecessors, Mendeleev had sufficient confidence in his periodic table to use it to predict several new elements and the properties of their compounds. He also corrected the atomic weights of some already known elements. Interestingly, Mendeleev admitted to having seen certain earlier tables, such as those of Newlands, but claimed to have been unaware of Lothar Meyer's work when developing his chart.

Although the predictive aspect of Mendeleev's table was a major advance, it seems to have been overemphasized by historians, who have generally suggested that Mendeleev's table was accepted especially because of this feature. These scholars have failed to notice that the citation from the Royal Society of London that accompanied the Davy Medal (which Mendeleev received in 1882) makes no mention whatsoever of his predictions. Instead Mendeleev's ability to accommodate the already known elements may have contributed as much to the acceptance of the periodic system as did his striking predictions. Although numerous scientists helped to develop the periodic system, Mendeleev receives most of the credit for discovering chemical periodicity because he elevated the discovery to a law of nature and spent the rest of his life boldly examining its consequences and defending its validity.

Defending the periodic table was no simple task—its accuracy was frequently challenged by subsequent discoveries. One notable occasion arose in 1894, when William Ramsay of University College London and Lord Rayleigh (John William Strutt) of the Royal Institution in London discovered the element argon; over the next few years, Ramsay announced the identification of four other elements—helium, neon, krypton and xenon—known as the noble gases. (The last of the known noble gases, radon, was discovered in 1900 by German physicist Friedrich Ernst Dorn.)

The name "noble" derives from the fact that all these gases seem to stand apart from the other elements, rarely interacting with them to form compounds. As a result, some chemists suggested that the noble gases did not even belong in the periodic table. These elements had not been predicted by Mendeleev or anyone else, and only after six years of intense effort could chemists and physicists successfully incorporate the noble gases into the table. In the new arrangement, an additional column was introduced between the halogens (the gaseous elements fluorine, chlorine, bromine, iodine and astatine) and the alkali metals (lithium, sodium, potassium, rubidium, cesium and francium).

A second point of contention surrounded the precise ordering of the elements. Mendeleev's original table positioned the elements according to atomic weight, but in 1913 Dutch amateur theoretical physicist Anton van den Broek suggested that the ordering principle for the periodic table lay instead in the nuclear charge of each atom. Physicist Henry Moseley, working at the University of Manchester, tested this hypothesis, also in 1913, shortly before his tragic death in World War I.

Moseley began by photographing the x-ray spectrum of 12 elements, 10 of which occupied consecutive places in the periodic table. He discovered that the frequencies of features called K-lines in the spectrum of each element were directly proportional to the squares of the integers representing the position of each successive element in the table. As Moseley put it, here was proof that "there is in the atom a fundamental quantity, which increases by regular steps as we pass from one element to the next." This fundamental quantity, first referred to as atomic number in 1920 by Ernest Rutherford, who was then at the University of Cambridge, is now identified as the number of protons in the nucleus.

Moseley's work provided a method that could be used to determine exactly how many empty spaces remained in the periodic table. After this discovery, chemists turned to using atomic number as the fundamental ordering principle for the periodic table, instead of atomic weight. This change resolved many of the lingering problems in the arrangement of the elements. For example, when iodine and tellurium were ordered according to atomic weight (with iodine first), the two elements appeared to be incorrectly positioned in terms of their chemical behavior. When ordered according to atomic number (with tellurium first), however, the two elements were in their correct positions.

Understanding the Atom

The periodic table inspired the work not only of chemists but also of atomic physicists struggling to understand the structure of the atom. In 1904, working at Cambridge, physicist J. J. Thomson (who also discovered the electron) developed a model of the atom, paying close attention to the periodicity of the elements. He proposed that the atoms of a particular element contained a specific number of electrons arranged in concentric rings. Furthermore, according to Thomson, elements with similar configurations of electrons would have similar properties; Thomson's work thus provided the first physical explanation for the periodicity of the elements. Although Thomson imagined the rings of electrons as lying inside the main body of the atom, rather than circulating around the nucleus as is believed today, his model does represent the first time anyone addressed the arrangement of electrons in the atom, a concept that pervades the whole of modern chemistry.

Danish physicist Niels Bohr, the first to bring quantum theory to bear on the structure of the atom, was also motivated by the arrangement of the elements in the periodic system. In Bohr's model of the atom, developed in 1913, electrons inhabit a series of concentric shells that encircle the nucleus. Bohr reasoned that elements in the same group of the periodic table might have identical configurations of electrons in their outermost shell and that the chemical properties of an element would depend in large part on the arrangement of electrons in the outer shell of its atoms.

Bohr's model of the atom also served to explain why the noble gases lack reactivity: noble gases possess full outer shells of electrons, making them unusually stable and unlikely to form compounds. Indeed, most other elements form compounds as a way to obtain full outer electron shells. More recent analysis of how Bohr arrived at these electronic configurations suggests that he functioned more like a chemist than has generally been credited. Bohr did not derive electron configurations from quantum theory but obtained them from the known chemical and spectroscopic properties of the elements.

In 1924 another physicist, Austrian-born Wolfgang Pauli, set out to explain the length of each row, or period, in the table. As a result, he developed the Pauli Exclusion Principle, which states that no two electrons can exist in exactly the same quantum state, which is defined by what scientists call quantum numbers. The lengths of the various periods emerge from experimental evidence about the order of electron-shell filling and from the quantum-mechanical restrictions on the four quantum numbers that electrons can adopt.

The modifications to quantum theory made by Werner Heisenberg and Erwin Schrödinger in the mid-1920s yielded quantum mechanics in essentially the form used to this day. But the influence of these changes on the periodic table has been rather minimal. Despite the efforts of many physicists and chemists, quantum mechanics cannot explain the periodic table any further. For example, it cannot explain from first principles the order in which electrons fill the various electron shells. The electronic configurations of atoms, on which our modern understanding of the periodic table is based, cannot be derived using quantum mechanics (this is because the fundamental equation of quantum mechanics, the Schrödinger equation, cannot be solved exactly for atoms other than hydrogen). As a result, quantum mechanics can only reproduce Mendeleev's origi-

POPULAR PERIODIC TABLE—known as the medium-long form—can be found in nearly every chemistry classroom and laboratory around the world. This version has the advantage of clearly displaying groups of elements that have similar chemical properties in vertical columns, but it is not particularly symmetrical. (The different colors of the table indicate elements with the same type of outer shell of electrons.)

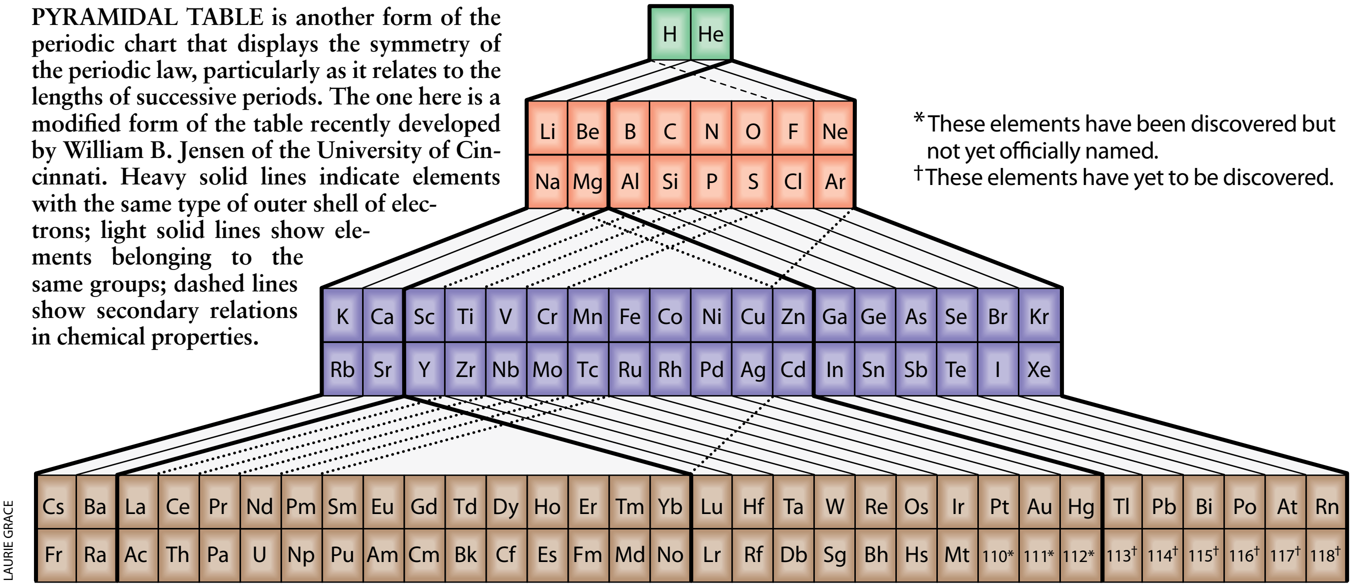
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110*	111*	112*	113†	114†	115†	116†	117†	118†

57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Td	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

*These elements have been discovered but not yet officially named.
†These elements have yet to be discovered.

LAURIE GRACE

PYRAMIDAL TABLE is another form of the periodic chart that displays the symmetry of the periodic law, particularly as it relates to the lengths of successive periods. The one here is a modified form of the table recently developed by William B. Jensen of the University of Cincinnati. Heavy solid lines indicate elements with the same type of outer shell of electrons; light solid lines show elements belonging to the same groups; dashed lines show secondary relations in chemical properties.



*These elements have been discovered but not yet officially named.
†These elements have yet to be discovered.

nal discovery by the use of mathematical approximations—it cannot predict the periodic system.

Variations on a Theme

In more recent times, researchers have proposed different approaches for displaying the periodic system. For instance, Fernando Dufour, a retired chemistry professor from Collège Ahuntsic in Montreal, has developed a three-dimensional periodic table, which displays the fundamental symmetry of the periodic law, unlike the common two-dimensional form of the table in common use. The same virtue is also seen in a version of the periodic table shaped as a pyramid, a form suggested on many occasions but most recently refined by William B. Jensen of the University of Cincinnati [*see illustration above*].

Another departure has been the invention of periodic systems aimed at summarizing the properties of compounds rather than elements. In 1980 Ray Hefferlin of Southern Adventist University in Collegedale, Tenn., devised a periodic system for all the conceivable diatomic molecules that could be formed between the first 118 elements (only 112 have been discovered to date).

Hefferlin's chart reveals that certain properties of molecules—the distance between atoms and the energy required to ionize the molecule, for instance—occur in regular patterns. This table has enabled scientists to predict the properties of diatomic molecules successfully.

In a similar effort, Jerry R. Dias of the University of Missouri at Kansas City devised a periodic classification of a type of organic molecule called benzenoid aromatic hydrocarbons. The compound naphthalene ($C_{10}H_8$), found in mothballs, is the simplest example. Dias's classification system is analogous to Döbereiner's triads of elements: any central molecule of a triad has a total number of carbon and hydrogen atoms that is the mean of the flanking entries, both downward and across the table. This scheme has been applied to a systematic study of the properties of benzenoid aromatic hydrocarbons and, with the use of graph theory, has led to predictions of the stability and reactivity of some of these compounds.

Still, it is the periodic table of the elements that has had the widest and most enduring influence. After evolving for over 200 years through the work of many people, the periodic table remains at the heart of the study of chemistry. It ranks as one of the most fruitful ideas in modern science, comparable perhaps to Charles Darwin's theory of evolution. Unlike theories such as Newtonian mechanics, it has not been falsified or revolutionized by modern physics but has adapted and matured while remaining essentially unscathed. SA

The Author

ERIC R. SCERRI studied chemistry at the universities of London, Cambridge and Southampton. He later earned a Ph.D. in history and philosophy of science from King's College, London. Since arriving in the U.S., he has been a research fellow at the California Institute of Technology and, most recently, has taught chemistry at Bradley University in Peoria, Ill. In January 1999 he will take up a position in the chemistry department at Purdue University. Scerri's research interests lie in the history and philosophy of chemistry, and he is the editor of the interdisciplinary journal *Foundations of Chemistry* (http://www.cco.caltech.edu/~scerri/ on the World Wide Web). His e-mail address is scerri@bradley.edu

Further Reading

The Periodic System of Chemical Elements: A History of the First Hundred Years. J. W. van Spronsen. Elsevier, 1969.

The Surprising Periodic Table: Ten Remarkable Facts. Dennis H. Rouvray in *Chemical Intelligencer,* Vol. 2, No. 3, pages 39–47; July 1996.

Classification, Symmetry and the Periodic Table. William B. Jensen in *Computing and Mathematics with Applications,* Vol. 12B, Nos. 1–2, pages 487–510; 1989.

Plus ça Change. E. R. Scerri in *Chemistry in Britain,* Vol. 30, No. 5, pages 379–381; May 1994. (Also see related articles in this issue.)

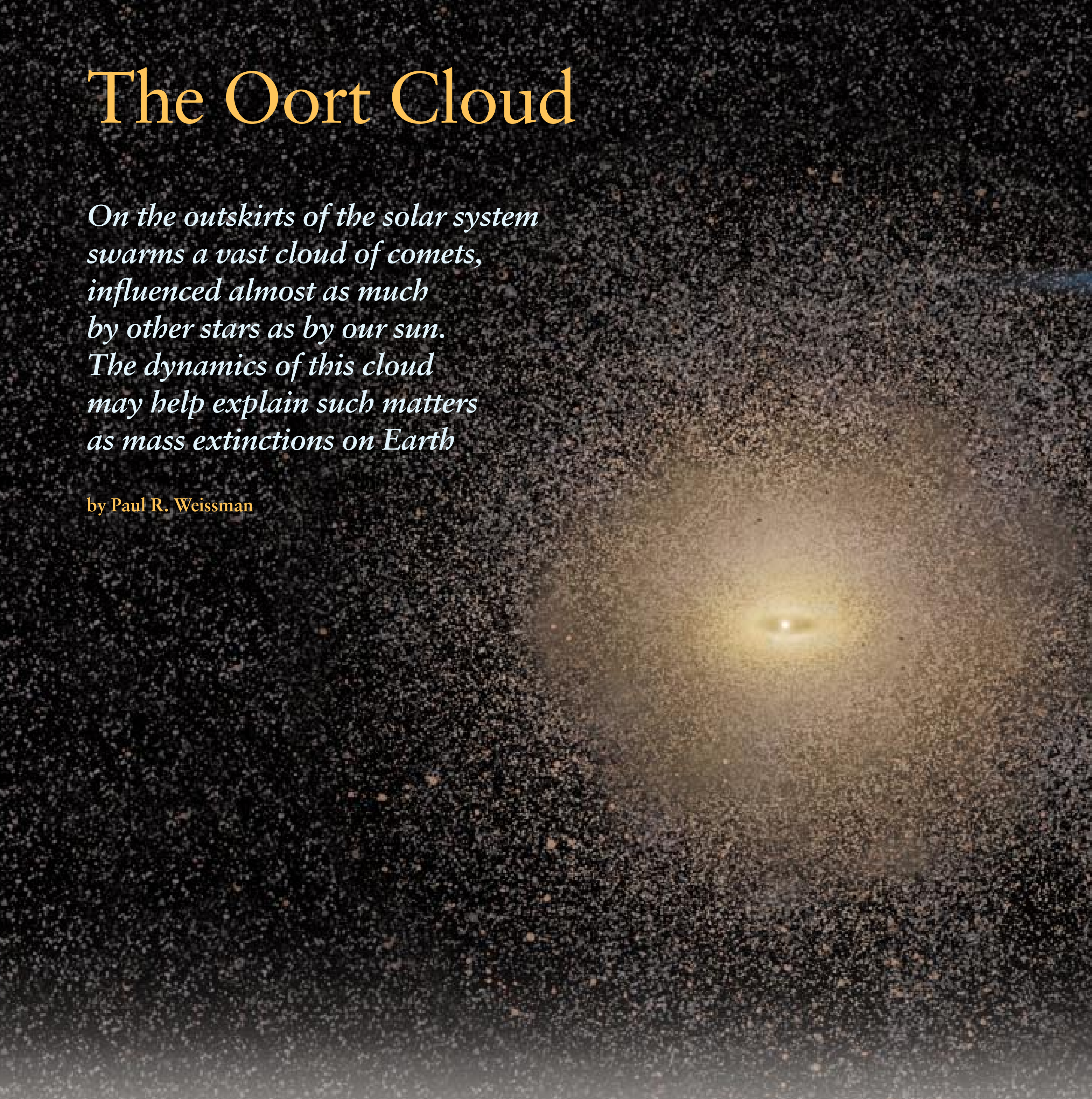
The Electron and the Periodic Table. Eric R. Scerri in *American Scientist,* Vol. 85, pages 546–553; November–December 1997.

For information on Fernando Dufour's three-dimensional periodic table, send e-mail to fernduf@aei.ca

The Oort Cloud

On the outskirts of the solar system swarms a vast cloud of comets, influenced almost as much by other stars as by our sun. The dynamics of this cloud may help explain such matters as mass extinctions on Earth

by Paul R. Weissman

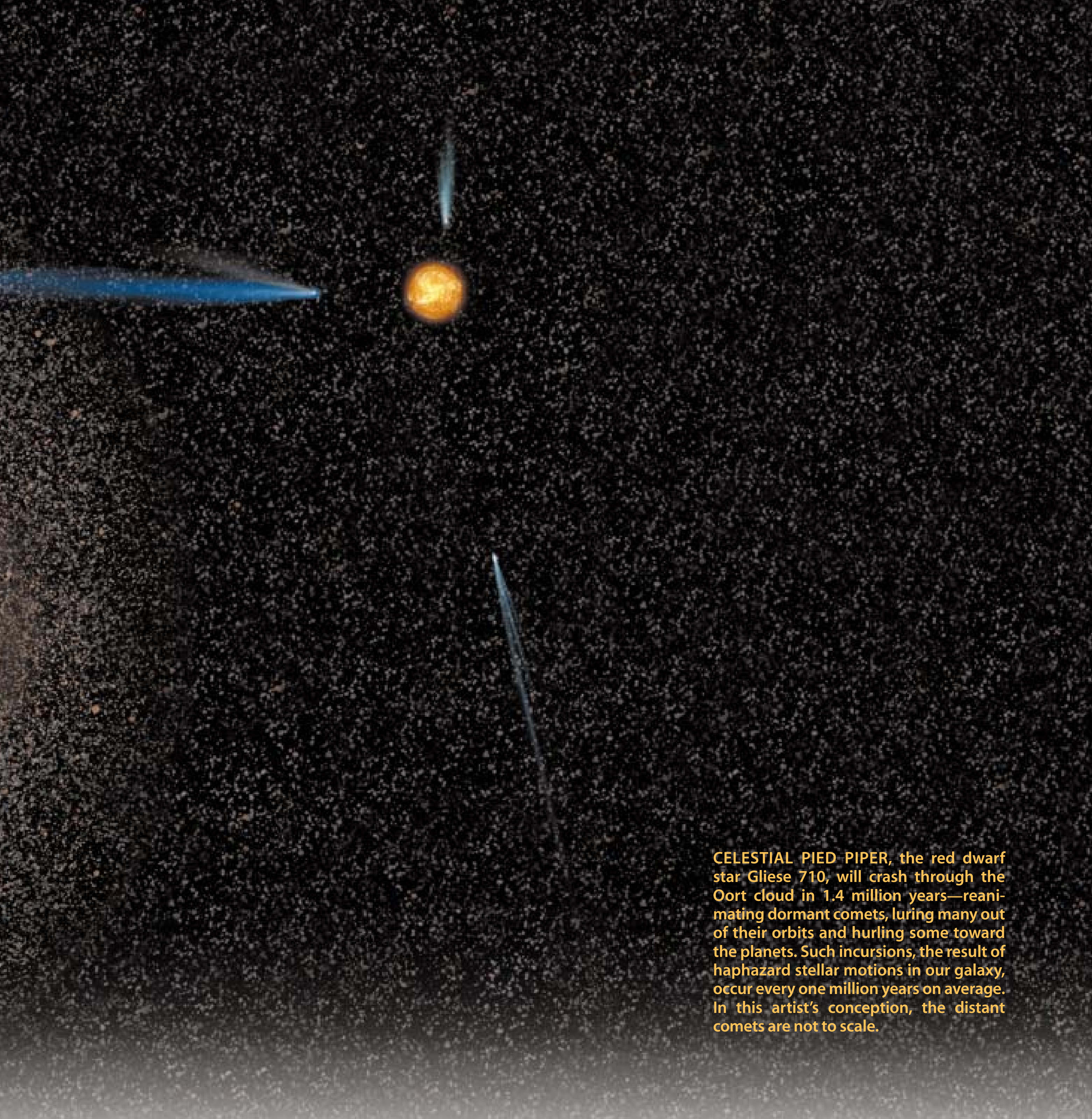


It is common to think of the solar system as ending at the orbit of the most distant known planet, Pluto. But the sun's gravitational influence extends more than 3,000 times farther, halfway to the nearest stars. And that space is not empty—it is filled with a giant reservoir of comets, leftover material from the formation of the solar system. That reservoir is called the Oort cloud.

The Oort cloud is the Siberia of the solar system, a vast, cold frontier filled with exiles of the sun's inner empire and only barely under the sway of the central authority. Typical noontime temperatures are a frigid four degrees Celsius above absolute zero, and neighboring comets are typically tens of millions of kilometers apart. The sun, while still the brightest star in the sky, is only about as bright as Venus in the evening sky on Earth.

We have never actually "seen" the Oort cloud. But no one has ever seen an electron, either. We infer the existence and properties of the Oort cloud and the electron from the physical effects we can observe. In the case of the former, those effects are the steady trickle of long-period comets into the planetary system. The existence of the Oort cloud answers questions that people have asked since antiquity: What are comets, and where do they come from?

Aristotle speculated in the fourth century B.C. that comets were clouds of lu-



CELESTIAL PIED PIPER, the red dwarf star Gliese 710, will crash through the Oort cloud in 1.4 million years—reanimating dormant comets, luring many out of their orbits and hurling some toward the planets. Such incursions, the result of haphazard stellar motions in our galaxy, occur every one million years on average. In this artist's conception, the distant comets are not to scale.

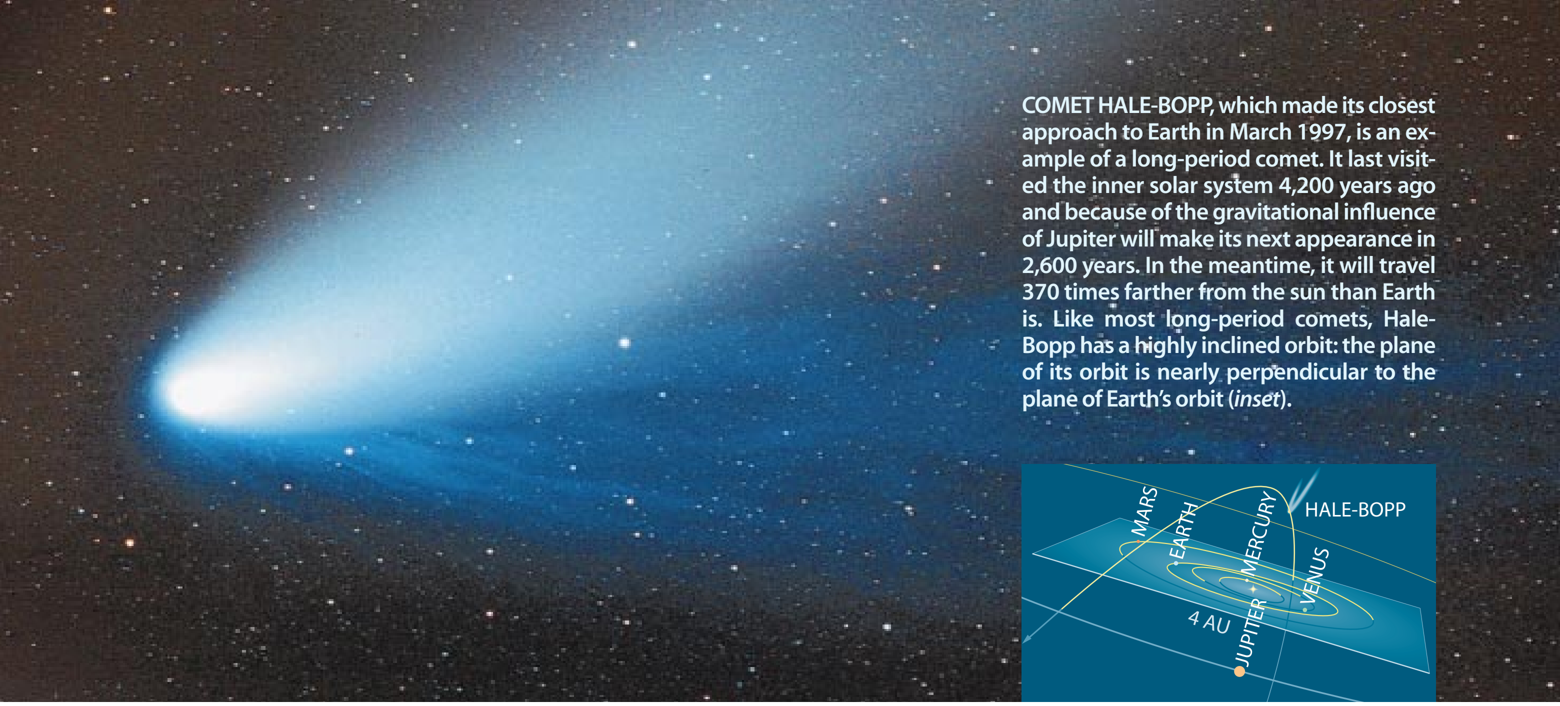
DON DIXON

minous gas high in Earth's atmosphere. But the Roman philosopher Seneca suggested in the first century A.D. that they were heavenly bodies, traveling along their own paths through the firmament. Fifteen centuries passed before his hypothesis was confirmed by Danish astronomer Tycho Brahe, who compared observations of the comet of 1577 made from several different locations in Europe. If the comet had been close by, then from each location it would have had a slightly different position against the stars. Brahe could not detect any differences and concluded that the comet was farther away than the moon.

Just how much farther started to become clear only when astronomers began determining the comets' orbits. In 1705 the English astronomer Edmond Halley compiled the first catalogue of 24 comets. The observations were fairly crude, and Halley could fit only rough parabolas to each comet's path. Nevertheless, he argued that the orbits might be very long ellipses around the sun:

> For so their Number will be determinate and, perhaps, not so very great. Besides, the Space between the Sun and the fix'd Stars is so immense that there is Room enough for a Comet to revolve, tho' the Period of its Revolution be vastly long.

In a sense, Halley's description of com-



COMET HALE-BOPP, which made its closest approach to Earth in March 1997, is an example of a long-period comet. It last visited the inner solar system 4,200 years ago and because of the gravitational influence of Jupiter will make its next appearance in 2,600 years. In the meantime, it will travel 370 times farther from the sun than Earth is. Like most long-period comets, Hale-Bopp has a highly inclined orbit: the plane of its orbit is nearly perpendicular to the plane of Earth's orbit (*inset*).

DENNIS DICICCO *Sky & Telescope*; MICHAEL GOODMAN (*inset*)

ets circulating in orbits stretching between the stars anticipated the discovery of the Oort cloud two and a half centuries later. Halley also noticed that the comets of 1531, 1607 and 1682 had very similar orbits and were spaced at roughly 76-year intervals. These seemingly different comets, he suggested, were actually the same comet returning at regular intervals. That body, now known as Halley's comet, last visited the region of the inner planets in 1986.

Since Halley's time, astronomers have divided comets into two groups according to the time it takes them to orbit the sun (which is directly related to the comets' average distance from the sun). Long-period comets, such as the recent bright comets Hyakutake and Hale-Bopp, have orbital periods greater than 200 years; short-period comets, less than 200 years. In the past decade astronomers have further divided the short-period comets into two groups: Jupiter-family comets, such as comets Encke and Tempel 2, which have periods less than 20 years; and intermediate-period, or Halley-type, comets, with periods between 20 and 200 years.

These definitions are somewhat arbitrary but reflect real differences. The intermediate- and long-period comets enter the planetary region randomly from all directions, whereas the Jupiter-family comets have orbits whose planes are typically inclined no more than 40 degrees from the ecliptic plane, the plane of Earth's orbit. (The orbits of the other planets are also very close to the ecliptic plane.) The intermediate- and long-period comets appear to come from the Oort cloud, whereas the Jupiter-family comets are now thought to originate in the Kuiper belt, a region in the ecliptic beyond the orbit of Neptune [see "The Kuiper Belt," by Jane X. Luu and David C. Jewitt; SCIENTIFIC AMERICAN, May 1996].

The Netherworld beyond Pluto

By the early 20th century, enough long-period cometary orbits were available to study their statistical distribution [*see illustration on page 89*]. A problem emerged. About one third of all the "osculating" orbits—that is, the orbits the comets were following at the point of their closest approach to the sun—were hyperbolic. Hyperbolic orbits would originate in and return to interstellar space, as opposed to elliptical orbits, which are bound by gravity to the sun. The hyperbolic orbits led some astronomers to suggest that comets were captured from interstellar space by encounters with the planets.

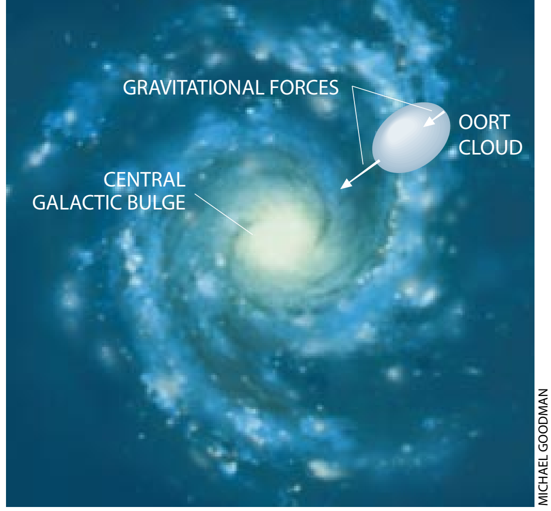
To examine this hypothesis, celestial-mechanics researchers extrapolated, or "integrated," the orbits of the long-period comets backward in time. They found that because of distant gravitational tugs from the planets, the osculating orbits did not represent the comets' original orbits [*see bottom illustration on opposite page*]. When the effects of the planets were accounted for—by integrating far enough back in time and orienting the orbits not in relation to the sun but in relation to the center of mass of the solar system (the sum of the sun and all the planets)—almost all the orbits became elliptical. Thus, the comets were members of the solar system, rather than interstellar vagabonds.

In addition, although two thirds of these orbits still appeared to be uniformly distributed, fully one third had orbital energies that fell within a narrow spike. That spike represented orbits that extend to very large distances—20,000 astronomical units (20,000 times the distance of Earth from the sun) or more. Such orbits have periods exceeding one million years.

Why were so many comets coming from so far away? In the late 1940s Dutch astronomer Adrianus F. van Woerkom showed that the uniform distribution could be explained by planetary perturbations, which scatter comets randomly to both larger and smaller orbits. But what about the spike of comets with million-year periods?

In 1950 Dutch astronomer Jan H. Oort, already famous for having determined the rotation of the Milky Way galaxy in the 1920s, became interested in the problem. He recognized that the million-year spike must represent the source of the long-period comets: a vast spherical cloud surrounding the planetary system and extending halfway to the nearest stars.

Oort showed that the comets in this cloud are so weakly bound to the sun that random passing stars can readily change their orbits. About a dozen stars pass within one parsec (206,000 astronomical units) of the sun every one million years. These close encounters are enough to stir the cometary orbits, randomizing their inclinations and sending



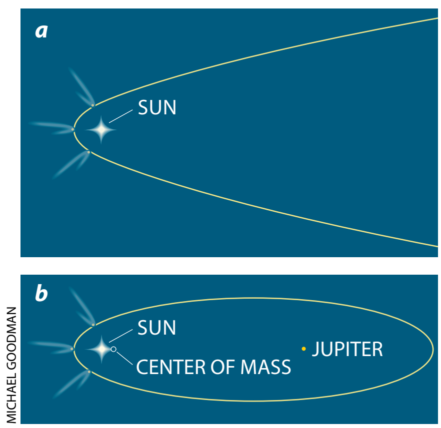
TIDAL FORCES arise because gravity becomes weaker with distance. Therefore, the central bulge of our galaxy—a concentration of stars at the hub of the spiral pattern—pulls more on the near side of the Oort cloud (*not to scale*) than on the far side. The galactic plane exerts a similar force in a different direction. The galactic tides are analogous to lunar tides, which arise because the side of Earth closest to the moon feels a stronger gravitational pull than the antipode does.

a steady trickle of comets into the inner solar system on very long elliptical orbits [*see illustrations on next page*]. As they enter the planetary system for the first time, the comets are scattered by the planets, gaining or losing orbital energy. Some escape the solar system altogether. The remainder return and are observed again as members of the uniform distribution. Oort described the cloud as "a garden, gently raked by stellar perturbations."

A few comets still appeared to come from interstellar space. But this was probably an incorrect impression given by small errors in the determination of their orbits. Moreover, comets can shift their orbits because jets of gas and dust from their icy surfaces act like small rocket engines as the comets approach the sun. Such nongravitational forces can make the orbits appear hyperbolic when they are actually elliptical.

Shaken, Not Stirred

Oort's accomplishment in correctly interpreting the orbital distribution of the long-period comets is even more impressive when one considers that he had only 19 well-measured orbits to work with. Today astronomers have more than 15 times as many. They now know that long-period comets entering the planetary region for the first time come from an average distance of 44,000 astronomical units. Such orbits have periods of 3.3 million years.



LONG-PERIOD COMET is so weakly bound to the sun that the planets have a decisive influence on it. Astronomers can usually see the comet only while it swings by the sun. When they apply Kepler's laws of celestial motion to plot its course—its "osculating," or apparent, orbit—the comet often seems to be on a hyperbolic trajectory, implying that it came from interstellar space and will return there (*a*). A more sophisticated calculation, which accounts for the planets (especially the most massive planet, Jupiter), finds that the orbit is actually elliptical (*b*). The orbit changes shape on each pass through the inner solar system.

Astronomers have also realized that stellar perturbations are not always gentle. Occasionally a star comes so close to the sun that it passes right through the Oort cloud, violently disrupting the cometary orbits along its path. Statistically a star is expected to pass within 10,000 astronomical units of the sun every 36 million years and within 3,000 astronomical units every 400 million years. Comets close to the star's path are thrown out to interstellar space, while the orbits of comets throughout the cloud undergo substantial changes.

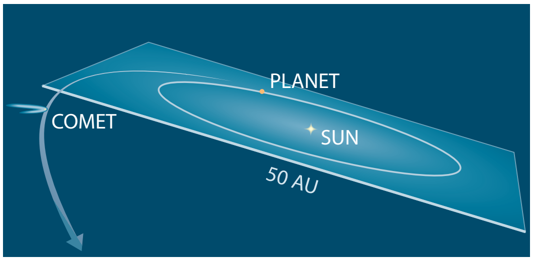
Although close stellar encounters have no direct effect on the planets—the closest expected approach of any star over the history of the solar system is 900 astronomical units from the sun—they might have devastating indirect consequences. In 1981 Jack G. Hills, now at Los Alamos National Laboratory, suggested that a close stellar passage could send a "shower" of comets toward the planets, raising the rate of cometary impacts on the planets and possibly even causing a biological mass extinction on Earth. According to computer simulations I performed in 1985 with Piet Hut, then at the Institute for Advanced Study in Princeton, N.J., the frequency of comet passages during a shower could reach 300 times the normal rate. The shower would last two to three million years.

Recently Kenneth A. Farley and his colleagues at the California Institute of Technology found evidence for just such a comet shower. Using the rare helium 3 isotope as a marker for extraterrestrial material, they plotted the accumulation of interplanetary dust particles in ocean sediments over time. The rate of dust accumulation is thought to reflect the number of comets passing through the planetary region; each comet sheds dust along its path. Farley discovered that this rate increased sharply at the end of the Eocene epoch, about 36 million years ago, and decreased slowly over two to three million years, just as theoretical models of comet showers would predict. The late Eocene is identified with a moderate biological extinction event, and several impact craters have been dated to this time. Geologists have also found other traces of impacts in terrestrial sediments, such as iridium layers and microtektites.

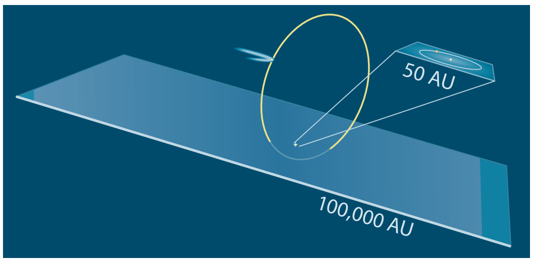
Is Earth in danger of a comet shower now? Fortunately not. Joan Garcia-Sanchez of the University of Barcelona, Robert A. Preston and Dayton L. Jones of the Jet Propulsion Laboratory in Pasadena, Calif., and I have been using the positions and velocities of stars, measured by the Hipparcos satellite, to reconstruct the trajectories of stars near the solar system. We have found evidence that a star has passed close to the sun in the past one million years. The next close passage of a star will occur in 1.4 million years, and that is a small red dwarf called Gliese 710, which will pass through the outer Oort cloud about 70,000 astronomical units from the sun. At that distance, Gliese 710 might increase the frequency of comet passages through the inner solar system by 50 percent—a sprinkle perhaps, but certainly no shower.

In addition to random passing stars, the Oort cloud is now known to be disturbed by two other effects. First, the cloud is sufficiently large that it feels tidal forces generated by the disk of the Milky Way and, to a lesser extent, the galactic core. These tides arise because the sun and a comet in the cloud are at slightly different distances from the midplane of the disk or from the galactic center and thus feel a slightly different gravitational tug [*see top illustration on preceding page*]. The tides help to feed new long-period comets into the planetary region.

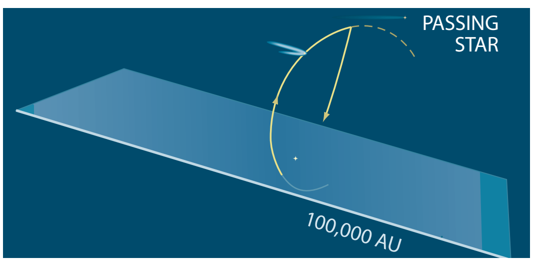
Second, giant molecular clouds in the galaxy can perturb the Oort cloud, as Ludwig Biermann of the Max Planck Institute for Physics and Astrophysics in Munich suggested in 1978. These massive clouds of cold hydrogen, the birthplaces of stars and planetary systems, are 100,000 to one million times as massive as the sun. When the solar system comes close to one, the gravitational perturbations rip comets from their orbits and fling them into interstellar space. These encounters, though violent, are infrequent—only once every 300 million to 500 million years. In 1985 Hut and Scott D. Tremaine, now at Princeton University, showed that over the history of the solar system, molecular clouds have had about the same cumulative effect as all passing stars.



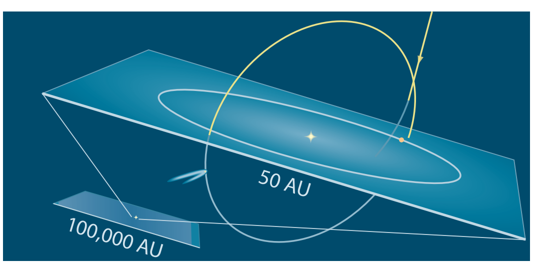
HISTORY OF A LONG-PERIOD COMET begins when it forms near the planets and is catapulted by them into a wide orbit.



There the comet is susceptible to the gravitational forces of random passing stars and giant molecular clouds, as well as the tidal forces of the galactic disk and core. These forces randomly tilt the orbital plane of the comet and gradually pull it farther out.



Beyond a distance of about 20,000 astronomical units (20,000 times the Earth-sun distance), the various outside influences are capable of throwing the comet back toward the planets.



MICHAEL GOODMAN

Once the comet reenters the inner solar system, the planets may pull it to a new orbit, so that it reappears on a regular basis.

Inner Core

Currently three main questions concern Oort-cloud researchers. First, what is the cloud's structure? In 1987 Tremaine, Martin J. Duncan, now at Queen's University in Ontario, and Thomas R. Quinn, now at the University of Washington, studied how stellar and molecular-cloud perturbations redistribute comets within the Oort cloud. Comets at its outer edge are rapidly lost, either to interstellar space or to the inner solar system, because of the perturbations. But deeper inside, there probably exists a relatively dense core that slowly replenishes the outer reaches.

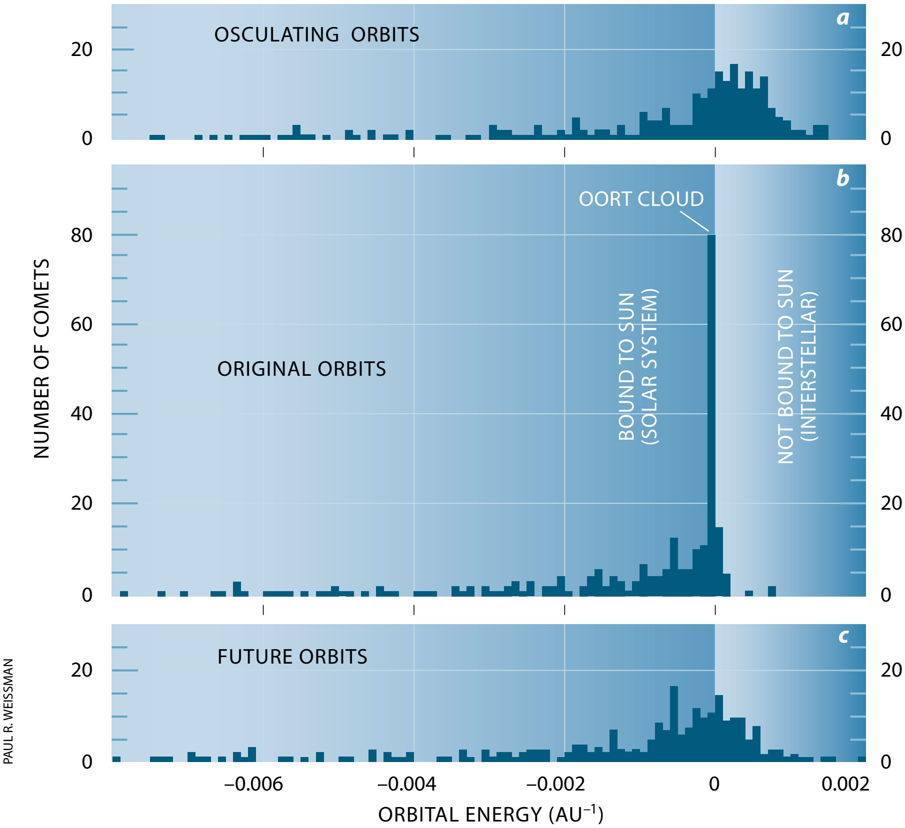
Tremaine, Duncan and Quinn also showed that as comets fall in from the Oort cloud, their orbital inclinations tend not to change. This is a major reason why astronomers now think the Kuiper belt, rather than the Oort cloud, accounts for the low-inclination, Jupiter-family comets. Still, the Oort cloud is the most likely source of the higher-inclination, intermediate-period comets, such as Halley and Swift-Tuttle. They were probably once long-period comets that the planets pulled into shorter-period orbits.

The second main question is, How many comets inhabit the Oort cloud? The number depends on how fast comets leak from the cloud into interplanetary space. To account for the observed number of long-period comets, astronomers now estimate the cloud has six trillion comets, making Oort-cloud comets the most abundant substantial bodies in the solar system. Only a sixth of them are in the outer, dynamically active cloud first described by Oort; the remainder are in the relatively dense core. If the best estimate for the average mass of a comet—about 40 billion metric tons—is applied, the total mass of comets in the Oort cloud at present is about 40 times that of Earth.

Finally, from where did the Oort-cloud comets originally come? They could not have formed at their current position, because material at those distances is too sparse to coalesce. Nor could they have originated in interstellar space; capture of comets by the sun is very inefficient. The only place left is the planetary system. Oort speculated that the comets were created in the asteroid belt and ejected by the giant planets during the formation of the solar system. But comets are icy bodies, essentially big, dirty snowballs, and the asteroid belt was too warm for ices to condense.

A year after Oort's 1950 paper, astronomer Gerard P. Kuiper of the University of Chicago proposed that comets coalesced farther from the sun, among the giant planets. (The Kuiper belt is named for him because he suggested that some comets also formed beyond the farthest planetary orbits.) Comets probably originated throughout the giant planets' region, but researchers used to argue that those near Jupiter and Saturn, the two most massive planets, would have been ejected to interstellar space rather than to the Oort cloud. Uranus and Neptune, with their lower masses, could not easily throw so many comets onto escape trajectories. But more recent dynamical studies have cast some doubt on this scenario. Jupiter and particularly Saturn do place a significant fraction of their comets into the Oort cloud. Although this fraction may be smaller than that of Uranus and Neptune, it may have been offset by the greater amount of material initially in the larger planets' zones.

Therefore, the Oort-cloud comets may have come from a wide range of solar distances and hence a wide range of formation temperatures. This fact may help explain some of the compositional diversity observed in comets. Indeed, recent work I have done with



ORBITAL ENERGY of known long-period comets, as shown in these histograms, reveals the Oort cloud. Astronomers first calculate the osculating orbits of the comets—the orbits they would take if their motion were entirely caused by the sun's gravity. One third of these orbits have a positive energy, making them appear interstellar (*a*). But when corrected for the influence of the planets and extrapolated backward in time, the energy is slightly negative—indicating that the comets came from the edge of the solar system (*b*). A few comets still seem to be interstellar, but this is probably the result of small observational errors. As the planets continue to exert their influence, some comets will return to the Oort cloud, some will escape from the solar system and the rest will revisit the inner solar system (*c*). Technically, the orbital energy is proportional to the reciprocal of the semimajor axis, expressed in units of inverse astronomical units (AU^{-1}).

Harold F. Levison of the Southwest Research Institute in Boulder, Colo., has shown that the cloud may even contain asteroids from the inner planets' region. These objects, made of rock rather than ice, may constitute 2 to 3 percent of the total Oort-cloud population.

The key to these ideas is the presence of the giant planets, which hurl the comets outward and modify their orbits if they ever reenter the planetary region. If other stars have giant planets, as observations over the past few years suggest, they may have Oort clouds, too. If each star has its own cloud, then as stars pass by the sun, their Oort clouds will pass through our cloud. Even so, collisions between comets will be rare because the typical space between comets is an astronomical unit or more.

The Oort clouds around each star may slowly be leaking comets into interstellar space. These interstellar comets should be easily recognizable if they were to pass close to the sun, because they would approach the solar system at much higher velocities than the comets from our own Oort cloud. To date, no such interstellar comets have ever been detected. This fact is not surprising: because the solar system is a very small target in the vastness of interstellar space, there is at best a 50–50 chance that people should have seen one interstellar comet by now.

The Oort cloud continues to fascinate astronomers. Through the good fortunes of celestial mechanics, nature has preserved a sample of material from the formation of the solar system in this distant reservoir. By studying it and the cosmo-chemical record frozen in its icy members, researchers are learning valuable clues about the origin of the solar system.

Several space missions are now being readied to unlock those secrets. The Stardust spacecraft, due for launch next year, will fly through the coma of comet Wild 2, collect samples of cometary dust and return them to Earth for laboratory analysis. A few years later the CONTOUR probe will fly by three comets and compare their compositions. The Deep Space 4/Champollion mission will send an orbiter and a lander to comet Tempel 1, and the Rosetta mission will do the same for comet Wirtanen. The new millennium is going to be a wonderful time for studying comets. SA

The Author

PAUL R. WEISSMAN is a senior research scientist at the Jet Propulsion Laboratory in Pasadena, Calif., where he specializes in studies of the physics and dynamics of comets. He is also the project scientist for NASA's Deep Space 4/Champollion mission, which is scheduled to land on short-period comet Tempel 1 in 2005. Weissman has written more than 80 refereed scientific papers and is one of three editors of the *Encyclopedia of the Solar System,* to be published by Academic Press this fall.

Further Reading

COMETS: A CHRONOLOGICAL HISTORY OF OBSERVATION, SCIENCE, MYTH AND FOLKLORE. Donald K. Yeomans. John Wiley & Sons, 1991.

COMETS IN THE POST-HALLEY ERA. Edited by Ray L. Newburn, Jr., Marcia Neugebauer and Jurgen Rahe. Kluwer Academic Publishers, 1991.

DYNAMICS OF COMETS: RECENT DEVELOPMENTS AND NEW CHALLENGES. Julio A. Fernandez in *Asteroids, Comets, Meteors 1993: Proceedings of the 160th Symposium of the International Astronomical Union.* Edited by A. Milani, M. Di Martino and A. Cellino. Kluwer Academic Publishers, 1994.

COMPLETING THE INVENTORY OF THE SOLAR SYSTEM. Edited by Terrence W. Rettig and Joseph M. Hahn. Astronomical Society of the Pacific Conference Series, Vol. 107; 1996.

Thermophotovoltaics

Semiconductors that convert radiant heat to electricity may prove suitable for lighting remote villages or powering automobiles

by Timothy J. Coutts and Mark C. Fitzgerald

Photovoltaics is a technology that typically transforms sunlight into electricity. Radiation from the visible part of the spectrum is, after all, abundant, nonpolluting and free. But photovoltaics can also provide useful amounts of electricity from infrared radiation—that is, radiant heat generated by a source of energy such as fuel oil.

This lesser-known approach, called thermophotovoltaics, offers a major advantage in certain settings: a generator can operate at night or when the sky is overcast, thereby eliminating any need for batteries to store electricity. The technology is also preferable in some ways to conventional electricity-generating technology based on burning fossil fuels. Its efficiency—the percent of fuel energy converted to electricity—can be substantially higher than that of electric generators powered by natural gas or another fossil fuel. Moreover, a semiconductor-based thermophotovoltaic system can be designed to minimize pollutants. And because it contains no moving parts, it will run silently and reliably, requiring little maintenance.

In spite of these advantages, thermophotovoltaic technology has not enjoyed the success of solar photovoltaics, which today constitutes a thriving, albeit specialized, segment of the energy market. But the divide may soon shrink. The technology, which has its roots in the same research that produced solar cells, appears to be coming into its own.

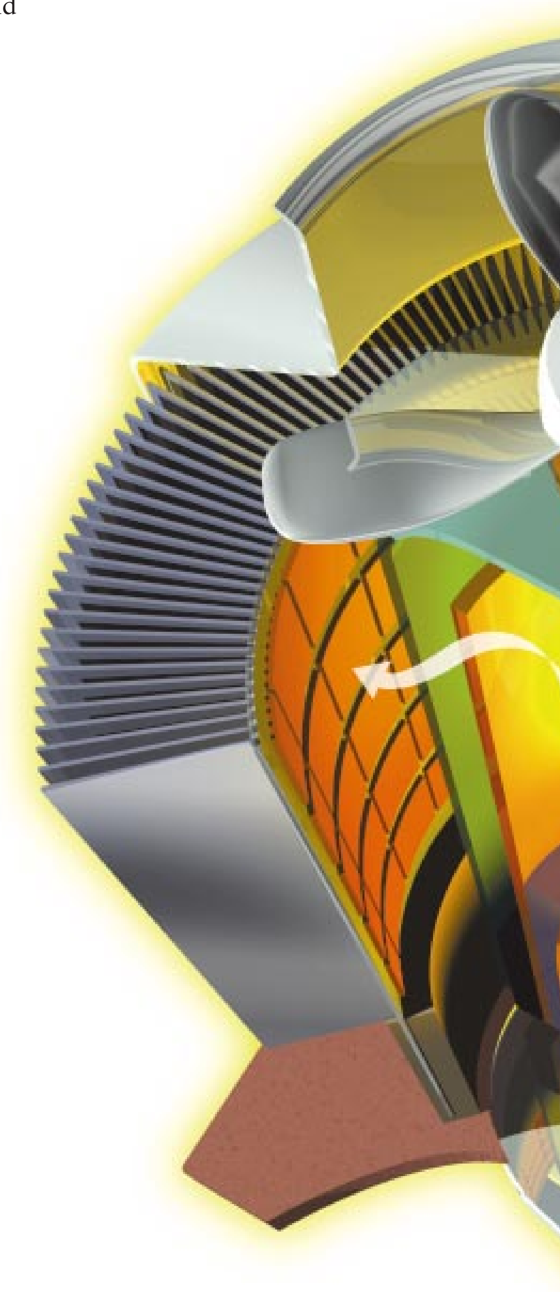
The ideas that underlie thermophotovoltaics stretch back 40 years. Pierre R. Aigrain of the École Normale Supérieure in Paris first described some of the basic concepts during a series of lectures in 1956. In the early 1960s U.S. Army researchers at Fort Monmouth, N.J., created the first documented prototype of a thermophotovoltaic generator. Its efficiency was less than 1 percent, however. Efficiencies of 10 to 15 percent were needed for a usable generator that could be deployed by troops operating in the field.

Throughout the late 1970s and into the early 1980s, research supported and carried out by the Electric Power Research Institute in Palo Alto, Calif., the Gas Research Institute in Chicago, Stanford University and other organizations achieved some improvement in performance. But the components of early systems could never channel enough heat to the units that convert infrared energy into electricity. Recently a new set of materials has taken the technology past the development stage.

Commercial Debut

Thermophotovoltaics is about to reach the commercial marketplace. A company in the Pacific Northwest plans to market a thermophotovoltaic generator to run electrical equipment on sailboats. Other applications under development include small power units that would supply electricity in remote areas or for roving military troops. The technology could also assist in running hybrid electric vehicles, in which an electric battery complements the power from an internal-combustion engine. Ultimately, thermophotovoltaics could produce megawatts of power, furnishing some of the requirements of utilities or helping industry meet its electricity needs by recovering the unused heat from industrial processes.

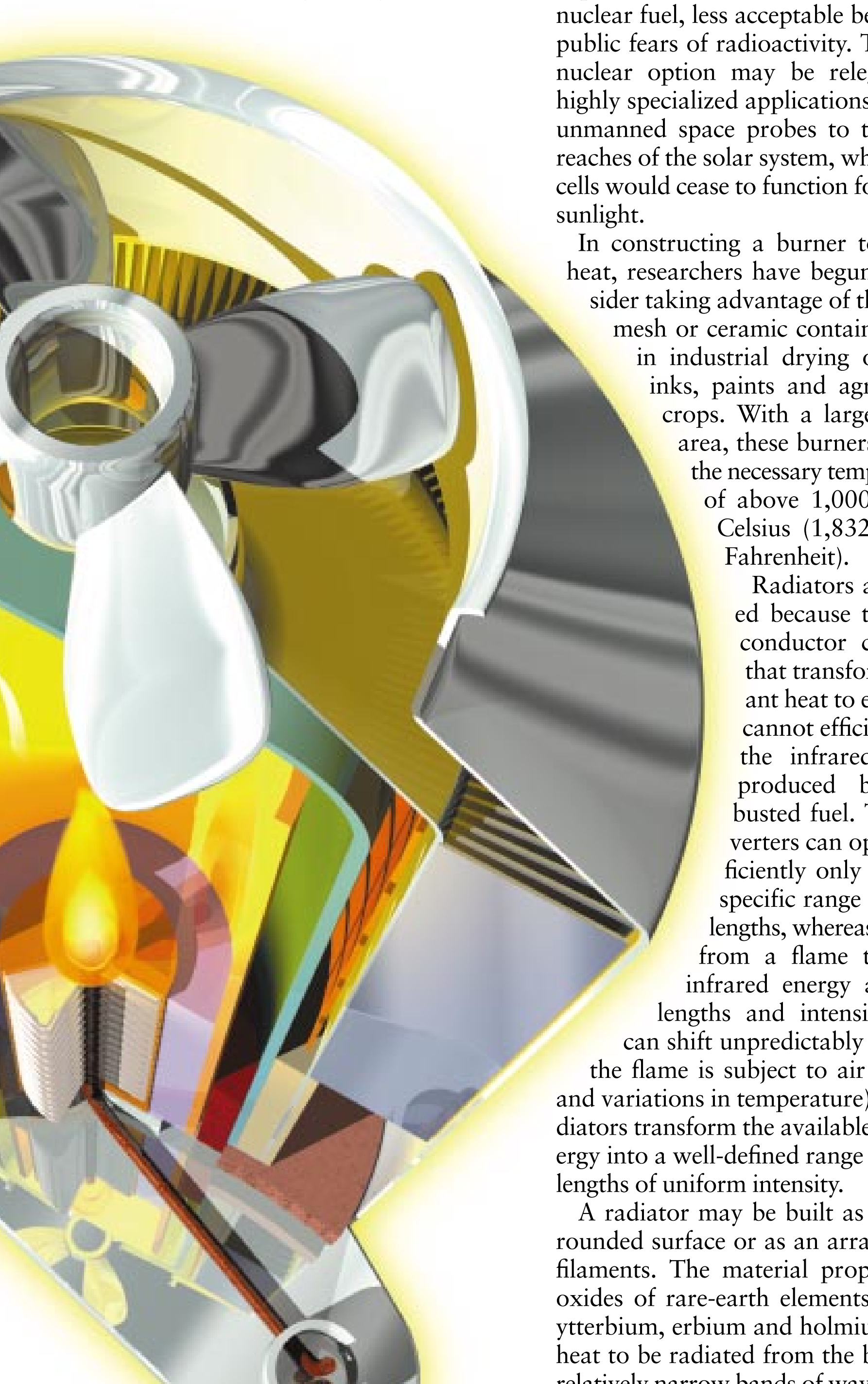
Production of electricity from radiant heat requires several functional elements. A source of heat must be coupled with a radiator, a material that emits desired wavelengths of infrared radiation. A semiconductor device consisting of an array of interconnected cells must be engineered to convert these se-



THERMOPHOTOVOLTAIC GENERATOR turns radiant heat into electricity. The burner (*center*) generates heat, whose infrared photons hit a radiator (*green enclosure*). The radiator channels certain infrared wavelengths to arrays of photovoltaic cells (*grid*) that convert the energy to electricity.

lect wavelengths to electricity, which is then sent through a circuit to perform useful work—running a refrigerator on a boat, for instance. Finally, to perform efficiently, a thermophotovoltaic system must rechannel unused energy back to the radiator. In some applications the waste energy can also be used for other purposes, such as heating a room.

Sources of heat for a thermophotovoltaic system might range from fossil-fuel burners to sunlight to a nuclear fission reaction. As a practical matter, most systems under development deploy fossil fuels. Solar energy boosted to high intensities by "concentrator" devices can be used to operate a thermophotovoltaic generator, but the room-size concentrators and devices needed to store heat for nighttime use are still in the early development stages. Many experts consider the other alternative, nuclear fuel, less acceptable because of public fears of radioactivity. Thus, the nuclear option may be relegated to highly specialized applications, such as unmanned space probes to the outer reaches of the solar system, where solar cells would cease to function for lack of sunlight.



SLIM FILMS

In constructing a burner to supply heat, researchers have begun to consider taking advantage of the metal-mesh or ceramic containers used in industrial drying of paper, inks, paints and agricultural crops. With a large surface area, these burners achieve the necessary temperatures of above 1,000 degrees Celsius (1,832 degrees Fahrenheit).

Radiators are needed because the semiconductor converter that transforms radiant heat to electricity cannot efficiently use the infrared energy produced by combusted fuel. The converters can operate efficiently only within a specific range of wavelengths, whereas the heat from a flame transmits infrared energy at wavelengths and intensities that can shift unpredictably (because the flame is subject to air currents and variations in temperature). The radiators transform the available heat energy into a well-defined range of wavelengths of uniform intensity.

A radiator may be built as a flat or rounded surface or as an array of tiny filaments. The material properties of oxides of rare-earth elements such as ytterbium, erbium and holmium allow heat to be radiated from the burner in relatively narrow bands of wavelengths; silicon carbide, in contrast, emits a broader spectrum.

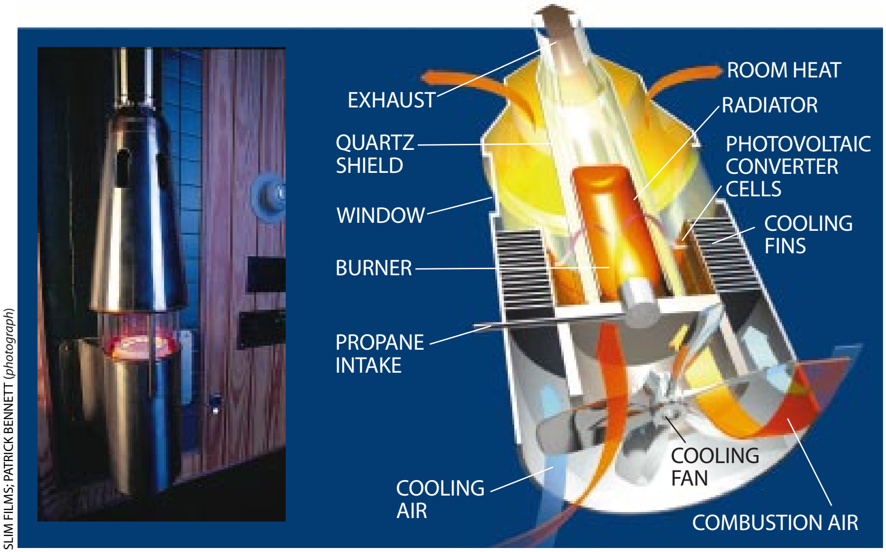
In constructing a converter, thermophotovoltaic researchers choose semiconductors that match the infrared spectrum of the wavelengths emitted by a radiator. These wavelengths correspond roughly to the energy needed to free electrons that would otherwise remain bound within the crystalline solid of the semiconductor converter [*see box on page 93*]. The most energetic immobile electrons reside in the so-called valence band of the semiconductor crystal, which describes the range of allowed energy levels of the outermost bound electrons. Electrons in the valence band are not free to move through the crystal. When an atom absorbs an infrared photon of just the right amount of energy, an electron is boosted to the conduction band, where it can flow in a current in the crystal. (A photon is a unit, or quantum, of electromagnetic energy. According to the laws of quantum mechanics, all electromagnetic energy has the properties of both a wave and a particle, and the energy of a photon determines the wavelength of its corresponding wave.) The photon energy required to move an electron from the valence to the conduction band is known as the band-gap energy and is expressed in electron volts.

Once in the conduction band, electrons cross a junction between two dissimilar areas of the semiconductor crystal, moving more freely in one direction than the other. The resulting congregation of negatively charged electrons on one side of the junction causes a buildup of negative electrical potential, which acts as a force pushing the electrons in a current that flows through the photovoltaic cell. Each converter consists of a series of cells wired together to increase the power output. The current generated from the converter can then move through a wire connected to a lamp or a household electrical system.

Advances in converter materials have enabled engineers to improve electrical output. They can select converters that because of their particular band gaps respond best to the range of wavelengths emitted by a given radiator. In past years, the absence of appropriate radiator-converter combinations had posed a key obstacle to furthering the technology.

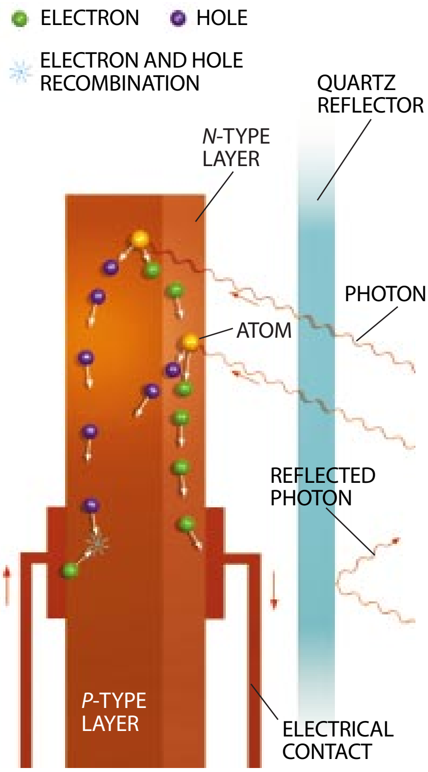
The first generation of thermophotovoltaic devices used radiators that produced a narrow band of wavelengths. Radiators made of the rare-earth material ytterbium oxide often accompanied silicon semiconductor converters with a band gap of 1.14 eV. In theory, a selec-

SLIM FILMS; PATRICK BENNETT (*photograph*)



GENERATOR (*photograph* and *diagram*) **consists of a burner that combines fuel and air to produce heat and of an infrared radiator surrounded by thermophotovoltaic converter cells with attached cooling fins. When the fuel ignites, the radiator heats up to 1,250 degrees Celsius or more. Then 48 gallium antimonide cells, connected in series, convert the infrared energy emitted by the radiator to electric power. At the same time, some of the current starts the fan, which blows air upward to cool the photovoltaic cells. Excess electric power produced by the generator is delivered to a battery for later use. At its optimal operating temperature, this circuit can supply 30 watts of power. Extra heat can be used to warm a room.**

SLIM FILMS



CONVERTER CELL transforms heat into electricity when infrared photons with the right energy penetrate the cell near the intersection of two unlike areas in a semiconductor crystal. When a photon encounters an atom there, it dislodges an electron and leaves a hole. The electron migrates into the *n*-type layer (one that has more electrons than holes), and the holes move into the *p*-type layer (one that has more holes than electrons). The electron then moves into an electrical contact on the cell and travels through an external circuit until it reappears in the *p*-type layer, where it recombines with a hole. If the photon has less than the desired energy, or band gap, it is reflected off the quartz shield and back to the radiator.

tive radiator—which emits a narrow band of wavelengths—should be more efficient than broader-spectrum (broadband) designs. The photons in a selective radiator should provide the minimum energy needed to lift an electron in the semiconductor converter into the conduction band; any excess energy would otherwise be lost as waste heat. Selective radiators should thus supply more power at a lower cost per watt of electricity. In practice, the systems have never performed as expected. The radiators fail to emit enough of the energy from the burning fuel at the precise wavelength needed by a material like silicon to make the conversion process efficient or to produce enough power.

Moreover, temperatures of 2,000 degrees C are required to provide sufficient intensity to achieve worthwhile power output. The heat can stress the material composing the radiator as well as other components, thereby shortening its life span. In addition, polluting nitrous oxide emissions can also result from fuel combustion at these torrid temperatures.

Thermophotovoltaics has advanced because researchers have learned how to pair radiators that transmit a relatively broad range of wavelengths with semiconductors able to cope with that wide spectrum. Broadband radiators, such as silicon carbide, are capable of working efficiently at cooler temperatures of up to 1,000 degrees C. Semiconductor materials developed for the solar energy industry from the third and fifth columns of the periodic table—so-called III-V materials, such as gallium antimonide and indium gallium arsenide—perform the photovoltaic conversion at the wavelengths emitted by these radiators. The band-gap energy required to generate power from III-V materials—0.5 to 0.7 eV—is much less than the 1.14 eV necessary for silicon.

No thermophotovoltaic system can convert all the infrared energy to electricity. Any photon with energy lower than the band gap of the converter cannot boost an electron from the valence to the conduction band and so does not generate electricity. These unused photons become waste heat, unless some means can be found to use them. A photon-recuperation system is a component of a thermophotovoltaic system that sends lower-energy photons back to the radiator, which reabsorbs them, helping to keep the radiator heated and to conserve fuel. Then, more of the emitted photons reach or exceed the band gap.

Retrieving Unused Photons

Investigators have explored several approaches to photon recovery, including the use of a grid of microscopic metal antennae. These antennae, which might be a thin metal film atop a converter cell, transmit the desired infrared wavelengths to the converter while reflecting other photons back to the radiator. Many photon-recovery schemes have fallen short: some systems detect too narrow a range of wavelengths; others are too expensive. The most promising option is the back-surface reflector—so named because the unabsorbed photons move entirely through the layers of the semiconductor converter and are then returned to the radiator by a highly reflective gold surface on the back of the converter.

Around the world, researchers are

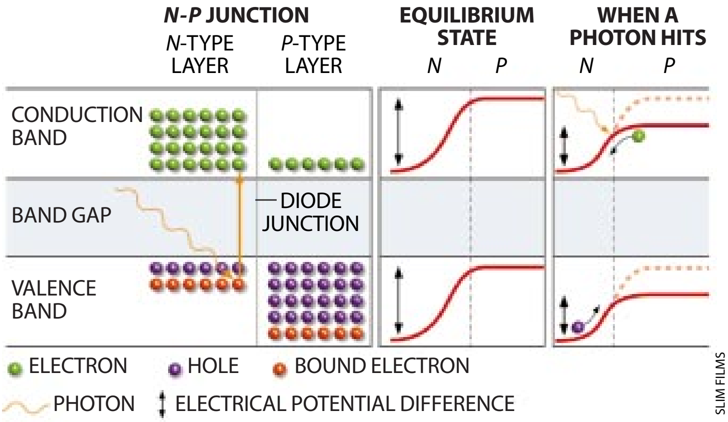
The Photovoltaic Effect

Thermophotovoltaics takes advantage of the semiconductor properties that allow a current to be generated when a photon is absorbed within the crystalline material. The electrons in a semiconductor crystal inhabit a range of defined energy levels, each one known as a band (*left panel in diagram*). Each band specifies the amount of energy needed to free an electron from the crystal. The outermost band in which electrons are still restricted from flowing freely in a current is known as the valence band. To move about, valence electrons must ascend into a higher-energy band (*arrow*), the conduction band. Between these two bands lies a range of energies that cannot be occupied by electrons, known as the forbidden gap, or band gap.

In a semiconductor the conduction band is partially filled with electrons, which constitute negative charge that originates from impurity materials added to the semiconductor. Meanwhile empty energy levels in the valence band—areas where electrons are missing, known as holes—are equivalent to positive electrical charge.

If the semiconductor has more electrons in the conduction band than holes in the valence band, it is called *n*-type, with *n* signifying negative. If it has more holes than electrons, it is called *p*-type, for positive.

When a contact is made between *n*- and *p*-type semiconductor material, a diode forms. At first, negatively charged electrons tend to move to the positively charged holes on the *p* side of the junction and the holes flow toward the *n* side. Negative charge builds up on the side of the junction where the electrons congregate, and positive charge is concentrated on the opposite side. An electrical potential difference, a steep energy barrier, opposes continued current flow (*"equilibrium state," in right panel*).



PHOTON absorption in a semiconductor triggers the photovoltaic effect.

The equilibrium condition is upset when photons—the fundamental particles of electromagnetic radiation—fall on the semiconductor junction. When photons with energies equal to or greater than the band gap are absorbed, they excite electrons from the valence band across the forbidden gap into the conduction band. The altered concentrations of electrons and holes then cause some electrons to move into the *n* side of the junction and some holes to cross to the *p* side. This reduces the potential difference across the junction (*far right*). This phenomenon is called the photovoltaic effect.The electrons may then move to an external circuit in the form of photogenerated electricity.

The band-gap energies needed to lift electrons into the conduction band are among the factors that distinguish thermophotovoltaic cells from solar cells. In a solar cell the band gap is about one to 1.5 eV. For the thermophotovoltaic cell the energies are typically in the range of 0.5 to 0.7 eV—the infrared rather than the visible spectrum.

The difference in band gap does not mean, however, that solar cells can necessarily produce more power. In the case of a solar cell, photons radiate from the sun at a temperature of about 6,000 degrees Celsius, whereas infrared photons destined for the semiconductor diode in a thermophotovoltaic cell are much cooler, in the range of 1,000 to 1,700 degrees C. The optical energy incident on the thermophotovoltaic cell is much larger than that falling on a solar cell, though, because the distance between the sun and the solar cell is 150 million kilometers. The separation between the thermophotovoltaic cell and the surface of the radiator is likely to be about two centimeters. Hence, the electric power output of the thermophotovoltaic converter is potentially far greater than that of a solar cell. *—T.J.C. and M.C.F.*

exploring various technical avenues for developing and commercializing thermophotovoltaics. Since 1994 they have gathered for three international conferences sponsored by the U.S. Department of Energy's National Renewable Energy Laboratory (NREL). The Defense Advanced Research Projects Agency (DARPA), the DOE and the U.S. Army Research Office all have funded programs to develop the technology.

Scientists are excited by laboratory modeling that suggests the promise of this technology. With a radiator that can operate at 1,500 degrees C, it appears possible for semiconductor cells with a single junction (the site where electric potential builds up) to obtain power-density outputs of three to four watts per square centimeter of converter area.

Cells with multiple junctions are also being considered, an approach borrowed from the solar photovoltaics industry. Multijunction cells would allow the converter to capture a wider spectrum of wavelengths, thereby making better use of broadband radiators. Each of the separate junctions would generate a current after absorbing photons in different energy ranges. Theoretically, multijunction devices might achieve five to six watts of energy per square centimeter. Compare these projections with those for the typical flat-panel solar-cell array, which typically achieves 15 milliwatts per square centimeter. Whereas these estimates derive from computer models, and actual power outputs will undoubtedly be smaller, prototypes have exhibited power densities greater than one watt per square centimeter.

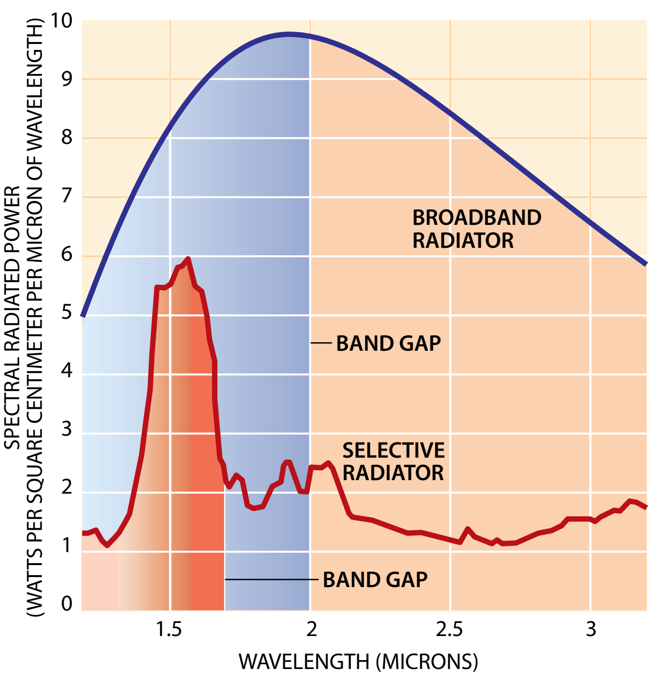
The slow and painstaking work of designing and integrating the component parts of a thermophotovoltaic system has begun at several private and government laboratories. To obtain usable amounts of power, each converter cell must be wired to other cells. Traditional semiconductor manufacturing methods can pattern, etch and wire together multiple cells on a single surface, called a wafer. Researchers at the NREL and those working separately at Spire Corporation in Bedford, Mass., and the National Aeronautics and Space Administration Lewis Research Center have demonstrated these techniques for arrays of thermophotovoltaic cells. One notable example comes from NREL researchers Scott Ward and Mark Wanlass, who have interconnected small thermophotovoltaic cells, each wired in series. The wiring connections move along the top of one cell and then run down along the back of an adjacent one in a configuration that reduces current, increases voltage and minimizes power loss.

This design could eventually reduce the converter to a wafer populated with cells that have only two external contacts—the ones needed to create a circuit—that could power a water pump or a cabin in the woods. Many such wafers could be connected to achieve desired power requirements. The integration of photovoltaic cells may reduce the high cost of the technology, because the cells can be manufactured using standard semiconductor manufacturing methods. In their prototype, Ward and Wanlass also devised a novel design for recirculating unused photons. The electrically active areas of the wafer rest atop an indium phosphide substrate that is semi-insulating. Because the material is nonconductive, and most electrons remain relatively tightly bound within the semiconductor crystal, low-energy photons can move through the substrate without getting absorbed by free electrons moving through the conduction band. The photons then reflect off a gold surface and return to the radiator. In other prototype designs, many unused photons are absorbed by the converters.

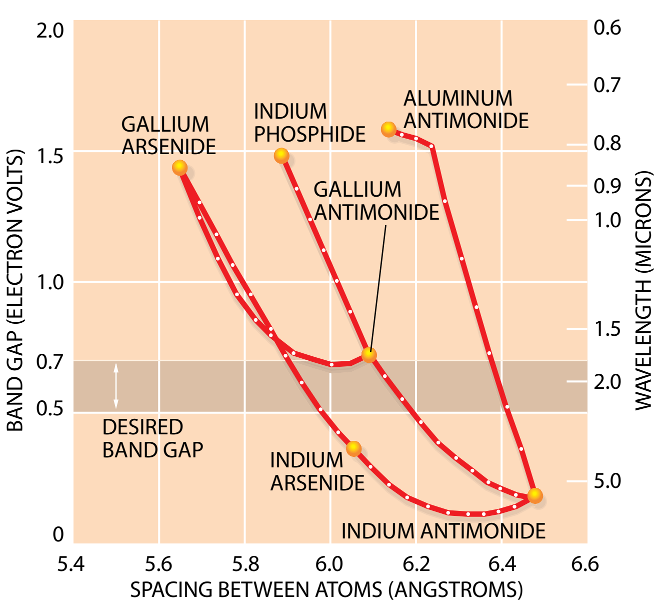
While developmental work continues elsewhere, the first thermophotovoltaic commercial product is about to reach the market. JX Crystals in Issaquah, Wash., has created a product—Midnight Sun—primarily for use on sailboats. The 14-centimeter-wide-by-43-centimeter-tall cylindrical heater, powered by propane gas, can produce 30 watts of electricity and is targeted as a means of recharging batteries that run navigation and other equipment. The unit not only provides electricity but acts as a co-generator, supplying space heating for the boat cabin. It uses a partially selective radiator made of magnesium aluminate and has gallium antimonide photovoltaic cells connected in series.

Although its current $3,000 price tag makes it more expensive than a conventional diesel generator, Midnight Sun runs silently and is expected to be more reliable, because it lacks any moving parts.

The product may also prove attractive to owners of recreational vehicles or wilderness homes, who could take advantage of a substantially less costly



PHOTON RADIATORS emit either a narrow band of wavelengths (selective) or a wider range (broadband) that can be transformed into electricity by a semiconductor converter. Selective radiators, made of rare-earth materials, produce a smaller number of photons at or above the band gaps at which electricity can be generated (*red area*). Broadband radiators, in contrast, can take advantage of more of the usable power reaching the cell (*blue area*). Lower-energy photons—those to the right of the band gap—become waste heat.



ILLUSTRATIONS BY SLIM FILMS

BUILDING CONVERTERS by alloying different materials from the third and fifth columns of the periodic table can achieve a desired band gap—the amount of energy needed to free an electron so it can flow in a current. Band gaps of 0.5 to 0.7 eV (*shaded area*) are ideal for thermophotovoltaic devices and correspond to increasingly shorter infrared wavelengths from the radiator. These band gaps can be obtained by combining compounds such as gallium antimonide with indium antimonide in increments represented by the dots along the red line. Calculating the spacing between atoms (*bottom axis*) allows designers to join together different compounds with distinct crystalline structures.

EXPERIMENTAL CAR at Western Washington University gains some of its power from a thermophotovoltaic generator (*inset*).

PHOTOGRAPHS BY TORE OFTENESS

unit than the stainless-steel and brass thermophotovoltaic generator necessary for the marine environment.

Despite the present drawbacks of selective radiator systems, researchers still study them. DARPA has funded Thermo Power's Tecogen division in Waltham, Mass., to create a gas-fired generator as a power supply for troop communications or for powering laptop computers in the field. In the civilian realm the same unit could keep a home furnace running in a power outage. The 150- and 300-watt-power modules use arrays of ytterbium oxide fibers at a wavelength of 980 nanometers. These selective radiators are matched with silicon photoconverters. A multiple-layer insulating filter recovers unused energy.

Beyond the Niche

Though still in its infancy—actually just barely out of the research laboratory—thermophotovoltaics holds great promise for many niche markets. Longer term, the technology could play a broader role in the global energy market. The terminology "niche market" can itself be deceiving, evoking the image of a few buyers served by a small cottage industry. Yet the world is replete with billion-dollar niche markets—many of which had humble beginnings. Thomas J. Watson, the founder of IBM, originally saw a market for only a small handful of computers worldwide but decided to pursue the opportunity anyway.

The recovery of industrial waste heat could establish a huge market for thermophotovoltaics. Many industrial sectors—manufacturers of glass, aluminum, steel and other products—generate enormous quantities of heat through their production processes. The glass industry estimates that two thirds of the energy it consumes emanates as waste heat, which could amount to a gigawatt of power. Thermophotovoltaic converters could possibly be used to generate electricity from waste heat, affording enormous savings in electricity costs. Another promising application is suggested by a hybrid electric vehicle introduced by Western Washington University. The vehicle supplements the power from electric batteries by supplying 10 kilowatts from a thermophotovoltaic generator.

Today the funding devoted to thermophotovoltaics is not more than \$20 million to \$40 million a year. That is the money researchers receive for their studies on converter devices, radiators and the integration of the various components into working generators. A recent market study financed by the NREL, however, indicates that thermophotovoltaics could produce \$500 million in commercial sales by 2005. Revenues for this nascent market would come from substituting this technology for diesel generators of less than two kilowatts used in the military and in recreational vehicle and boating markets. Thermophotovoltaics may lead to the possibility of cleaner, more efficient and inexpensive solutions for a number of alternative energy markets. A technology that had its roots in the 1950s may finally get a chance to prove itself at the turn of the new millennium. SA

The Authors

TIMOTHY J. COUTTS and MARK C. FITZGERALD have worked on renewable energy and related technologies for many years. Coutts, who has studied solar-cell technology since 1971, is a research fellow in the National Center for Photovoltaics of the U.S. Department of Energy's National Renewable Energy Laboratory (NREL) in Golden, Colo. He was educated at Newcastle Polytechnic in Britain, where in 1968 he received his doctorate in thin-film physics. At the NREL, he has broad-ranging responsibilities in both solar and thermophotovoltaics technologies. Fitzgerald is president of Science Communications, a photovoltaic consulting firm in Highlands Ranch, Colo. He has 19 years of experience providing education services about renewable energy.

Further Reading

The First NREL Conference on Thermophotovoltaic Generation of Electricity. Timothy J. Coutts and John P. Benner in *AIP Conference Proceedings*, Vol. 321. AIP Press, 1995.

The Second NREL Conference on Thermophotovoltaic Generation of Electricity. Edited by John P. Benner, Timothy J. Coutts and David S. Ginley in *AIP Conference Proceedings*, Vol. 358. AIP Press, 1996.

The Third NREL Conference on Thermophotovoltaic Generation of Electricity. Edited by Timothy J. Coutts, Carole S. Allman and John P. Benner in *AIP Conference Proceedings*, Vol. 401. AIP Press, 1997.

by Shawn Carlson

Spooling the Stuff of Life

If you're a man who's looking to get married, here's some friendly advice. Only consider women who are smarter than you are. I followed this prescription four years ago when I married Michelle Tetreault, a charming and brilliant biophysicist. Now my wife always intrigues me with her insights and never lets me get away with anything dubious at home.

At work Michelle employs the latest techniques in biochemistry to unravel the secrets of photosynthesis. By manipulating the smallest units of inheritance, the individual base pairs on a single strand of DNA, she can change one by one the amino acids that make up a key protein and then study how well this altered molecule can do its job.

Hearing about Michelle's research so often at the dinner table recently prompted me to try my hand at molecular biology. Although the many cutting-edge techniques she uses are probably beyond the range of amateur dabbling, recent advances have opened up intriguing avenues for informal explorations into biotechnology. To help clear the way, this column explains how anyone can do what biotechnologists do routinely: extract and purify DNA.

DNA is the largest molecule known. A single, unbroken strand of it can contain many millions of atoms. When released from a cell, DNA typically breaks up into countless fragments. In solution, these strands have a slight negative electric charge, a fact that makes for some fascinating chemistry. For example, salt ions are attracted to the negative charges on DNA, effectively neutralizing them, and this phenomenon prevents the many separate fragments of DNA from adhering to one another. So by controlling the salt concentration, biologists can make DNA fragments either disperse or glom together. And therein lies the secret of separating DNA from cells.

The procedure is first to break open the cells and let their molecular guts spill into a buffer, a solution in which DNA will dissolve. At this point, the buffer contains DNA plus an assortment of cellular debris: RNA, proteins, carbohydrates, and a few other bits and pieces. By binding up the proteins with detergent and reducing the salt concentration, one can separate the DNA, thus obtaining a nearly pristine sample of the molecules of inheritance.

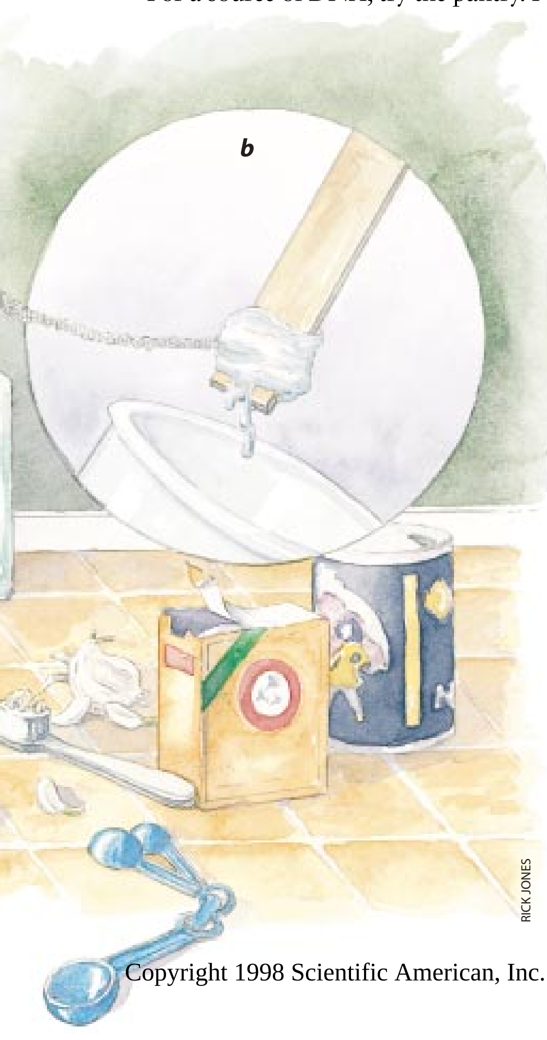
My profound thanks go to Jack Chirikjian and Karen Graf of Edvotek

KITCHEN LABORATORY includes most of the items needed to isolate DNA. A drinking straw, for example, can be used to add alcohol to the solution (a), *and a coffee stirrer serves to spool the DNA* (b).

(www.edvotek.com), an educational biotech company in West Bethesda, Md., for showing me how anyone can purify DNA from plant cells right in the kitchen. You'll first need to prepare a buffer. Pour 120 milliliters (about four ounces) of water into a clean glass container along with 1.5 grams ($^1/_4$ teaspoon) of table salt, five grams (one teaspoon) of baking soda and five milliliters (one teaspoon) of shampoo or liquid laundry detergent. These cleaners work well because they have fewer additives than hand soaps—although do feel free to try other products as well.

The detergent actually does double duty. It breaks down cell walls and helps to fracture large proteins so they don't come out with the DNA. The people at Edvotek recommend using pure table salt and distilled water, but I have used iodized salt and bottled water successfully, and once I even forgot to add the baking soda and still got good results. In any case, try to avoid using tap water. To slow the rate at which the DNA degrades, it's best to chill the buffer in a bath of crushed ice and water before proceeding.

For a source of DNA, try the pantry. I got great results with an onion, and the folks at Edvotek also recommend garlic, bananas and tomatoes. But it's your experiment: choose your favorite fruit or vegetable. Dice it and put the material into a blender, then add a little water and mix things well by pulsing the blades in 10-second bursts. Or, even simpler, just pass the pieces through a garlic press. These treatments will break apart some of the cells right away and expose many cell walls to attack by the detergent.



RICK JONES

Place five milliliters of the minced vegetable mush into a clean container and mix in 10 milliliters of your chilled buffer. Stir vigorously for at least two minutes. Next you'll want to separate the visible plant matter from the molecule-laden soup. Use a centrifuge if possible. (To learn how to build a centrifuge, see the January Amateur Scientist column.) Spin the contents at low speed for five minutes and then delicately pour off at least five milliliters of the excess liquid into a narrow vessel, such as a clean shot glass, clear plastic vial or test tube. If you do not have a centrifuge, strain the material through an ordinary coffee filter to remove most of the plant refuse. With luck, any stuff that leaks through should either sink or float on top, so it will be a simple matter to pour off any solids into the sink and then pour the clear liquid into a clean vessel.

The solution now contains DNA fragments as well as a host of other molecular gunk. To extract the DNA, you will need to chill some isopropyl alcohol in your freezer until it is ice-cold. Most drugstores sell concentrations between 70 and 99 percent. Get the highest concentration (without colorings or fragrances) you can find. Using a drinking straw, carefully deposit 10 milliliters of the chilled alcohol on top of the DNA solution. To avoid getting alcohol in your mouth, just dip the straw into the bottle of alcohol and pinch off the top. Allow the alcohol to stream slowly down along the inside of the vessel by tilting it slightly. The alcohol, being less dense than the buffer, will float on top. Gently insert a narrow rod through the layer of alcohol. (Edvotek recommends using a wooden coffee stirrer or a glass swizzle stick.)

Gingerly twirl back and forth with the tip of the stick suspended just below the boundary between the alcohol and the buffer solution. Longer pieces of DNA will then spool onto the stick, leaving smaller molecules behind. After a minute of twirling, pull the stirrer up through the alcohol, which will make the DNA adhere to the end of the stick and appear as a transparent viscous sludge clinging to the tip.

Although these results are impressive, this simple and inexpensive procedure does not yield a pure product. Professionals add enzymes that tear apart the RNA molecules to make sure they do not get mixed up with the coveted DNA.

Even after the most thorough extraction, some residual DNA typically lingers in the vessel, forming an invisible cobweb within the liquid. But with a little more effort, you can see that material, too. Some dyes, like methylene blue, will bind to charged DNA fragments. A tiny amount added to the remaining solution will thus stain tendrils of uncollected DNA. I don't know whether any household dyes, like food coloring or clothing or hair dyes, will also work, so I invite you to find out. Add only a drop: you want all the dye molecules to bind to the DNA, with none left over to stain the water.

Exciting as it may be, extracting an organism's DNA is only the first step in most biological experiments. You'll probably want to learn what further investigations you can do—for example, sorting the various DNA fragments according to their lengths. This department hasn't described suitable methods for separating large organic molecules in decades. (See the Amateur Scientist columns for August 1955, July 1961 and June 1962.) But I'll be revisiting some of these techniques in coming months to help you bring your kitchen lab into the modern age of biotechnology. SA

The Society for Amateur Scientists and Edvotek have joined forces to create a kit containing cell samples, lab ware, enzymes, buffers and detergents that can help you create higher-quality preparations. To order, send $35 to the Society for Amateur Scientists at 4735 Clairemont Square, Suite 179, San Diego, CA 92117. For more information about this and other amateur science projects, visit the forum hosted by the society at http://web2.thesphere.com/SAS/WebX.cgi on the World Wide Web. You may also write the society at the above address or call 619-239-8807.

by Ian Stewart

Counting the Pyramid Builders

THE GREAT PYRAMID
at Giza **(above, in center)** *could have been built by as few as 10,000 workers* **(below).**

The pyramids of ancient Egypt rank among the most enigmatic of archaeological mysteries. The largest, the Great Pyramid at Giza, was built by Egyptian king Khufu in about 2500 B.C. and is still mostly intact. Its original height was almost 147 meters (482 feet), and it weighed more than seven billion kilograms (7.7 million tons). The pyramid was made from huge blocks of stone, which were quarried, trimmed to a fairly regular shape, transported to the construction site and then piled on top of one another with astonishing precision.

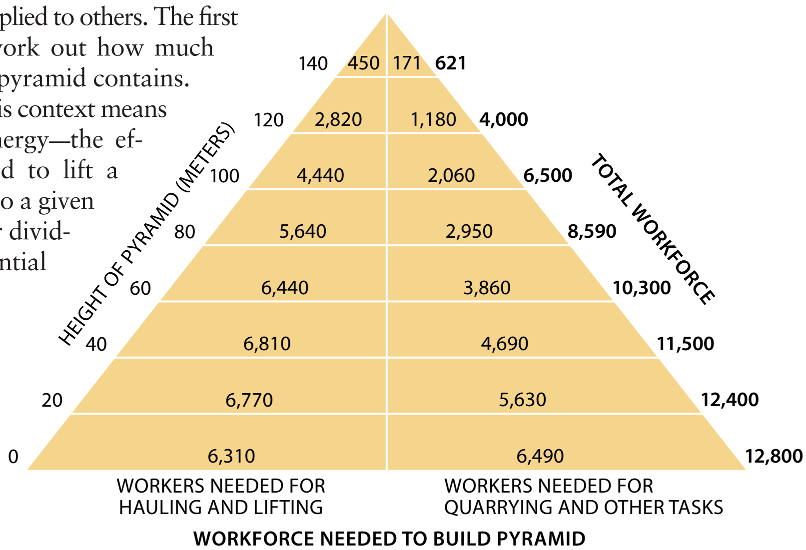
How did the Egyptians put together such giant edifices? And why were they built? Historians and archaeologists do not know for certain. Many pyramids were used as tombs for kings, and there are plenty of theories about their construction. And now historians can make a good guess at the size of the workforce, thanks to some mathematical detective work by Stuart Kirkland Wier, an amateur Egyptologist who belongs to a study group sponsored by the Denver Museum of Natural History. His calculations appeared in the *Cambridge Archaeological Journal* (April 1996).

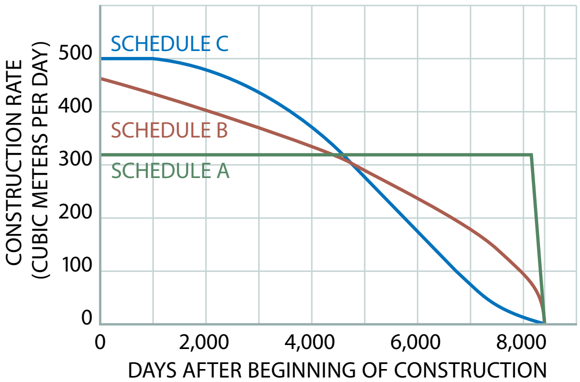
Some two millennia after the Great Pyramid was completed, Greek historian Herodotus reported that it took 100,000 men to build the structure. Herodotus, however, was not a reliable source on ancient Egypt. It now seems that he overestimated the workforce by an order of magnitude. According to Wier, the true figure was surprisingly small: only about 10,000. How did Wier come up with this number without knowing how the pyramid was built? He estimated the total amount of work needed to construct the pyramid, then made a few assumptions about how the Egyptians scheduled the project.

Wier performed his calculations for Khufu's pyramid, but the same method can be applied to others. The first step is to work out how much "energy" a pyramid contains. Energy in this context means potential energy—the effort required to lift a given mass to a given height. Wier divided the potential energy of the pyramid by the number of days spent building it. This calculation yielded the amount of work required each day to lift the blocks of stone. The sole source of power for the Egyptian workforce was human muscle. Dividing the required work by the energy output of a typical Egyptian laborer, Wier determined the number of workers needed to perform the lifting tasks.

The biggest uncertainty is time. How long did it take to build Khufu's pyramid? Khufu reigned for 23 years. Construction of his pyramid probably did not begin before his reign, but it could have ended near the time of his death or many years before. Wier assumed that Khufu's pyramid took 23 years to build—the same length as the king's reign. That amounts to some 8,400 days if the work continued year-round. The actual time might have been half as long; because of this fundamental uncertainty, we can make only a rough estimate of the size of the workforce.

Khufu's pyramid, when new, was 146.7 meters high, and the sides of its square base were 230.4 meters long. The volume of a pyramid of height h and base side s is $s^2 \times {}^{h}/_{3}$, which works out to be 2.6 million cubic meters for the Great Pyramid. The material was limestone with a density d of 2,700 kilograms per cubic meter, so the mass was about seven billion kilograms. The potential energy of a pyramid is $h^2 \times d \times$





DMITRY KRASNY

THREE POSSIBLE SCHEDULES lead to the completion of the Great Pyramid in 8,400 days.

$s^2 \times g/12$, where g is the acceleration caused by gravity (9.81 meters per second squared). For the Great Pyramid, this amounts to 2.52 trillion joules, a joule being the standard unit of energy.

The average amount of useful work provided by a man in a day is about 240,000 joules. Assuming perfect efficiency, 1,250 men would be needed to raise the blocks into position over 8,400 days. This estimate is clearly too low; Wier multiplied it by a fudge factor of 1.5 to account for the inefficient use of muscle power. He also factored in the amount of work needed to raise the stones from the bottom of the Khufu quarry to the level of the pyramid site—a distance known to be 19 meters.

Of course, pyramids are not just featureless heaps of stone: they have passages and chambers, some of which are remarkable feats of engineering in their own right. But the major part of the construction work was simply piling up the blocks, so we can ignore the structural details. The pyramids were built in layers from bottom to top. Certainly no block could be added to the structure until the blocks below had been put in place. Nobody knows how the blocks were raised. Some experts believe the Egyptians hauled the blocks up vast ramps of sand. Others say the laborers used a cunning arrangement of levers.

Workers transported the stone blocks from the quarry to the construction site by dragging them on wooden sleds. (Egyptian carvings show this being done.) The Khufu quarry was several hundred meters from the base of the pyramid. To calculate the number of men required for hauling, Wier estimated the friction between a loaded wooden sled and the ground, then determined the amount of work needed to overcome the friction. Manpower was also needed for other tasks: quarrying the stone blocks, trimming them to shape, installing them in the pyramid. Wier assumed that all such tasks together required a maximum of 14 men per cubic meter of stone per day.

These calculations give the average number of workers at Giza. But did the pyramid builders employ a workforce of fixed size? Or did they enlist extra workers when the project demanded it and dismiss them when they were no longer needed? The records don't say, but several possibilities can be inferred.

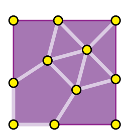
Dividing the volume of the Great Pyramid by the time available to build it shows that an average of 310 cubic meters of stone must be put in place each day. As the pyramid grows, however, the stone blocks must be raised higher, so more work is required to lift them. And the amount of workspace available to the lifters—the area at the top of the unfinished pyramid—decreases as the structure gets higher. These factors make it clear that a steady rate of 310 cubic meters a day is not sensible. Instead the construction rate should be greater when the pyramid is low and should drop off as it gains in height.

Wier considered three possible construction schedules [*see illustration at left*]. Under schedule A, the workers install the blocks at a constant rate of 315 cubic meters each day until day 8,110. At that point, the construction rate plummets because of the limited space at the top of the pyramid. Under schedule B, the construction rate starts at 462 cubic meters a day and declines steadily until day 8,000, after which it falls off more rapidly. Under schedule C, the construction rate starts at 500 cubic meters a day but drops sharply after day 2,000 and tails off to zero after day 7,000. All three schedules result in the completion of the pyramid in 8,400 days, but they require different numbers of workers as the job progresses.

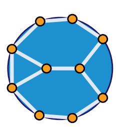
The bottom illustration on page 98 shows the number of workers required at each stage of the pyramid's construction under schedule B. At the start, about half the workers are hauling or lifting the blocks, and half are quarrying the stone, installing the blocks and undertaking other tasks. By the end, though, nearly three quarters of the workers are moving the blocks. At no stage does the number of workers exceed 12,800 men—about 1 percent of the estimated Egyptian population at that time. Schedules A and C lead to similar results.

Perhaps the simplest schedule is to employ a workforce of constant size, except near the end when there is not enough room at the top for more than a few workers. Under this schedule, 10,600 men would suffice to build Khufu's Great Pyramid. Even fewer workers would be needed to build the other large pyramids at Giza and Dahshûr. Despite the tremendous size of these monuments, there were more than enough workers in Egypt to build them. SA

FEEDBACK



In "Tight Tins for Round Sardines" [February], the column on packing circles into regions of various shapes, I wrote that "nearly all the information we have about such questions dates from 1960 or later." One reader misunderstood this remark, complaining that I had cavalierly disregarded the great classical work of Gauss, Lagrange and others. By "such questions" I meant the packing of objects into finite regions, like 11 circles in a square (*above*) or 10 circles in a larger circle (*left*). (The circles are represented by points marking their centers.) The classical work is on packings of the infinite plane, and it also assumes that the objects form a regular array. In the problems discussed in the column, the region is of limited extent, and the arrangement of circles need not be regular.



Martin Trump of Newbury, England, has explored the analogous problem of distributing charged particles inside a sphere. His World Wide Web site—http://www.geocities.com/Paris/3142—has pictures of these distributions. —*I.S.*

DAVID FIERSTEIN

REVIEWS AND COMMENTARIES

MATHEMATICS FOR THE MANY

Review by Donald J. Albers

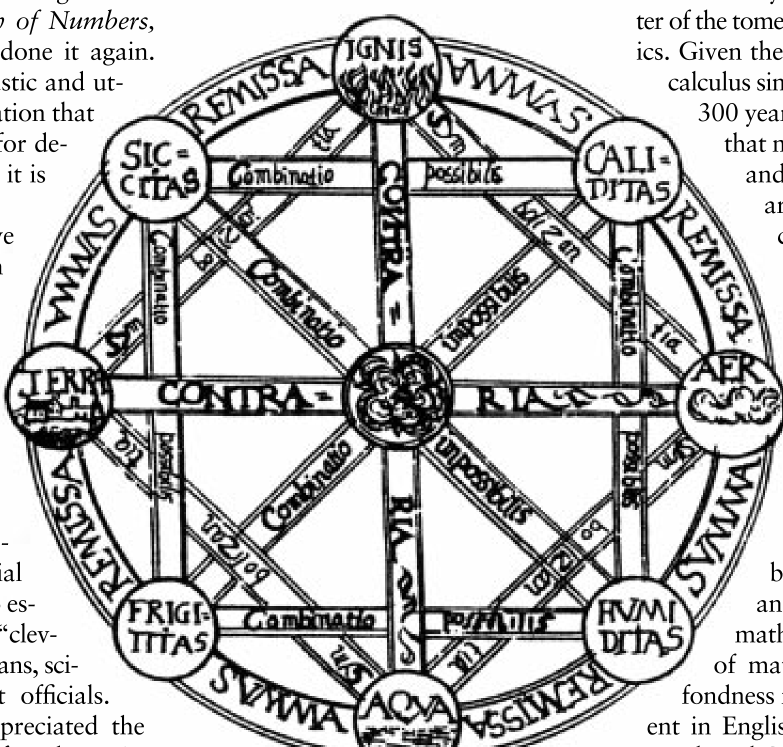
Mathematics: From the Birth of Numbers
BY JAN GULLBERG
W. W. Norton and Company, New York, 1997 ($50)

Sixty-one years ago W. W. Norton and Company published Lancelot Hogben's *Mathematics for the Million,* one of the truly great popularizations of this century. It has sold more than 600,000 copies and is still selling thousands of copies a year. With the publication of Jan Gullberg's *Mathematics: From the Birth of Numbers,* Norton seems to have done it again. The book is an enthusiastic and utterly amazing popularization that promises to be in print for decades. After 18 months, it is in its sixth printing.

The two books have much in common. Both were written by non-mathematicians: Hogben was a zoologist and Gullberg a surgeon. Neither was an American, but each ended up being published by an American company. Both saw mathematics as a social equalizer and essential for helping the masses to escape control by society's "clever people"—mathematicians, scientists and government officials. Both authors richly appreciated the historical development of mathematics.

And both books are big: Hogben at 649 pages and Gullberg at 1,093 pages. Hogben cuts a broad mathematical swath, from the development of counting to calculus and statistics. Gullberg covers nearly all that and much more. When Hogben's classic was published in 1937, it was praised for its illustrations. The number and variety of pictures in Gullberg's opus dwarf those of Hogben. Both books are serious, although Gullberg also makes generous use of humor. But here is a sharp contrast: Hogben claimed to have written his book in six weeks; Gullberg took 10 years for his.

An itinerant surgeon (or "Dr. Sawbones" as he sometimes referred to himself), Gullberg seemed to be drawn to small towns, each a bit off the beaten track. In the years that he worked on the book, he lived in Angola, Ind.; Forks, Morton, Moses Lake and Sequim in Washington State; Moab, Utah; Møsjen and Rjukan in Norway; and Karlskoga, Sweden. He claimed that his peripatetic way of life was crucial to the gathering of data and the development of ideas. Certainly his book project provided something of an anchor during that period. Gullberg died at age 62 in May 1998 of the consequences of a stroke suffered a few days earlier while at work at a hospital near his home in Nordfjordeid, Norway. Fortunately, he lived long enough to see his book become a publishing success.



MATHEMATICS'S DEEP HISTORY pervades Gullberg's book. This illustration is from the frontispiece of Leibniz's Dissertatio de arte combinatoria, *a work that strove to reduce all truths of reason to a simple system of arithmetic.*

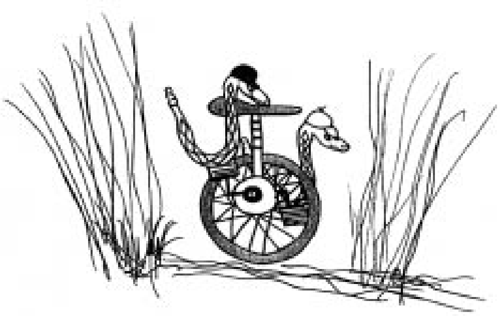
The scope of the book is huge—numbers and language, combinatorics, logic, set theory, theory of equations, geometry, trigonometry, fractals, sequences and series, calculus, harmonic analysis, probability and differential equations. Few mathematicians would undertake to write a survey so ambitious. A full quarter of the tome is devoted to calculus topics. Given the tremendous influence of calculus since its invention more than 300 years ago, perhaps it deserves that much space. Bits of humor and other ancillary material are present in the calculus chapters, though not with the same frequency as in other chapters.

The book's large pages are loaded with eye-catching graphics and illustrations, frontispieces from original works, cartoons and jokes. The illustrations reveal much about Gullberg's remarkably broad and deep curiosity about mathematics, his appreciation of mathematics in history, his fondness for languages (he was fluent in English, Swedish, Norwegian, French and German, and he could read Greek and Latin) and his humorous side. The opening illustration is a photograph of a sculpture by Swedish sculptor Axel Ebbe. Depicting a young man emerging from an imposing mass of stone, it is entitled *Breaking Away from the Darkness of Ignorance,* to which Gullberg forcefully adds, "and then there was Mathematics."

Within the first five pages, as the author explores "The Origins of Reckoning," we encounter lovely whimsical drawings by his wife, Ann, and twin sons, Kalin and Kamen, then nine years old, and apt quotations from Thomas Hobbes, Umberto Eco and Carl von Linné (Linnaeus). And he never lets up throughout the almost 1,100 pages—history, brief profiles, more than 1,000 technical illustrations done by his son Pär, and excerpts from an astonishingly wide range of sources, from daily newspapers to Newton to *Winnie the Pooh*. He is even bold enough to include some of his own verse and jokes, most of which are quite good.

The illustrations alone make the book worth the price—and at less than five cents per page, it is a bargain. Gullberg searched through an amazingly large number of classics and reproduced from them a wonderful collection. Among them are the frontispiece of Gottfried Leibniz's *Dissertatio de arte combinatoria* (1666); Pascal's original triangle from his *Traité du triangle arithmetique* (1665); Johannes Kepler's drawings of 13 semiregular polyhedra in his *Harmonices mundi* (1619); the frontispiece and table of contents of the first printed textbook on calculus, *Analyse des infiniment petits* by L'Hospital (1696); brief selections from John Napier's *A Description of the Admirable Table of Logarithms with a Declaration of the Most Plentiful, Easy, and Speedy Use Thereof in Both Kinds of Trigonometry, as Also in All Mathematical Calculations;* and a map of Königsberg circa 1740 with its seven bridges that inspired the famous problem of the same name (solved by Leonhard Euler in 1736), which helped to launch graph theory.

A third of the way through, Gullberg presents a chapter entitled "Overture to the Geometries." He begins the chapter with a chilling quotation from the book of Roman civil law (circa A.D. 650) that serves as a reminder that mathematicians have not always been held in high regard: "The study and teaching of the science of geometry are in the public interest, but whosoever practices the damnable art of mathematical divination shall be put to the stake." In the same chapter, and on a decidedly lighter note, we find instructions on "How to Catch Elephants—An Elephantasy," which appeared in the Copenhagen newspaper *Berlingske Tidende*.



PLAYFUL TAKE ***on circular functions in the chapter on trigonometry was drawn by Gullberg's wife, Ann.***

By the time we have finished the chapter, we have been treated to wonderfully readable sections on history, geometric abstraction, perspective and projection, form and shape, survey of geometries, topology and Euclidean and non-Euclidean geometries. The next chapter continues with more geometry, which is followed by chapters on trigonometry (plane and spherical), hyperbolic fractions, analytic geometry, vector algebra and fractals.

Gullberg gives the reader plenty of motivation and a multitude of worked-out examples, which are especially helpful. He derives many familiar formulas and provides many proofs, but he is careful not to overwhelm the reader with detail. He was particularly keen on reading the masters, as is reflected in the opening remarks to his extensive "Works Cited" section: "The strength of those who have gone before should not be gleaned solely from histories or random quotations. True inspiration and power lie in the original works."

Not only did Gullberg write *Mathematics: From the Birth of Numbers,* he built the book using his Macintosh Plus. He set the type and lovingly designed the pages, deciding where to place each illustration, drawing and cartoon. His son Pär, now a computer graphics specialist, provided vital technical assistance, including the computer-generated illustrations. After the manuscript was reviewed by several specialists in different areas of mathematics, corrected, copy edited and the pages put into final form by Gullberg, all the publisher had to do was start the presses. The volume is remarkably clean for a book of this size and scope. Typos and mathematical glitches can be found, but they are rare. Few of us, whether amateurs like Gullberg or professional mathematicians, could write (or even read) a manuscript of this size without occasional lapses.

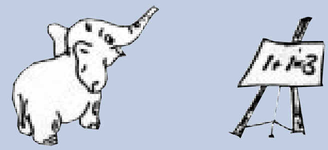
How to Catch Elephants – an Elephantasy

If you want to catch elephants, this is how you do it according to the Copenhagen newspaper Berlingske Tidende.

All you need is a blackboard, a piece of chalk, a mariner's telescope, a pair of tweezers and an empty jampot.

You begin by writing 1 + 1 = 3 on the blackboard and set it up in a place rich in elephants, and hide yourself in a nearby tree.

Very soon an inquisitive little baby elephant will draw near to muse upon this remarkable statement and by and by several of her elders will join her.



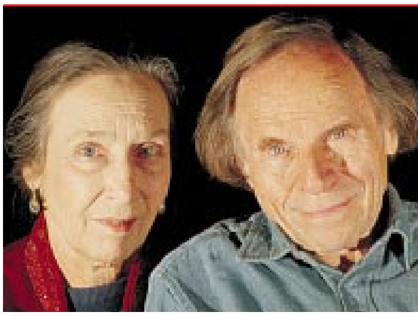
When there are enough of them, you just turn your telescope the wrong way around, so that the assembled elephants become quite small. Then you pick the small elephants up with your tweezers, one by one, and put them in the jampot.

WHIMSY from the discussion of perspective and projection

About a year ago, when I asked Gullberg why he wrote *Mathematics: From the Birth of Numbers,* he responded: "So that people might develop confidence in doing mathematics and an appreciation for how long the development of mathematics has taken." In the same interview, when asked about his target group, he said, "I can't write for a target group—only for myself—as a means of exploring mathematics and its history."

Readers of *Scientific American,* students, teachers of mathematics and all others who care about my favorite subject will want to own this book. It is an important reference and a book that is plain fun to dip into. If a family is to have only one mathematics book on the reference shelf, then this is the one. It is a great present for anyone who likes mathematics. But if you buy this extraordinary tome for someone, there is considerable danger that you will end up keeping it for yourself—and then you will need to buy another copy.

DONALD J. ALBERS is director of publications of the Mathematical Association of America. His profiles and interviews of mathematicians have appeared in two books, Mathematical People *and* More Mathematical People.



WONDERS

by Philip and Phylis Morrison

A Minor Shift in Point of View

Scientists have discovered some eight or 10 planetary systems beyond our own. So far the homiest example attends a dim, single star in the constellation of Ursa Major, the Great Bear. That sunlike star is designated as 47 UMa, entered long ago as just one more star by the first ever astronomer royal, John Flamsteed. (You can see it on a clear, dark night, as one of a sprinkle of faint stars. Extend the line of the Pointers backward from the North Star by about three times the sky separation between the two Pointers to find the little grouping.)

The distance of 47 UMa from Earth, about 45 light-years, is known by trigonometric measurement; from its spectrum we find that it is nearly the size and mass of our sun, though somewhat older, brighter and hotter. An unseen giant planet circles 47 UMa every 1,100 days: a super-Jupiter, yet too faint for the best telescopes. Our indirect evidence is compelling; tiny, elliptical motions of that sun-star in response to the wide swing of its planet have been mapped in detail since 1990. Alas, we lack the sensitivity to detect any sun's minute response to an Earth-like orbiter, which would be hundreds of times lighter than a giant gas planet.

But imagination stands ready. Suppose (here we leave fact to voyage on a single fantastic journey) our space probe has toted a good video camera into an Earth-like annual orbit near 47 UMa. That would shift our point of view about 50 light-years out, a minor hop on a galactic scale, although a tedious journey, lasting some hundred thousands of years, with any of NASA's 20th-century solar system vehicles. Out in the Big Bear system from our small probe we peer at an airless, star-studded sky, night-black even close to the brilliant disk of 47 UMa itself. That sunlit orb moves annually—don't ever point your camera at it!—around a belt in the sky, and a single shimmering planet rounds there, too. Expect comets and asteroids, but they will probably be too faint to notice. The Milky Way will look familiar. But what constellations will enliven that dark sky?

This is no alien sky, for we have not yet left our galactic neighborhood. The conspicuous figure of the brilliant constellation Orion will be quite recognizable from the Bear system. Orion's stars are so far from us here at the sun that our shift to 47 UMa will simply edit the shape of the hunter a little, his belt and sword hardly changed. Other constellations—such as the Big Dipper (still far away from 47 UMa even though we have journeyed almost straight toward it), Scorpius and Cygnus—will be recognizable as well, although one of the Swan's half a dozen defining stars is much diminished.

Our own sun will be an obscure, southerly object for binoculars.

Yet the very brightest star in our home port, Sirius, or Dog Star, will no longer be a skymark in the Bear world. It will dwindle to resemble 50 or more stars of the second magnitude, known out there only to adept stargazers. Our own sun will present an obscure, southerly object for binoculars. Canopus is the second brightest star in our night sky, easily viewed from Florida any winter evening. Out in the Great Bear it is Canopus—not Sirius—that is the brightest of all.

The explanation of this spotty change is neat. Stars appear faint or bright for two wholly differing reasons. Even a dim star can appear bright if it is close by. (After all, although our sun is a dwarf among stars, for us here it is not merely bright but sends life-giving heat, at the tiny distance of only 0.0000158 light-year, or 500 light-seconds.) A really distant star can be visually bright if it is sufficiently luminous; ample optical power overcomes the distance. Among the 100 brightest stars as seen from Earth, our sun compares in light output only with its two nearest neighbors, both in Centaurus far south: one star a fair match for the sun and its reddish companion rather fainter; all the rest are at least 10 times more luminous than the sun.

A shift of 40 or 50 light-years will reduce to the commonplace any neighbor under 10 light-years away, like Sirius, but would not even halve the powerful gleam of Canopus, a massive, bright giant around 300 light-years out that is intrinsically thousands of times more luminous than the modest but enduring sun. Among the well-seen stars from Earth a dozen

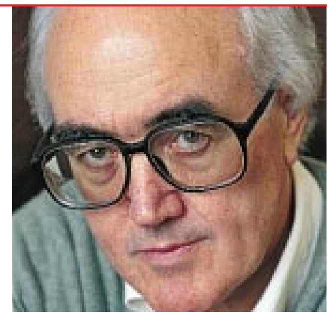
Continued on page 108

DUSAN PETRICIC

DUSAN PETRICIC

CONNECTIONS

by James Burke



Rebellious Affairs

I was in one of London's oldest coffeehouses the other day, sipping a cup (of Folger's, as it happens) and thinking about the first such watering holes. Like the one in St. Paul's churchyard, where those who supported the American colonists' cause used to meet and greet during the 1760s. One such dangerous coffee-drinking liberal was mathematician Richard Price, who put actuarial studies on a modern basis with his analysis of the Northampton Bills of Mortality. From which he was able to deduce life-and-death matters well enough for the new insurance firms to be able to charge realistic premiums and not go broke before their clients did.

Price's statistical work attracted the French finance minister, Anne-Robert-Jacques Turgot, who had the unenviable task of balancing the books at a time when the French economy was heading rapidly toiletward. A pal and adviser to Turgot was the Marquis de Condorcet, whose claim to fame was the way he took the new statistical view of things to the next stage and invented what he called "social mathematics." Thus armed, he intended to predict all aspects of behavior and thereby set the study of society onto a scientific footing.

Turgot and Condorcet both came out of the physiocrat school of economic thought, which looked to the English example of a lightly regulated agricultural market as the only way to save France from ruin. For them, the price of a loaf of bread held the key to political stability. Well, they never made it out of the bakery in time to save their heads. But en route to his eventual death in a French Revolutionary prison, Condorcet tried hard to popularize the agronomic ideas of an English lawyer, Jethro Tull. In 1711 Tull had gone to the south of France for his health and had seen peasants in the vineyards around Frontignan hoeing between furrows and (although he didn't know it) in this way aerating the dirt and making it easier for water to permeate. As a result of which they were getting great crop yields without having to resort to the use of expensive cow dung. When Tull, on native soil, found that with this trick he could grow corn in the same, unmanured field for 13 years, he was in like Flynn with his farmer friends. His "how-to" book, *The New Horse-Hoeing Husbandry,* became an instant hit.

When Tull's groundbreaking work went into a second version, the editor was a journalist so famed for his prickly style he was known as Peter Porcupine. William Cobbett by name, he had begun his writing career while teaching English to French immigrants in Philadelphia. With remarkable disregard for reality, he wrote diatribes against upstart American ideas about democracy. In 1794 the eminent English scientist Joseph Priestley arrived, on the run from anti-American lab-burning mobs in England, and the loyalist Cobbett greeted him with an article lambasting the liberalism of both Priestley and his pro-American science cohorts back home.

> *Banks* did *have an impeccable scientific background—that is, he was born rich.*

Priestley himself had started life as a church minister and at one time headed a Sunday school. A member of his staff was a guy named Rowland Hill, who would go on to renown when in later life he reformed the English mail service and introduced the first lick-it-and-stick-it, prepaid postage stamp: the Penny Black. Hill's teaching career took off when he became master at the radical new Hazelwood School in Birmingham. The place was gaslit and had central heating as well as a swimming pool; the curriculum included such off-the-wall stuff as applied mathematics and modern languages. In 1822 Hill and his brother published a modest work entitled *Public Education,* which brought Hill to the attention of the left-leaning great and good, including Robert Owen, enlightened mill owner and founder of what would end up being the British Socialist Party.

Like many of his stripe before him, Owen went to America to establish one of the many (temporary) utopian communities fashionable at the time. His was in New Harmony, Ind. When the commune failed, Owen left, but his sons stayed behind and became U.S. citizens. One son, Robert Dale Owen, became a leading figure in Indiana politics, spending two terms in the House of Representatives (where he introduced the bill to set up the Smithsonian). In 1888 his daughter, Rosamond, married (for one week, whereupon the new husband inconsiderately died) fervent Zionist supporter and spiritualist Laurence Oliphant.

Earlier in life Oliphant had set up as a travel writer for the *Times* of London, covering the Crimean War. He also worked as a British spy. At one point he took a job as secretary to Lord Elgin, who was governor general of Canada but who went down in history as the son of the man who, in 1803, snitched the Elgin Marbles. Or, as Elgin pater would have put it: "Removed them for their own good." The marbles (today in the British Museum) were mostly large bits of the Acropolis frieze, dating from the fifth century B.C. In Elgin's time they belonged to the Turkish occupying power, who could not have cared less about

Continued on page 108

Wonders, *continued from page 106*

or more are as far away as a couple of thousand light-years. Like Canopus, they will remain as interesting out there as they are to observers here. (The new Hipparcos survey allows for a more precise mapping of the faint stars out there; anyone who would draw us a map of bright stars from 47 UMa will be gratefully noticed here.)

The apparent path against the stars that our sun and planets trace during the year is the zodiac. It is the relic of the flat, whirling disk of gas and dust out of which observation and theory together suggest that planets are formed. Their orbits still recall that common motion, even though most material was lost to space or to the sun. It is natural to set our imaginary orbit into the zodiacal plane of 47 UMa. That far-off sky has its differences, but it is only a revised version of the naked-eye sky on Earth. The super-Jupiter in the Bear (we don't know whether it is the only wanderer visible in that night) circles in three years, not in 12 as Jupiter does here. From time to time, it will shine about as bright as Venus shines in these parts.

But shift our viewpoint out 10 times as far as the 47 UMa system, and our starry neighborhood is left behind. Stars are plentiful, but only the most improbable stellar events—near collisions or supernova explosions of unprecedented proximity—much affect planets. Only that central sun holds real power over its small siblings, the planets. Given a sun similar to our own, the orbits and chemistry of the planets and the chancy interactions among them are what determine events. Big, nearby planets are influential. They can eject a smaller planet from orbit by rhythmic gravitational disturbances. During the youth of a star—the sun's, too, we believe—planets, once formed, were often expelled as orphans into space or into a fiery merger.

The most likely proposition is that this Great Bear system, with its super-Jupiter close in, has never been home to any keen skywatchers. But we estimate that there are tens of millions of other planetary systems in this galaxy, a metropolis of a few thousand star neighborhoods like our own, a quarter-million candidates in each. We do not know the range of variety among those as yet undetected planetary systems in our vast spiral. We have plenty to learn. SA

Connections, *continued from page 107*

the building and certainly not one being gradually pulverized for lime.

It took 13 years to transport the marbles to England and sell them to the British government for a bargain price that would ruin the Elgin family for two generations. But because Elgin was in dire straits, he was in no position to haggle. A promoter of the acquisition deal was Sir Thomas Lawrence, painter to the king and general artistic biggie. He had started life as a child prodigy and by this time, 1816, was charging such eminences as the Duke of Wellington and the Prince Regent an arm and a leg to do their likeness. Already famous back in 1792, when he was an associate of the Royal Academy, he had also been elected Painter to the Dilettanti Society, entry to which required that you were (a) an aristo and (b) had crossed the Alps in search of culture. Lawrence wasn't and hadn't, but he *was* a friend of Sir Joseph Banks.

Banks was president of the Royal Society at the time and pals with the king, so the Dilettanti rules were waived for Lawrence (you get the idea). But Banks *did* have what was considered an impeccable scientific background—that is, born rich with good family connections and an obsession for botany. So early on, in 1768, it wasn't hard for him to win the coveted post of naturalist on Captain James Cook's first voyage of exploration to the Pacific. Cook's various landfalls inspired Banks later to promote the notion of Australia as the absolutely perfect spot to dump convicted criminals as well as to organize another kind of transplantation: that of Tahiti breadfruit trees to the West Indies.

The first ship sent on this food-aid mission was the HMS *Bounty,* of mutiny fame. After the mutineers had set Captain Bligh adrift in a rowboat, they sailed the *Bounty* off into the blue, and nothing further was heard of them until their descendants were discovered 30 years later on Pitcairn Island, where local scuttlebutt maintained that their leader—the *Bounty*'s first mate, Fletcher Christian—had secretly made it back to England and a life of obscurity, thanks to a passing American ship and its obliging master. He was Captain Folger, the great-uncle of the man who founded the company that made the coffee I was enjoying at the start of this column. SA

WORKING KNOWLEDGE

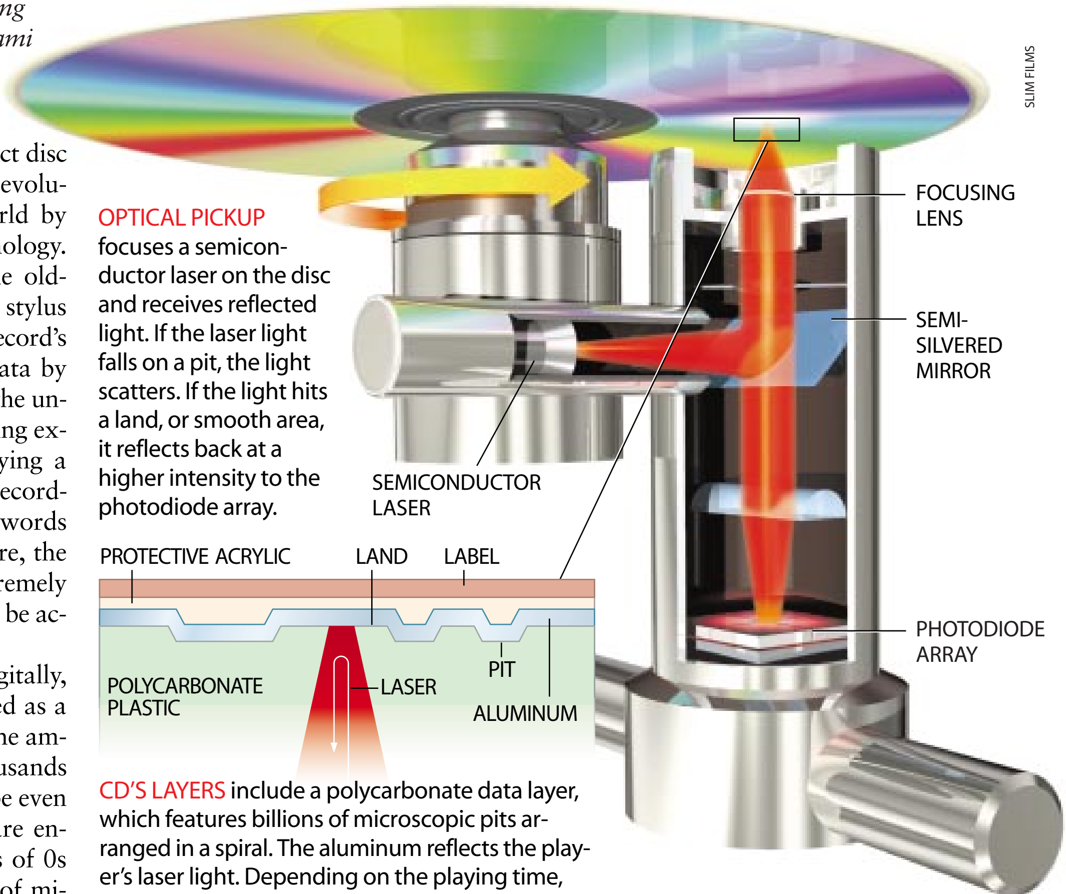
COMPACT-DISC PLAYER

by Ken C. Pohlmann
Professor of Music Engineering Technology, University of Miami

The release of the compact disc and its player in 1982 revolutionized the audio world by introducing optical digital technology. Unlike the analog format of the old-fashioned LP record, in which a stylus traces a groove pressed into the record's surface, a CD player retrieves data by shining a focused laser beam on the underside of the disc. Because nothing except light touches the disc, playing a CD causes no more wear to the recording than reading does to the words printed on this page. Furthermore, the audio quality of a CD is extremely high, and any part of the disc can be accessed quickly for playback.

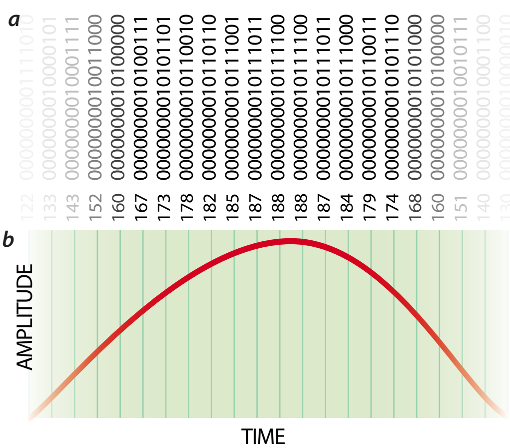
When music is recorded digitally, sound is sampled and represented as a series of numbers that measure the amplitude of the source signal. Thousands of numbers are needed to describe even a brief sound. These numbers are encoded in binary form (as strings of 0s and 1s) and stored in the form of microscopic pits and smooth areas called lands on the disc's data surface. During playback, the CD player's laser light reflects back from the rotating disc at varying intensities as it strikes the pits and lands. A photodiode array detects these fluctuations, which are then translated into 0s and 1s. This binary stream is decoded through demodulation and error correction and converted back into a variable electric signal, which is amplified and played through headphones or loudspeakers.

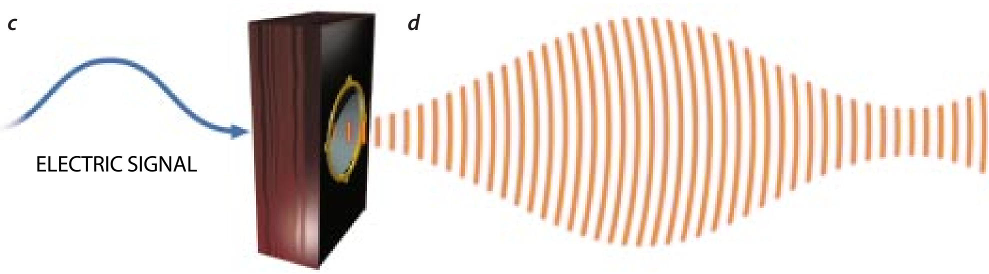
Other types of CDs have also been introduced, requiring specialized players. CD-ROM (compact disc read-only memory) stores data primarily for readout through a computer, and other formats allow one to record and erase information on a disc. The DVD format is the successor to the CD and is already widely used to store feature films with multichannel sound tracks. Both DVD and CD have the same basic optical storage technology, but DVD offers dramatically increased storage capacity.



OPTICAL PICKUP focuses a semiconductor laser on the disc and receives reflected light. If the laser light falls on a pit, the light scatters. If the light hits a land, or smooth area, it reflects back at a higher intensity to the photodiode array.

CD'S LAYERS include a polycarbonate data layer, which features billions of microscopic pits arranged in a spiral. The aluminum reflects the player's laser light. Depending on the playing time, a data spiral might contain three billion pits and stretch about five kilometers (three miles) long.





DECODING THE DISC begins with the data track's pits and lands, which are translated into 0s and 1s (*a*). This binary stream consists of 16-bit strings, each coding for a numerical value between –32,768 and +32,767. These numerical values represent amplitude at points along a wave (*b*). A digital-to-analog converter takes the numbers and produces a correspondingly variable electric signal (*c*). This signal is amplified and fed into a loudspeaker, where it is changed into sound waves (*d*).