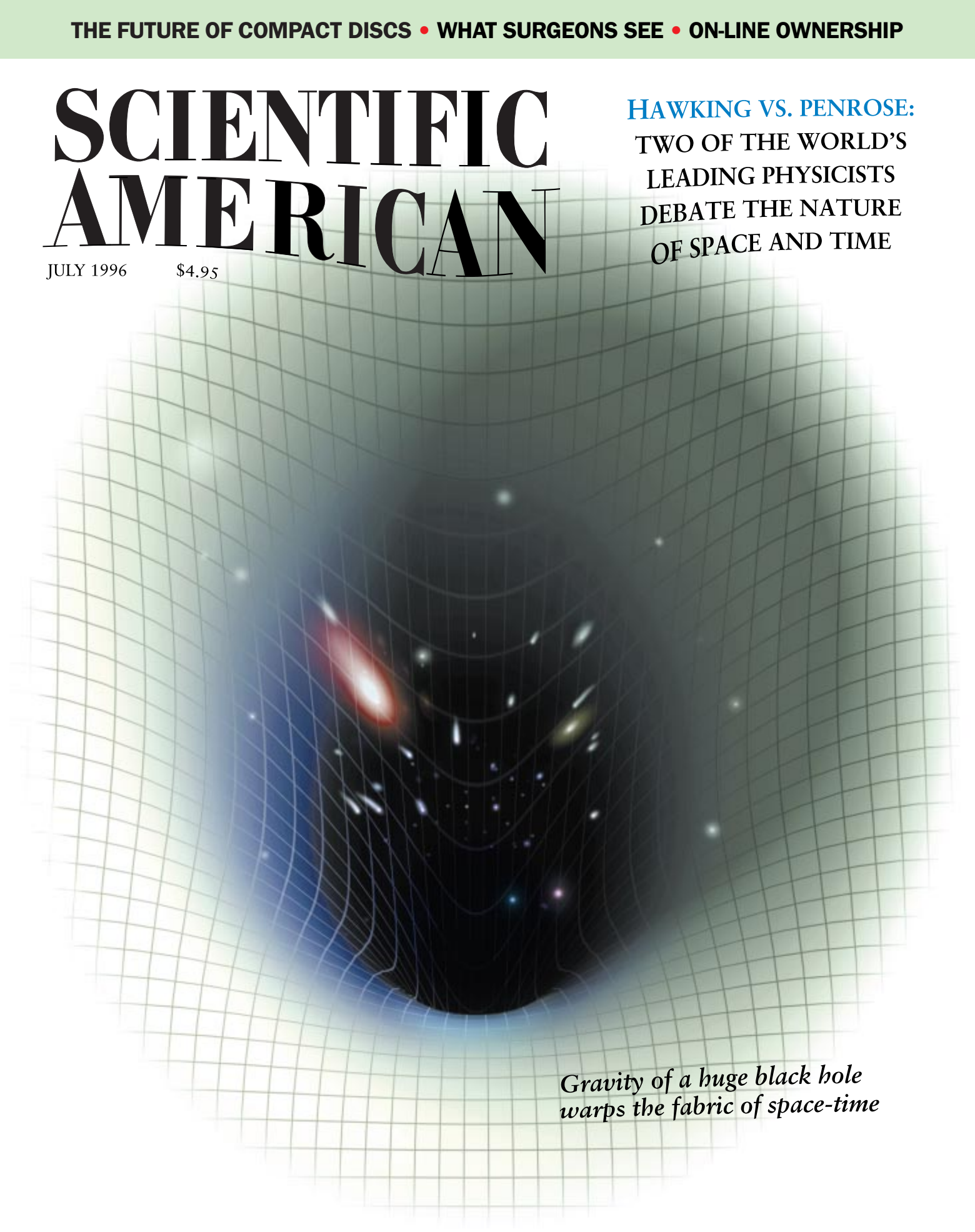
THE FUTURE OF COMPACT DISCS • WHAT SURGEONS SEE • ON-LINE OWNERSHIP

SCIENTIFIC AMERICAN

JULY 1996 $4.95

HAWKING VS. PENROSE: TWO OF THE WORLD'S LEADING PHYSICISTS DEBATE THE NATURE OF SPACE AND TIME

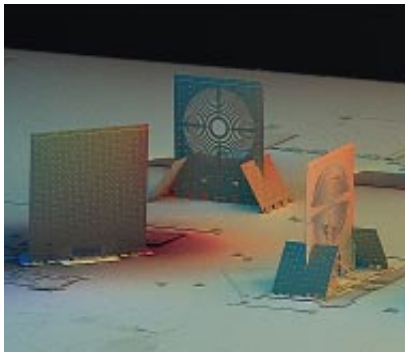


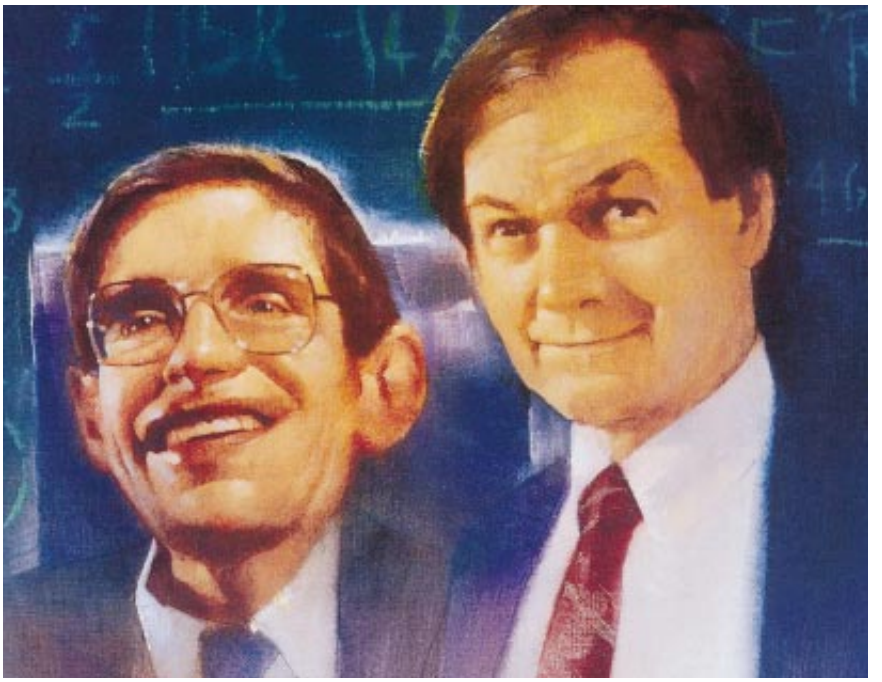
Gravity of a huge black hole warps the fabric of space-time

SCIENTIFIC AMERICAN

July 1996 Volume 275 Number 1

NEWS AND ANALYSIS





About the Cover
Matter and energy falling into a black hole disappear, but what of the information they carry? Stephen W. Hawking and Roger Penrose disagree about its fate. Drawing by Slim Films.

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian GST No. R 127387652; QST No. Q1015332537. Subscription rates: one year $36 (outside U.S. and possessions add $11 per year for postage). Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. Reprints available: write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; fax: (212) 355-0408 or send e-mail to info@sciam.com Visit our World Wide Web site at http://www.sciam.com/ **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.**

FROM THE EDITORS

Glimpses of the Familiar but Unknown

Who can resist a peek behind a door marked "Keep Out! This Means You"? For most of us, the domain of surgeons and their craft is off-limits in just this way. Although scenes set in operating rooms are a staple of movies and television, the focus is always on the interplay of the actors, not on the work itself. Photographer Max Aguilera-Hellweg switches that emphasis, however, revealing the true drama in operating theaters.

Max's photographs have graced *Scientific American* articles several times in the past. Our photography editor, Nisa Geller, introduced him to me two years ago, and thumbing through his portfolio, I became an instant fan. Whereas most medical photography is flat and sterile, his moody use of shadow and rich color evokes memories of Rembrandt; think of the *Anatomy Lesson* paintings. His technique sustains the tension of the surgical moment—and the wonder. Aside from their value as art, these photographs also succeed as materia medica, documents of brave medical accomplishment, of lives saved and futures repaired.

MAX AGUILERA-HELLWEG

A SURGEON'S VIEW is explored in astonishing photographs.

A word of warning for the understandably squeamish: the very power of Max's photographs can make them unsettling. I hope nonetheless that you will find in them a renewed appreciation of life, its frailty and its resilience. We are honored to publish the first significant portfolio of his work, with accompanying notes on the procedures, in "The Hidden World of Surgery," beginning on page 66.

Law, like medicine, is a vast, specialized world unto itself, but some areas of it—sorry, counselors—seem almost unfathomably dull to outsiders. Copyright law in particular occupies one of the grayer zones on the map. That's a pity, because as Ann Okerson notes in "Who Owns Digital Works?" (page 80), some contemplated revisions and amendments to it, now taking the form of pending legislation, could crimp the information revolution.

Digitally copying text, images or other products without permission can intrude on the right of a writer, artist or other creator to control and benefit from those works. Yet some copying may be a fair and reasonable extension of the privileges already enjoyed in libraries, galleries, bookstores and newsstands. Moreover, for technical reasons, duplication of files is unavoidable in many applications, and so a strictly literal enforcement of the suggested laws might make many routine on-line activities impossible. Devotees of the World Wide Web, or any other medium, for that matter, would do well to pay attention to how this dispute is settled.

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

SCIENTIFIC AMERICAN®
Established 1845

John Rennie, EDITOR IN CHIEF

Board of Editors
Michelle Press, MANAGING EDITOR
Marguerite Holloway, NEWS EDITOR
Ricki L. Rusting, ASSOCIATE EDITOR
Timothy M. Beardsley, ASSOCIATE EDITOR
John Horgan, SENIOR WRITER
Corey S. Powell, ELECTRONIC FEATURES EDITOR
W. Wayt Gibbs; Kristin Leutwyler; Madhusree Mukerjee; Sasha Nemecek; David A. Schneider; Gary Stix; Paul Wallich; Philip M. Yam; Glenn Zorpette

Art
Edward Bell, ART DIRECTOR
Jessie Nathans, SENIOR ASSOCIATE ART DIRECTOR
Jana Brenning, ASSOCIATE ART DIRECTOR
Johnny Johnson, ASSISTANT ART DIRECTOR
Carey S. Ballard, ASSISTANT ART DIRECTOR
Nisa Geller, PHOTOGRAPHY EDITOR
Lisa Burnett, PRODUCTION EDITOR

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

Production
Richard Sasso, ASSOCIATE PUBLISHER/ VICE PRESIDENT, PRODUCTION
William Sherman, DIRECTOR, PRODUCTION
Carol Albert, PRINT PRODUCTION MANAGER
Janet Cermak, MANUFACTURING MANAGER
Tanya DeSilva, PREPRESS MANAGER
Silvia Di Placido, QUALITY CONTROL MANAGER
Rolf Ebeling, ASSISTANT PROJECTS MANAGER
Carol Hansen, COMPOSITION MANAGER
Madelyn Keyes, SYSTEMS MANAGER
Carl Cherebin, AD TRAFFIC; Norma Jones

Circulation
Lorraine Leib Terlecki, ASSOCIATE PUBLISHER/ CIRCULATION DIRECTOR
Katherine Robold, CIRCULATION MANAGER
Joanne Guralnick, CIRCULATION PROMOTION MANAGER
Rosa Davis, FULFILLMENT MANAGER

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

Marketing Services
Laura Salant, MARKETING DIRECTOR
Diane Schube, PROMOTION MANAGER
Susan Spirakis, RESEARCH MANAGER
Nancy Mongelli, ASSISTANT MARKETING MANAGER

International
EUROPE: Roy Edwards, INTERNATIONAL ADVERTISING MANAGER, London; Peter Smith, Peter Smith Media and Marketing, Devon, England; Bill Cameron Ward, Inflight Europe Ltd., Paris; Karin Ohff, Groupe Expansion, Frankfurt; Mariana Inverno, Publicosmos Ltda., Parede, Portugal; Barth David Schwartz, DIRECTOR, SPECIAL PROJECTS, Amsterdam
SEOUL: Biscom, Inc. TOKYO: Nikkei International Ltd.

Administration
John J. Moeling, Jr., PUBLISHER
Marie M. Beaumonte, GENERAL MANAGER
Constance Holmes, MANAGER, ADVERTISING ACCOUNTING AND COORDINATION

Chairman and Chief Executive Officer
John J. Hanley

Corporate Officers
John J. Moeling, Jr., PRESIDENT
Robert L. Biewen, VICE PRESIDENT
Anthony C. Degutis, CHIEF FINANCIAL OFFICER

Program Development
Linnéa C. Elliott, DIRECTOR

Electronic Publishing
Martin Paul, DIRECTOR

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

PRINTED IN U.S.A.

LETTERS TO THE EDITORS

OVER THE RAINBOW

Gary Stix's vicious little squib ["The Rainbow Majority," Science and the Citizen, February] exactly confirms the contention of my book *Alien Nation* that antiracism hysteria has paralyzed discussion of the unexpected consequences of the pivotal 1965 Immigration Act to the point where the current situation can be regarded as Hitler's posthumous revenge on America. Stix argues that the shifting ethnic makeup of this country is "inexorable," suppressing the fact that it will not happen without continued immigration. He then concludes by suggesting that opposition to continued immigration is "neoapartheid." It would be fairer to say that his peculiar zeal to see the present American nation displaced by an immigrant "rainbow majority" is a species of treason.

PETER BRIMELOW
Senior Editor, *Forbes*

AIDS AND CIRCUMCISION

John C. Caldwell and Pat Caldwell propose in "The African AIDS Epidemic" [March] that lack of male circumcision has sustained the heterosexual AIDS epidemic in sub-Saharan Africa. Although several studies have documented that lack of circumcision does contribute to heterosexual transmission of the virus, it is doubtful that it is the leading cause of the epidemic in sub-Saharan Africa. Lack of male circumcision may increase the likelihood of HIV transmission from females to males but not from males to females. Hence, it is unlikely that lack of circumcision plays a role in the majority of transmission events.

Also, recent studies have established that a variety of sexually transmitted diseases, not only chancroid, are probably the major determinant in the transmission of HIV from men to women as well as from women to men, regardless of circumcision status. By considering only the role of chancroid in increasing susceptibility to HIV, the Caldwells give these data short shrift. High levels of other diseases, including chlamydia, gonorrhea, herpes and syphilis, undoubtedly go much farther to explain the African AIDS epidemic.

NANCY PADIAN
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

In Africa, male circumcision and female circumcision occur in the same communities. Another researcher with a bias in favor of female circumcision could just as easily suggest that it is *female* circumcision that protects against HIV infection. Amputating parts of the reproductive organs of either sex will not prevent venereal infections. Only education will accomplish this goal.

PAUL M. FLEISS
FREDERICK HODGES
Los Angeles, Calif.

The Caldwells respond:

The heterosexual AIDS epidemic in sub-Saharan Africa has been sustained by an unfortunate concurrence of circumstances: high levels of multiple sexual partners and prostitution, poor medical care (resulting in a high incidence of untreated sexually transmitted diseases) and a large, contiguous population of uncircumcised men. Lack of male circumcision is not the critical factor promoting the spread of HIV, but it is the additional one that distinguishes the AIDS belt from the rest of Africa.

Furthermore, recent work with more than 4,000 women in Nairobi shows that lack of male circumcision trebles the likelihood of male-to-female transmission: if a large number of uncircumcised men, who are more likely to be infected by female prostitutes, bring the disease home to their female partners, more women will contract HIV.

And we do not mean to diminish the role that other sexually transmitted diseases play in increasing susceptibility to HIV infection. The AIDS belt has experienced not merely the usual burden of sexually transmitted diseases but also the added burden of extremely high levels of chancroid, which is so common precisely because most men are uncircumcised. Finally, although many communities practice both male and female circumcision in northern Africa, this is not true in the south, where the epidemic retains its intensity in Malawi, Zambia, Zimbabwe and Botswana.

BRIGHT LIGHTS, BIG CITY

In "Urban Planning in Curitiba" [March], Jonas Rabinovitch and Josef Leitmann write, "As late as the end of the 19th century, even a visionary like Jules Verne could not imagine a city with more than a million inhabitants." What an impoverished imagination poor Verne had! All he had to do was imagine something that already existed. By 1850 the population of London was over two million, and Parisians numbered more than one million. Closer to (our) home, greater New York City's population in 1900 was over three million.

GARETH PENN
San Rafael, Calif.

THE MIRACLE OF MICROBIOLOGY

I respect James Randi's fight to escape medieval superstition through scientific inquiry ["Investigating Miracles, Italian-Style," Essay, February] and do not fault his citation of my work on *Serratia marcescens*. But I wish to clarify that my research did not conclude that the "most celebrated miracle of the 13th century 'may be more microbiological than metaphysical.' "

It was not the miracle but the physical manifestations that occurred during the event that I determined were microbiological in origin. Arguably, my research did not so much disprove the miracle as it supported a sacramental view of nature—one in which God worked through nature to resolve the incredulous priest's doubts and bring him to faith.

JOHANNA C. CULLEN
Georgetown University
Medical Center

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

50, 100 AND 150 YEARS AGO

SCIENTIFIC AMERICAN

JULY 1946

Radio transmission between two points on the earth after reflection from the moon may enter the realm of practicality, now that it has been demonstrated by radar that the space surrounding the earth is not impassable to radio waves. If an ultra-high-frequency pulse was beamed at the moon from an antenna, the receiver could be located at any place on earth where the moon could be 'seen' at the same time as at the transmitter. Hence, the blocking action of the earth's curvature to high-frequency line-of-sight transmission would be eliminated, and nation-wide television broadcasts from a central station might become practical."

"Demand is zooming for prefabricated houses built on mass-production principles. Designs for prefabs range from conventional practice to weird hemispherical structures of aluminum alloy and steel with a central steel mast, the whole unit built on suspension-bridge principles. Some are so radical in appearance that there is a serious question as to whether the public will accept them. No matter what the design, the problem of materials still haunts. Until shortages are relieved, and home builders can perfect their plans, prefabs as well as conventional hand-built homes will suffer delays."

JULY 1896

The Biological Survey will be the name of a brand new government institution to go into existence the first of next month. Besides indicating the sections of the country in which valuable animal and plant life can be raised with success, Dr. C. Hart Merriam says that his survey will determine the zones in which injurious insects, animals and weeds abound, or are likely to migrate when certain species are introduced. This will further save our country, it is thought, many thousands of dollars."

"The ravages of the rinderpest in South Africa are said to be more appalling than any cattle plague which has affected the region within living memory. As an instance of the devastation wrought in Bechuanaland, it is reported that Khama, the paramount chief, who recently visited England, has lost from his private herds alone, 8,000 head of cattle. At Pitsani, at last advices, the cattle were dying by the hundred daily. To the South African native, cattle are a medium of exchange and a staple of the diet."

"According to a report published by the U.S. Department of Agriculture, examinations of milk made at various places yielded numbers varying from 330,000 to 9,000,000 microbes per ounce. The milk supply of Boston was found to be particularly rich in microbes, as many as 135,000,000 germs being found per ounce. Although much has been accomplished in our country of late to improve the sanitary conditions surrounding public milk supplies, a great deal still remains to be done. There cannot be a doubt that the next important step will be the distribution by our dairies of 'pasteurized' milk and butter."

"Our illustration presents a view of a gigantic land turtle from the Egmont Islands, located to the northeast of Madagascar. The length, in a straight line, of the animal's carapax is 4.33 feet. The view of the back, showing a metric measure and four men holding the animal, gives a perfect conception of the size of this gigantic reptile, whose weight is 528 pounds. This turtle probably belongs to the species Testudo Daudinii."

JULY 1846

Simultaneous and instantaneous ignition of gas lamps in towns by means of electricity, states a correspondent, will ere long be substituted for the present slow and irregular method. He further states, 'I confess that I am astonished that electricity has never been enlisted into the service of the *steam engine,* when every clear intellect must perceive that it must ultimately do away with the present employment of fuel and boilers, and their auxiliaries.' "

"It has recently been discovered that there is constantly issuing from the bottom of the Monongahela river, at a point opposite Pittsburg, a highly noxious gas, composed in part of ammonia. Several persons have been drowned while bathing in the river at this place, supposed to have been occasioned by inhaling this gas; and recently a small alligator, having breathed some of this gas, floundered to the shore and immediately died."

Gigantic land turtle of the Indian Ocean islands

NEWS AND ANALYSIS

IN FOCUS

WAKING UP

Finding a purpose for sleep has been as elusive as rest to an insomniac, but researchers are getting much closer

LOUIS PSIHOYOS *Matrix*

SLEEP RESEARCHERS *hope to understand the mechanisms of sleep disorders, which afflict millions of people.*

Sleep may well be "a gentle thing, beloved from pole to pole," as Samuel Taylor Coleridge observed. For physiologists, it remains a biological mystery of the first order. Why should mammals and birds spend such a large part of their lives unresponsive and, worse, vulnerable? Although denying an animal sustenance produces bodily changes that are readily measured, nobody understands what harm is done to an animal—or a person—deprived of sleep. Yet something clearly goes terribly wrong. Researchers have known for more than a decade that a rat prevented from sleeping will lose the ability to maintain body heat and die in about three weeks, leaving no clues in the form of physiological damage. For humans, sleep deprivation undermines thinking, but science has no explanation.

There are, however, plenty of theories—and thus plenty of enmity in the field. Sleepers lower their metabolic rate, thereby conserving energy. But this does not explain why we lose consciousness. Most researchers believe sleep benefits the brain, perhaps by giving neurons a chance to recuperate. Some, pointing to the fervid neuronal activity during the bouts of REM (rapid-eye movement) sleep that punctuate our nights, suggest we doze to consolidate memories. Others propose that dreams are mental junk being eliminated: we sleep to forget. Although it is too soon to proclaim the conundrum of sleep solved, findings are illuminating processes that seem to control it. At the same time, investigators are refining their ideas about the benefits of slumber for the brain. Understanding its purposes may ultimately help the millions of people who suffer from sleep disorders, which range in severity from the merely irritating to the fatal.

The starting point for many investigations into the control of sleep has been the hypothalamus, a platformlike structure in the brain that has long been known to have an important role. Damage to the back part of the hypothalamus causes somnolence, suggesting that when intact, it maintains alert-

ness. Damage near the front part, in contrast, induces insomnia, indicating that the spur to sleep is there. Investigators have long looked for a controlling circuit for slumber that operates between the two halves of the hypothalamus.

The hypothalamus also plays a part in temperature regulation, and some physiologists have speculated that sleep evolved out of a more primitive thermostat. Last year M. Noor Alam, Dennis McGinty and Ronald Szymusiak of the Department of Veterans Affairs Medical Center in Sepulveda, Calif., found the first evidence of neurons that fill both functions. The team discovered neurons in the front part of the hypothalamus of cats that fire more rapidly when they are warmed by two degrees Celsius—and automatically increase their firing rate while the animal sleeps. The researchers suggest that these neurons are part of the body's thermostat and that they are responsible for controlling naturally occurring non-REM sleep.

A related discovery was reported earlier this year by Jonathan E. Sherin, Priyattam J. Shiromani, Robert W. McCarley and Clifford B. Saper of Harvard Medical School. These workers uncovered evidence that clusters of neurons in part of the front hypothalamus of rats—a site called the ventrolateral preoptic (VLPO)—seem to be activated when the animal is not awake. The researchers tracked the levels of a gene product that appears to be present whenever a cell is busy: the busy signal in these neurons was greater in animals that had slept more.

Sherin and his colleagues then took another step. They had previously suspected that neurons in the VLPO region send extensions to the rear part of the hypothalamus. By injecting what is called a retrograde tracer into the suspected target region in the rear of the hypothalamus and then following the diffusion of the tracer, they proved that the sleep-active neurons in the VLPO area did indeed project to the back part of the hypothalamus, where they wrap around their target cells. The pathway "probably is playing a major role and may play a critical role in helping sleep," according to Saper.

JEROME M. SIEGEL

NARCOLEPTIC DOBERMAN is helping scientists comprehend sleep.

Evidence from two quite different avenues of inquiry is consistent with the idea that a crucial piece of the puzzle resides in that region. One is narcolepsy, which affects 250,000 Americans, causing them suddenly and unpredictably to lose muscle control and fall asleep. Any emotionally laden event—even hearing a joke—can trigger such attacks. Neurologists have supposed that some specific type of brain damage must underlie the condition, but nobody has been able to pinpoint it.

Until now. Jerome M. Siegel of the University of California at Los Angeles studied the brains of narcoleptic Doberman pinschers and found destruction of cells in the amygdala, a region involved in emotional responses. Damage to these areas could explain the symptoms of narcolepsy, Siegel suggests. Moreover, neurons run from the amygdala to the front part of the hypothalamus. It is therefore possible, others observe, that cell death in the amygdala might somehow influence the VLPO, bringing on drowsiness and the loss of muscle control characteristic of REM sleep.

Another VLPO clue comes from studies of circadian rhythms, described roughly as a 24-hour cycle of sleep and waking. Recognized as providing one cue for sleep in animal studies, the circadian clock resides in a part of the hypothalamus called the suprachiasmatic nucleus. And the suprachiasmatic nucleus sends neuronal projections to the VLPO, Saper reports. This pathway could be what directs signals about the time of day from the suprachiasmatic nucleus to the VLPO region.

Details of the neural circuitry that turn on sleep beg the question of what sleep is ultimately for. No damage to the brain prevents sleep indefinitely, notes James M. Krueger of the University of Tennessee. Therefore, Krueger argues, the final explanation must involve a benefit to neural functioning. And he asserts that the benefit is closely linked to the immune system.

Krueger points to experiments conducted by Carol A. Everson, also at Tennessee, showing that rats deprived of sleep have high numbers of bacterial pathogens that are normally suppressed by the immune system. Everson says there is little doubt that the bacteria eventually kill the rats. The exhausted, dying rats fail to develop fever, which would be the normal response to infection. Prolonged sleep deprivation, then, apparently dangerously suppresses the immune system. In humans, even moderate sleep deprivation has a detectable influence on immune system cells.

Further, the effect of sleep on the immune system is not a one-way street: the immune system affects sleep in return. Infections are well known to cause sleepiness, and Krueger has shown that several cytokines, molecules that regulate immune response, can by themselves induce slumber. In addition, cytokines have direct effects on neural development. Krueger and his colleagues have recently demonstrated that in rats, a gene for one cytokine becomes more active in the brain during sleep. He suggests that cytokine activity during sleep reconditions the synapses, the critical junctions between neurons, thereby solidifying memories. The cytokines also keep the immune system in shape. Neural pathways like the one in the VLPO region, according to Krueger, may simply coordinate a process that arises at the level of small groups of neurons.

Many physiologists still regard Krueger's ideas as speculative—but later this year Krueger says he will present hard data indicating that cytokines are involved in normal sleep. Genetically engineered mice that lack receptors for two important cytokines, interleukin-1 and tumor necrosis factor, sleep less than usual, Krueger says. So these and related cytokines may well trigger normal sleep in healthy animals, not just the sleepiness of infection and fever.

Whether cytokines, heat-sensitive neurons and the VLPO area indeed hold the key to understanding sleep is a question for the future. But one thing is clear: sleep researchers have never before had so many tantalizing leads or such a full agenda. *—Tim Beardsley in Washington, D.C.*

ARCHAEOLOGY

POT LUCK

Linear A, an ancient script, is unearthed in Turkey

Throughout this century, scholars studying the ancient civilizations of the Mediterranean have pondered a vexing puzzle. The mystery unfolded soon after 1900, when the English archaeologist Sir Arthur J. Evans began excavating the buried palace of Minos at Knossos on the island of Crete. Among the many artifacts found were clay tablets bearing two related forms of unintelligible writing that Evans termed Linear A and Linear B script.

Evans, along with many other classicists, struggled for decades to decode the enigmatic symbols. It was an amateur—a young English architect named Michael G. Ventris—who finally deciphered Linear B in 1952, concluding correctly that the language it represented was archaic Greek. The older and more rarely preserved Linear A code seemed obviously of a different origin, but the identity of that language remained unknown. Now an archaeological discovery in Turkey links the authors of that script—the so-called Minoans—with lands to the east.

There are many thoughts about what language the far-ranging Minoans spoke. Some scholars believe Linear A inscriptions may be in the language of the Hittites, who some 4,000 years ago dominated what is now Turkey. Others suggest that Linear A transcribes Luwian, a more obscure ancient language of that area. Some have proposed that Linear A symbols spell out Semitic words. It also may be completely possible that the mysterious dialect of the Minoans is not related to any known language at all.

Because there is so little certainty about the origin or extent of Minoan civilization, scholars have been particularly intrigued by the recent findings: Wolf-Dietrich Niemeier of the University of Heidelberg's Archeological Institute has discovered Minoan artifacts bearing Linear A script on mainland Turkey, marking a strong connection between the ancient inhabitants of Crete and the mainland to the east.

Niemeier's work began in 1994, at the ruins of Miletus. He had returned to excavations made there by German teams during the 1950s and 1960s. Niemeier installed powerful pumps to lower the water table so that he could explore even deeper levels. Although his initial discovery of Linear A was made during the first season of fieldwork, he did not realize the significance of the find. He thought the curious marks incised on a shard of pottery were just a graffito, a mere doodle. But in the second year his team uncovered two additional pieces with similar inscriptions. At that point, Niemeier remarks, "I recognized it immediately as Linear A." He remembered the earlier discovery: "We pulled out the box with the shard, the so-called graffito, and it matched."

According to Thomas G. Palaima, chairman of the department of classics at the University of Texas at Austin, "There's absolutely no doubt that this is Linear A." With only small fragments of pottery bearing three signs found so far, there is not much to read—even if one knew how. Still, this cryptic message helps to paint a picture of the Minoans who lived some 36 centuries ago.

Because Minoan artifacts have been found on several of the Aegean Islands, experts have wondered whether these people presided over a maritime empire that stretched beyond Crete. Did they, for example, rule overseas colonies, or was it just that they exported their wares? (To make an analogy, one might find Chinese porcelain among items from Victorian England, yet it would be wrong to conclude that China had dominated the British Isles.)

From the type of clay used, it is apparent that the pottery in Miletus was made locally. It is also clear that these Linear A symbols were inscribed before the pot on which they were written was fired. According to Palaima, these facts (and the observation that one of the signs is rather rare) suggest that Minoan speakers must have been there—probably as members of a Minoan colony.

Greater insight into Minoan society would come from reading Linear A inscriptions, but decoding remains elusive, in part because so few examples have been available to scrutinize. Perhaps archaeologists as determined as Niemeier will eventually recover sufficient text to make decipherment possible. But for the time being, the mystery of Linear A endures. —*David Schneider*

COURTESY OF WOLF-DIETRICH NIEMEIER

MINOAN POTTERY
(left) ***recovered from mainland Turkey includes an example of as yet undeciphered Linear A script*** **(upper left).**

PHYSICS

MIRROR, MIRROR

A whiff of supersymmetry at Fermilab

A single piece of data rarely draws much serious attention—except when it comes from the Fermi National Accelerator Laboratory. Earlier this year the lab reported that its Tevatron collider produced an anomaly suggesting that quarks might have structure, a violation of cherished conventional wisdom. More recently, it noted that the collider generated an event hinting that supersymmetric particles were formed.

Supersymmetry, or SUSY for short, is the theory that extends the Standard Model, currently a successful, albeit conceptually incomplete, view of subatomic particles and the forces that act on them. SUSY unifies those two aspects by postulating that every particle of matter has a force partner and that every force particle has a matter partner. An electron has a SUSY mate called a selectron, the photon has a photino, and so on. (Supersymmetry is also a natural outgrowth of superstring theory, currently the most popular "theory of everything.")

For years, though, SUSY remained just a nice idea, for most SUSY particles are hypothesized to be extremely heavy, and the energy needed to create them lay beyond the power of current accelerators. But calculations indicated that on rare occasions particle collisions might give birth to the lightest SUSY particles.

That's what the Tevatron may have created. The world's most powerful accelerator, it smashes protons and antiprotons together, yielding a burst of energy that can form other particles. In April 1995, during the Tevatron's second round of collisions to confirm the discovery of the top quark (the last of the six quarks to be found), unusual byproducts emerged from one collision. Two electrons, two photons and a shortfall of energy were found, which could not be explained by the Standard Model.

The missing energy suggests that a SUSY particle may have emerged from the wreckage but went undetected. More than 10 years ago Gordon L. Kane and his colleagues at the University of Michigan calculated that a proton-antiproton collision could first produce a selectron and its antimatter twin. The selectron would then decay into a photon and a photino, and the photino into a photon and a Higgsino, the superpartner of the hypothetical Higgs boson. The antiselectron would decay in a similar way. With more generic assumptions, Michael Dine and his collaborators at the University of California at Santa Cruz figured that the event may have produced a superlight gravitino, the partner of the graviton, the particle of gravity.

"Anything can happen once," remarks Henry J. Frisch of the University of Chicago, a member of one of the Fermilab detection teams. He notes that the data, which have yet to be published, are still being analyzed. "What one can do is take such an event as a signpost and ask, 'What is it trying to tell me?' "

Possibly quite a bit, at least according to Kane. Higgsino production would not only explain the Fermilab event but would also account for some inexplicable results obtained at the Large Electron-Positron (LEP) collider at CERN, the European laboratory for particle physics near Geneva. There researchers found that so-called Z particles decayed in unexpected ways. Moreover, the Higgsino could also be the invisible, cold dark matter thought to permeate the universe. A Higgsino may weigh about 40 billion electron volts, some 40 times a proton's heft. Calculating the number of Higgsinos that would have been left over from the big bang, Kane finds that Higgsinos provide just the right amount of matter to account for the "missing mass" of the universe.

Although Kane's proposal sounds too good to be true, he argues that several consistency checks in the calculations all seem to point to the same numbers. Barring some technical glitch, Kane figures that the chance of the event happening under the auspices of the Standard Model is slim: the Tevatron would have to run for 20,000 years.

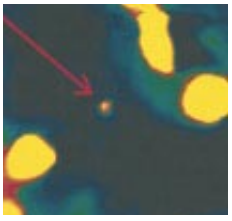
That fact would seem at least to put the SUSY proposal on much firmer ground than the idea of quark structure. Given that other detectors have never recorded evidence of what would be a major aspect of matter, most physicists dismiss the explanation of quark parts.

Whether supersymmetric particles were created may be settled as early as September. This past month LEP began a new round of collisions, and researchers will be on the lookout for supersymmetric particles. "One is certainly not out of line to be optimistic," Kane states. —*Philip Yam*

IN BRIEF

Vive la Francium

Physicists have for the first time trapped atoms of francium, the rarest naturally occurring element. Because francium decays very quickly, the team needed to make one million atoms of it each second. Workers used six laser beams and a magnetic field to hold some 10,000 francium atoms in a space the size of a pinhead. Francium has a very simple structure, and so it may enable scientists to make precise measurements of the weak nuclear interaction.



S.U.N.Y. STONY BROOK

Allergy Relief

Soon peptides and DNA vaccines, rather than antihistamines, may stop the sniffles and itching of an allergy attack. Scientists have created two peptides that block the activity of IgE, an immunoglobulin molecule that normally attaches itself to allergens—say, pollen—and in doing so starts the allergic response. Another group is developing vaccines containing DNA that encodes other kinds of allergens. Through exposure to irritating proteins, such as those made by dust mites, the body may become tolerant.

Forecast on Venus

Venus was long considered the best example of a runaway greenhouse effect: its atmosphere consists mostly of carbon dioxide, and its surface is about 850 degrees Fahrenheit. Earth, too, it was thought, would continually warm if its atmosphere were saturated with carbon dioxide. But according to a new study by a team at the University of Colorado at Boulder, Venus, like Earth, may have an unstable climate system that could suddenly change.

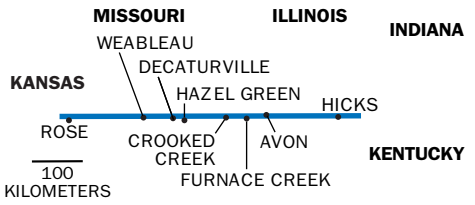
Disease-Free Mosquitoes

The dengue virus, which can cause deadly hemorrhagic fevers, often spreads by way of mosquitoes. Netting and vaccines have failed to keep transmission in check. So scientists have searched out a new solution: through genetic engineering, they have made insects that cannot carry the virus from one victim to the next. The altered genes block viral replication.

Continued on page 24

In Brief, *continued from page 21*

It's a Bird, It's a Plane…
It's a new comet, caught by NASA's Near-Earth Asteroid Tracking (NEAT) system camera. True to its name, NEAT found four new asteroids as well—two of which, scientists say, may someday hit Earth. A recent report notes that a



MICHAEL RAMPINO *New York University*

similarly hazardous asteroid or comet probably touched down in Kansas, Missouri and Illinois 320 million years ago. A string of craters across these states resembles the marks that Comet Shoemaker-Levy 9 made when it bombarded Jupiter for a week in July 1994.

Anatomy Update
A year ago scientists working on a cadaver uncovered an unknown muscle used for mastication. Now confirmation has come from several sources, including *The Dissectible Human,* an interactive CD-ROM by Engineering Animation. The disc, based on images from the National Library of Medicine's Visible Human Project, gives anatomists more to chew on.

Staying Afloat
Scientists now have all the facts about how some fish control the depth at which they swim. Workers knew that many species rise by pumping gas into a swim bladder and sink by sucking gas out of that same organ, regulating the flux by way of lactic acid. The acidity prompts hemoglobin in the blood to release oxygen gas. But what they didn't know until now was that the lower pH radically alters hemoglobin's shape, making it less able to bind to oxygen.

Crystallization Made Easy
To study the structure of proteins, scientists must first arrange them into crystalline layers. It is not a simple task, but a new technique from the California Institute of Technology and the University of Washington should make it easier. Researchers designed special lipids adorned with copper ions. The lipids align themselves into thin films, and the ions bind to an amino acid that nearly all proteins contain. Thus, like tugboats, the lipids grab proteins and pack them into a dense sheet, seeding a two-dimensional crystal in the process.

Continued on page 26

AQUACULTURE

PINK GOLD

The trials and tribulations of shrimp farming

Shrimp farming is an economic mainstay for some countries, but it has also proved a mainstay of controversy. Both wealth and conflict were in evidence at a recent meeting at the United Nations. And although the International Shrimp Tribunal was run informally—no seating assignments, for instance—all the important issues were touched on.

Jason Clay of the World Wildlife Fund opened the session by reading a litany of woes. Shrimp fishing generates enormous bycatch, he stated: 10 pounds of finfish are killed for each pound of shrimp. Space for Third World shrimp farms is cleared by cutting down the mangroves in which many marine fish breed. Catching wild shrimp larvae and rearing them with intense concentrations of food and antibiotics leave coastal waters poisoned and adjacent lands saline. As a consequence, fishermen and farmers are uprooted. Further, the harvest is almost entirely exported; the farm owners are generally wealthy city dwellers who enjoy government support.

The chair, Jacob Scherr of the Natural Resources Defense Council, turned next to the countries slated to make presentations. But Ecuador was awaiting visual-display aids, and India had gone for a walk, so the U.S. described the positive results of attaching turtle-excluder devices to shrimping nets. Shrimp farming in the U.S. accounts for a mere 1,300 metric tons a year, but Alfredo Quarto of the Mangrove Action Project pointed out that in September 1995 a class-action suit was filed against shrimp farmers in Texas for polluting estuaries.

The lights were dimmed. Saraswadi Plod Prasop of Thailand said his government has allocated more than $40 million for mangrove conservation. His video sputtered on, displaying miles of shrimp ponds and shrimp fry reared by the Department of Fisheries wiggling off into the wild. But Quarto stood up again. "More than 50 percent of the ponds on Thai coasts are closed," he maintained. "The pollution problems are not resolved. The industry came to Thailand from Taiwan and is moving to Mexico, Cambodia, India. Behind it lies ruin." A report by Thailand's National Committee on Mangrove Resources estimated that 65,000 hectares of mangrove forests had been lost to shrimp farming by 1993, leaving only 168,700 hectares.

Ecuador—which produces $500-million worth of cultured shrimp a year—was up next and gave a slick presentation. The water flushed from shrimp farms, asserted Juan X. Cordovez of the National Chamber of Aquaculture, an industry group, is cleaner than the water taken in. The farms account for 67 percent of the country's exports; 9 percent of the population is supported by the industry. And the government is trying to stop mangrove destruction.

Gina Chavez of Acción Ecológica, however, had a different perspective. Bribes of $100 per hectare and connections between the industry and government ensure that permits for new farms keep flowing, she alleged; many people are displaced, and few benefit. Cordovez raised his eyebrows: "I don't know where she got her information." From


FULVIO ECCARDI *Bruce Coleman Inc.*

***SHRIMP FARM** in Mexico is just one in a growing industry.*

In Brief, *continued from page 24*

Modeling Life from Clay

James P. Ferris and his colleagues at the Rensselaer Polytechnic Institute and at the Salk Institute for Biological Studies have recently revised the recipe for cooking up life from scratch. Since Darwin's day, molecular chefs had mixed organic and inorganic ingredients in solution, hoping the resulting nucleic acid–based polymers would become long enough to establish a genetic system. But, alas, the polymers always remained too short—suggesting that life did not spring forth from some primordial soup. Now Ferris has found that by letting the polymers condense on claylike minerals, they can become long enough to, in principle, self-replicate and evolve. His finding lends strong support to the notion that clay contributed to the start of life.

Don't Tell…

But two fellows of the American Physical Society are seeking whistle-blower protection from the DOE. Both published classified information that they had obtained from public sources. Hugh E. DeWitt of Livermore National Laboratory quoted congressional debates and received a category A infraction for it. Alex De Volpi of Argonne National Laboratory did clear his paper with the lab before publication but then lost his clearance anyway. Shh.

FOLLOW-UP

It's Not Easy Being Green

Although researchers have been slow to accept the decline of amphibians, new data show that all seven frog species found around Yosemite National Park are threatened—and three varieties have vanished. Suspected causes include increases in fish populations, ultraviolet radiation, chemical pollution and disease. (See April 1995, page 52.)

ALAN BLANK *Bruce Coleman Inc.*

Neutrinos Weigh in…

At two tenths of an electron volt, at least, report scientists at Los Alamos National Laboratory. They detected oscillations from neutrinos, which suggests that the particles have mass after all. (See August 1994, page 22.)

—Kristin Leutwyler

documents of the Ministry of Agriculture, Chavez retorted.

Australia spoke, and half the room wandered off—the country was later congratulated for reducing its fishing fleet. The audience perked up when Japan took the floor and was asked if it cares how its consumption of shrimp—the world's largest after the U.S.—is affecting the environment. Japan said it recognizes its responsibilities but cannot take unilateral measures. Jacob D. Raj of the Indian nongovernmental organization PREPARE registered a protest: he stated that although Japan may discourage shrimp farms on its own coasts, it provides technical assistance to other countries to set them up.

The chair called Ghana, last on the roster. Edwin Barnes said Ghana does very little shrimp farming. A recent U.N. report recommended the pristine African coastline as the next prime location. "But I am beginning to wonder if this is an area we really want to get into," he said. "I need to have a second thought." For the first time, the room broke into applause. *—Madhusree Mukerjee*

FIELD NOTES

Headshrinker Convention

The first thing one notices on entering New York City's cavernous Jacob Javits Center, site of the 149th annual meeting of the American Psychiatric Association, is the Eli Lilly exhibit. The golden, shrinelike tower emblazoned with "Prozac" in Day-Glo red stands amid interactive video screens and fiercely cheerful Lilly salespeople touting the wonders of the best-selling antidepressant (sales topped $2 billion last year).

Some 16,000 people—including psychiatrists, psychotherapists, researchers and drug-company representatives—have gathered here in early May for lectures on everything from "Kids Who Kill" and "The Psychobiology of Binge Eating" to emerging markets for psychiatric services. One "area of opportunity," reveals Melvin Sabshin, medical director of the APA, is forensic psychiatry. "We have more people with psychiatric disorders in jails and prisons than in hospitals," he explains.

JASON GOLTZ

A big buzzword is "parity"—the principle that insurance companies should provide the same coverage for mental disorders as they do for physical ones. A bill calling for mental-health parity won approval from the Senate in April after heavy lobbying by the APA but still has to run the gauntlet of the House. "This is about fairness," declares Marge Roukema—a Republican representative from New Jersey and a fierce advocate of parity—to a cheering audience. Most people who see therapists, argues Roukema (who happens to be married to a psychoanalyst), are not self-absorbed neurotics like the ones depicted in Woody Allen films but people with a real need.

Psychiatrists here voice concern about the encroachment of psychologists and social workers, who usually charge less than psychiatrists do. On the other hand, psychiatrists are M.D.'s and can prescribe drugs, which are cheaper than protracted talk therapy. And psychiatrists flock to breakfasts and dinners featuring lectures on the latest drugs for insomnia and depression—meals sponsored by Pfizer, SmithKline Beecham and other pharmaceutical firms.

Not every attendee embraces the better-living-through-chemistry philosophy. At a session entitled "The Future of Psychotherapy," which is attended by only 20 or so people, Gene L. Usdin, a psychiatrist at the Ochsner Clinic in New Orleans, frets that "we are selling our souls" to the drug companies. Another dissenter is a sales rep for Somatics, which has a modest booth in the shadow of the Prozac pavilion. His company, he claims, provides a far more effective treatment for severely depressed patients: electroconvulsive therapy. *—John Horgan*

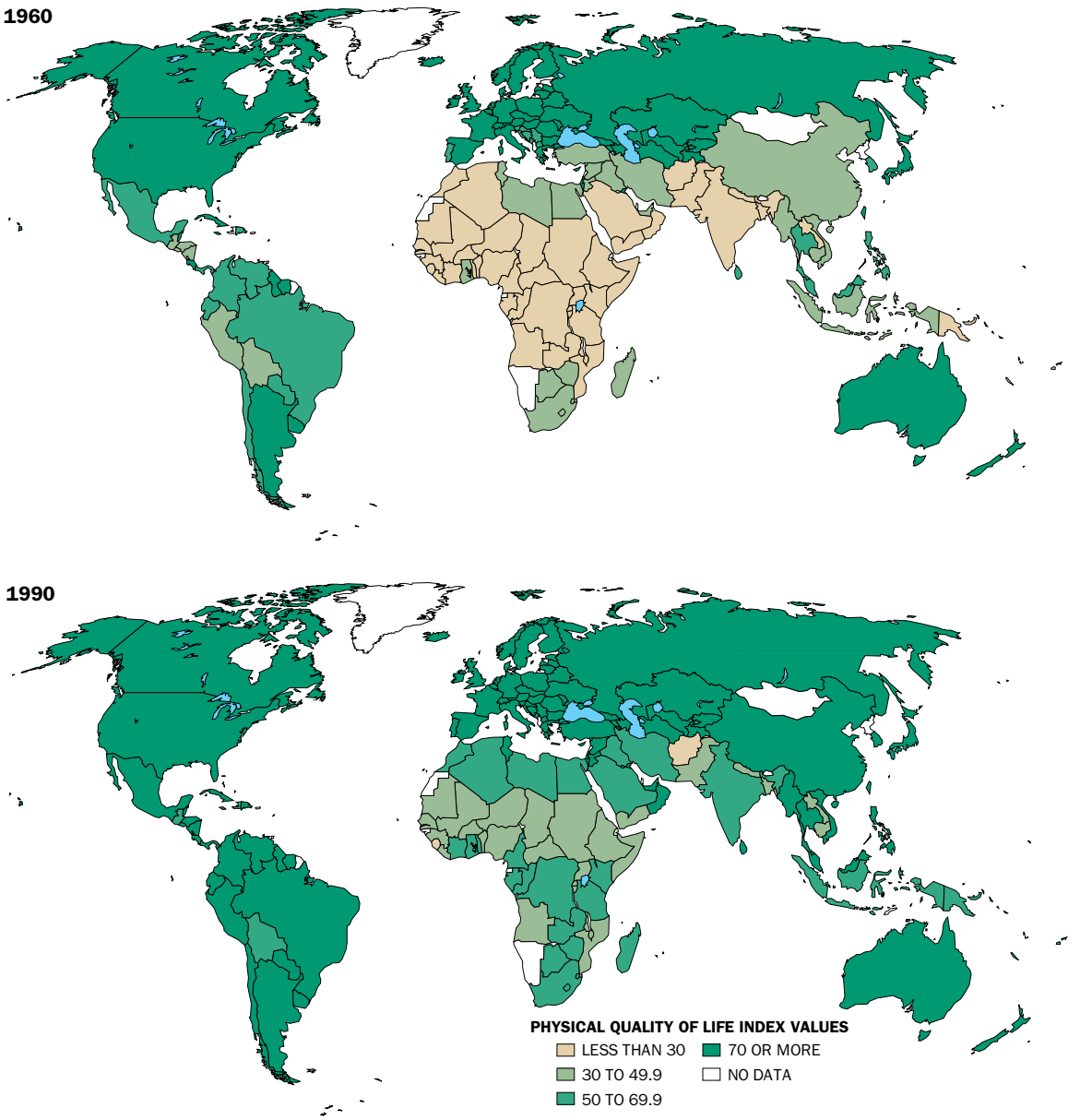
BY THE NUMBERS

The Changing Quality of Life

These maps show the Physical Quality of Life Index (PQLI), developed by Morris David Morris of Brown University to measure progress among the poorer countries. The PQLI is based on life expectancy at age one and rates of literacy and infant mortality. Values range from a low of 6.3 in the West African nation of the Gambia in 1960 to a high of 94 in Japan in 1990. Because the PQLI is based on end results, it has advantages over other methods. Per capita gross national product in Iran, for example, is less than one third that of Saudi Arabia, yet the 1990 PQLI scores of the two countries are identical, indicating that income and wealth are more evenly distributed in Iran.

The most important conclusion to be drawn from the maps is that despite a huge global increase in population, there was considerable improvement in the quality of life among developing nations, including those in sub-Saharan Africa, the poorest region on the earth. Preliminary data for 1993 show further progress in most areas, a major exception being 13 countries in sub-Saharan Africa that suffered drops in PQLI scores averaging three points, which came as a result of decreased life expectancy and increased infant mortality. Losses are caused, at least in part, by the spread of AIDS, which has affected this area more severely than any other.

But the long-term prospect is not necessarily bleak, for the AIDS epidemic may subside, perhaps as early as the next decade. Furthermore, the historical record has registered a more or less steady improvement in the PQLI. Other countries that once had scores as low as those in the sub-Saharan region have shown remarkable change: Sri Lanka had a score of only 19 in 1921, but by 1993 it had reached 85. And 100 years ago the U.S. had about the same PQLI score as the sub-Saharan countries do today. *—Rodger Doyle*



EVOLUTIONARY BIOLOGY

GROUP THINK

A previously rejected theory about natural selection makes a comeback

Slowly, tentatively, the once shunned idea of group selection is creeping back into evolutionary theory. The concept posits that natural selection can operate not only on genes and individual organisms, as most mainstream theorists hold, but on hives, herds, clans and other aggregations of organisms.

Charles Darwin himself speculated that natural selection might favor groups whose members engaged in altruistic behavior, even if the individual do-gooders harmed their own fitness. But in 1966 George C. Williams of the State University of New York at Stony Brook attacked that proposal in his classic book *Adaptation and Natural Selection*. Williams asserted that genes encouraging truly altruistic behavior—defined as acts that increase the fitness of others while decreasing the fitness of the benign individual—would almost certainly vanish over time.

Other analysts argued that altruism could be explained by the concepts of inclusive fitness, in which an individual's apparently selfless act enhances the fitness of close relatives, and reciprocal altruism, which posits that generosity toward non-kin occurs only on a tit-for-tat basis. In the 1970s Richard Dawkins of the University of Oxford pounded a few extra nails into the coffin of group selection with his "selfish gene" model, which built on the work of Williams.

In large part through the determination of David Sloan Wilson of S.U.N.Y. Binghamton, discussions of group selection are sprouting up again. In fact, Wilson has been publishing papers on the topic for some 20 years. Perhaps his most influential paper, however, is one co-authored with his frequent collaborator Elliott R. Sober, a philosopher at the University of Wisconsin, and published in the December 1994 issue of *Behavioral and Brain Sciences*.

Wilson and Sober argue that just as separate organisms can be viewed as collections of mutually dependent genes, so groups can be like "individuals in the harmony and coordination of their

ANTI GRAVITY

Wonderful Town

No wonder it's "the city that never sleeps." A study in the May issue of *Fertility and Sterility* showed that New York City leads the nation in sperm counts. Actually, the study found that the Big Apple outdoes only two other cities. But, more important, the findings contradict previous studies suggesting a global decline in sperm counts.

Unless you are one of those people who thinks Testicles was a hero of the Trojan War, you have probably read about the possible link between falling sperm counts and chemicals that may behave like estrogens. A 1992 paper by Danish researcher Niels Skakkebaek noted that studies done around the world indicated that sperm counts had fallen from about 113 million per milliliter in 1938 to 66 million per milliliter in 1990. Combine that with a rise in testicular cancer and genetic reproductive abnormalities in some countries, and experts began to worry that we were on our way to a future of infertility. Accounts of the controversy appeared in this magazine (which is published in New York).

The new study, by Harry Fisch and colleagues at the Columbia-Presbyterian Medical Center (which is in New York), reports that what Skakkebaek took to be a worldwide decline may have been a misinterpretation of natural geographic variations. "There are geographic variations in everything—cancer and heart disease, for example," Fisch says. "I would be more surprised if sperm counts were the same everywhere."

Fisch looked at counts for about 1,300 men who had donated at sperm banks in New York, Roseville, Minn., and Los Angeles between 1970 and 1994. Rather than diminishing, counts rose in New York and Roseville. The differences among cities, however, were striking. Los Angeles came in at 73 million per milliliter, Roseville at 101. Start spreading the news that New York, N.Y., came in at the top of the heap with a whopping 132.

MICHAEL CRAWFORD IN NEW YORK CITY

This New York talent could account for a misperception of an international decline—apparently, it's not true that if you can make it there, you can make it anywhere. When Fisch examined the 1992 Danish paper, he found that 94 percent of the men studied before 1970 were from the U.S., 87 percent of them from New York. But after 1970, only half the subjects came from the U.S.—and only 25 percent of them were New Yorkers. If geographic variations do exist in sperm production, then what appeared to be a ubiquitous decline may have been merely the result of a shift in study sites.

None of which explains New York's explosive ability for sperm production. "We don't know why New York sperm counts are highest," Fisch admits. In what may or may not be a related story, New York was recently shown to lead the nation in obsessive-compulsive disorder. With so much sperm to count, this was perhaps obvious. *—Steve Mirsky in New York City*

SCIENTIFIC AMERICAN

CORRESPONDENCE

Reprints are available; to order, write Reprint Department, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, or fax inquiries to (212) 355-0408. E-mail: info@sciam.com

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

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

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

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

Advertising correspondence should be addressed to Advertising Manager, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, or fax (212) 754-1138. E-mail: advertising@sciam.com

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

Visit our Web site at http://www.sciam.com/

parts." In the eyes of Wilson and Sober, a group may be not only a close-knit family but also a community of unrelated individuals and even a pair of different species locked in a symbiotic relationship. When these groups compete with one another, natural selection can favor one group over another and so exert a strong influence on the group's characteristics.

The article provoked a guardedly positive response. Thomas D. Seeley of Cornell University is "sympathetic" to the idea that group selection can explain the behavior of certain social insects—namely, bees. Within the same hive, bees certainly compete with one another, Seeley explains, but they may also band together to compete with other hives for resources, thus acting like one organism. But Seeley thinks Wilson "goes overboard" in ascribing the behavior of more complex animals—namely, humans—to group selection.

"There is some value in [Wilson's] terminology and focus," says Paul W. Ewald of Amherst College, an authority on the evolution of diseases, "because it helps us think about the survival of packets of organisms" in fresh ways. For example, Ewald notes, the selfish-gene model implies that microorganisms inhabiting a larger animal should tend to replicate as rapidly as possible—even if that means they kill off their host. The fact that certain parasitic microorganisms settle into a more benign and stable relationship with their host makes sense from the perspective of group selection, according to Ewald.

On the other hand, he observes, avirulence can be explained by the kin selection hypothesis, which is a major component of the selfish-gene model but which Wilson and Sober consider to be a type of group selection. "So much of this argument comes down to semantics," Ewald says. Williams, the original slayer of the group selection concept, has a similar complaint, arguing that Wilson and Sober "define group selection in such a way as to make everything group selection."

Undeterred, Wilson is organizing symposia on group selection for the annual meeting of the Human Behavior and Evolution Society, to be held June 26 to 30 at Northwestern University, and for the American Society of Naturalists, convening August 11 to 15 in Providence, R.I. Wilson and Sober also attempt to explain human morality in group selection terms in *Unto Others: The Evolution of Altruism,* which is scheduled to be published next year by Harvard University Press. *—John Horgan*

PALEONTOLOGY

Triassic Bug

Ancient insects do not often survive death well enough to be preserved in the geologic record. Nevertheless, many such creatures from the Cretaceous period (145 to 65 million years ago) have been recovered because they became entombed in fossil tree sap, or amber. Some much older Triassic specimens (which lived some 220 million years ago) were also delicately captured, as silvery carbon films in sediments deposited at the bottom of a large lake. This fossil of a tipulid fly was recently found on rocks collected at the Solite Quarry in southern Virginia by Nicholas C. Fraser of the Virginia Museum of Natural History and his colleagues.

DAVID A. GRIMALDI

—David Schneider

CYBER VIEW

New Stars for the New Media

Something intriguing is happening in American homes. Computers seem to be luring people away from the television set. It's still too early to tell if this is the long-heralded end of a 50-year obsession with the "idiot box." But it does seem to be the beginning of an affair with CD-ROMs and the World Wide Web, and as it heats up, the door is thrown open for another generation of stars. Just as the salad days of TV were defined by Edward R. Murrow, Lucille Ball and Ed Sullivan, so the Web, too, is ready for characters to bring this emerging world to life.

Every six months, a San Francisco–based market-research firm called Odyssey interviews 5,000 American consumers about their tastes in TV, computers and other electronic media. The most recent interviews, completed this past April, found that home usage of new media was exploding, and as this novel way of spending time became more popular, it appeared to be eroding Americans' loyalty to TV. In general, the more access consumers had to CD-ROMs and the Web, the more disenchanted they were with TV.

So far, fairly few Americans play with CD-ROMs or cruise on the Web: only 15 and 8 percent, respectively. But Web penetration has doubled in the past six months, and use of CD-ROMs doubled over the past year. Both look set to keep growing: over 75 percent of U.S. households are aware of what the Web offers—a percentage double that of six months ago—and ability to link to the Web has become the most important factor in choosing an on-line service provider.

Odyssey found that families surfing the Internet were half as likely as the average household to say they could usually find something on TV they wanted to view (15 percent as opposed to 31 percent). And Internet households were nearly twice as likely to say they would take time away from watching TV in order to use a computer at home (62 percent as opposed to 36 percent). On average, usage of computers for personal work and amusement climbed to 11.4 hours a week, from 8.5 hours a week a year earlier.

As the firm is quick to point out, computers are not the only medium competing to grab American eyeballs. People are also subscribing to more pay-TV on cable stations, tuning in to more direct-broadcast satellite shows and playing more videocassettes. Further, not all new-media enthusiasts eschew TV altogether. Others seem to love everything on a screen—any screen will do.

As computer-based entertainment increasingly becomes a part of day-to-day life, a vernacular is being invented. In the same way that the first years of TV created the sitcoms and newsmagazine programs that still fill the airwaves, so, too, are the early years of the new media starting to produce tomorrow's clichés today—but in very different format. A shift in point of view lies at the center of this change. In the early days of film, the crucial realization was that the camera could be moved, so that the audience's viewpoint need not be fixed somewhere around row G, seat 29. With the power of computers, more can move than the camera: the entire information landscape can be re-created for each user. The trick lies in making a do-it-yourself point of view comprehensible.

DAVID SUTER

To date, there are three leading approaches to customization. Probably the most popular are malls-cum-magazines, which allow the on-line visitor to stroll across the links of information, wandering from site to site like a shopper in a sprawling mall. Pathfinder, Time Warner's version, which brings together all of that firm's products on a single home page, is one of the most frequented sites on the Web. Microsoft seems to be pursuing a similar course for its revamped Microsoft Network. Although these forums have business advantages—they bring brand-name goods to a sole site, which can sell advertising through one sales force—they cannot change the nature of the content they house.

That honor goes to the search engines AltaVista, Lycos, Yahoo, Excite and their ilk—which allow surfers to rummage through billions of pages in search of those they want to read. In effect, they serve up an entire newsstand by ripping the covers off electronic magazines. So far, though, these engines are stronger on completeness than discretion. Most do little more than simply grab all the pages with given words in them—leaving a vast pile to sort through.

Firefly, the Web site created by Agents, Inc., an entrepreneurial spin-off of the Massachusetts Institute of Technology's Media Lab, is evolving a third approach. The basic idea is to match up profiles of the musical tastes and preferences of those using its Web site. Then it can recommend recordings that are enjoyed by people just like you—except that they seem to know about this music that you haven't heard of yet. The recommendations—which could, of course, extend far beyond music—are succinct and customized. But they risk sinking to the lowest common denominator. Hence the opportunity for celebrities in cyberspace.

Until there is the technology to customize the Web to be exactly what you want, an obvious interim measure is to browse according to the tastes of someone you would like to be. Instead of an anonymous network of musicians, why shouldn't Firefly base its recommendations on the tastes of Yo-Yo Ma? And why shouldn't AltaVista filter the results of its raw page searches on new computer technology through the information tastes of Media Lab director Nicholas Negroponte?

All the early stars and characters of TV rendered the medium more approachable. From Beaver Cleaver to Mr. Ed, these characters made instantly recognizable a whole suite of tastes and preferences that would be too hard to explain as disembodied words and pictures. Although the interactions that cyberspace allows with its personality cults will be more complex, a strong character could make the intricacies of new media more manageable. It could also move the conversation between people and machines onto more human ground.

—John Browning in London

ENGINEERING

LIGHT WORK

Micromechanics helps to integrate electronics and optical technologies

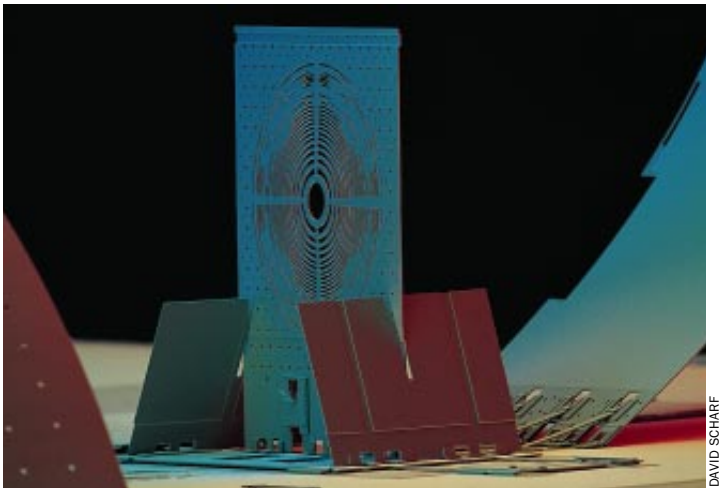
The image of spinning motors no more than a hair's width graced the science pages of newspapers in the late 1980s. But what exactly do you do with a rotor 70 microns wide? The rotating elements on these tiny motors produced only a few trillionths of a newton-meter of torque, enough to get a flagella to its destination, but not much else.

Nearly a decade after the headlines, a few intrepid researchers have found compelling applications for miniature motors and other microscopic machinery. These mite-size actuators may prove ideal for manipulating something as light as light. Vanishingly small machines can move tiny mirrors, lenses and other elements so that photons reflect, transmit or diffract. In their best techno-gibberish, engineers are calling this nascent field micro-opto-mechanical systems, often lovingly reduced to the acronym "MOMS."

Semiconductor manufacturer Texas Instruments has perhaps best highlighted the prospects for this natural coupling of optics with electronics. The company has begun to find a market for an electronic display whose picture elements, or pixels, are controlled by micromechanical structures. The display contains 500,000 aluminum mirrors, each of which measures 16 microns on a side. The mirrors tilt into an "on" or "off" position in response to an electrical impulse. When the display is turned on, pulses of colored light are reflected onto a larger screen. Texas Instruments has sold its display to 13 companies that have begun to market it for projectors or for use in computers [see "Engineering Microscopic Machines," by Kaigham J. Gabriel; September 1995].

But Texas Instruments is no longer alone. A number of universities and small companies, some of which receive funding from the Defense Advanced Research Projects Agency, have begun to explore the use of micromechanics as a technology for processing light. For instance, Mehran Mehregany and his colleagues at Case Western Reserve University have put a scanner on a chip, fashioning a miniaturized version of the kind of supermarket checkout system that reads a bar code with a beam of light. The scanner was made by etching a diffraction grating into the middle of the spinning micromotors. To operate it, an infrared laser diode focuses its light on the grating (a series of parallel grooves in the surface of the rotor). The grating diffracts the beam onto a bar code, and the movement of the rotor scans the beam across the various lines. The light reflected off the bar code hits a photodetector, which translates the signal into an electrical impulse for processing by computing circuitry.

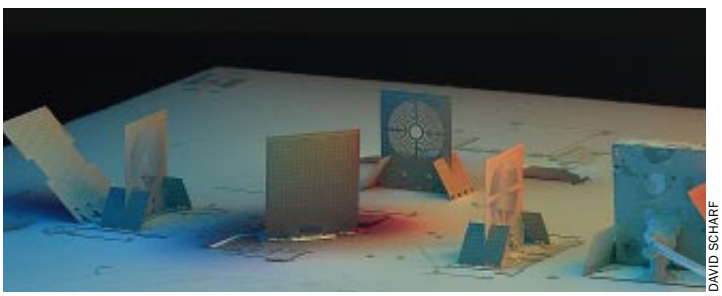
Such microscanners could become small and cheap. The rotating scanning element, which is 500 microns in diameter, is larger than the earliest micromotors. But it is minuscule enough that many spinning wheels can work simultaneously. Such an array could read complex bar codes that contain more than just the product name and price. In addition, a scanner could be used by telecommunications companies to diffract light from an incoming signal, allowing information from one optical fiber to be channeled to several other fibers. Mehregany and his team have also developed a process to create mirror-like finishes on silicon. A silicon mirror, etched onto a cantilevered beam that rises off the surface of a microchip, can tune the wavelength of a laser or serve as a filter in a device capable of converting optical signals into electronic impulses.



DAVID SCHARF

MICROSCOPIC LENS
is 280 microns in diameter and sits atop a silicon microchip.

Researchers at the University of California at Los Angeles have begun to integrate a broad range of optical devices on a silicon microchip. The group, headed by Ming Wu, has made an optical light bench—that is, a miniaturized version of the collection of lenses, lasers and filters used in university physics departments. Aligning and positioning optical components is usually a task painstakingly carried out by skilled technicians. In a large-scale optical bench, components are attached to sophisticated mountings on stabilized tables. In Wu's setup, all the optical elements, except the lasers, can be made on a microchip using conventional lithographic and etching techniques. This approach avoids the expensive and labor-intensive assembly of optoelectronic components by hand.



DAVID SCHARF

SENSING HEAD for an optical disk was made at U.C.L.A.

The U.C.L.A. workers

have fashioned a lens attached to a microscopic hinge, developed by electrical engineering professor Kristofer S. J. Pister. The hinge allows the lens to rotate until it sits at right angles to the chip's silicon surface. Once the lens is raised off the flat silicon, a beam parallel to the microchip surface can pass through it. Although the lens still requires some manual adjustment—a step that researchers are trying to eliminate—the final alignment of the device can be made by micromechanical motors, small plates that turn or move back and forth underneath the lens.

This method might be used to build a read-head (or sensing head)—the mechanical sensor that moves across the surface of an optical disk to decode data. The negligible weight of a micromechanical head would let it move more quickly over the disk surface, speeding up the rate at which data can be read. Wu has used the same techniques to fabricate a mechanical switch that can route an optical signal so that it bypasses a faulty computer.

Other research has targeted micro-opto-mechanical systems for adaptive optics that can adjust light entering a telescope to compensate for distortions introduced by the presence of atmospheric turbulence. The ideas—and the acronyms—are bound to keep flowing. And if all goes well, engineers should learn to love their MOMS. *—Gary Stix*

ASTRONOMY

TREBLE VISION

Combining telescopes makes seeing easier

If one telescope does not provide a clear enough view of a distant star or galaxy, why not use two, three or more? The idea of blending light from different instruments to improve the quality of images has tantalized astronomers for decades. Modern high-speed light detectors and computers are putting the goal within reach for arrays of telescopes a few tens of meters apart.

Earlier this year a group at the Cambridge Optical Aperture Synthesis Telescope (COAST) in England achieved a notable first when it published a test image of a double star obtained by optical interferometry. This technique has proved challenging: in order to form a clear image, the lengths of the light beams from the observed object have to be within a few wavelengths of one another as they come through the different instruments. Unfortunately, atmospheric turbulence continually changes the path length above the respective telescopes; the light beams must be equalized by equipment that monitors fluctuations and rapidly compensates for them. The COAST instruments, for example, bounce their beams off computer-controlled mirrors on moving, rail-mounted trolleys.

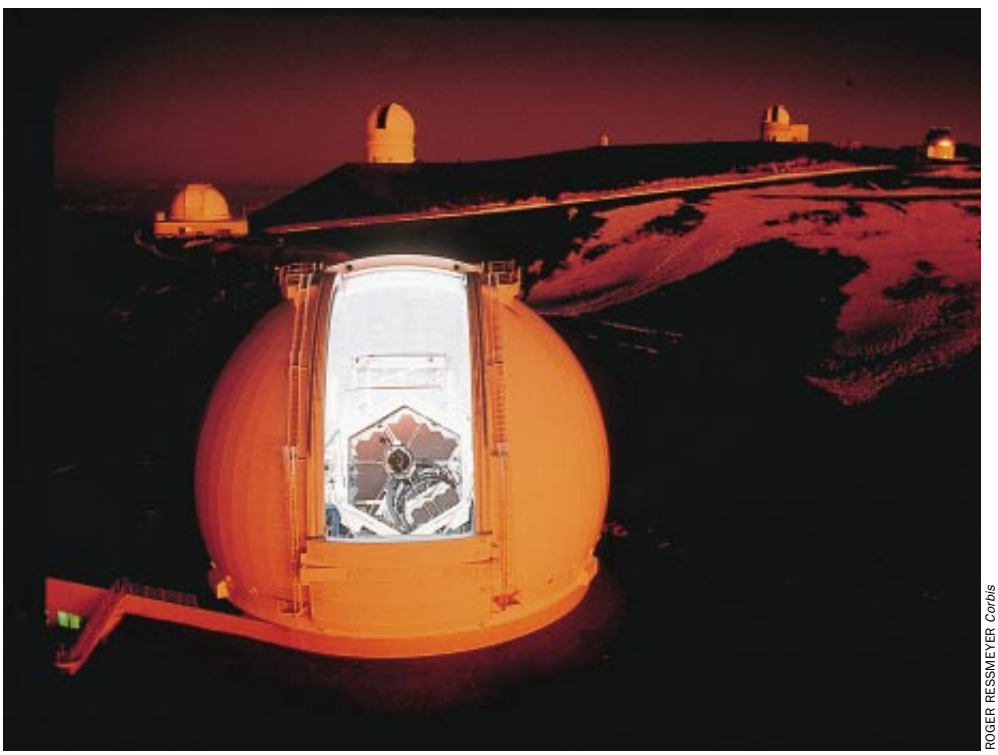
The equipment of the Navy Prototype Optical Interferometer (NPOI) near Flagstaff, Ariz., works similarly. Not to be outdone by the English team, J. Thomas Armstrong of NPOI notes that his colleagues also now have data that indicate a double star. COAST and NPOI both use three small telescopes for their observations, and each will soon introduce a fourth.

Several large planned projects—on Mount Wilson in California, in Hawaii and in Chile—will also use optical interferometry to boost resolution. The Hawaii project may be the most ambitious of the three. It envisages linking the two 10-meter Keck telescopes on Mauna Kea. This double-headed instrument will break new ground because it will use the light from a bright "guide star," which resides close to the object of interest, to equalize the path lengths of the relevant light beams.

The upshot of the innovation is that the Keck interferometer should be able to detect oscillations in the positions of stars. These variations would, in turn, suggest the presence of orbiting planets one thirtieth the mass of Jupiter in systems 30 light-years away. A prototype has proved that the concept, at least, works, according to Michael Shao of the Jet Propulsion Laboratory in Pasadena, Calif.

Although ground-based interferometry can provide much higher resolution than the *Hubble Space Telescope* for small, bright objects, it is not so good with larger, complex ones. The latter require more intricate calculations. So *Hubble* has nothing to fear, at least for a few years.

—Tim Beardsley in Washington, D.C.



ROGER RESSMEYER *Corbis*

KECK I TELESCOPE
is now part of a pair that will use optical interferometry to improve visibility.

ECONOMICS

PANACEA LOST?

Pity the economist who tries to market social insights

Crime, substance abuse and other social ills are generally considered to have a relatively indirect connection to economic theory. But lately economists have been asserting that these problems, too, are subject to the inexorable laws of supply and demand. Raise taxes on beer, and highway deaths will fall, asserts the National Bureau of Economic Research (NBER). Lower drug prices (say, by decriminalizing marijuana or cocaine), and the number of addicts will climb. This spring a paper by Steven Levitt of Harvard University—reporting that increased police hiring is effective against crime—ended up on President Bill Clinton's desk during budget debates.

Levitt's work was based on the observation that mayors tend to beef up police forces during election years to make voters think they are "tough on crime." As elections otherwise have relatively little effect on the rate of mugging, burglaries or petty theft, Levitt reasoned, any reduction in crime that showed an election-year cycle could probably be attributed to increased police activity rather than to other, extraneous causes. Levitt freely concedes, however, that his work probably traveled as far as it did because of its conclusion rather than its impeccable logic.

In general, economists' work has been of less use to policymakers' preconceptions. Christopher J. Ruhm of the University of North Carolina at Greensboro, who did the alcohol-related research that the NBER relies on, notes that 30 of the 48 contiguous states did not change the nominal tax on a case of beer at all during the period in the mid-1980s for which he analyzed data. Because of inflation, that translates to a significant decline in the effective tax rate.

Nevertheless, traffic deaths related to alcohol declined, apparently thanks to other measures, such as more stringent enforcement of drunk-driving laws, the increase in the minimum legal drinking age, grassroots antidrinking campaigns and improvements in automotive safety. Ruhm and his colleagues are left to argue, based on the differing rates of de-

crease in the states that raised alcohol taxes compared with those that kept them the same, that the taxes are more effective than other laws alone. Michael Grossman of the NBER notes that the most recent federal alcohol-tax increase, in 1991, was passed in order to raise revenue, and the legislators rejected efforts to hike the tax rate to a point that might reduce consumption (and thus government income) significantly.

The current test case for economic theory appears to be the war on drugs. Henry Saffer of the NBER and Frank J. Chaloupka of the University of Illinois are among the few who have undertaken studies of the laws of supply and demand for illegal drugs. They correlated a database of prices from 15,000 undercover purchases of heroin and 23,000 of cocaine by the Drug Enforcement Administration with responses to drug-use questions from the Census Bureau's National Household Survey on Drug Abuse given between 1988 and 1991 (the survey does not cover homeless people, prisoners or college students).

After adjustments for the "nonlinear" relation between price and quality, Saffer says, a detectable relation between drug prices and consumption emerged. A 10 percent increase in the price of cocaine led to about a 3 percent decrease in the number of people using the drug more than once a month. (Although, as Chaloupka points out, the effects of cost are hard to disentangle from those of the stepped-up police work usually responsible for price increases.)

Curiously, the Drug Enforcement Administration does not release similarly detailed information on the price of marijuana, even though it commands a far larger market share than harder drugs. Chaloupka says he and Saffer tried to make do with the limited data they could get, but the analysis "did not pay off," yielding relations between price and demand that made little apparent sense. As with any heavily regulated market, it appears that simple economic arguments may be inadequate.

For certain issues, researchers may want to emulate Levitt: while Democrats were citing his conclusion that police hiring reduces crime, Republicans were brandishing another of his papers, more to their taste, showing that increasing the number of convicts in prison has a similar effect. —*Paul Wallich*

NUCLEAR POWER

KEEPING THE "TIGER" AT BAY

With fewer experts and facilities, the DOE is trying new ways of preventing nuclear accidents

One of the most dreaded mishaps involving nuclear materials is an accidental criticality—the inadvertent accumulation of enough fissile materials, typically plutonium or highly enriched uranium, to cause a spontaneous shower of deadly radiation. Since the mid-1940s, there have been 34 known cases of accidental criticality in the U.S. Five of them killed a total of seven people; those five and others also injured workers or subjected them to significant doses of radiation.

Seven of the U.S. incidents, and one that happened in England in 1970, were "process" accidents, meaning they occurred at places where plutonium or highly enriched uranium was being created or processed. Last autumn it was revealed that 12 such episodes took place in the Soviet Union between 1953 and 1978. In the late 1980s, after the cold war, the U.S. and the Soviet Union stopped making plutonium and enriched uranium for weapons. Ironically, however, some experts say the threat of criticality is perhaps greater now than ever.

"The eight previous criticality accidents [in the U.S. and the U.K.] in processing plants were in situations very similar to what we're in today," notes Shirley J. Olinger of the U.S. Department of Energy's Rocky Flats complex near Denver. For the past seven years, the DOE, which oversees the national weapons complexes, has been decommissioning and decontaminating some plants. At certain sites, those tasks involve extracting fissile nuclear materials from ducts or air filters, from decommissioned weapons or from pipes or tanks.

"In the past, process accidents almost always occurred during nonstandard, cleanup-type activities and with materi-

SPACE TECHNOLOGY

The Underwater Lightness of Being

What's bulky and white and wet all over? An astronaut getting used to zero gravity. To prepare for the next *Hubble Space Telescope* maintenance mission, Mark Lee and Steve Smith spent time this past March bobbing inside a 5,600-cubic-meter tank at the National Aeronautics and Space Administration Marshall Space Flight Center in Huntsville, Ala. They evaluated tools, equipment and procedures that will enable them to install an infrared camera and upgrade a guidance sensor, spectrograph and other instruments on the telescope. The sessions in the 12-meter-deep tank helped accustom Smith and Lee to laboring weightlessly in space suits as well as to using foot restraints and handholds. A mock-up of the telescope is shown at the right; part of a dummy space-shuttle cargo bay can be seen at the bottom. Two additional series of dives in the 27-year-old national historic landmark—officially known as the Neutral Buoyancy Simulator—are scheduled before the *Hubble* servicing mission is launched next February. —*Glenn Zorpette*

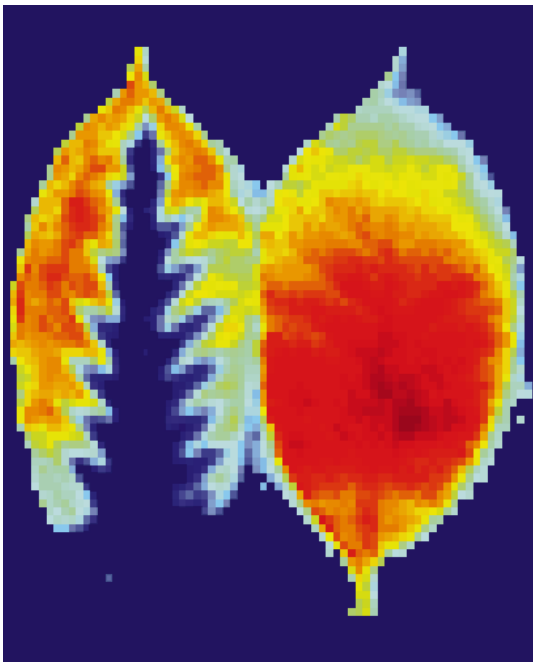
DAN BURTON

MONITORING

Leaf It to Them

It's not quite the handheld medical scanner you've seen on *Star Trek,* but simply push a button, and this spectrophotometer can quickly diagnose what's finishing off your foliage. The device, designed by a team of scientists led by Larry S. Daley and Li Ning at Oregon State University, monitors the fluorescence of living plants. It can detect subtle changes in the vegetation's health long before problems become apparent to the naked eye.

Daley and Ning's instrument relies on the fact that during photosynthesis, leaves produce strong fluorescence signals and that the intensity of the signal coming from a leaf depends on how healthy it is. A plant damaged by frost barely emits a signal; a leaf treated with an herbicide shows an unusual pattern of fluorescence (*left*) as compared with a healthy leaf (*right*). Typical plant monitors do not pick up variations that are restricted to small areas because they look at averages for an entire leaf or plant.



COURTESY OF APPLIED SPECTROSCOPY

Botanists such as Gary A. Strobel of Montana State University see exciting uses for the new instrument in agriculture, environmental monitoring and even home horticulture. Daley cites a project the team has recently begun in collaboration with companies interested in remote-sensing technology: because the cost of coffee goes up after a hard freeze, researchers would like to be able to assess the amount of damage so that prices can be set accordingly. Such measurements might someday be possible from satellites—permitting scientists to see the forest for its leaves. *—Sasha Nemecek*

als in solution," explains Ronald A. Knief of Ogden Environmental & Energy Services in Albuquerque, N.M. Criticalities in solutions have been triggered by accumulations of as little as 1.5 kilograms of plutonium or 2.5 kilograms of highly enriched uranium.

There are other complications now, too. In recent years the DOE has lost to retirement many of its most experienced operators, engineers and criticality-safety specialists, most of whom spent their entire careers at the plants. Also, the DOE has only one laboratory, at Los Alamos in New Mexico, performing safety-related experiments on critical mass. As recently as the late 1980s there were three; the others were shut down for budgetary reasons (the one at Los Alamos narrowly escaped the budget ax in 1993).

There have been no such accidents in the U.S. since 1978. But over the past few years, there have been several serious violations of rules established to prevent them. These procedures are conservative, and in none of these cases was a criticality imminent. In the fall of 1994 an infraction took place at the Y-12 facility of the Oak Ridge complex in Tennessee; another occurred around the same time at Rocky Flats. Each site had been hit hard by the loss of experienced personnel. At Oak Ridge, some drums containing highly enriched uranium were stored in a manner inconsistent with safety rules. At Rocky Flats, workers overstepped their authority by draining some solution from a tank, leaving an excessively high concentration of plutonium. Operations have been partly or entirely suspended, as experts review safety procedures and retrain workers.

Tara O'Toole, the DOE's assistant secretary for environment, safety and health, states that in the past, important knowledge about criticality and other subjects "lived in the individual and was passed person to person." That system is being supplanted by one dependent on worker training, supervision and an emphasis on rules. "We're trying to translate an individual-expertise-based culture to a standards-based one," O'Toole says. "This is a major and essential transformation for the DOE, and it has been painful," she adds. "We're talking about changing hearts, minds and human behavior, and it doesn't happen overnight."

Of course, authoritative information on criticality will still be needed. But with fewer labs, "we don't have the ability to get data on different operations and different configurations," Knief asserts. "In cleanup, what we're finding is fissile material mixed with other materials—it could be iron, silicon or components from steel or ceramic crucibles—and we don't have good experimental data on the neutronic, or chain-reaction, behavior of plutonium or uranium mixed with these residual materials."

Robert E. Wilson of Safe Sites of Colorado, the safety contractor at Rocky Flats, notes that the lack of specific data from critical-mass experiments requires specialists to "develop ways of building added margins of safety into the control system" governing some cleanup tasks. At the same time, the country's sole remaining critical-mass lab is working through a prioritized list of experiments suggested by other DOE sites, many of which need the data for cleanup.

Olinger and others maintain that even the best information and procedures are of little value if workers become complacent. Instilling vigilance can be especially challenging at Rocky Flats, she explains, because it is being decontaminated and decommissioned and because it is one of the few large DOE weapons sites that has never had a criticality accident. Keith Klein of the DOE's Rocky Flats organization agrees. "One of the problems we're acutely aware of," he says, "is when you've lived next to the tiger for so long, you can forget it's a dangerous animal." *—Glenn Zorpette*

PROFILE: GURSARAN PRASAD TALWAR

Pushing the Envelope for Vaccines

Gazing at the cream interior of Gursaran Prasad Talwar's office in New Delhi, I idly count about 40 framed certificates and medals—the French Legion of Honor, the Padmabhushan from India's president, a dancing Shiva with a citation at its base. The medals are displayed in velvet-lined cases laid open on the bookshelves, flanking brightly colored volumes on immunology and contraception. After some 30 minutes, Talwar turns around from reviewing a student's paper to ask what documents I need. The 70-year-old man projects an aura of power and vigor. His accented, measured speech is touched with an edge of wariness; I wonder why. I collect a volume and leave. The real interview is the following day, a Saturday, at Talwar's home.

Talwar, declares Sheldon Segal of the Population Council in New York City, is one of the top three scientists from the developing countries—"maybe the top one." In the 1970s Talwar pioneered a contraceptive vaccine that induces antibodies against part of the reproductive process in women. As the founder of India's National Institute of Immunology (NII) in 1986, he is credited with creating a world-class institute and training a generation of scientists. The ventures flowing from his fertile brain include a vaccine against leprosy, a topical contraceptive derived from the neem tree, a male contraceptive vaccine and others against prostate and lung cancer.

The next morning it takes fully 13 landmarks, sketched by Talwar on a map, for my taxi driver to find the whitewashed house at the end of a labyrinthine road. A uniformed guard opens the gate, escorting me into a living room hung with canvases by prominent local artists. Talwar is elegant in casual *khurta-pajamas;* he pours himself a drink while I settle into the silk cushions and request a coffee to clear my head. I note that I am intimidated by Talwar—as an Indian woman half his age, I automatically take on a respectful tone, and he, a paternal one. We begin, cautiously.

Talwar started his career by studying immunology at the Pasteur Institute in Paris. In 1956, after completing a Humboldt Fellowship in Germany, he joined the brand-new All India Institute of Medical Sciences in New Delhi. "For the first six years, we could do nothing," he recalls. "There were no buildings." After the facilities were built, he worked on ovarian hormones, figuring out an essential way in which estradiol promotes the growth of the uterus. "Talwar was a pioneer in demonstrating these steps," Segal attests. "He has not gotten enough credit."

The research was exciting. "But, you see, living in a country where you are surrounded by so many problems, you cannot remain immune from what is happening around you," Talwar explains. One ailment that caught his attention was leprosy. Only 1 percent of those who are exposed to the disease contract it. Talwar discovered a bacterium, called *Mycobacterium w,* that enhanced their immune response and speeded up the treatment regime. In one of his papers, I find astonishing before-and-after pictures, showing a woman's features transformed from a grotesque glob to a smooth, shiny face.

The vaccine is now undergoing clinical trials in India, along with rival vaccines from the World Health Organization and from Madhav G. Deo of the Cancer Research Institute in Bombay. Later I learn that Deo has alleged that *Mw* is actually the Bombay bacterium, which Talwar acquired and renamed: the two organisms are reported to be almost identical. Talwar admits to receiving Deo's culture but insists that *Mw* is different and that Deo has refused to provide his bacterium for comparison.

A six-year-old girl with dark eyes and a long braid, Talwar's granddaughter Nayana, shyly comes in to display her colorful sketch. Duly admiring it, Talwar continues as she climbs all over him. "You know, in India you have to work on more than one problem at a time. Partly because, at that time, we were very dependent on chemical reagents imported from abroad. You could be held up for months for lack of one chemical." In choosing his second problem, Talwar drew from his visits to the plains of the Ganges. "Even coming from Delhi, I

RAGHU RAI *Magnum*

found the people in Benares to be diminutive, to be like those Japanese trees—what are they called?" I supply the name. "Like bonsai. You have all the features there, but somehow they are more stunted. Why were they nutritively so undernourished?" Concluding that the problem was overpopulation, Talwar observed that the available contraceptives required too much motivation.

Talwar decided to develop a vaccine. His target was human chorionic gonadotropin, or hCG, the hormone that allows an embryo to be implanted in a uterus. Although hCG consists of two subunits, alpha and beta, it was considered safest to stimulate antibodies to just the beta subunit. To induce an immune reaction, Talwar coupled the beta hCG to something the body would recognize as an enemy: the tetanus toxin. The result was a vaccine against pregnancy and tetanus.

"In India you have to work on more than one problem at a time."

In the early 1970s the WHO decided to fund research on contraceptive vaccines, supporting a similar program headed by Vernon Stevens at Ohio State University. The WHO argued that because parts of beta hCG resembled the beta subunit of the luteinizing hormone (LH), Talwar's vaccine caused antibodies to LH to be developed as well, raising fears of complications. Stevens's vaccine, based on a unique fragment of beta hCG, was deemed safer. "If we were both given funding, it would have been okay, stimulated healthy competition," Talwar says with some bitterness.

But Talwar did have supporters. The Population Council stepped in, conducting trials in Finland, Chile and the Dominican Republic, and in 1976 the International Development Research Center in Canada started to fund research in India. The longer beta hCG chain used by Talwar, it turned out, was more efficacious in producing antibodies (although it generated enough in only 80 percent of the women). "Surprisingly, LH was not a problem," says Nancy Alexander of the National Institutes of Health. Trials of the WHO vaccine, on the other hand, were suspended in 1994 after several women developed reactions at the injection site; the trials may resume next year with a reformulated vaccine. Still, Talwar's vaccine needs more work: its effects wear off in three months, necessitating repeated injections.

Talwar has his critics on this front as well. Autar Singh Paintal, a prominent cardiologist who heads the Society for Scientific Values in New Delhi, charges that in 1974 Talwar injected women with a contraceptive vaccine before trying it out on animals—a claim that Talwar says is absurd. But the WHO's David Griffin also reports that Talwar had apparently vaccinated women without adequate animal studies. Whatever the truth may be, women's organizations have strongly opposed the vaccine. One group, Saheli, advises women to avoid the vaccine, warning that they may be "tested upon." Part of the problem, explains Saroj Pachauri of the Population Council in New Delhi, is the adversarial relationship that the Indian population-control program has traditionally had with the women who are its targets.

The vaccine faces other barriers as well. The Population Council has discontinued its trials because of a lack of funds. Although most doctors consider pregnancy to begin when an embryo attaches to the uterus, the U.S. Congress, among others, deems it to begin when the sperm attaches to the egg. By the second definition, Talwar's and Stevens's vaccines are abortifacients and cannot receive American funding.

Talwar pours himself a second drink. "I am feeling in a holiday mood," he explains. "I am having a man come over to give me a massage. A luxury that is affordable in India." The caffeine and the alcohol, I note with relief, have both taken effect. Nayana comes in to play again; we take a break, with Talwar showing off his art collection. I venture to ask more about his life.

Talwar's mother died eight days after he was born—tetanus, he guesses, a major killer of women in childbirth. That realization led him to choose tetanus as the conjugate in his pregnancy vaccine. Brought up in Lahore, Talwar was athletic in college, the captain of his rowing team. Then came the partition: independence from Britain, in 1947, gave bloody birth to two nations, as Hindus and Muslims slaughtered one another. Talwar was in India at the time and suddenly found his home to be in Pakistan. "I did not know where my parents were," he relates. He joined a military convoy put together by the Indians to rescue Hindus in Lahore but, arriving there, found his house stripped, empty.

Talwar, who is Hindu, mercifully ran into a Muslim poet, Hafeez Jullundary, who risked his own life by hiding Talwar in the secluded women's quarters of his home. Sometimes Talwar, who grew a beard and donned a cap to pass as a Muslim, would venture into the city to persuade the few remaining Hindus to leave. With one father and son, he did not succeed; later he found them shot. "Funny that people who looked very much alike, spoke the same language, had to kill each other," he muses dispassionately. Talwar eventually found his father, safe but broken, in India.

If it had not been for the partition, Talwar shrugs, he would most likely have been running his father's factory in Lahore. The only thing he misses about his life there, he confesses, is his room full of rowing trophies. For an instant, I have a vision of his office at the NII, an institution that Talwar built from scratch, planting the trees at its periphery even before construction began.

Talwar has just retired from the NII, to join the International Center for Genetic Engineering and Biotechnology. The contraceptive vaccine project remains with the NII: "I have left all the legacy—grants, money, patents, science and a base. If it materializes or not, I have no say." Talwar is concentrating on the neem cream and his anticancer vaccines.

Yet another controversy surrounding Talwar has to do with TALSUR, his male contraceptive vaccine for animals. Maneka Gandhi, a former minister of environment—perhaps best known as the rebellious daughter-in-law of the late prime minister Indira Gandhi—charges that Talwar's vaccine has killed "a great many" dogs. "He kept saying nothing was wrong with the vaccine, the application was wrong," Gandhi attests. So she had him inject her own pet: "I had to have him operated to save him." But Talwar protests that he does not know of any confirmed deaths.

Nayana is hungry, and so am I. Talwar graciously invites me to share their light lunch—which he says is one secret of his vitality, along with yoga. Tucking the napkin under Nayana's chin, Talwar tells me of his concern for the many problems Indian women face: he has set up a trust that allows women to resume scientific careers interrupted by marriage. Soon I take my leave, walking out into the hot sun to wake up the taxi driver, who is curled up across the front seat. Turning to wave good-bye to Talwar, I realize that the man remains an enigma. —*Madhusree Mukerjee*

Next-Generation Compact Discs

A novel agreement among competing electronics companies has delivered an innovative plan for compatible "DVD" products—the first are due out this fall

by Alan E. Bell

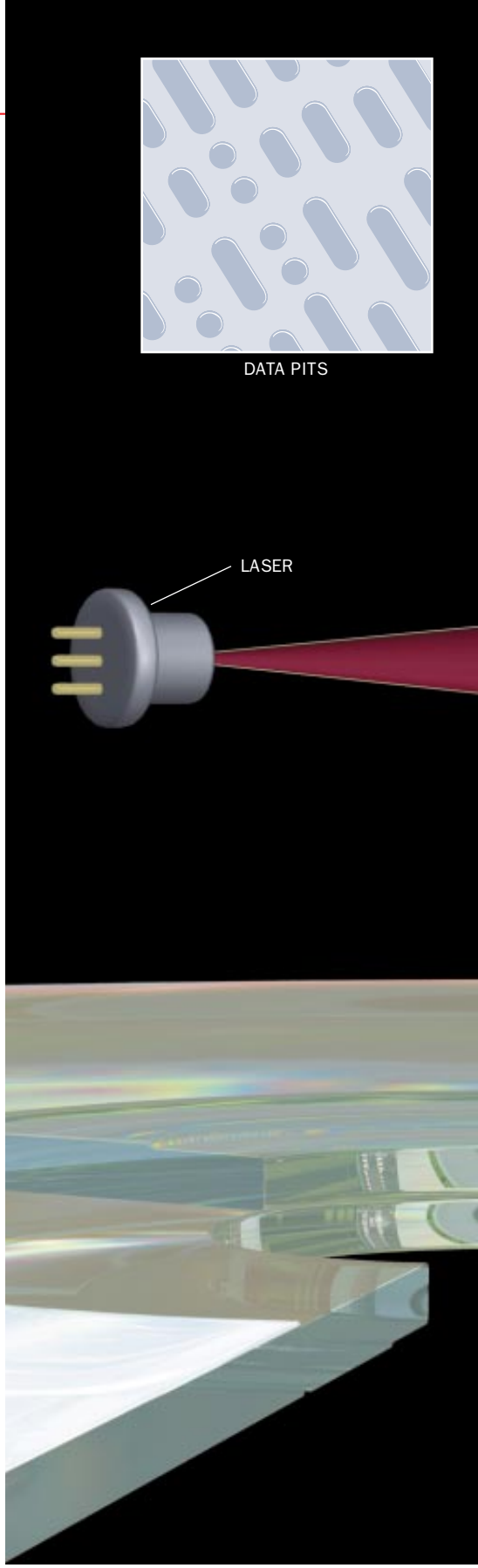
Compact discs and their players are among the world's most successful consumer electronic products. Since CD audio products were introduced in 1982, more than 400 million players and six billion discs have been sold. The CD-ROM (read-only memory), an extension of the original CD audio format, has proved to be equally successful in personal computer applications; forecasts predict that more than 35 million CD-ROM drives will be sold in 1996 alone. The second generation of this technology—based on the DVD, or digital versatile disc, format—will soon make its debut. Ten major electronics companies plan to unveil a range of DVD products, including DVD movie players and DVD-ROM drives, later this year and early in 1997.

The DVD format is the result of an unprecedented agreement reached in late 1995 among rival groups of international companies. The competing groups combined the best features of their individual approaches, which had been developed separately. The new breed of optical-disk reader prescribed by the agreement will play both existing CDs and DVDs—discs that, thanks to a variety of design innovations, can store about 14 times more information than current CDs can. In addition, the rate at which the first-generation DVD player plays back data—11 million bits per second—matches that of a fast 9X CD-ROM player, setting a new benchmark for performance.

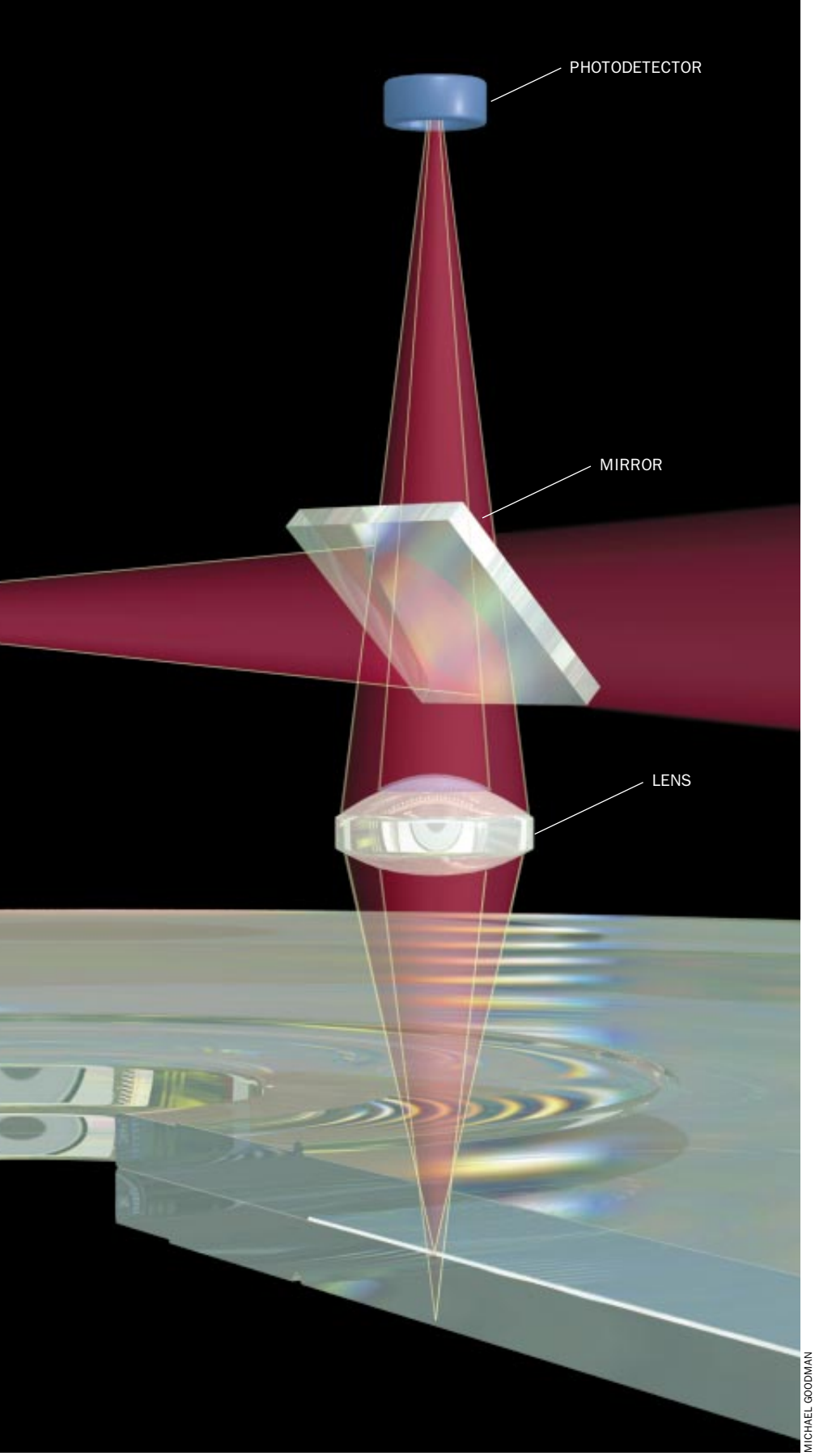
As might be expected, such high capacity and performance lend themselves to an impressive range of applications. DVDs can store music, films, games and other multimedia packages, as CDs already do. But DVDs will contain far more and will play it back with better quality. In addition, DVDs should inspire entirely new products. A video made for a DVD player, for example, might not only store an entire movie but also offer viewers choices between camera angles, plots or sound-track languages as well. Interactive karaoke programs are expected to be popular. And within a few years, recordable DVD-RAM (random-access memory) and DVD-R (recordable) discs and players should reach the marketplace. Looking even farther ahead, I believe digital camcorders based on the recordable DVD-RAM disc should become available.

More Pits

The DVD and CD formats share the same basic optical storage technology: information is represented by microscopic pits, formed on the surface of the plastic disc when the material is injected into a mold. The pitted side of the disc is then coated with a thin layer of aluminum, followed, in the case of a CD, by a layer of protective lacquer and a label. To read the data, the player shines a small spot of laser light through the disc substrate onto the data layer as the disc rotates. The intensity of the light



NEW OPTICAL STORAGE DISKS, unlike CDs, contain two layers of data pits (*shown enlarged in box*). As the disk rotates, these minute indentations pass underneath a beam of laser light (*red*) and



MICHAEL GOODMAN

cause variations in the amount of light the disk reflects. A photodetector and other electronics translate this variation into the 0's and 1's of digital data. Adjusting the position of the lens permits the player to read information from either the upper or lower information layer of a DVD (digital versatile disc). The light passing through the hologram in the center of the lens focuses to a second spot, suitable for reading existing CDs.

reflected from the disc's surface varies according to the presence (or absence) of pits along the information track. When a pit lies directly underneath the "readout" spot, much less light is reflected from the disc than when the spot is over a flat part of the track. A photodetector and other electronics inside the player translate this variation into the 0's and 1's of the digital code representing the stored information.

There are two essential physical differences between CD and DVD discs. First, the smallest DVD pits are only 0.4 micron in diameter; the equivalent CD pits are nearly twice as large, or 0.83 micron wide. And DVD data tracks are only 0.74 micron apart, whereas 1.6 microns separate CD data tracks. So although a DVD is the same size as a CD, its data spiral is upward of 11 kilometers long—more than twice the length of a CD's data spiral. To read the smaller pits, a DVD player's readout beam must achieve a finer focus than a CD player's does. In order to do this, it uses a red semiconductor laser that has a wavelength of 635 to 650 nanometers. In contrast, CD players use infrared lasers with a longer wavelength of 780 nanometers. Also, DVD players employ a more powerful focusing lens—one having a higher numerical aperture than the lens in a CD player. These differences, together with the additional efficiencies of the DVD format described below, account for the huge 4.7-gigabyte capacity of each DVD information layer.

A DVD's capacity can be doubled to 9.4 gigabytes—and nearly doubled again to about 17 gigabytes—by two more innovations. Although DVDs and CDs have the same overall thickness—1.2 millimeters—DVDs possess two substrates that can carry information, whereas CDs have one. A DVD's substrates are bonded together so that their pitted surfaces face each other in the center of the disc. This setup shields the surfaces from the damaging effects of dust particles and scratches. In the simplest design, the second DVD side is accessed by physically removing the disc from the player, turning it over and reinserting it. Another variation—the multilayer design—enables both information surfaces to be played from the same side of the disc.

In a multilayer disc, the upper substrate is coated with a partially reflective, partially transmissive layer. The reflectivity of the upper layer is sufficient to enable the laser beam to read the pits in the upper substrate; its transmissivity

also permits the laser beam to focus on the lower substrate and read the pits in that layer. When the laser focuses on pits in the lower information layer, the pits in the upper information layer are out of focus and so do not interfere. (To accommodate the small but unavoidable loss of playback quality in this approach, a slight capacity reduction to 8.5 gigabytes is necessary—which explains why a double-sided, double-layer DVD would hold about 17 gigabytes.) An optical glue of superior quality must be used to bond the two substrates, and the thickness of that bond must be controlled with precision to avoid excessive aberrations in the focused readout spot.

The two-substrate DVD design offers advantages in addition to increased capacity: it reduces errors caused by disc tilt and warping. All compact discs are prone to warping, and when a disc's surface tilts so that it is no longer perpendicular to the laser beam, reading errors can result. The degree to which tilt degrades the readout spot is directly proportional to the substrate's thickness. The DVD substrate is only 0.6 millimeter thick and so benefits the overall design. This thin substrate makes the DVD less sensitive to tilt than the CD, which has a substrate that is 1.2 millimeters thick. For other reasons, the DVD is less susceptible to certain kinds of warping and tilt in the first place.

For instance, sudden changes in temperature or humidity can cause swelling or shrinkage in the DVD's plastic substrates. But because of the DVD's symmetric construction, changes in one layer tend to counteract those in the other, reducing the overall effect of environmental changes and minimizing the resulting amount of tilt.

Because consumers have already invested a good amount of money in their CD audio and CD-ROM collections, it was considered a top design priority that DVD players read existing CDs as well as new discs. To make DVD players with this ability required specific optical design features. The simplest design is to mount two lenses in a single optical head—one optimized for a 1.2-millimeter-thick substrate and another for a 0.6-millimeter-thick one—and then to switch mechanically from one to the other as needed.

A more elegant solution that emerged uses a single molded optic with a holographic element at its center. The light passing through the outer annulus of the lens is unaffected by the hologram and focuses to a reduced spot size in the plane small enough to be suitable for DVD readout. About one third of the readout beam incident on the central part is focused by both the lens and the hologram to a spot suitable for reading pits on the thicker CD.

More Bits

Besides having more pits than CDs do, DVDs also pack more information into those pits, thanks to improvements in two aspects of format coding efficiency. Whatever the original form of the information—data, text, image, audio or video—the digital 0's and 1's that directly represent the content, called user bits, must be protected from the effects of errors introduced during playback. These errors arise from such factors as dust, scratches or corrosion. Error correction and control (ECC) techniques minimize such problems by way of special algorithms that compute additional data bits to be stored along with the user data. These additional bits, though es-

One Format for All

Eighteen months ago it seemed unlikely that a single approach for second-generation compact discs would soon emerge. Two teams of consumer electronics companies—one led by Sony Corporation and Philips Electronics (co-inventors of the original CD) and the other led by Toshiba Corporation, Matsushita Electric Industrial Corporation and Time Warner—had independently created their own designs. Sony and Philips proposed the multimedia compact disc (MMCD), which would store 3.7 gigabytes on a single-sided disc much like the current CD. But the group led by Toshiba had plans for a radically new format, called SD for super density, capable of holding five gigabytes on each side of a double-sided disc.

Toshiba initially won the support of many in the motion-picture industry, who felt SD was the only format that could accommodate full-length digital movies with CD-quality sound tracks. On the other hand, many manufacturers were more comfortable with MMCD—a conservative extension of the tried-and-true, if overextended, CD format. Neither the studios nor the computer companies, however, wanted two incompatible formats and, possibly, a replay of the VHS/Betamax debacle in videocassette recording. That would slow the introduction of either format until one became dominant. Then those unfortunate consumers who invested in the unsuccessful approach would suffer as content providers switched over to the dominant format.

An ad hoc group of computer data storage experts, originally convened by Sony and Philips to review their format, expanded its scope to review the Toshiba proposal as well. As chairman of this group, I have been at the center of these reviews as they have unfolded during the past 18 months. From the start, the Technical Working Group (TWG), as the team called itself, made clear its foremost recommendation and goal: arriving at one format supporting a family of highly compatible read-only and recordable disc types.

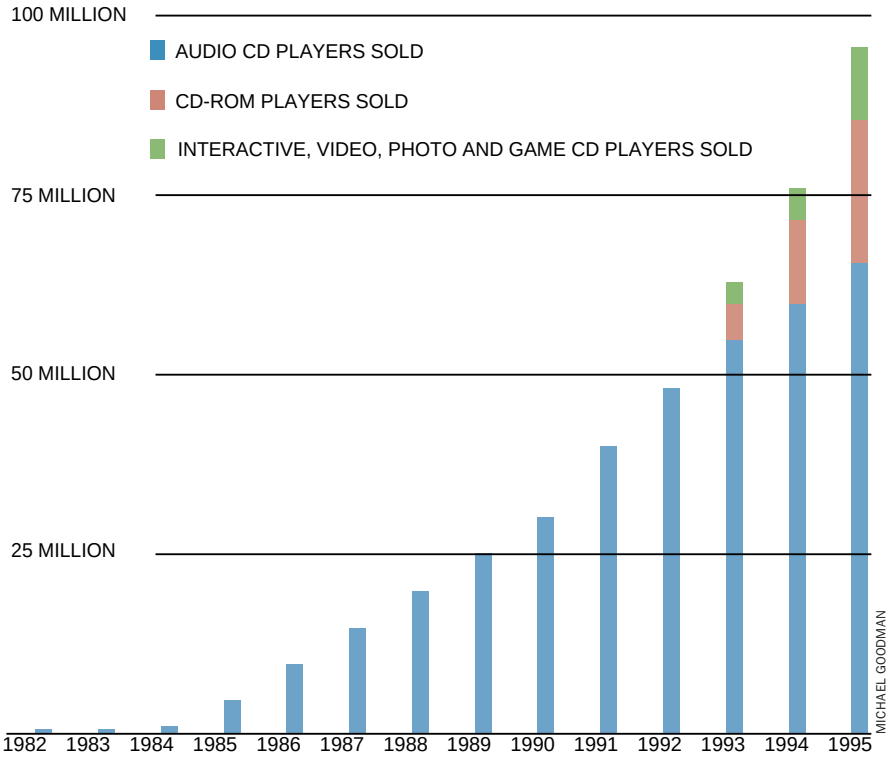
In August 1995 the TWG deemed each format acceptable for computer applications, but the team found it unacceptable that both formats had not been unified. It became clear that some computer companies were ready to choose one format over the other, and shortly thereafter serious talks began to combine the best elements of each design. The format sponsors eventually accepted IBM's offer to arbitrate any technical differences and on September 15 agreed to a single format: it borrowed the thin substrate and error-correction code from the Toshiba/Time Warner proposal and the signal modulation code from the Sony/Philips format.

On December 8 the remaining details of the unified format were confirmed, and the new format for read-only discs was finally announced and christened DVD, for digital versatile disc. Since then, the TWG has continued its work with the DVD Alliance companies, focusing on the technical aspects of the erasable DVD-RAM and write-once DVD-R recordable formats. The TWG now aims to encourage the DVD format companies to maximize the compatibility between DVD-ROM, DVD-RAM and DVD-R media and their products—even across various application environments, including personal computers and entertainment systems. —*A.E.B.*

sential, reduce the fraction of the total disc capacity available for actual content.

Nevertheless, the DVD ECC is extremely powerful. It can, for example, correct an error burst up to 2,000 bytes long, which corresponds to about four millimeters in length along the track. In the DVD format, the ECC data account for about 13 percent of the disc's capacity. In contrast, similar data occupy a full third of a CD's total capacity. The improved efficiency of the DVD ECC, at no cost in error-correcting abilities, results in large part from the increased computing power of today's silicon chips. Such power was not available when the CD format was designed.

During recording, the combined user data and ECC data must be converted into so-called modulation code bits, which are the actual streams of binary bits represented by the pits in the disc surface. This step is necessary to control the range of pit sizes required to represent the data, an important aspect of assuring reliable data detection and tracking during playback. The CD format's coding method transforms eight user bits into 17 modulation code bits. The DVD uses an improved method that transforms eight user bits into only 16 modulation code bits, while preserving the benefits inherent in the original CD method. Because fewer modulation code bits are needed to represent the user data bits, the DVD can hold more of them at once. This feature leads to an efficiency improvement of about 6 percent over the CD format.



SALES OF CD PLAYERS have risen steadily since 1982, when the devices were introduced. The chart shows the number of players sold worldwide over the past 14 years.

More Possibilities

The DVD format will allow consumers, for the first time, to purchase and view full-length motion pictures—complete with a theater-quality sound track—that appear virtually indistinguishable from the original digital master tapes. This capability raises important issues of copyright protection in the digital domain—a subject of ongoing discussion. But it is a fact that the capacity and performance of the DVD was selected so that movie playing could be one of its primary uses.

Despite the huge 4.7-gigabyte capacity of each DVD surface, a digital motion picture must be stored in a compressed form to fit on the disc. Variable data-rate compression, called the MPEG2 standard, is used. With this method, the degree of compression is optimally matched to the instantaneous complexity of the scene: visually detailed or fast-action sequences tend to be more difficult to compress—and are therefore allocated more bits—than the less complex parts of the film. As a result, the data rate of the compressed stream varies during playback. This variable data-rate compression optimizes the overall quality of the picture even when the total capacity is constrained. Another interesting feature is that the user can choose to play a DVD movie in either the standard television aspect ratio (a horizontal to vertical length ratio of 4:3) or the more movielike letter-box form (16:9).

Digital enhancements will benefit the audio quality of DVD movies as well. As anyone who has seen *Jurassic Park* or *Apollo 13* in a state-of-the-art movie theater knows, high-quality, multichannel audio greatly improves the experience. The DVD movie format specifies either Dolby Laboratories AC-3 multichannel (5.1) compressed audio or pulse-code modulation (PCM) audio-encoding methods, with MPEG audio as an option. The Dolby system uses a compressed audio-bit rate as high as 448,000 bits per second to create five independent audio channels having CD-quality audio, as well as a sixth channel dedicated to low-frequency sound effects. (This breakdown explains the system's designation as 5.1.) The DVD specifications for pure audio discs—successors to CD audio discs—have not been completely worked out, pending input from the music recording industry and independent authorities such as the Audio Engineering Society.

The increased capacity of DVDs will make possible richer video games, entertainment products and educational references at about the same disc manufacturing cost as current CDs. Multimedia producers clearly anticipate that the pint pot of today's CD will turn into a gallon container. Rather than spending their time cutting content down to a mere 680 megabytes, multimedia authors will soon have the luxury of considering how to fill with engaging material two-layer discs of 8.5-gigabyte capacity or even four-layer DVDs of 17-gigabyte capacity.

More Ahead

Read/write versions of the DVD drive should appear in 1998, and both the write-once DVD-R and erasable DVD-RAM promise to be much more capable and useful than the CD-R and CD-E (erasable) formats. Until now, optical recording systems have relied predominantly on magneto-optical tech-

How the DVD and CD Compare

Feature	New Format	Old Format
Disc diameter	120 millimeters	120 millimeters
Disc structure	Two substrates, each 0.6 millimeter thick	One substrate, 1.2 millimeters thick
Minimum pit length	0.4 micron	0.83 micron
Laser wavelength	635 to 650 nanometers	780 nanometers
Capacity	Two layers, one on each side, 9.4 gigabytes total Two layers, both on one side, 8.5 gigabytes total Four layers, two on each side, 17 gigabytes total	One layer on one side, 0.68 gigabyte total
Numerical aperture	0.60	0.45
Track density	34,000 tracks per inch	16,000 tracks per inch
Bit density	96,000 bits per inch	43,000 bits per inch
Data rate	11 megabits per second	1.2 to 4.8 megabits per second
Data density	3.28 gigabits per square inch	0.68 gigabits per square inch

LISA BURNETT

nology. But for the DVD-RAM, the recording medium is more likely to use so-called phase-change materials. In this scenario, a thin, ultrafine-grained polycrystalline film is deposited on the surface of the RAM substrate. To store each bit, an intense but fleeting recording pulse from the laser melts a submicron-size region of the film. Because such a small spot of material cools very quickly, the molten region cannot recrystallize. Rather it remains frozen in a disordered, amorphous state that happens to reflect much less light than the crystalline phase does. This reflection difference means that a low-intensity reading beam—one that does not remelt any material—can decode the data.

Considerable research has identified and optimized materials that are sensitive enough to record data, stable enough to preserve those data and durable enough to withstand hundreds of thousands of record/erase cycles. And phase-change recording has several key attractions. Because playback depends simply on the intensity of the light reflected back from the disc, one can use the same optical heads as those found in a DVD-ROM player. In contrast, magneto-optical playback requires unique components that can sense small changes in the polarization of the reflected readout beam. So the cost of a DVD-RAM drive based on phase-change media may not be much more than its read-only cousins. Moreover, DVD-RAM discs can be played by DVD-ROM readers—assuming that all the other details of the format are designed with this compatibility in mind.

At present, a confusing variety of removable data storage devices exist: floppy disks of low and high capacity, removable hard-disk cartridges, magneto-optical media and various magnetic-tape cartridge options. No one expects future recordable DVD-R and DVD-RAM drives to replace the ubiquitous and higher-performance magnetic hard disk. But they certainly have the potential to provide a single solution for the various applications requiring removable read/write media. With the cost of writable DVD media expected to fall to 0.2 to 0.5 cent per megabyte—low enough to satisfy even the most eager Internet-downloading aficionado—recordable DVD drives can offer high-capacity, good random-access performance and low cost as well as compatibility with what is expected to become the pervasive format for information distribution—DVD-ROM. A single DVD-RAM device could thus accommodate many current and future applications of the personal computer system, including desktop publishing, multimedia authoring, content distribution, data transfer, and data archive and backup.

The DVD format was conceived with extensions in mind. For example, the development of reliable shorter wavelength lasers emitting green or blue light can perhaps double data density again [see the accompanying article "Blue-Laser CD Technology," by Robert L. Gunshor and Arto V. Nurmikko, page 48]. We can imagine that derivatives of the basic DVD technology may someday dish out 50 gigabytes or more on one 1.2-millimeter-thick platter—essentially a small library on a single disc.

A collection of experts are now completing the overall specifications for the unified DVD format family. They must strive to make technical choices that will provide the most functional basis for future applications, including many that have yet to be conceived. Just as the medium of television rapidly evolved to become far more than "radio with pictures," so, too, applications based on the new compact-disc format are likely to evolve in surprising and unpredictable ways for years to come. SA

The Author

ALAN E. BELL began his career in 1974 at RCA Corporation, where he invented several key designs that are now used in optical recording products. He joined IBM in 1982 and has held a variety of positions there. During 1995, he was chairman of the Technical Working Group, an ad hoc team of representatives from computer systems companies that helped to unify competing formats into the DVD format adopted for the next generation of compact-disc products. He received his Ph.D. in physics at Imperial College, London University.

Further Reading

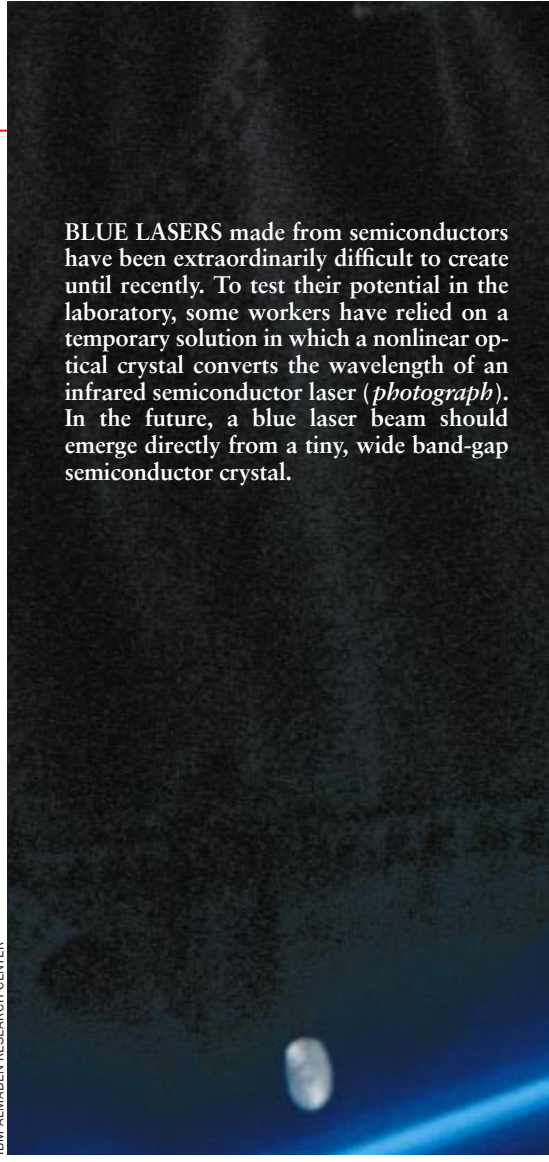
THE COMPACT DISC HANDBOOK. Ken C. Pohlmann. The Computer Music and Digital Audio Series, Vol. 5. A–R Editions, Madison, Wisc., 1992.

A list of frequently asked questions about digital versatile discs is available at http://www.ima.org/forums/imf/dvd/faq.html

Blue-Laser CD Technology

Coaxing semiconductor crystals into lasing blue light is no easy task, but the rewards—among them, greater storage space on optical disks—are well worth the wait

by Robert L. Gunshor and Arto V. Nurmikko



IBM ALMADEN RESEARCH CENTER

BLUE LASERS made from semiconductors have been extraordinarily difficult to create until recently. To test their potential in the laboratory, some workers have relied on a temporary solution in which a nonlinear optical crystal converts the wavelength of an infrared semiconductor laser (*photograph*). In the future, a blue laser beam should emerge directly from a tiny, wide band-gap semiconductor crystal.

The key device inside a compact-disc or CD-ROM player is a tiny but potent laser, which serves as an exceptionally sharp optical stylus. It enables the player to read information stored on the CD's surface in the form of tiny pits. (In the next generation of optical-disk players, the laser also writes information onto the disk by making reversible changes in the material.) The wavelength of the laser's light limits the number of pits—and so the amount of data—that can be stored on the disc: the shorter the wavelength, the smaller the pits it can read.

Currently the lasers inside CD players are made from gallium arsenide (GaAs) and related semiconductors—compounds that, once energized, emit light having a wavelength of approximately 820 nanometers (billionths of a meter). This infrared light can read pits no smaller than about a micron in size, which is roughly one fiftieth the diameter of a human hair. As described in the preceding article by Alan E. Bell, new optical-disk formats are being developed that take advantage of breakthroughs in red semiconductor lasers to increase this information density. But blue diode lasers—emitting light at a wavelength of 460 nanometers—could do even better, because they could read far smaller pits. Marked with these smaller pits, an audio CD could store, say, all nine of Beethoven's symphonies, instead of just one. Multimedia applications would also stand to benefit enormously.

Despite their potential, it has been difficult to create blue lasers for CD players. To see why, we must consider how these semiconductor devices work. The lasers inside CD players or CD-ROM readers are tiny specks of highly perfect, atomically engineered crystals, divided into two main regions. Each side of this diode has a distinct electrical character. On the so-called *n*-type side, an excess of electrons carry electric current. On the *p*-type side, abundant holes—positively charged particles marking the absence of electrons—play the equivalent role. When a positive voltage is applied to the *p*-side and a negative voltage is applied to the *n*-side, the electrons and holes rush toward one another.

The particles meet in an ultrathin trench, a no-man's-land termed the quantum well. There they recombine, annihilating one another and, under favorable conditions, emitting photons, the smallest units of light. When the emission is coupled with a feedback mechanism—namely, a highly reflective pair of mirrors embedded in the device that "recycle" the photons—lasing results: all the photons are coherent and endow the emerging pencil-sharp beam with its exceptional qualities.

The energy of the photons, equivalent to the wavelength of the emitted light, is determined by a fundamental electronic yardstick: the band-gap energy, or roughly, the "electronic elasticity," of the semiconductor material in which the recombination event occurs [*see illustration on page 51*]. For diode lasers made from GaAs, the band-gap energy is about 1.45 electron volts. To produce blue wavelength light, the band-gap energy needs to be nearly doubled. Thus, researchers must turn to another class of semiconductors, called the wide band-gap materials, examples of which include zinc selenide (ZnSe), a II-VI compound, so named for its placement on the periodic chart, and a III-V compound, gallium nitride (GaN). These materials might be better described as semi-insulators, a name that highlights just one basic problem in using them in electrical devices.

In the early 1980s a handful of scientists began trying to form II-VI semiconductor crystals using a technique called molecular beam epitaxy. In this method, atomic "showerheads" in an ultrahigh-vacuum chamber gently rain constituent elements on a substrate, creating an atomic "skyscraper."(The process resembles erecting a building brick

by brick, in this case with atomic precision.) Using this approach, scientists quickly devised "designer materials" of high crystalline quality. From this work came a better understanding of how quantum-well layers worked and how researchers could coax blue and green photons from sophisticated man-made materials. (The wavelength of green light is only slightly longer than that of blue light.) But the quest for useful blue and green diode lasers was not ended. These early devices emitted light only when they were energized by another, desktop-size excitation laser.

For nearly a decade, investigators could find no way to energize blue and green lasers electrically. Although they could easily dope ZnSe crystals with an excess of electrons, they could not similarly provide them with an overabundance of holes. Thus, they could not fabricate a *pn*-junction. Then, in 1990, researchers perfected means to incorporate nitrogen atoms into the process and finally succeeded in doping ZnSe with extra holes. The next summer 3M laboratories and, independently, our joint team at Brown and Purdue universities described the first blue and green diode lasers. These lasers could not be operated continuously and had to be cooled to the temperature of liquid nitrogen (77 kelvins). Shortly thereafter, though, the 3M and the Brown-Purdue groups had improved their designs, achieving both continuous operation at 77 kelvins and pulsed operation at room temperature.

In the fall of 1993 researchers from both Sony laboratories and the Brown-Purdue group attained the continuous operation of a diode laser—energized with only five volts—for up to tens of seconds at room temperature. The very first blue and green laser devices needed as much as 30 volts. The dramatic improvement resulted from some further crafty atomic engineering of the materials. Most recently, Sony has created a green diode laser that emits light at a wavelength of 520 nanometers and operates continuously for about 100 hours at room temperature. The Brown-Purdue team has demonstrated a blue laser at 460 nanometers.

Of course, 100 hours is still insufficient for practical applications. But knowing that the now ubiquitous GaAs-based infrared diode lasers suffered from essentially the same problems at their infancy, we are optimistic. Very recently, Nichia Chemical Industries in Japan has made progress working with gallium nitride materials. Researchers there have produced beautiful blue light–emitting diodes and made preliminary demonstrations of a blue laser, though under pulsed, high-voltage conditions.

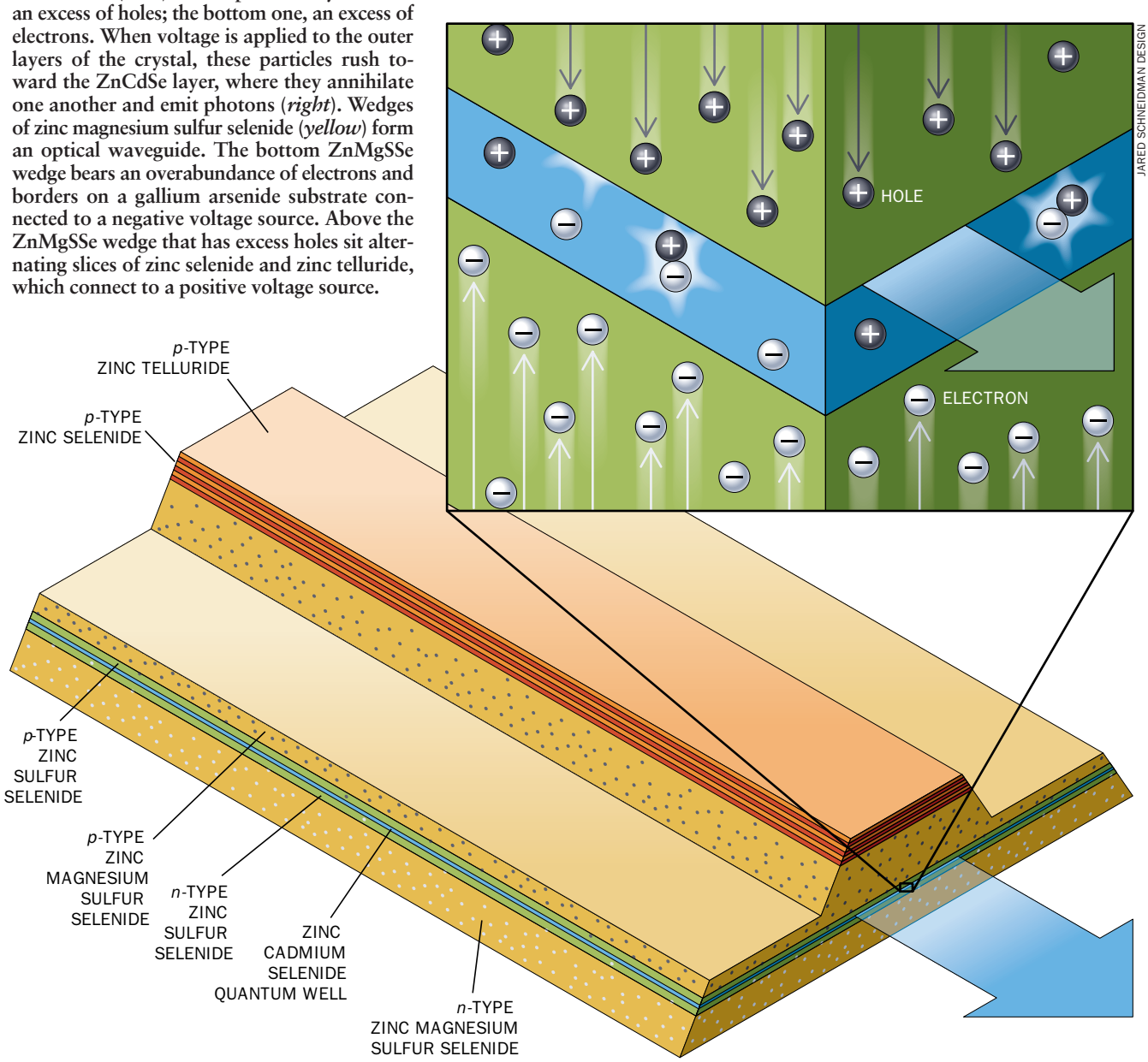
The Bright, Blue Future

Despite their limitations, blue and green diode lasers can be improved in a number of ways. By paying detailed attention to the atomic arrangements in

the crystal, engineers can make crystals that can better withstand the very high current densities required to energize a laser. The illustration below shows a schematic of a laser device constructed from a combination of II-VI semiconductors, mostly derived from ZnSe. The specific sequence of precision-engineered layers optimizes the delivery of the input electrical energy into the ultrathin zinc cadmium selenide (ZnCdSe) quantum well, the electronic and optical heart of the laser. The arrangement also provides the proper optical housing for guiding the emitted photons and generating the feedback mechanism. The quantum well—which is about one one-hundredth of a micron in thickness—is formed by confining electrons and holes between the two zinc sulfur selenide (ZnSSe) layers, which have a wider band-gap energy.

Although the recombination of one electron and one hole in a semiconductor may seem fairly intuitive, quantum mechanics strips these particles of their individuality. In fact, the electrons and holes pair off by electrostatic (coulombic) attraction and form pairs called excitons. Typically, at room temperature, vibrations in the crystal lattice break these entities apart. So, too, if many particles are packed into a small area, the scattering between them breaks excitons. But in a thin layer of a wide band-gap semiconductor, such as ZnCdSe, the excitonic pairs are squeezed so much that they stay correlated even at room temperature or in an overcrowded laser device. These longer-lived pairs are actually more likely to emit photons, and so the device requires less current to sustain lasing.

Another design feature corrects a problem that arises because the wavelength for blue and green emissions exceeds the thickness of the quantum well by nearly two orders of magnitude. Yet another material, zinc magnesium sulfur selenide (ZnMgSSe), is used to define an optical waveguide, which is used to trap the light in the vertical direction.



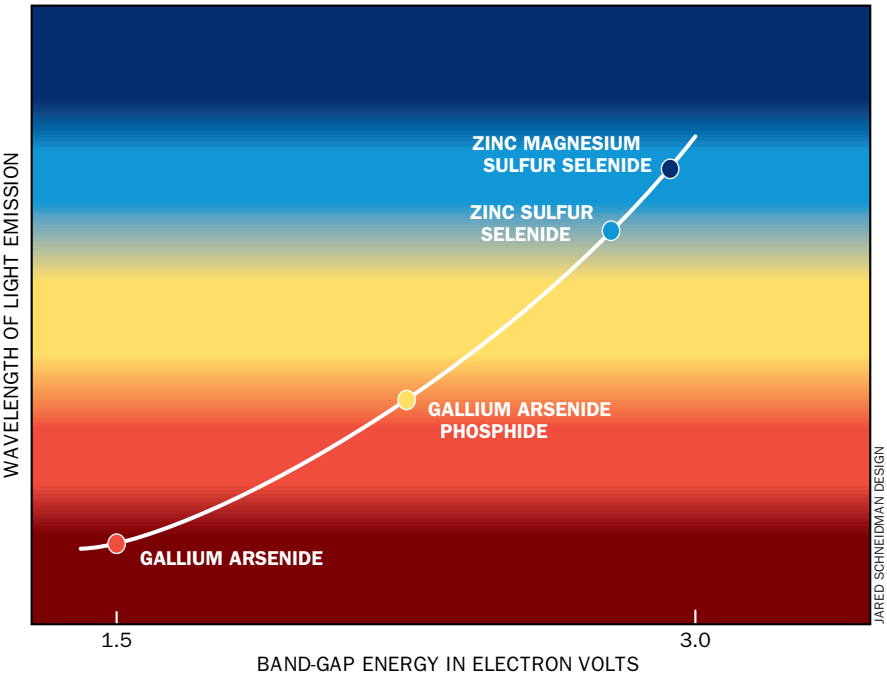
SEMICONDUCTOR CRYSTALS must be arranged in an exact manner to produce blue light (*below*). Two thin layers of zinc sulfur selenide (*green*) surround a sliver of zinc cadmium selenide (*blue*). The top ZnSSe layer carries an excess of holes; the bottom one, an excess of electrons. When voltage is applied to the outer layers of the crystal, these particles rush toward the ZnCdSe layer, where they annihilate one another and emit photons (*right*). Wedges of zinc magnesium sulfur selenide (*yellow*) form an optical waveguide. The bottom ZnMgSSe wedge bears an overabundance of electrons and borders on a gallium arsenide substrate connected to a negative voltage source. Above the ZnMgSSe wedge that has excess holes sit alternating slices of zinc selenide and zinc telluride, which connect to a positive voltage source.

In 1993 researchers at Philips Laboratories showed how this light guiding improved the laser's operation. Lithography and other similar processing techniques physically remove material in the lateral direction, creating an optical waveguide with a "mesa" effect. As a result, the entire structure channels blue and green light in predominantly one (axial) direction; the radiation bounces between the crystallographically cleaved end facets that form nearly perfect mirrors, then exits through one of them as well. An analogous arrangement, based on the III-VI nitride compounds, has given rise to recent reports of blue laser emission from these semiconductors; the quantum well is composed of indium gallium nitride (InGaN), with GaN and aluminum gallium nitride (AlGaN) forming the electronic and optical confinement layers, respectively.

The topmost part, above the *p*-type ZnSSe layer, is called the Zn(Se,Te) graded gap contact, made from alternating layers of zinc selenide and zinc telluride. One problem with the first blue and green diode lasers was that although they could be operated continuously at room temperature, they would overheat in less than a second if connected directly to a battery source. This heating happened at the electrical contact, where a metal electrode met a *p*-doped semiconductor. A huge electronic mismatch, known as a Schottky barrier, existed between the two and inhibited conduction. It was not until late in 1992 that researchers from the Brown-Purdue group and, independently, from Sony found a way around the dilemma. The solution involves a graded alloy of ZnSe and ZnTe that replaces the sharp barrier to conduction between the metal and ZnSe with a gentle incline.

Most research worldwide now focuses on increasing the lifetime of continuously operated II-VI semiconductor blue and green diode lasers and on improving the quality of the nitride compound materials. The present obstacle for the ZnSe materials—which otherwise provide an extraordinary display of prowess in modern materials science—is a crystalline defect that occurs when the first layers of ZnSe are nucleated on a GaAs substrate, which shares roughly the same lattice dimensions. A few atoms here and there (perhaps one in a million or so) sometimes align themselves incorrectly, triggering the growth of an extra plane of material. The defect, called a stacking fault, expands as the crystal grows. When it passes through the quantum well, it offers incoming electrons and holes sites where they can recombine nonradiatively—that is, without producing a photon. Each of these nonradiative events releases approximately 2.5 electron volts of energy to the nearby crystal lattice—an amount comparable with that holding the crystal together. Hence, additional defects often result, leading to more heat and, eventually, to the laser's failure.

A number of diagnostic techniques—including high-resolution optical microscopy and transmission electron microscopy—have revealed in part how this failure takes place in a II-VI semiconductor diode laser. With such information, the leading laboratories are exploring a number of schemes to prevent stacking faults—relying again on the many techniques available for engineering the laser material so that it possesses extra toughness. The challenges with the nitride compounds are different; these materials are very tough because of their chemical bond arrangements. Thus, strong electric current is less likely to cause stacking faults. But at present, the defects in nitride compounds are so numerous that they inhibit lasing. Still, solutions to these problems cannot be far off. Given their promise in high-density optical storage applications, as well as a host of other uses from projection television to medical diagnostics, wide band-gap semiconductor lasers should have a bright, blue future. SA



BAND-GAP ENERGY of a semiconductor crystal determines the energy of the photons it emits and thus the wavelength of the light those photons produce. A material having a high band-gap energy, such as zinc sulfur selenide, yields light having a short wavelength, somewhere in the blue-green range of the spectrum.

The Authors

ROBERT L. GUNSHOR and ARTO V. NURMIKKO have collaborated for 12 years. Gunshor is the Thomas Duncan Distinguished Professor of Microelectronics at Purdue University. Nurmikko is the L. Herbert Ballou University Professor of Engineering and Physics at Brown University and director of the Center for Advanced Materials Research there.

Further Reading

PHYSICS AND DEVICE SCIENCE IN II-VI SEMICONDUCTOR VISIBLE LIGHT EMITTERS. A. V. Nurmikko and R. L. Gunshor in *Solid State Physics* (Academic Press), Vol. 49, pages 205–282; 1995.
SEMICONDUCTOR LASERS: PAST, PRESENT AND FUTURE. Govind Agrawal. American Institute of Physics Press, 1995.

Sunlight and Skin Cancer

Although most skin cancers appear in older people, the damage often begins decades earlier, when the sun's rays mutate a key gene in a single cell

by David J. Leffell and Douglas E. Brash

In 1775 the British physician Percivall Pott reported a curious prevalence of ragged sores on the scrotums of many chimney sweeps in London. Other doctors might have concluded that the men were afflicted with a venereal disease that was then rampant throughout the city. But Pott was more astute. He realized they were in fact suffering from a type of skin cancer. Pott's discovery was a medical milestone. By observing that men continually exposed to coal tar were "peculiarly liable" to this form of cancer, he documented for the first time that cancer could be caused by an external agent rather than by internal factors.

ANTI-CANCER COUNCIL OF VICTORIA

YOUNG BATHERS, such as these Australian children, may predispose themselves to skin cancer as they play. Yet only one youngster here is taking precautions.

More recently, investigators have identified another link between the environment and skin cancer, but this time the agent is much more ubiquitous. It is nothing less than light from the sun. The painstaking efforts of dozens of researchers have revealed a great deal about how solar rays contribute to the development of an astonishingly high number of skin cancers every year.

In the U.S. alone, about a million new cases occur annually, rivaling the incidence of all other types of cancer combined. Skin cancer typically takes one of three forms corresponding to the three major types of skin cells: basal cells, squamous cells and melanocytes. Cancer of melanocytes, called malignant melanoma, is the most lethal variety—and perhaps the most mysterious to researchers attempting to understand how these tumors are triggered. Fortunately, it is also the least common. In the U.S. there will be about 38,000 new cases of melanoma this year and approximately 7,000 deaths from the disease. The two other forms, together called nonmelanoma skin cancer, account for the balance of the cases but kill a much smaller percentage of the affected population. A few thousand people are expected to die in the U.S. during 1996 from nonmelanoma (almost exclusively squamous cell) skin cancer.

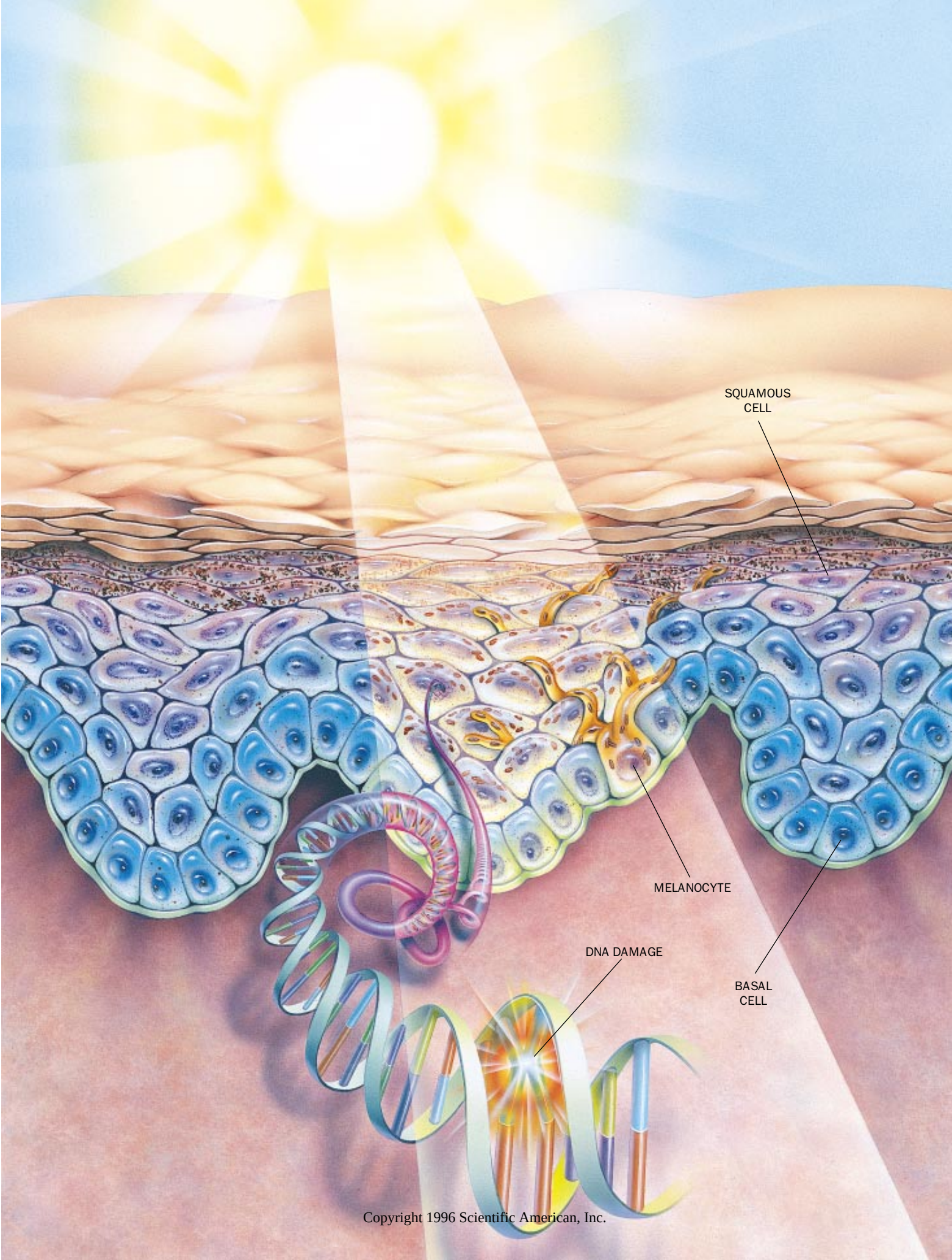
If caught early, most cases of nonmelanoma skin cancer are easily treated in a doctor's office under local anesthesia. Such cancers can be cured by a variety of simple techniques, including scraping, burning, freezing or surgically excising the malignant tissue. Even melanoma, if diagnosed when the tumor is still less than one millimeter thick, can usually be cured by simple excision. But because skin cancer plagues members of all age groups, and because it can become disfiguring and deadly if left untreated, medical researchers have mounted an immense scientific effort over the years to unravel the mechanisms that cause this disease. Curiously, an accident of history contributed much to that quest.

An Accidental Experiment

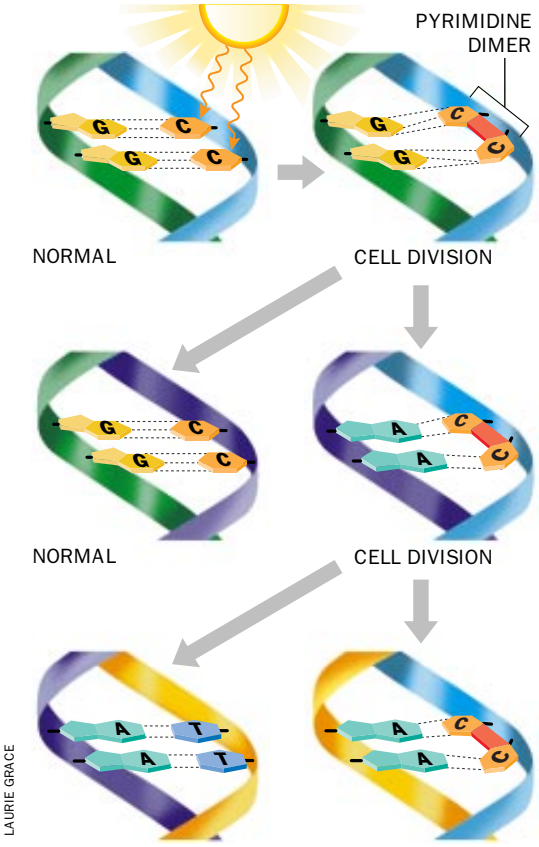
At the time Pott was studying scrotal cancer, Georgian England had a legal system that inflicted severe punishments for petty crimes: forgery or thievery often resulted in a death sentence. But a backlash against the harshness of execution for such misdemeanors soon led to milder sentences—and thus to the overcrowding of jails. To unburden the country's prisons, the House of Commons voted to banish criminals to remote locales beginning in the 1780s. The destination of choice was a little known shore bordering the South Pacific Ocean. Within a few decades, the east coast of Australia was populated with British and Irish men and women. Those early colonists often shared the Celtic features of fair skin and light hair, and today their descendants predominate on that southern continent.

HUMAN SKIN includes three major cell types, all of which are susceptible to sunlight-induced cancer. Near the base of the epidermis lie round, basal cells. Closer to the surface are flattened, squamous cells. Melanocytes (cells that produce the protective pigment melanin) are interspersed in the basal layer and have numerous extensions that reach outward. Solar rays, which can penetrate well below the surface of the skin, damage segments of a cell's DNA that are particularly vulnerable to ultraviolet light. Damage to a gene called *p53* appears crucial to basal cell and squamous cell skin cancers.

AUDRA GERAS



How Sunlight Can Cause a Permanent Mutation



LAURIE GRACE

ULTRAVIOLET LIGHT can break chemical bonds in adjacent pyrimidine bases, often at a point on the DNA strand where two cytosines (*C*) occur. New bonds then form (*red*), linking the disrupted bases together in a so-called pyrimidine dimer.

REPLICATION requires that a cell separate the paired DNA strands (*green* and *blue*), each of which is used as a template to construct a new strand (*purple*)—by mating guanine (*G*) with cytosine and adenine (*A*) with thymine (*T*). The strand unaffected by sunlight produces normal DNA (*left*), but the strand containing the pyrimidine dimer pairs disturbed *C*s with *A*s instead of matching them properly with *G*s.

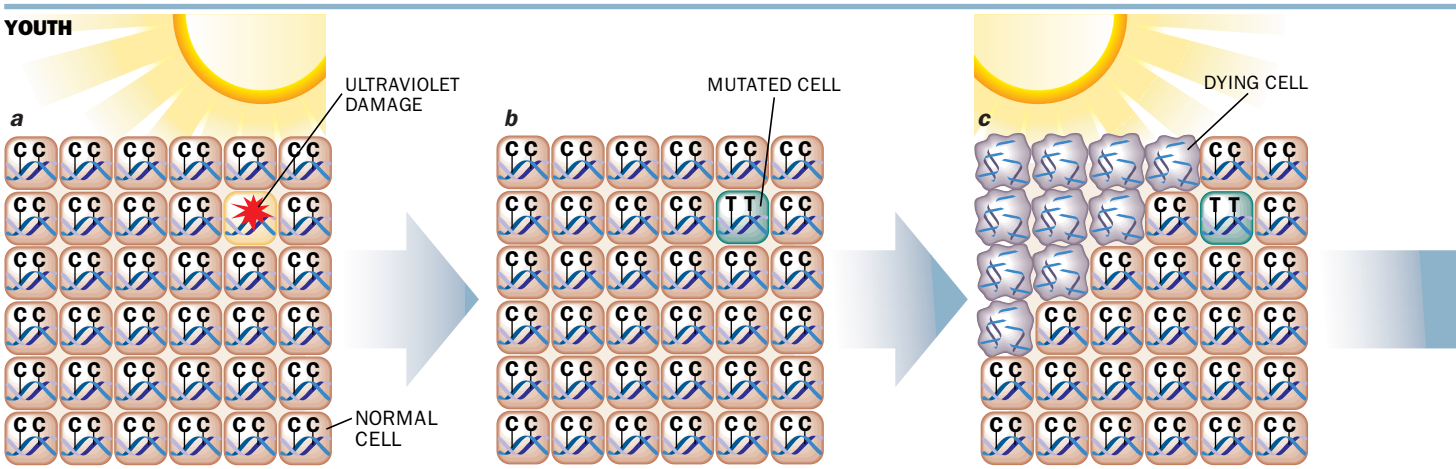
CONTINUED REPLICATION repeats the error, mating the dimer once more with a pair of *A*s (*right*). On the opposite strand (*left*), these *A*s are matched with *T*s, creating a genetic mutation. The dimer may eventually be eliminated by "excision repair," but the *C*-to-*T* mutation is permanent. When such a mutation falls within a cancer-related gene, the cell becomes prone to malignancy.

What began as an 18th-century attempt at penal reform ultimately culminated in a de facto large-scale experiment on the links between complexion, solar radiation and skin cancer. With their fair skin continually exposed to intense sun, whites in Australia now have the highest rate of all kinds of skin cancer of any people in the world. Their British relatives, who live under cloudy northern skies, are more fortunate. They have a relatively low risk of acquiring these malignancies—as do Australian Aborigines, who with much darker skin are rarely affected by sun-induced cancers of the skin.

Investigators recognized as early as 50 years ago that the Australian experience implicated strong sun and fair skin as important risk factors for skin cancer. But for decades scientists were unable to explain what the sun was actually doing to skin cells to make them become cancerous. Clarifying that mystery required more than an accidental experiment on a sun-drenched continent. It took years of study in research laboratories of molecular biologists around the world before the details of that process began to be uncovered.

When the two of us started to attack this problem in the late 1980s, two types of insults from the sun seemed equally suspect. In one category were mutations of specific genes within skin cells. A cell may reproduce excessively if a mutation either turns a normal gene into an overzealous growth promoter (an oncogene) or inactivates a gene that normally limits cell growth (a tumor suppressor gene). The other class of causes we considered at the outset included more widespread events—ones that would affect every sun-exposed cell. For example, the sun's radiation might suppress the skin's immune response (reducing its natural ability to eliminate tumor cells) or directly stimulate cell division. With such diverse explanations possible, we knew that isolating the causes of skin cancer would not be easy.

But we were guided by the knowledge that the damaging effects of sunlight can occur many years before tumors appear. Such delayed effects were most clearly demonstrated in studies undertaken by Anne Kricker, then at the University of Western Australia, Robin Marks of the



GROWTH OF A NONMELANOMA SKIN TUMOR is thought to involve sunlight altering the *p53* gene in a squamous or basal cell of the skin (*a*). The mutation that results (*b*) destroys the ability of genetically injured cells to delay replication until they have repaired their DNA. The *p53* mutation also prevents such cells from killing themselves when damaged beyond

Anti-Cancer Council of Victoria and their colleagues. They noted that people who had emigrated from cloudy England to sunny Australia before the age of 18 acquired the higher Australian incidence of skin cancers, but if they moved when they were older, they retained the native risk.

These findings indicated that Australian skin cancer patients must have received a critically high dose of sunlight years before the appearance of tumors (which rarely occurred before middle age). Widespread events, such as immunosuppression, last for only a few days after the injurious radiation ceases. But genetic changes persist (being passed from one generation of cells to another). Looking for genetic changes therefore seemed a more promising avenue for our research. So we began a hunt for sunlight-induced mutations that could occur early in life and set the stage for the development of skin cancer much later on.

A Signature Mutation

That search was daunting. The DNA in a human cell contains as many as 100,000 genes, and each gene typically includes thousands of nucleotides (the building blocks of DNA)—only some of which would be likely to bear traces of sun-induced damage. And even if we managed to identify mutations in skin cancer samples, how could we be sure that sunlight had caused them? Fortunately, other investigators had given us a useful clue by finding that ultraviolet B radiation—long suspected to be the carcinogenic factor in sunlight—had a characteristic signature.

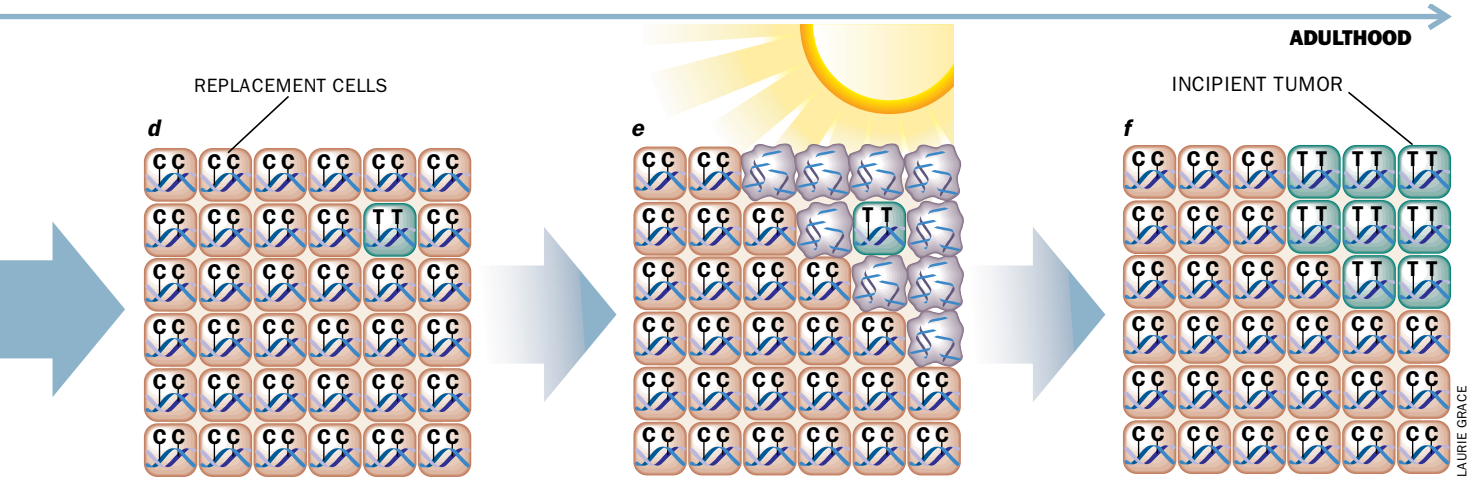
After studying everything from viruses to human cells, groups of researchers from Switzerland, France, Canada and the U.S. had shown that ultraviolet light causes mutations at points on a DNA strand containing specific nucleotide bases. Bases are the variable parts of nucleotides and go by the names adenine (*A*), guanine (*G*), cytosine (*C*) and thymine (*T*). Ultraviolet light creates mutations where a so-called pyrimidine base—cytosine or thymine—lies adjacent to another pyrimidine. About two thirds of these mutations are *C*-to-*T* substitutions, and about 10 percent of these changes occur at two adjacent *C*s, with both bases changing to *T*s. These features of the mutations created by ultraviolet light constitute a fingerprint of sorts, because they are made by no other agents.

We thus had a good idea of the kinds of distinctive mutations that should result from exposure to sunlight. But we needed to pinpoint which of the vast number of human genes mutated to produce a carcinogenic effect. Our best guess was that the solution lay with the handful of human genes already known to be involved in cancer.

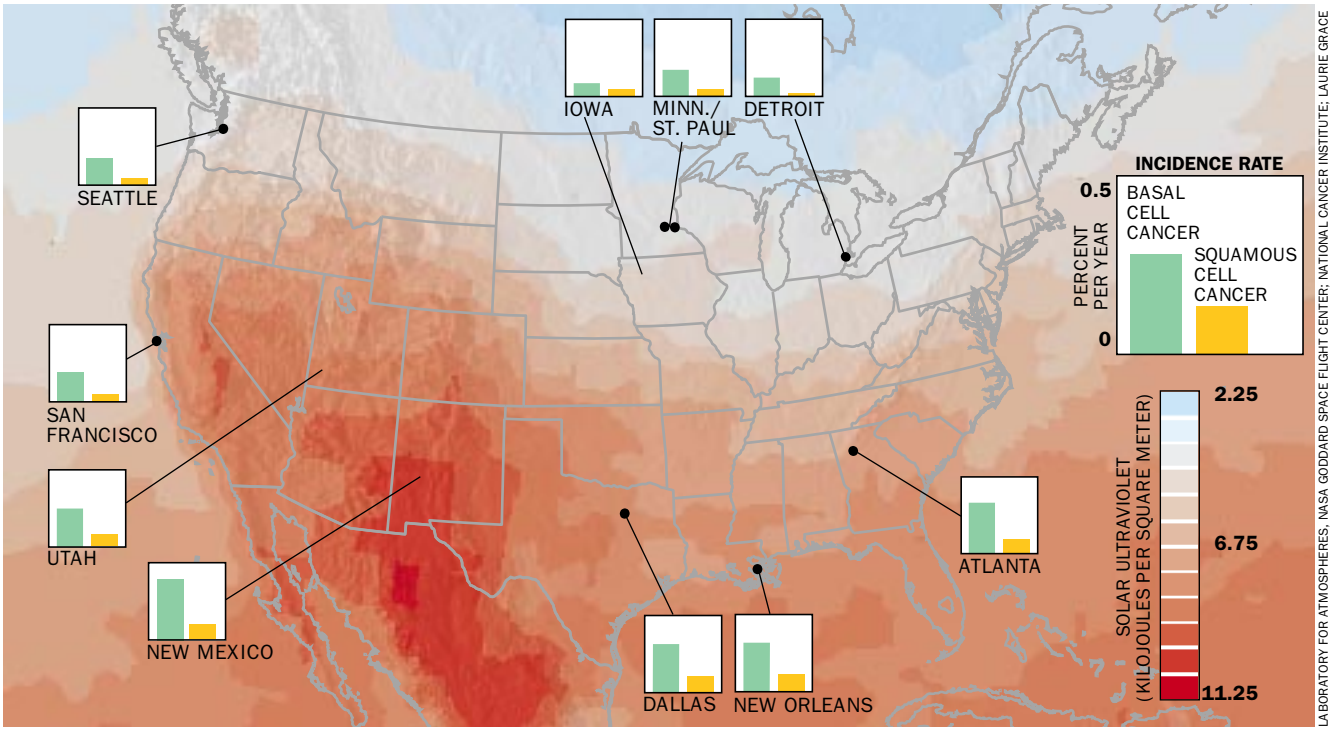
Of the recognized oncogenes and tumor suppressor genes, we chose to examine a tumor suppressor gene called *p53*, which is now known to be mutated in more than half of all people's cancers. At the time, we suspected that *p53* might be involved in many cases of skin cancer because of an intriguing connection between nonmelanoma skin cancer and a rare affliction (epidermodysplasia verruciformis) that causes wartlike growths to appear on the skin.

Previous research had revealed that such growths contain DNA from the human papillomavirus and that when these growths are located on sun-exposed skin, they can progress to basal cell or squamous cell cancer. Peter M. Howley and his colleagues at the National Cancer Institute had further shown that one of the proteins made by the papillomavirus inactivates the p53 protein. (Genes give rise to proteins, and the p53 protein, as might be expected, is the product of the *p53* gene.) So all indications were that *p53* might play a special role in nonmelanoma skin cancer. But we needed solid confirmation.

To find that proof, we studied squamous cell carcinomas, tumors unquestionably linked to sunlight (they occur on the face and hands, especially among whites living in the tropics). In collaboration with Jan Pontén of Uppsala University Hospital in Sweden, we discovered that more than 90 percent of the squamous cell carcinomas from a set of samples collected in the U.S. had a mutation somewhere in the *p53* tumor suppressor gene. These mutations occurred at sites with adjacent pyrimidine bases, and they had the distinctive *C*-to-*T* pattern associated with ultraviolet exposure. Our research group, along with several others, later pinpointed sunlight-related *p53* mutations in basal cell carcinomas as well. (Melanoma does not appear to be associated with alterations to *p53*. Researchers are still studying cancerous melanocytes for genes affected by sunlight.) After examining samples in our laboratory, Annemarie Ziegler found that precancerous skin also contains mutations of *p53*, indicating that the genetic changes occur long before



repair. If sunlight later burns unaltered cells (*c*), massively damaged cells will commit "cellular suicide" and be replaced by cells derived from healthy skin nearby (*d*). But if sunlight burns tissue near a *p53*-mutated cell that cannot self-destruct (*e*), the mutated cell may replace the dying, sunburned cells with its own progeny (*f*), thereby promoting growth of a tumor.



SKIN CANCER RISK depends greatly on exposure to solar ultraviolet radiation. Satellite measurements of ozone and cloud cover allow atmospheric scientists to estimate the amount of DNA-damaging ultraviolet light that reaches Earth's surface (shown here as an average for July 1992). Light-skinned people living in parts of the U.S. exposed to intense ultraviolet rays during the summer months are most prone to skin cancer, because they produce less of the melanin pigment that protects dark skin from ultraviolet damage. The bar graphs show incidence rates for nonmelanoma skin cancer in whites.

tumors appear. But were these mutations truly the cause of nonmelanoma skin cancer, or were they simply an irrelevant indicator of lifetime exposure to sunlight?

We could rule out this last possibility by the particular way the genetic code had been altered. The nucleotides in genes are arranged in well-defined codons—groups of three bases that specify different amino acids. The sequence of codons in a gene determines the sequence of amino acids that are strung together to construct a protein. But different codons can sometimes specify the same amino acid—as if the name of the amino acid could be spelled any of several ways. Typically the amino acid does not change when the first two bases of the codon are constant and only the third varies. Hence, if the *p53* mutations found in skin cancer were just a random effect of exposure to the sun, we would expect to find changes in the third position occurring as often as in the first or second. That is, there would be plenty of examples where the codon mutated (underwent a nucleotide base substitution) without altering its corresponding amino acid. Yet studies of this gene in skin cancers from around the globe had consistently revealed mutations that modified one or more amino acids in the p53 protein. These genetic changes to *p53*, then, were not just a side effect of ultraviolet exposure. They were in fact causing the skin cancers.

To better understand how the *p53* gene was affected in nonmelanoma skin cancer, we investigated whether certain segments of the *p53* gene were particularly prone to the mutation by sunlight of adjacent pyrimidine bases (that is, *C*s or *T*s). Biologists have found so-called mutation hot spots (places on a DNA strand where mutations tend to occur) whenever they expose living cells to carcinogens. After analyzing many tumors, we determined that the *p53* gene in nonmelanoma skin cancer contains about nine hot spots. In cancers unrelated to sunlight (such as colon or bladder cancer), five codons of *p53* are most often mutated, three of which are among the hot spots in skin cancers. At the two hot spots found only in the other cancers, the mutating *C* is flanked on either side by a *G* or *A* but never by a *T* or another *C*. Lacking a pair of pyrimidine bases, equivalent sites on the DNA of skin cells are protected from mutation by ultraviolet light.

Of the hundreds of places on the *p53* gene with adjacent pyrimidines, why do only a few sites act as hot spots when cells are exposed to sunlight? Several researchers have recently helped answer that question by building on a discovery made more than three decades ago at Oak Ridge National Laboratory by Richard B. Setlow and William L. Carrier. Setlow and Carrier determined that cells can reverse ultraviolet damage to their DNA by an enzymatic process called excision repair. Cells essentially snip out disrupted bases and replace them with intact ones. Working in our lab in 1992, Subrahmanyam Kunala showed that cells repair damage particularly slowly at some pyrimidine pairs. Subsequently, Gerd P. Pfeifer and his colleagues at the City of Hope Beckman Research Institute in Duarte, Calif., found that cells repair the *p53* sites mutated in nonmelanoma skin cancer more sluggishly than they do many other sites in the gene. Hence, it seems quite likely that the hot spots we found for skin cancer owe their existence to an inability of skin cells to mend these sites efficiently.

Cellular Proofreading

Even after we had identified the relevant *p53* mutations, the story of carcinogenesis remained woefully incomplete. After all, genes do not get cancer—cells do. It was clear enough that the p53 protein must operate in normal

skin cells to prevent cancer, but how? One hint was available from Michael B. Kastan of Johns Hopkins Hospital. He found that cells subjected to x-rays stepped up production of the p53 protein, which in turn prevented the cells from dividing. Peter A. Hall and David P. Lane of the University of Dundee and Jonathan L. Rees of the University of Newcastle have shown a similar effect on the p53 protein in skin cells exposed to ultraviolet radiation. Cancer researchers speculate that the p53 protein normally stops a DNA-damaged cell from reproducing until it has had time to make repairs.

Moshe Oren and his colleagues at the Weizmann Institute of Science in Israel have proposed another function for the p53 protein as well: it can prevent cancer in situations where the DNA damage is too extensive to be repaired. They find that elevated levels of the p53 protein in a cell lead to apoptosis—programmed cell death. (Such cell death is a normal part of many biological processes, including embryonic development.) In this case, "suicide" of a sun-damaged cell would prevent it from becoming cancerous by permanently erasing its genetic mistakes. Such apoptosis could be called cellular proofreading. Because the skin sheds cells routinely, we surmised that skin cells often used *p53* in this way. But even before we began to test our idea, some evidence was already available to support it.

Dermatologists have recognized for a long time that when skin is sunburned, some cells come to resemble apoptotic cells. By 1994 we could show that sunburned cells contained breaks in their DNA similar to those in other apoptotic cells. The sunburned cells thus appeared to be in the process of committing cellular suicide, and we began immediately to wonder whether cells that had lost *p53* could undergo such self-inflicted death.

At about the time we arrived at this investigative juncture, Tyler Jacks and his colleagues at the Massachusetts Institute of Technology had developed mice lacking the *p53* gene. When Alan S. Jonason and Jeffrey A. Simon irradiated the skin of these so-called *p53* knockout mice in our laboratory, they found far fewer sunburned, apoptotic cells than in normal mice exposed to the same ultraviolet radiation. Mice in which the *p53* gene had been only partially inactivated had only a moderate tendency to undergo light-induced cell suicide. These results suggested that programmed cell death was important for preventing nonmelanoma skin cancer and that loss of *p53* could block this process.

Double Punch from Sunlight

It is now possible to envision how the failure of cellular proofreading would lead to skin cancer. Normal skin exposed to sunlight will accumulate DNA damage caused by the ultraviolet B part of the solar spectrum. Cells unable to repair their DNA in a timely fashion die through apoptosis. But if the *p53* gene in a cell has mutated during a previous episode of exposure to sunlight, that cell will resist such self-destruction—even if it has been badly injured.

The situation is actually much worse. A cell on the verge of becoming cancerous is surrounded by normal cells that undergo apoptosis when damaged. The dying cells thus must be leaving some space into which the *p53*-mutated cell can grow. By inducing healthy cells to kill themselves off, sunlight favors the proliferation of *p53*-mutated cells. In effect, sunlight acts twice to cause cancer: once to mutate the *p53* gene and then afterward to set up conditions for the unrestrained growth of the altered cell line. These two actions, mutation and tumor promotion, are the one-two blows of carcinogenesis. Although mutation and promotion are carried out by separate agents in other tumors, in skin cancer ultraviolet radiation appears to throw both punches.

There are undoubtedly other genes involved in the development of skin cancer as well as other effects of sunlight that researchers do not yet fully understand. For example, medical researchers know that Gorlin syndrome (a disease in which patients have multiple basal cell cancers) is caused by an inherited mutation in a different tumor suppressor gene. With further investigation, the various mechanisms of carcinogenesis will become even more clear, and scientists may find clever ways to interrupt the progression of normal skin cells to cancerous ones.

It is not beyond reason to hope that the detailed understanding researchers are gaining of nonmelanoma skin cancer will yield new kinds of therapies. Perhaps drugs that restore normal function to a mutated p53 protein will allow doctors to offer their patients an effective remedy that does not involve surgery. Such a cure, perhaps administered as a simple skin cream that is absorbed by the affected cells, might be available within the next decade or two. If so, it will be of great benefit to countless aging members of the sun-loving baby-boom generation—a group to which we both admittedly belong. SA

The Authors

DAVID J. LEFFELL and DOUGLAS E. BRASH have worked together for nearly a decade to understand the role of the sun in causing skin cancer. Leffell, a professor of dermatology and surgery at the Yale School of Medicine, has brought to their research collaboration the experience gained in clinical practice. He earned his M.D. at McGill University in 1981 and trained at Cornell Medical School, Memorial Sloan-Kettering Cancer Center and the University of Michigan before taking a position on the faculty at Yale in 1988. Brash, too, is on the medical school faculty at Yale, and his credentials include a bachelor's degree in engineering physics from the University of Illinois. He shifted from engineering to the study of biophysics at Ohio State University, where he received his Ph.D. in 1979. Thereafter Brash pursued postdoctoral training in microbiology (at the Harvard School of Public Health) and pathology (at Harvard Medical School) until 1984. He spent the next five years at the National Cancer Institute before moving to Yale.

Further Reading

A Role for Sunlight in Skin Cancer: UV-Induced p53 Mutations in Squamous Cell Carcinoma. D. E. Brash, J. A. Rudolph, J. A. Simon, A. Lin, G. J. McKenna, H. P. Baden, A. J. Halper and J. Pontén in *Proceedings of the National Academy of Sciences U.S.A.,* Vol. 88, No. 22, pages 10124–10128; November 15, 1991.

Sunburn and p53 in the Onset of Skin Cancer. A. Ziegler, A. S. Jonason, D. J. Leffell, J. A. Simon, H. W. Sharma, J. Kimmelman, L. Remington, T. Jacks and D. E. Brash in *Nature,* Vol. 372, pages 773–776; December 22–29, 1994.

Cancer Free: The Comprehensive Cancer Prevention Program. Sidney J. Winawer and Moshe Shike. Simon & Schuster, 1996.

Sunlight, Ultraviolet Radiation and the Skin. NIH Consensus Statement. Vol. 7, No. 8, pages 1–29; May 8–10, 1989. Available at http://text.nlm.nih.gov/nih/cdc/www/74txt.html

The Nature of Space and Time

Two relativists present their distinctive views on the universe, its evolution and the impact of quantum theory

by Stephen W. Hawking and Roger Penrose

In 1994 Stephen W. Hawking and Roger Penrose gave a series of public lectures on general relativity at the Isaac Newton Institute for Mathematical Sciences at the University of Cambridge. From these lectures, published this year by Princeton University Press as *The Nature of Space and Time,* SCIENTIFIC AMERICAN has culled excerpts that serve to compare and contrast the perspectives of the two scientists. Although they share a common heritage in physics—Penrose served on Hawking's Ph.D. thesis committee at Cambridge—the lecturers differ in their vision of quantum mechanics and its impact on the evolution of the universe. In particular, Hawking and Penrose disagree on what happens to the information stored in a black hole and on why the beginning of the universe differs from the end.

One of Hawking's major discoveries, made in 1973, was that quantum effects will cause black holes to emit particles. The black hole will evaporate in the process, so that ultimately perhaps nothing of the original mass will be left. But during their formation, black holes swallow a lot of data—the types, properties and configurations of the particles that fall in. Although quantum theory requires that such information must be conserved, what finally happens to it remains a topic of contentious debate. Hawking and Penrose both believe that when a black hole radiates, it loses the information it held. But Hawking insists that the loss is irretrievable, whereas Penrose argues that the loss is balanced by spontaneous measurements of quantum states that introduce information back into the system.

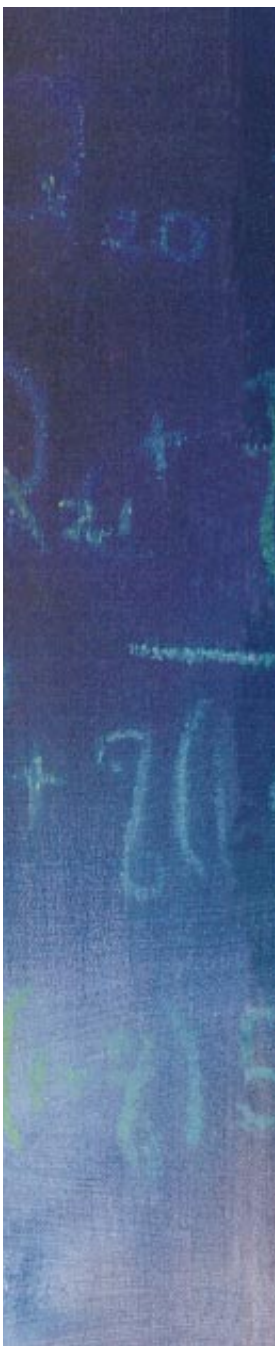
Both scientists agree that a future quantum theory of gravity is needed to describe nature. But they differ in their view of some aspects of this theory. Penrose thinks that even though the fundamental forces of particle physics are symmetric in time—unchanged if time is reversed—quantum gravity will violate time symmetry. The time asymmetry will then explain why in the beginning the universe was so uniform, as evinced by the microwave background radiation left over from the big bang, whereas the end of the universe must be messy.

Penrose attempts to encapsulate this time asymmetry in his Weyl curvature hypothesis. Space-time, as Albert Einstein discovered, is curved by the presence of matter. But space-time can also have some intrinsic bending, a quantity designated by the Weyl curvature. Gravitational waves and black holes, for example, allow space-time to curve even in regions that are empty. In the early universe the Weyl curvature was probably zero, but in a dying universe the large number of black holes, Penrose argues, will give rise to a high Weyl curvature. This property will distinguish the end of the universe from the beginning.

Hawking agrees that the big bang and the final "big crunch" will be different, but he does not subscribe to a time asymmetry in the laws of nature. The underlying reason for the difference, he thinks, is the way in which the universe's evolution is programmed. He postulates a kind of democracy, stating that no point in the universe can be special; therefore, the universe cannot have a boundary. This no-boundary proposal, Hawking claims, explains the uniformity in the microwave background radiation.

The physicists diverge, ultimately, in their interpretation of quantum mechanics. Hawking believes that all a theory has to do is provide predictions that agree with data. Penrose thinks that simply comparing predictions with experiments is not enough to explain reality. He points out that quantum theory requires wave functions to be "superposed," a concept that can lead to absurdities. The scientists thus pick up the threads of the famous debates between Einstein and Niels Bohr on the bizarre implications of quantum theory.

—The Editors



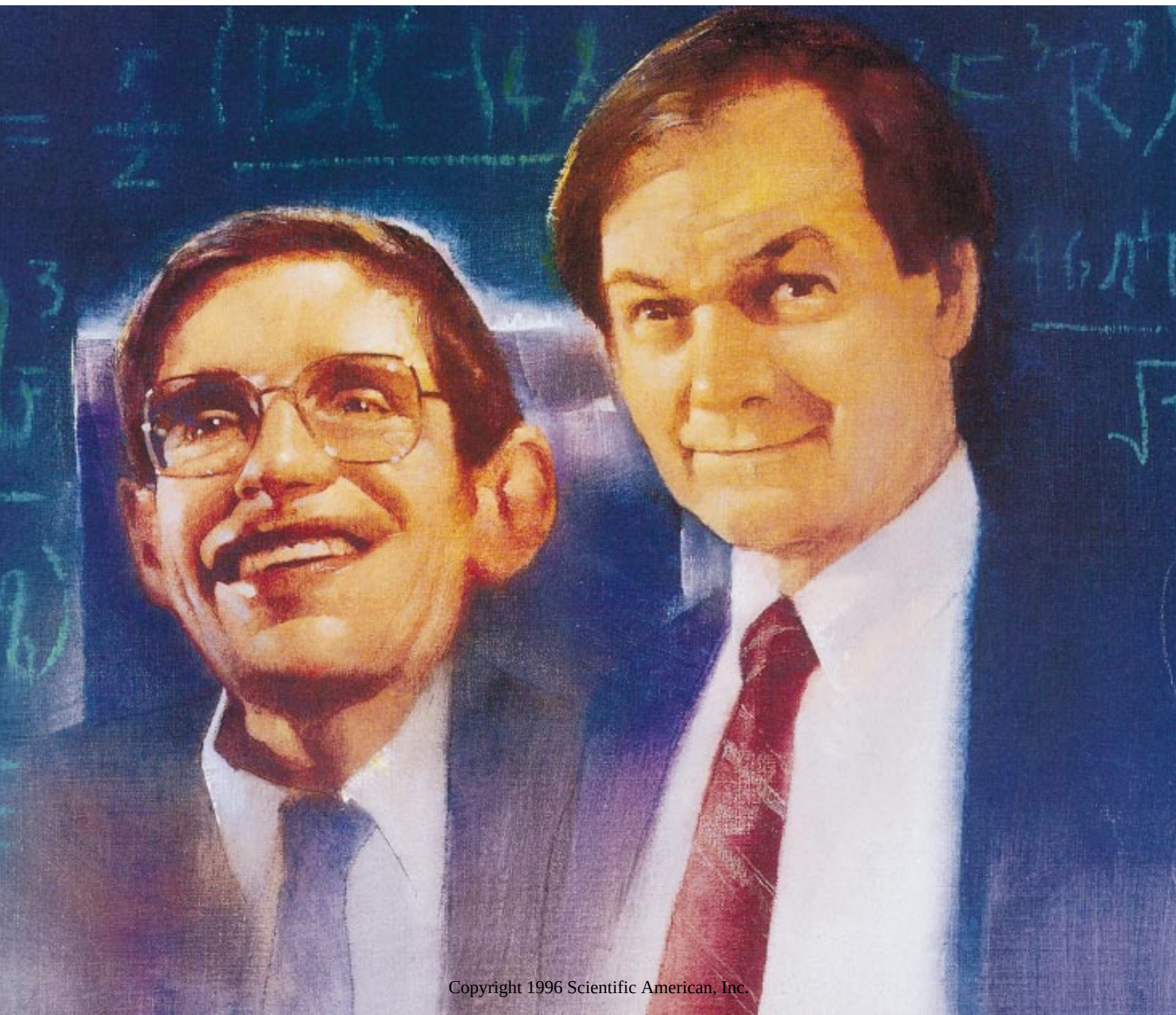
PAINTING BY ROBERT HUNT

Stephen Hawking on quantum black holes:

The quantum theory of black holes…seems to lead to a new level of unpredictability in physics over and above the usual uncertainty associated with quantum mechanics. This is because black holes appear to have intrinsic entropy and to lose information from our region of the universe. I should say that these claims are controversial: many people working on quantum gravity, including almost all those who entered it from particle physics, would instinctively reject the idea that information about the quantum state of a system could be lost. However, they have had very little success in showing how information can get out of a black hole. Eventually I believe they will be forced to accept my suggestion that it is lost, just as they were forced to agree that black holes radiate, which went against all their preconceptions…

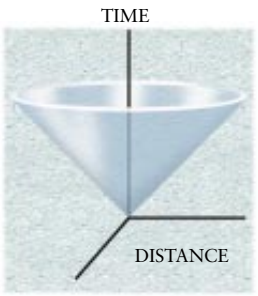
The fact that gravity is attractive means that it will tend to draw the matter in the universe together to form objects like stars and galaxies. These can support themselves for a time against further contraction by thermal pressure, in the case of stars, or by rotation and internal motions, in the case of galaxies. However, eventually the heat or the angular momentum will be carried away and the object will begin to shrink. If the mass is less than about one and a half times that of the Sun, the contraction can be stopped by the ***degeneracy pressure*** of electrons or neutrons. The object will settle down to be a white dwarf or a neutron star, respectively. However, if the mass is greater than this limit there is nothing that can hold it up

DEGENERACY PRESSURE
No two electrons or neutrons can occupy the same quantum state. Thus, when any collection of these particles is squeezed into a small volume, those in the highest quantum states become very energetic. The system then resists further compression, exerting an outward push called degeneracy pressure.



LIGHT CONES

To depict space-time, physicists routinely plot time on a vertical axis and space on a horizontal. In this scheme, light rays emanating from any point in space fan out along the surface of a vertical cone. Because no physical signal can cover more distance in a given time than light can, any signals originating at that point are confined within the volume of the light cone.



NULL SURFACE

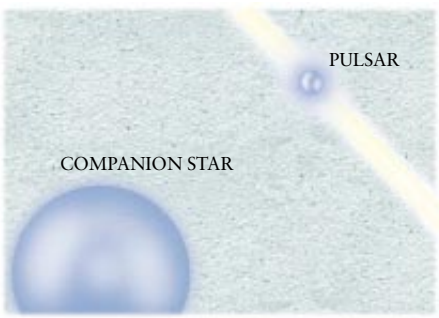
A surface in space along which light travels is known as a null surface. The null surface surrounding a black hole, called an event horizon, has the shape of a spherical shell. Nothing that falls inside the event horizon can come back out.

MULTIPOLE MOMENTS

The dynamics of an object can be summarized by determining its multipole moments. Each moment is calculated by dividing an object into tiny elements, multiplying the mass of each element by its distance from the center zero, one or more times, then adding these terms for all the elements. A sphere, for example, has a monopole moment, whereas a dumbbell has a dipole moment, which allows it to acquire angular momentum easily.

PULSARS

Some dying suns collapse into neutron stars, massive objects made entirely of densely packed neutrons. Rapidly rotating neutron stars become pulsars, so called because they emit pulses of electromagnetic radiation at astonishingly regular millisecond-to-second intervals. A pulsar sometimes orbits another neutron star, forming a binary pair.



ILLUSTRATIONS BY LAURIE GRACE

and stop it continuing to contract. Once it has shrunk to a certain critical size the gravitational field at its surface will be so strong that the ***light cones*** will be bent inward.... You can see that even the outgoing light rays are bent toward each other and so are converging rather than diverging. This means that there is a closed trapped surface....

Thus there must be a region of space-time from which it is not possible to escape to infinity. This region is said to be a black hole. Its boundary is called the event horizon and is a ***null surface*** formed by the light rays that just fail to get away to infinity....

[A] large amount of information is lost when a body collapses to form a black hole. The collapsing body is described by a very large number of parameters. There are the types of matter and the ***multipole moments*** of the mass distribution. Yet the black hole that forms is completely independent of the type of matter and rapidly loses all the multipole moments except the first two: the monopole moment, which is the mass, and the dipole moment, which is the angular momentum.

This loss of information didn't really matter in the classical theory. One could say that all the information about the collapsing body was still inside the black hole. It would be very difficult for an observer outside the black hole to determine what the collapsing body was like. However, in the classical theory it was still possible in principle. The observer would never actually lose sight of the collapsing body. Instead it would appear to slow down and get very dim as it approached the event horizon. But the observer could still see what it was made of and how the mass was distributed.

However, quantum theory changed all this. First, the collapsing body would send out only a limited number of photons before it crossed the event horizon. They would be quite insufficient to carry all the information about the collapsing body. This means that in quantum theory there's no way an outside observer can measure the state of the collapsed body. One might not think that this mattered too much, because the information would still be inside the black hole even if one couldn't measure it from the outside. But this is where the second effect of quantum theory on black holes comes in....

[Quantum] theory will cause black holes to radiate and lose mass. It seems that they will eventually disappear completely, taking with them the information inside them. I will give arguments that this information really is lost and doesn't come back in some form. As I will show, this loss of information would introduce a new level of uncertainty into physics over and above the usual uncertainty associated with quantum theory. Unfortunately, unlike Heisenberg's uncertainty principle, this extra level will be rather difficult to confirm experimentally in the case of black holes.

Roger Penrose on quantum theory and space-time:

The great physical theories of the 20th century have been quantum theory, special relativity, general relativity and quantum field theory. These theories are not independent of each other: general relativity was built on special relativity, and quantum field theory has special relativity and quantum theory as inputs.

It has been said that quantum field theory is the most accurate physical theory ever, being accurate to about one part in about 10^{11}. However, I would like to point out that general relativity has, in a certain clear sense, now been tested to be correct to one part in 10^{14} (and this accuracy has apparently been limited merely by the accuracy of clocks on Earth). I am speaking of the Hulse-Taylor binary ***pulsar*** PSR 1913 + 16, a pair of neutron stars orbiting each other, one of which is a pulsar. General relativity predicts that this orbit will slowly decay (and the period

shorten) because energy is lost through the emission of gravitational waves. This has indeed been observed, and the entire description of the motion…agrees with general relativity (which I am taking to include Newtonian theory) to the remarkable accuracy, noted above, over an accumulated period of 20 years. The discoverers of this system have now rightly been awarded Nobel Prizes for their work. The quantum theorists have always claimed that because of the accuracy of their theory, it should be general relativity that is changed to fit their mold, but I think now that it is quantum field theory that has some catching up to do.

Although these four theories have been remarkably successful, they are not without their problems.... General relativity predicts the existence of space-time ***singularities.*** In quantum theory there is the "measurement problem"—I shall describe this later. It may be taken that the solution to the various problems of these theories lies in the fact that they are incomplete on their own. For example, it is anticipated by many that quantum field theory might "smear" out the singularities of general relativity in some way....

I should now like to talk about information loss in black holes, which I claim is relevant to this last issue. I agree with nearly all that Stephen had to say on this. But while Stephen regards the information loss due to black holes as an extra uncertainty in physics, above and beyond the uncertainty from quantum theory, I regard it as a "complementary" uncertainty.... It is possible that a little bit of information escapes at the moment of the black hole evaporation…but this tiny information gain will be much smaller than the information loss in the collapse (in what I regard as any reasonable picture of the hole's final disappearance).

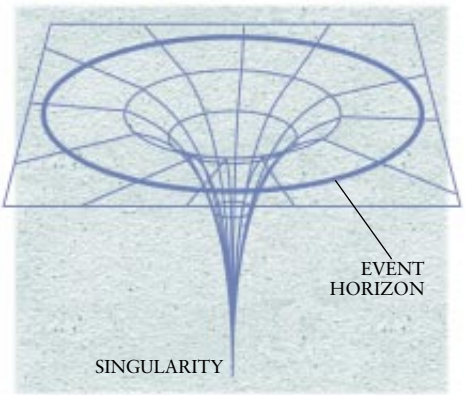
If we enclose the system in a vast box, as a thought experiment, we can consider the phase-space evolution of matter inside the box. In the region of ***phase space*** corresponding to situations in which a black hole is present, trajectories of physical evolution will converge and volumes following these trajectories will shrink. This is due to the information lost into the singularity in the black hole. This shrinking is in direct contradiction to the theorem in classical mechanics, called Liouville's Theorem, which says that volumes in phase space remain constant....Thus a black hole space-time violates this conservation. However, in my picture, this loss of phase-space volume is balanced by a process of "spontaneous" quantum measurement in which information is gained and phase-space volumes increase. This is why I regard the uncertainty due to information loss in black holes as being "complementary" to the uncertainty in quantum theory: one is the other side of the coin to the other....

[Let] us consider the ***Schrödinger's cat*** thought experiment. It describes the plight of a cat in a box, where (let us say) a photon is emitted which encounters a half-silvered mirror, and the transmitted part of the photon's wave function encounters a detector which, if it detects the photon, automatically fires a gun, killing the cat. If it fails to detect the photon, then the cat is alive and well. (I know Stephen does not approve of mistreating cats, even in a thought experiment!) The wave function of the system is a superposition of these two possibilities....But why does our perception not allow us to perceive macroscopic superpositions, of states such as these, and not just the macroscopic alternatives "cat is dead" and "cat is alive"?...

I am suggesting that something goes wrong with superpositions of the alternative space-time geometries that would occur when general relativity begins to become involved. Perhaps a superposition of two different geometries is unstable and decays into one of the two alternatives. For example, the geometries

SINGULARITIES

According to general relativity, under certain extreme conditions some regions of space-time develop infinitely large curvatures, thus becoming singularities where the normal laws of physics break down. Black holes, for example, should contain singularities hidden inside the event horizon.

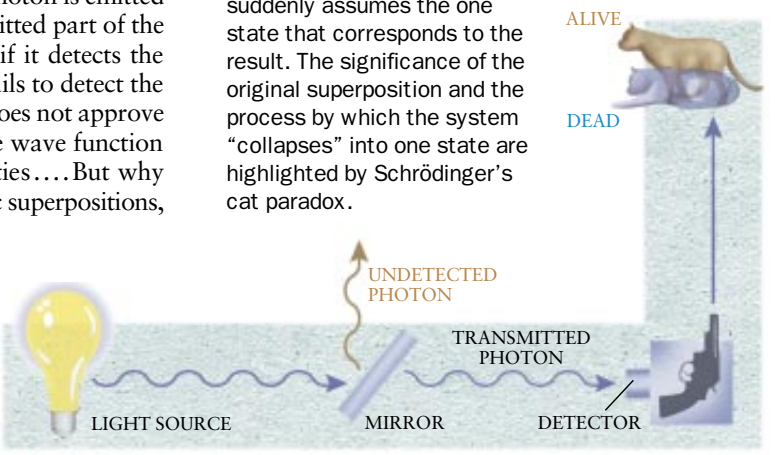


PHASE SPACE

A phase-space diagram is a mathematical volume of many dimensions formed when coordinate axes are assigned to each of the distance and momentum values of each particle. The motion of a group of particles can then be represented by a moving element of volume in phase space.

SCHRÖDINGER'S CAT

Penrose invokes a thought experiment originally invented by Einstein and used by Erwin Schrödinger to study the conceptual knots tied by wave functions. Prior to a measurement, a system is assumed to be in a "superposition" of quantum states or waves, so that the value of, say, the momentum is uncertain. After a measurement, the value of a quantity becomes known, and the system suddenly assumes the one state that corresponds to the result. The significance of the original superposition and the process by which the system "collapses" into one state are highlighted by Schrödinger's cat paradox.



might be the space-times of a live cat, or a dead one. I call this decay into one **or** the other alternative objective reduction, which I like as a name because it has an appropriately nice acronym (OR). How does the Planck length 10^{-33} centimeter relate to this? Nature's criterion for determining when two geometries are significantly different would depend upon the ***Planck scale,*** and this fixes the timescale in which the reduction into different alternatives occurs.

Hawking on quantum cosmology:

I will end this lecture on a topic on which Roger and I have very different views—the arrow of time. There is a very clear distinction between the forward and the backward directions of time in our region of the universe. One only has to watch a film being run backward to see the difference. Instead of cups falling off tables and getting broken, they would mend themselves and jump back on the table. If only real life were like that.

The local laws that physical fields obey are time symmetric, or more precisely, ***CPT (charge-parity-time) invariant.*** Thus, the observed difference between the past and the future must come from the boundary conditions of the universe. Let us take it that the universe is spatially closed and that it expands to a maximum size and collapses again. As Roger has emphasized, the universe will be very different at the two ends of this history. At what we call the beginning of the universe, it seems to have been very smooth and regular. However, when it collapses again, we expect it to be very disordered and irregular. Because there are so many more disordered configurations than ordered ones, this means that the initial conditions would have had to be chosen incredibly precisely.

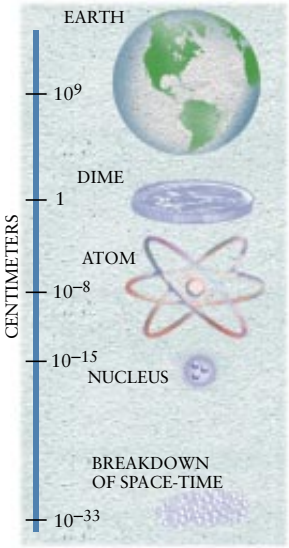
It seems, therefore, that there must be different boundary conditions at the two ends of time. Roger's proposal is that the ***Weyl tensor*** should vanish at one end of time but not the other. The Weyl tensor is that part of the curvature of space-time that is not locally determined by the matter through the Einstein equations. It would have been small in the smooth, ordered early stages but large in the collapsing universe. Thus, this proposal would distinguish the two ends of time and so might explain the arrow of time.

I think Roger's proposal is Weyl in more than one sense of the word. First, it is not CPT invariant. Roger sees this as a virtue, but I feel one should hang on to symmetries unless there are compelling reasons to give them up. Second, if the Weyl tensor had been exactly zero in the early universe, it would have been exactly homogeneous and isotropic and would have remained so for all time. Roger's Weyl hypothesis could not explain the fluctuations in the background nor the perturbations that give rise to galaxies and bodies like ourselves.

Despite all this, I think Roger has put his finger on an important difference between the two ends of time. But the fact that the Weyl tensor was small at one end should not be imposed as an ad hoc boundary condition but should be deduced from a more fundamental principle, the ***no-boundary proposal****.*...

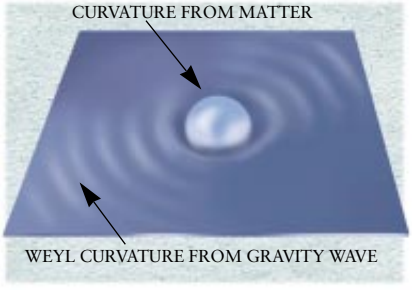
How can the two ends of time be different? Why should perturbations be small at one end but not the other? The reason is there are two possible complex solutions of the field equations.... Obviously, one solution corresponds to one end of time and the other to the other.... At one end, the universe was very smooth and the Weyl tensor was very small. It could not, however, be exactly zero, for that would have been a violation of the uncertainty principle. Instead there would have been small fluctuations that later grew into galaxies and bodies like us. By contrast, the universe would have been very irregular and chaotic at the other end of time with a Weyl tensor that was typically large. This would explain the observed arrow of time and why cups fall off tables and break rather than mend themselves and jump back on.

PLANCK SCALE
The Planck scale is an unattainably small distance—related, by quantum mechanics, to an impossibly small time span and high energy—that emerges when the fundamental constants for gravitational attraction, the velocity of light and quantum mechanics are appropriately combined. The scale represents the distance or energy at which current concepts of space, time and matter break down, and a future theory, quantum gravity, presumably takes over.



CPT (CHARGE-PARITY-TIME) INVARIANCE
This powerful principle requires that theories describing particles must remain true even when the charge, parity (or handedness) and time simultaneously reverse. In other words, the behavior of a negatively charged electron with clockwise spin moving forward in time must be identical to that of a positively charged positron with anticlockwise spin moving backward in time.

WEYL TENSOR
The curvature of space-time has two components. One derives from the presence of matter in space-time; the other, recognized by the German mathematician Hermann Weyl, occurs even in the absence of matter. The mathematical quantity that describes this curvature is called the Weyl tensor.



NO-BOUNDARY PROPOSAL
Hawking suggests that the evolution of the universe is explained by the no-boundary proposal, put forth in 1983 by him and James B. Hartle of the University of California at Santa Barbara. The idea that the universe has no boundary places constraints on how the equations of cosmology are solved. Hawking believes these conditions will lead to the ends of the universe being different, thereby determining the direction of time's arrow.

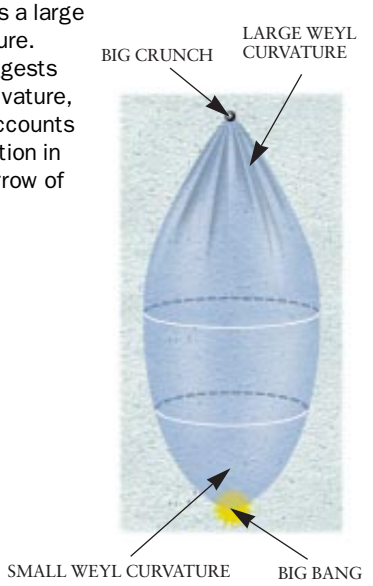
Penrose on quantum cosmology:

From what I understand of Stephen's position, I don't think that our disagreement is very great on this point [the ***Weyl curvature hypothesis***]. For an initial singularity the Weyl curvature is approximately zero.... Stephen argued that there must be small quantum fluctuations in the initial state and thus pointed out that the hypothesis that the initial Weyl curvature is zero at the initial singularity is classical, and there is certainly some flexibility as to the precise statement of the hypothesis. Small perturbations are acceptable from my point of view, certainly in the quantum regime. We just need something to constrain it very near to zero....

Maybe the no-boundary proposal of [James B.] Hartle and Hawking is a good candidate for the structure of the *initial* state. However, it seems to me that we need something very different to cope with the *final* state. In particular, a theory that explains the structure of singularities would have to violate [CPT and other symmetries] in order that something of the nature of the Weyl curvature hypothesis can arise. This failure of time-symmetry might be quite subtle; it would have to be implicit in the rules of that theory which goes beyond quantum mechanics.

WEYL CURVATURE HYPOTHESIS

The universe just after the big bang has a small Weyl curvature, whereas near the end of time it has a large Weyl curvature. Penrose suggests that this curvature, therefore, accounts for the direction in which the arrow of time points.



Hawking on physics and reality:

These lectures have shown very clearly the difference between Roger and me. He's a Platonist and I'm a positivist. He's worried that Schrödinger's cat is in a quantum state, where it is half alive and half dead. He feels that can't correspond to reality. But that doesn't bother me. I don't demand that a theory correspond to reality because I don't know what it is. Reality is not a quality you can test with litmus paper. All I'm concerned with is that the theory should predict the results of measurements. Quantum theory does this very successfully....

Roger feels that...the collapse of the wave function introduces CPT violation into physics. He sees such violations at work in at least two situations: cosmology and black holes. I agree that we may introduce time asymmetry in the way we ask questions about observations. But I totally reject the idea that there is some physical process that corresponds to the reduction of the wave function or that this has anything to do with quantum gravity or consciousness. That sounds like magic to me, not science.

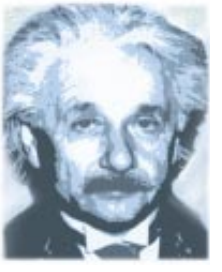
Penrose on physics and reality:

Quantum mechanics has only been around for 75 years. This is not very long if one compares it, for example, with Newton's theory of gravity. Therefore it wouldn't surprise me if quantum mechanics will have to be modified for very macroscopic objects.

At the beginning of this debate, Stephen said that he thinks that he is a positivist, whereas I am a Platonist. I am happy with him being a positivist, but I think that the crucial point here is, rather, that I am a realist. Also, if one compares this debate with the famous debate of ***Bohr and Einstein,*** some 70 years ago, I should think that Stephen plays the role of Bohr, whereas I play Einstein's role! For Einstein argued that there should exist something like a real world, not necessarily represented by a wave function, whereas Bohr stressed that the wave function doesn't describe a "real" microworld but only "knowledge" useful for making predictions.

Bohr was perceived to have won the argument. In fact, according to the recent biography of Einstein by [Abraham] Pais, Einstein might as well have gone fishing from 1925 onward. Indeed, it is true that he didn't make many big advances, even though his penetrating criticisms were very useful. I believe that the reason why Einstein didn't continue to make big advances in quantum theory was that a crucial ingredient was missing from quantum theory. This missing ingredient was Stephen's discovery, 50 years later, of black hole radiation. It is this information loss, connected with black hole radiation, which provides the new twist. SA

NIELS BOHR



ALBERT EINSTEIN

CORBIS-BETTMANN; LAURIE GRACE

The Editors acknowledge the assistance of Gary T. Horowitz of the University of California at Santa Barbara.

The Hidden World of Surgery

In his finely resolved images of surgery, a photographer sees clues to who and what we are

Text and photographs by Max Aguilera-Hellweg

The first thing I saw when I walked into the room with my camera was the patient, hanging in midair, suspended from the ceiling by a clamp screwed into his skull. His eyes were taped shut, and the sound of his heart, amplified and projected from speakers, filled the room.

"Stenosis," the surgeon said, the word for a narrowing. The spinal column had calcified and was applying enough pressure on the spinal cord within the column to paralyze the man. Hanging him from the ceiling allowed the vertebrae to stretch out completely so the surgeon could cut and cauterize with precision; one false move could paralyze him forever.

I was there to take pictures of the surgeon, Frances K. Conley, but at one point she stepped aside and said, "Here, photograph this." I was awestruck. Exposed before me, awash in light, was the spinal cord, the thick strand of nervous tissue that extends from brain stem to coccyx, or tailbone. It was the most intimate, inviolate and vulnerable thing I had ever seen.

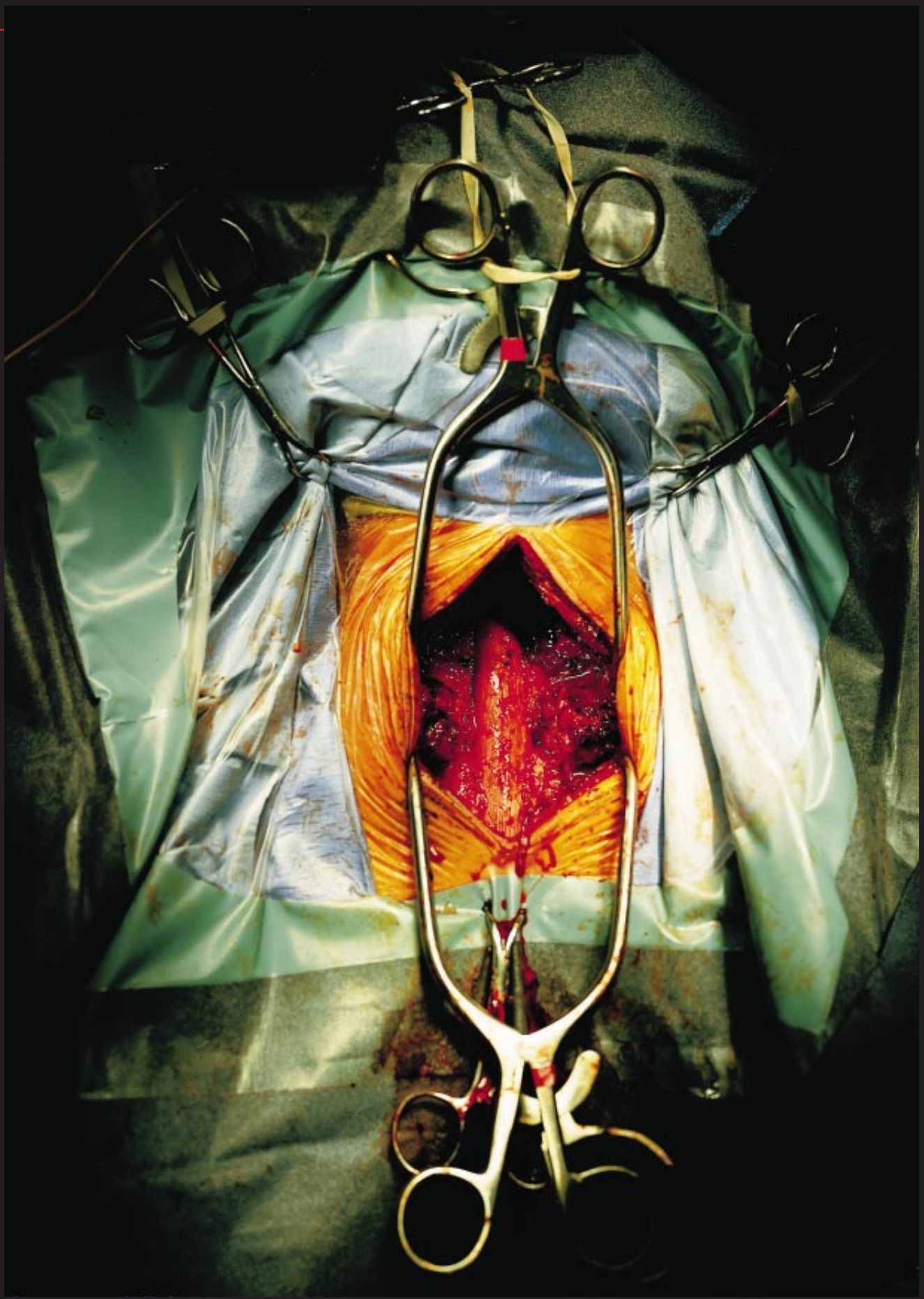
Positioning my camera, I set the exposure for the dome lamps, which mimic the spectrum of daylight and are so powerful that the fluorescent illumination of the rest of the room is dim in comparison. These lamps bathe the surgical field in light while the rest of the room fades to black. The special lighting and the large-format camera that I use contribute to resolution so fine and colors so vivid that the images recorded on film seem almost three-dimensional.

In the seven years since that revelatory day, I have photographed more than 50 different surgical procedures in nine different hospitals, including Johns Hopkins Hospital, Yale–New Haven Hospital, Columbia-Presbyterian Medical Center, New York University Medical Center and the University of California at San Francisco Medical Center.

These photographs reveal what is hidden to all but a few: the "priesthood" of physicians and surgeons. Some people are unable to look at these pictures, finding them too disturbing. Others see in them a means to address questions they have about their own bodies, their memories of disease or ongoing experiences with illness and treatment, or even profound questions about mortality and the soul. Like a Rorschach test, the meaning of the photographs is in the eye of the beholder.

A Tibetan monk once told me that you can slice a body into thin sections, centimeter by centimeter from head to toe, but never find the soul. Likewise, these photographs, both clinical and poetic, open the body to us and yet raise deep questions about who and what we are.

SPINAL CORD is revealed by the removal of the back sides of seven vertebrae. The dorsal pieces had become overly calcified, applying pressure to the cord and paralyzing the patient. Frances K. Conley performed the five-hour procedure, as the patient hung by his skull from a hook assembly in the ceiling. The successful operation was carried out at the Veterans Administration Medical Center in Palo Alto, Calif.



KNEE REPLACEMENT puts a steel-and-plastic joint in the leg to replace a natural one, typically crippled by arthritis. The "cutting jig," the metal-framed device surrounding the knee, enables surgeons to make precision cuts into the damaged bone and cartilage to remove them. The prosthetic joint must be precisely mounted to ensure smooth, natural motion of the resurfaced knee. (Translucent, orangish material visible here and in other photographs is a surgical drape, coated with Betadine, an iodine-based disinfectant.) Walter R. O'Brien, an orthopedic surgeon, performed this operation at Saint John's Hospital and Health Center in Santa Monica, Calif. The patient was under general anesthesia; in some cases, however, a spinal anesthetic is used to eliminate sensation only below the waist.

LIVER TRANSPLANT is seen shortly before completion, as the surgeons prepare to close the abdominal cavity. The patient, an Asian woman in her middle sixties, had been infected with the hepatitis C virus, which is relatively common in parts of Asia. Some 50 percent of patients who suffer liver damage from hepatitis and undergo transplants have a recurrence of symptoms in the new liver. The transplant gave this patient, in whom cirrhosis had developed from the disease, one more year of life. The surgeon was Nancy Ascher of the University of California at San Francisco Medical Center.

EYELIDS are sewn shut to prevent scratches or other accidental damage to the eyes before a particularly involved form of craniofacial surgery. This child patient had a bone disorder in which the plates of the skull grew at irregular rates, deforming it and leaving the cranium too small. In the subsequent procedure, surgeons fold back the scalp and restructure part of the underlying skull, including the facial part. Because the surgery is performed so close to the patient's brain, a neurosurgeon removes the skull, and a plastic surgeon does the reconstructive procedure. The doctors in this case, at Johns Hopkins Children's Center in Baltimore, were a pediatric neurosurgeon, Benjamin S. Carson, and a pediatric plastic surgeon, Craig Vander Kolk. The surgery was successful, resulting in a 9 percent increase in the young patient's cranial volume and a marked improvement in the shape of her head.

MOMENTS AFTER BIRTH, a boy rests on his mother's abdomen, umbilical cord still attached. The birth, by cesarean section, was carried out by Gerardo Del Zalle at University Hospital in Albuquerque, N.M. Fluid accumulation often causes a swelling of the testicles, which subsides soon after delivery.

COCHLEAR IMPLANT is inserted into a two-year-old girl born with profound nerve deafness. Noel L. Cohen of New York University Medical Center implanted the electronic device into the girl's inner ear, where it would directly stimulate the auditory nerve, sending signals to the auditory cortex of the brain. The white tube visible in the left side of the photograph is a surgical drain, which prevents the collection of fluid under the incision. It is removed the morning after the surgery, and the child is sent home. The auditory sensations possible with such an implant are different from those of normal hearing; to those who once were able to hear, the voices are said to sound initially like a record at the wrong speed or like Donald Duck. In time, however, the sound seems more normal. The implants help many adults and even congenitally deaf children to understand speech but are controversial in the deaf community, where their use is perceived as a threat to deaf culture.

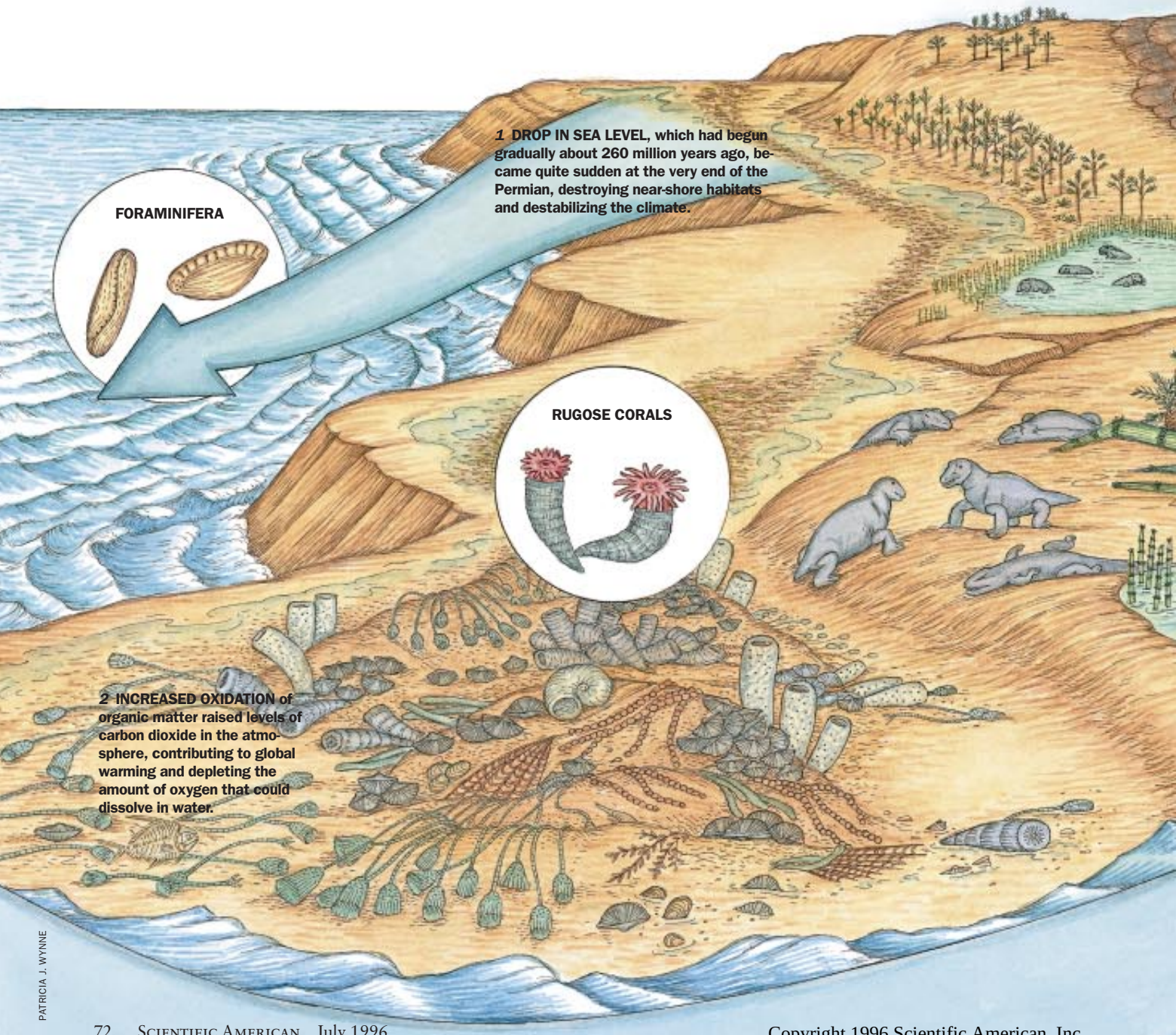
The Author

MAX AGUILERA-HELLWEG is a photographer and photojournalist. His work has appeared in many national magazines in the U.S., and one of his photographs is part of the permanent collection of the Museum of Modern Art in New York City. Recently he enrolled in undergraduate studies at Columbia University to begin fulfilling the requirements for admission to medical school. He is currently at work on a book of his surgical photographs, *The Sacred Heart: An Atlas of the Body Seen through Invasive Surgery.*

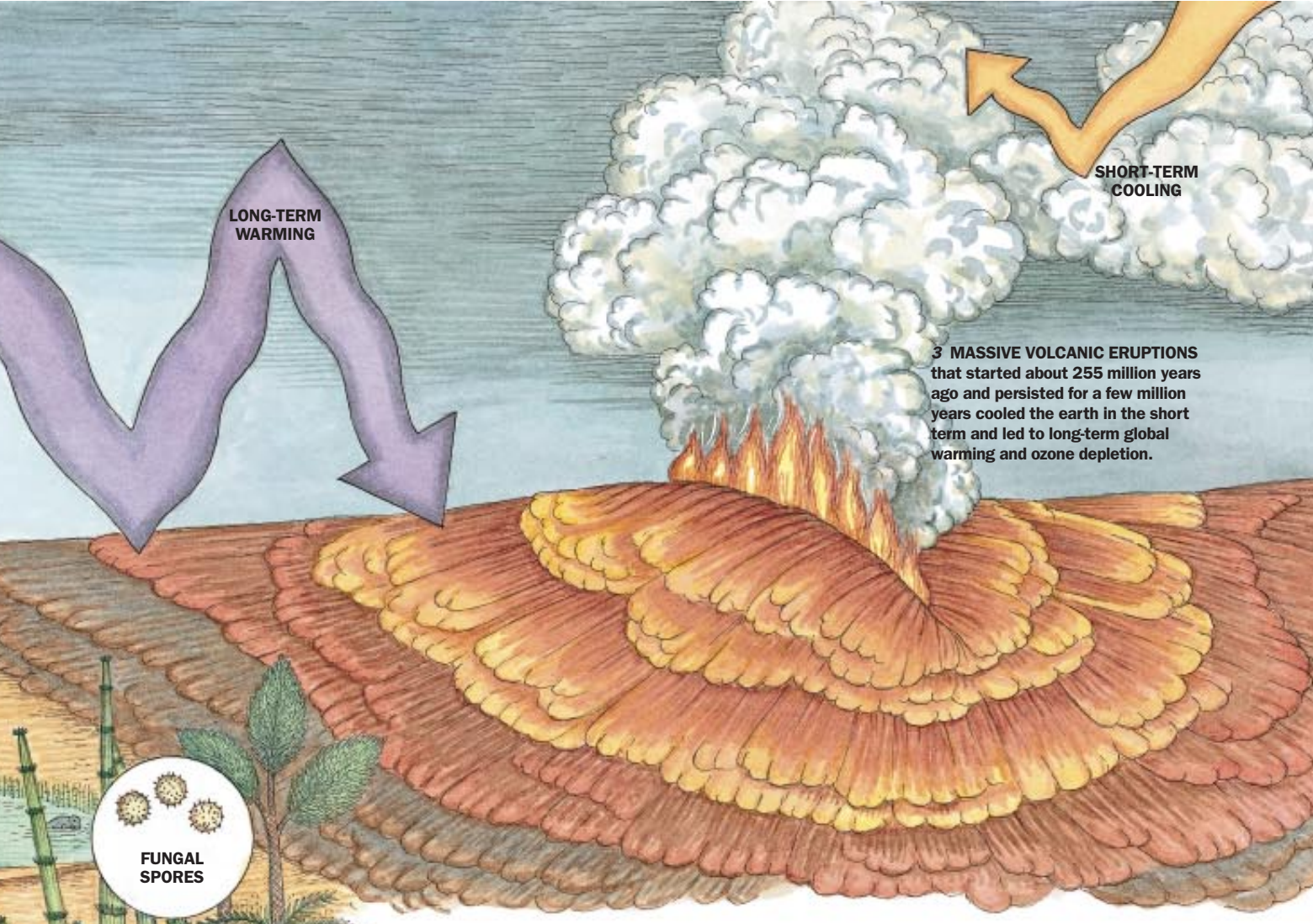
The Mother of Mass Extinctions

Disaster struck 250 million years ago, when the worst decimation in the earth's history occurred. Called the end-Permian mass extinction, it marks a fundamental change in the development of life

by Douglas H. Erwin



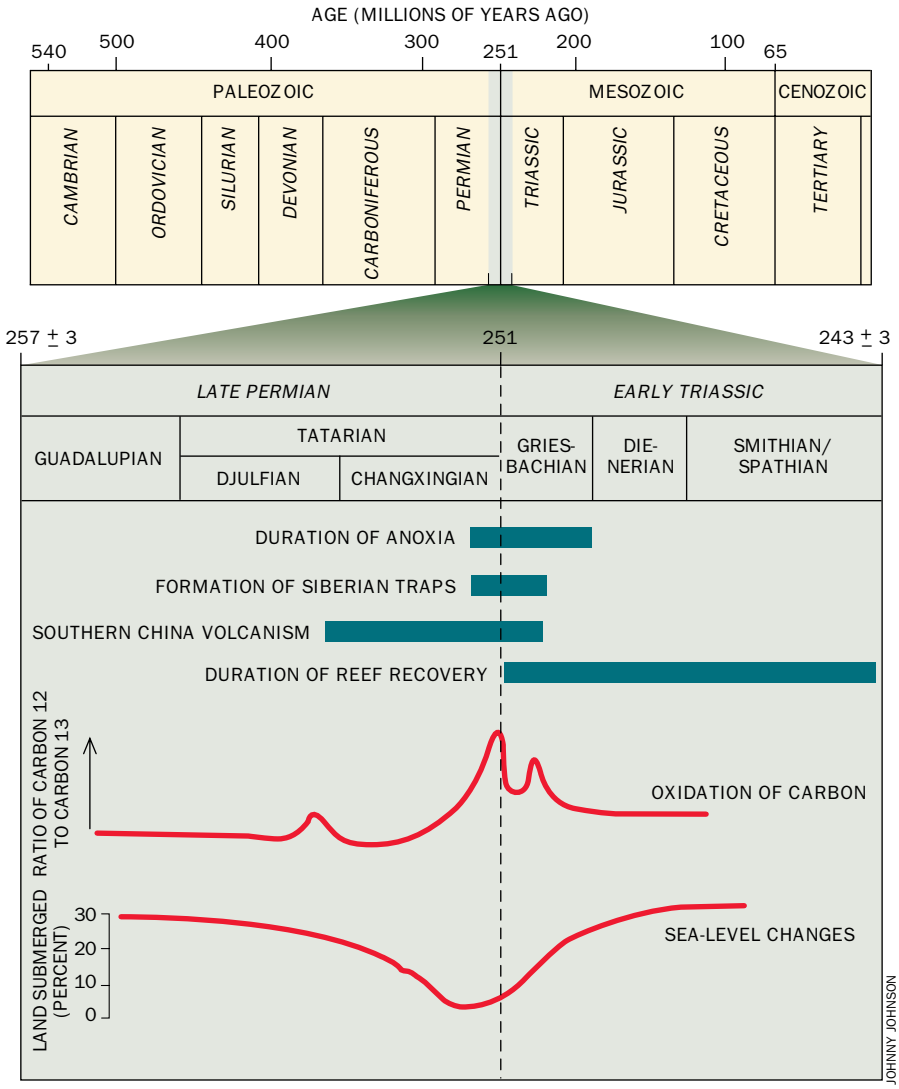
PATRICIA J. WYNNE



The history of life on the earth is replete with catastrophes of varying magnitudes. The one that has captured the most attention is the extinction of the dinosaurs and other organisms 65 million years ago—between the Cretaceous and Tertiary periods—which claimed up to half of all species. As severe as that devastation was, it pales in comparison to the greatest disaster of them all: the mass extinction, some 250 million years ago, at the end of the Permian period. Affectionately called "the mother of mass extinctions" among paleontologists (with apologies to Saddam Hussein), it yielded a death toll that is truly staggering. About 90 percent of all species in the oceans disappeared during the last several million years of the Permian. On land, more than two thirds of reptile and amphibian families vanished. Insects, too, did not escape the carnage: 30 percent of insect orders ceased to exist, marking the only mass extinction insects have ever undergone.

But from catastrophes, opportunities arise. For several hundred million years before the end-Permian event, the shallow seas had been dominated by life-forms that were primarily immobile. Most marine animals lay on the seafloor or were attached to it by stalks, filtering the water for food or waiting for prey. In the aftermath of the extinction, many once minor groups—active, predatory relatives of modern-day fish, squids, snails and crabs—were able to expand. Some completely new lineages appeared. This ecological reorganization was so dramatic that it forms a fundamental boundary in the history of life. Not only does it demarcate the Permian and Triassic periods, it also establishes the close of the Paleozoic era and the start of the Mesozoic era. The modern tidal pool reflects

DEADLY CATASTROPHES combined to wipe out most of life on the earth at the end of the Permian period, about 250 million years ago. Reef and shallow-water communities, consisting of corals, sea lilies, foraminifera and other organisms, were especially hard hit. On land, more than two thirds of reptiles and amphibians and nearly one third of insects disappeared. An increase in fungal spores suggests that plants, too, suffered.



DEGRADED ENVIRONMENT during the late Permian is revealed by geochemical and fossil evidence. At that time, more carbon was oxidized, the sea began dropping, and volcanism in what is now Siberia and China took place. Some layers in the oceans may also have become anoxic. Reefs did not recover fully until the middle Triassic.

what lived and what died 250 million years ago.

Over the past few years, exciting new insights into the causes and consequences of the end-Permian mass extinction have poured in from virtually every branch of the earth sciences. Some of these findings include detailed studies of rapid changes in ocean chemistry, more thorough documentation of extinction patterns and new analyses showing that large volcanic eruptions occurred at the Permo-Triassic boundary [see "Large Igneous Provinces," by Millard F. Coffin and Olav Eldholm; Scientific American, October 1993].

My own research during the past decade has been driven by curiosity about the events that structure life's evolutionary patterns. How much do mass extinctions contribute to the evolution of a group, as compared with long-term adaptive trends? For example, sea urchins are ubiquitous in modern oceans but were relatively uncommon during the Permian. Only a single genus, *Miocidaris,* is known for certain to have survived the extinction. Did *Miocidaris* survive by pure chance, or was it better adapted? Would sea urchins today look any different had it not been for the end-Permian extinction?

A Few Good Rocks

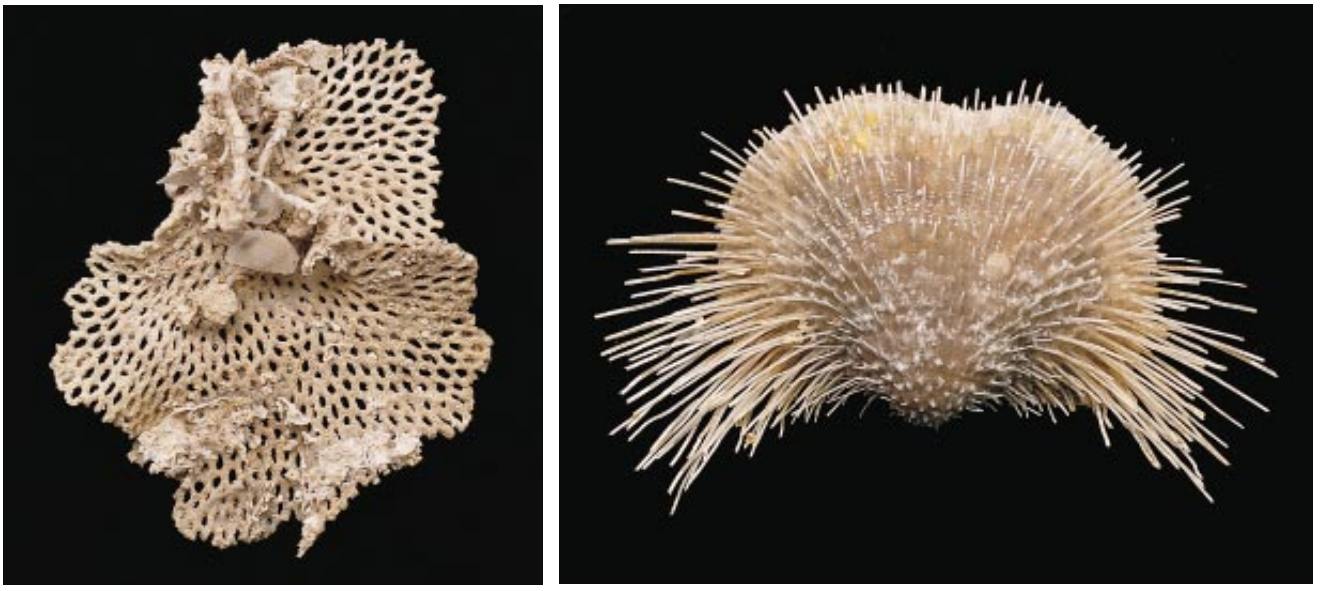
To resolve such questions, we need to learn more about the causes of the catastrophe and how those species that survived differed from those that disappeared. The key sources for this information are rock layers and fossils. Unfortunately, samples from the late Permian and early Triassic are notoriously difficult to come by. The fossil record across the boundary is plagued by poor preservation, a lack of rock to sample and other problems, including access. An extensive drop in sea level during the late Permian limited the number of marine rocks deposited on land, and many areas where the best rocks were preserved (most notably, in southern China) have been relatively hard for some geologists to reach.

As such, it has proved difficult to ascertain just how quickly life was snuffed out or if the deaths were subject to any regional variations. Some creatures, especially those sensitive to changes in the environment, died off rapidly, as shown by Erik Flügel and his colleagues at the University of Erlangen, who arrived at this conclusion after examining reefs in southern China and Greece. Other evidence indicates more gradual loss of life. For example, in studying the incredibly diverse and beautifully preserved fauna in the limestone outcroppings of western Texas and adjacent New Mexico and Arizona, I have found that many snails began vanishing late in the middle of the Permian, well before the main pulse of extinction.

Intensive studies of newly found and critical boundary layers in Italy, Austria and southern China have helped our understanding. They indicate that the duration of the extinction is shorter than previously thought, implying that abruptly calamitous environmental conditions must have set in. Only a few years ago, I believed the extinction period may have persisted five to 10 million years. It now appears that the final pulse may have lasted less than one million years.

Steven M. Stanley of Johns Hopkins University theorizes that the extinction may have consisted of two brief episodes, one occurring at the end of the middle Permian and the second at the end of the late Permian. Jin Yugan of the Nanjing Institute of Geology and Palaeontology, Samuel A. Bowring of the Massachusetts Institute of Technology and I are collaborating on a project to date volcanic ash beds in southern China and soon should have a better sense of the length of the extinction period. In any case, the rate appears to be about as rapid as many other mass extinctions.

In this geologically brief interval, the Permian oceans experienced a complex pattern of life and death. Quantifying the taxonomic extent of disappearance—from order to family to genus to spe-



KATHERINE LAMBERT

PERMIAN LIFE recorded in fossilized remains includes the bryozoans (*left*) and the brachiopods (*right*), two closely related phyla that were major components of marine life before the extinction. Each fossil is about 40 millimeters long.

cies—can be difficult. It is believed that on a global scale, 49 percent of families and 72 percent of genera were wiped out. Because species are harder to identify, that loss is far more difficult to pin down, and researchers have come up with varying estimates. Yin Hongfu and his colleagues at the China University of Geosciences examined the rock layer demarcating the Changxing stage in southern China. (This stage, along with the Djulfian stage, marks the later of the two subdivisions of the late Permian; each stage is named for the part of the world where the fossil record is clearest.)

Yin and his colleagues reported that out of 476 late Permian invertebrate species, 435 (or 91 percent) vanished. (Other estimates of global species extinction range from 80 to 95 percent, but the lower end of this range is probably most realistic.) By way of comparison, the event that occurred at the end of the Ordovician period, 439 million years ago, eliminated 57 percent of marine genera. The Cretaceous-Tertiary extinction, which killed off the dinosaurs, claimed up to 47 percent of existing genera.

The end-Permian devastation hit some animals harder than others. Groups that lived attached to the seafloor and filtered organic material from the water for nourishment suffered the greatest extinction. They included corals, articulate brachiopods (a kind of shelled invertebrate), some bryozoans (filter feeders that lived in colonies) and a variety of echinoderms (sea lilies). Other decimated marine groups included the last few trilobites, shallow-water foraminifera (a type of zooplankton) and ammonoids (distant relatives of the nautilus). Snails, bivalves and nautiloids came through the period fairly well, suffering just a few group losses. The only marine group that was truly indifferent to the mounting chaos was the conodonts, primitive chordates whose easily preserved mouthparts serve as important markers of time.

Things were not much better on land. Terrestrial vertebrates and insects both experienced substantial losses. Among vertebrates 78 percent of reptile and 67 percent of amphibian families disappeared during the late Permian, although how rapidly this occurred remains a subject of debate. Earlier studies from the magnificent fossils found in the Karroo region of South Africa suggested the decline took place over several million years, perhaps with two peaks in the extinction rate. Some recent work, though, suggests a more rapid decline, similar to the pace of marine extinction.

The extinction of insect species marks a major transformation of fauna. Of the 27 orders of insects known from the Permian, eight became extinct near the Permo-Triassic boundary, four suffered severe decimation but recovered, and three more barely survived into the Triassic before becoming extinct. This is the only significant insect extinction event yet identified, and it serves as a testament to the severity of the environment at the time.

Terrestrial flora suffered as well. To what extent, however, is impossible to say, for the evidence on the magnitude is, at the moment, not especially solid. In examining Australian leaf fossils, Greg J. Retallack of the University of Oregon showed this past year that plant extinctions were far more dramatic than had been thought and led to a rapid shift in the dominant floral types in Australia. (The loss of plant life may have also contributed to the disappearance of insects that fed on the flora.)

The record of pollen and spores more accurately reflects the effects on plants. In late Permian strata, pollen from gymnosperms (woody plants such as conifers) is almost absent, and succeeding layers harbor only fungal cells and some additional organic detritus. Last year Henk Visscher and his colleagues at the University of Utrecht in the Netherlands found that this so-called fungal spike seems to have begun in a latter part of the Permian—specifically, in the late Changxingian stage—before reaching its climax at the Permo-Triassic border.

Extinction Traps

Given the evidence of marine and terrestrial fossils, clearly the late Permian was a period when almost everything went wrong—at least if a species wanted to stay alive. What could have caused the wholesale loss of life? About the only thing that did not happen, or at least for which we have no evidence, was an extraterrestrial impact, an event that most likely killed off the

KATHERINE LAMBERT

EXTINCTION VICTIMS include fauna from the Karroo region of South Africa, where this 36-centimeter-long *Oudendon* (*left*), a therapsid (mammallike reptile), was recovered. The nine-centimeter-wide ammonoid (*right*) was found in Texas.

dinosaurs. In the mid-1980s a group of geologists claimed to have found vanishingly small traces of iridium, a critical indicator of a collision, at the Permo-Triassic boundary layer in southern China. But despite many attempts, no one has substantiated these assertions.

There is no shortage of murder suspects, however. One possibility is volcanism. A key piece of evidence is the Siberian traps, solidified layers of ancient lava. The traps (after the Swedish word for "stairs," which describes the steplike edges of the deposits) include at least 45 separate flows and range from 400 to 3,700 meters in thickness. They cover at least 1.5 million cubic kilometers, and perhaps more, for they may extend westward under younger rocks to the Ural Mountains. (In comparison, the eruption of Mount Pinatubo in 1991 was a mere puff, one that spewed ash but no magma. Perhaps a better comparison is the Laki eruption in Iceland, which in 1783 produced 15 cubic kilometers of lava.)

Recent radiometric dating suggests that all the lava of the Siberian traps may have erupted within a period of less than one million years, and perhaps in only 600,000 years, beginning near the Permo-Triassic boundary and extending into the earliest Triassic. Paul R. Renne of the Berkeley Geochronology Center has found that this estimate matches well with other large eruptions that deposited volcanic ash in southern China.

Could extensive volcanism have purged the earth during the late Permian? Eruptions have a variety of short-term effects, including cooling from both dust and sulfates ejected into the stratosphere (remember Pinatubo), acid rain, wildfires, release of poisonous trace elements and an increase in ultraviolet radiation from ozone-layer depletion. And over a longer timescale, the carbon dioxide emitted may lead to warming.

As appealing as this hypothesis is, killing some 90 percent of the species in the oceans is remarkably difficult. By itself, volcanism, even as rapid and extensive as that which produced the Siberian traps, is not up to the task. My colleague Thomas A. Vogel of Michigan State University and I examined volcanic ash sheets deposited by eruptions during the past 100 million years. These eruptions were similar in magnitude to those that produced the ash in southern China at the end of the Permian. We found that none of these events greatly affected the diversity of regional and global life on land or in the oceans.

Moreover, the environmental damage produced by any eruption depends on several factors. Many volcanic effects, such as the amount of sulfate ejected into the stratosphere, are difficult to infer from eruptions that occurred 250 million years ago. So eruptions may have been involved in the extinctions, but only as part of a more complex process.

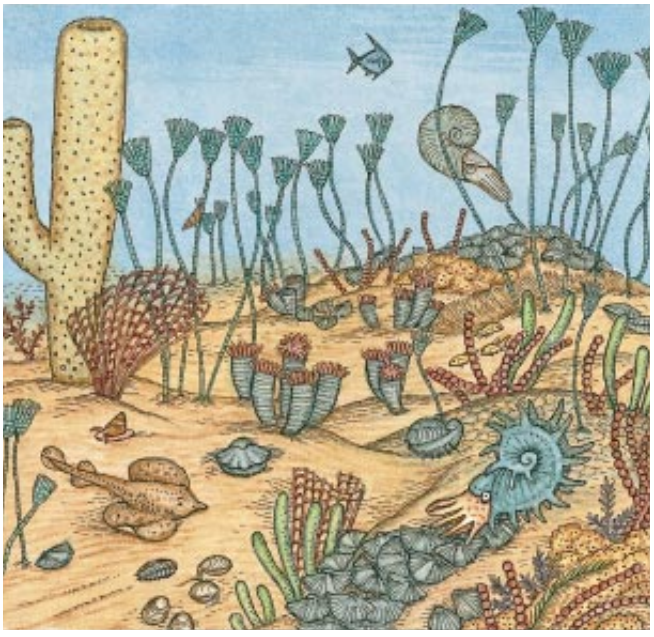
Geochemistry and the Stagnant Sea

The most intriguing new evidence about the end-Permian mass extinction comes from the field of geochemistry. Perhaps the most relevant geochemical changes are the shifts in carbon isotopes found in rocks (specifically, the ratio of carbon 12 to carbon 13). This fact indicates that, apparently, more organic matter was being buried during the late Permian than in previous times.

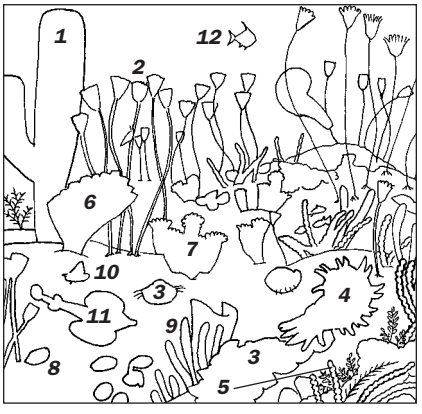
Although this burial of carbon is telling us something about geochemical changes during the end-Permian extinction, it is not entirely clear what. It may have to do with the sudden, deadly drop in sea level. During the early Permian, the continents merged to form the single supercontinent Pangaea. Around the continental shelves, reefs and other shallow-water communities thrived. Then, near the end of the Permian, the sea level fell. (No one knows exactly why, but it may have been caused by changes in the earth's mantle that enlarged the ocean basins.) The drop disrupted the habitats along the shore. With more of Pangaea's continental shelf exposed, greater erosion and oxidation of organic matter probably occurred. This oxidation reduced the oxygen and increased the carbon dioxide in the atmosphere, which may have humidified the planet and warmed it by as much as two degrees Celsius.

Further disruption occurred when sea levels rose again, perhaps several hundred thousand years later. The rising ocean waters engulfed near-shore habitats and swept inland. Such intrusions undoubtedly killed off many coastal communities.

Decreased amounts of atmospheric oxygen might also have exacerbated the hostile conditions already developing. Less oxygen would have dissolved into



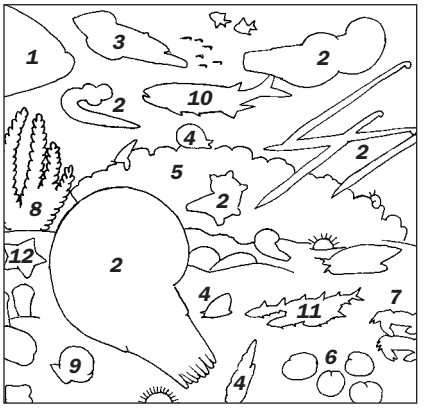
PATRICIA J. WYNNE



PATRICIA J. WYNNE

PERMIAN PERIOD

1 SPONGE
2 CRINOID
3 BRACHIOPOD
4 NAUTILOID
5 BEADED SPONGE
6 BRYOZOAN
7 CORAL
8 TRILOBITE
9 ALGAE
10 SNAIL
11 FISH (*Janessa*)
12 FISH (*Dorypterus*)



CRETACEOUS PERIOD

1 COELACANTH
2 AMMONOID
3 BELEMNOID
4 SNAIL
5 BIVALVES (RUDISTS)
6 SEA URCHIN
7 CRAB
8 ALGAE
9 SCALLOP
10 FISH (*Thrissops*)
11 FISH (*Davichthys*)
12 STARFISH

STRUCTURE OF MARINE LIFE was dramatically changed after the mass extinction. In the middle Permian (*left*), the seas contained mostly immobile animals, with some fish and a few trilobites. But by the Cretaceous period (*right*), the ocean resembled modern-day seas, with mobile bivalves, gastropods, swimming fish and cephalopods.

the oceans, causing anoxia, which could have suffocated some marine life. The evidence for anoxic waters is reflected by geochemical anomalies. Several researchers have recently made the intriguing, though not completely compelling, argument that patterns of extinction among various species reflect the abilities of the organisms to withstand anoxia.

What actually caused the extinctions? There may not be a single smoking gun, but rather all the possibilities mentioned may have contributed. None of them alone would have caused an extinction of this size, but it was the bad luck of the exquisite Permian faunas to have all of them interact at about the same time.

I believe the extinction consisted of three phases. The first began with the drop in sea level around much of Pangaea, which led to a loss of habitat, climatic instability and the elimination of many narrowly distributed species. As the oceanic regression continued, phase two began, with volcanic eruptions and the release into the atmosphere of large volumes of carbon dioxide, which increased climatic instability and facilitated ecological collapse. The rise in sea level and subsequent floods of possibly anoxic waters at the very end of the Permian and into the early Triassic initiated the third phase. It destroyed near-shore terrestrial habitats and contributed to the extinction of many surviving taxa.

Life after Death

The aftermath of the end-Permian extinction is at least as interesting as the event itself. After other mass extinctions, life began to recover within about a million years. In this case, though, it took perhaps five million years. The lag stems from the fact that biological communities had been so severely disrupted that millions of years were required for them to re-form and flourish. (There is also the possibility that the recovery appears longer than it actually was because of the poor preservation of fossils.)

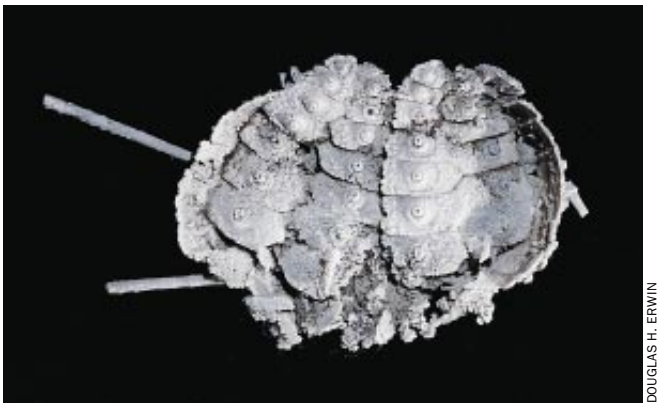
Regardless of how long it took to become reestablished, life on the earth had changed dramatically. As I noted earlier, largely immobile animals had dominated the Permian seas: brachiopods, bryozoans and echinoderms. They sat on the bottom, filtering the water for food or waiting for prey to swim by. The mobile animals—fish, bivalves, cephalopods (squids and their relatives) and gastropods (snails)—were around but formed only a small part of the community. A few trilobites remained.

Shortly after the extinction, during what is called the survival stage of the early Triassic, the few remaining species tended to be abundant and widespread. Earliest Triassic faunas consist of some

clams, ammonoids and a few gastropods. The reptilelike *Lystrosaurus,* an ancestor to mammals, was the most common vertebrate on land and was found throughout Pangaea. The clam *Claraia* thrived in the oceans. By the middle of the Triassic, though, some 25 million years later, sea urchins and other groups that were relatively sensitive to their environments began to reemerge, marking the start of a return to more normal marine conditions. These "Lazarus taxa" (as David Jablonski of the University of Chicago refers to them, for their return from the dead) began displacing the surviving fauna.

Indeed, by this time the seas were bearing some resemblance to the modern oceans. The more mobile creatures, such as bivalves, gastropods and crabs, dominated. A greater diversity of cephalopods and other predatory swimmers also emerged. Burrowing became fashionable, perhaps reflecting a need for more protection from predation. An evolutionary arms race between predator and prey ensued, driving changes in skeletal architecture on through the Mesozoic era that yielded fauna that had more flesh than Paleozoic creatures had. Such changes produced more complex and sophisticated ecosystems—there was more to eat and more of a menu selection, too.

Details of evolutionary changes on land for the same period are still somewhat sketchy, because the fossils have yet to receive a detailed, bed-by-bed sampling. Those proposed studies promise to boost our understanding of the terrestrial extinctions. We do know that several reptile and amphibian groups came to an end. Also, insects shifted from a variety of dragonflylike groups, which have wings that were fixed in the flight position and that could not be folded over the body, to forms that could fold their wings. These newer forms, which make up 98 percent of the insects of today, also had separate larval and adult stages. The adaptations may reflect an ability to exploit new habitats and to withstand severe seasonal swings and other climatic instabilities.



DOUGLAS H. ERWIN

PERMIAN SEA URCHIN *Miocidaris,* four centimeters long, was the only genus of echinoid to have survived the extinction.

Survival of the Fittest?

The changes that took place among insects raise the general question of whether species that made it successfully out of the Permian had specific adaptations that enabled them to survive or whether their survival was more random. Fossils of *Claraia* are found in rocks that harbor evidence of anoxic conditions. The large numbers and widespread distribution of *Claraia* may indicate that this species could survive on little oxygen. Another example is *Miocidaris,* the sole echinoid (sea urchin) survivor (although a close relative very likely survived as well). *Miocidaris* has only two columns of interambulacral plates (roughly speaking, the areas between the "petals" on the shell of a sea urchin); other Permian echinoids had anywhere from one to eight columns of plates. Because *Miocidaris* was the only genus to survive, the predominant form of echinoids shifted from those that had highly variable numbers of column plates to those that have only two. Some paleontologists have argued that an echinoid skeleton is stronger if it is composed of only two columns of plates and, thus, perhaps better adapted to survive predation in the post-Permian world.

Unfortunately, it is nearly impossible to say whether the Permo-Triassic extinction selected for certain features. All modern echinoids might have developed two-column plates even if the end-Permian extinction had never occurred. The surviving fauna may simply consist of groups that were the most abundant and widely distributed before the extinction and thus had the best chance to survive. Distinguishing between these two possibilities has turned out to be quite tough.

The only thing we can say for certain is that the end-Permian mass extinction had the greatest effect on the history of life of any event since the appearance of complex animals. Without this episode, there is little doubt that the composition of a modern tidal pool would look vastly different. Children would have grown up learning about crinoids and brachiopods instead of starfish and sea urchins, perhaps even looking in pools to catch a fleeting glimpse of a passing trilobite. SA

The Author

DOUGLAS H. ERWIN is a research paleobiologist and curator of Paleozoic gastropods and of the Burgess Shale in the department of paleobiology at the National Museum of Natural History of the Smithsonian Institution. His research focuses on the evolutionary history of Paleozoic gastropods and on large-scale evolutionary patterns, particularly the explosive spread of animals during the Cambrian and after the end-Permian mass extinction. He has studied Cambrian and Permian rocks in Siberia, China, Newfoundland and throughout the western U.S. In addition to one solo effort, he wrote with Derek Briggs and Fred Collier *Fossils of the Burgess Shale* (Smithsonian Press, 1994) and edited with Robert Anstey *New Approaches to Speciation in the Fossil Record* (Columbia University Press, 1995).

Further Reading

EXTINCTION: BAD GENES OR BAD LUCK? David M. Raup. W. W. Norton, 1991.

PERMO-TRIASSIC EVENTS IN THE EASTERN TETHYS: STRATIGRAPHY, CLASSIFICATION, AND RELATIONS WITH THE WESTERN TETHYS. Edited by Walter C. Sweet, Yang Zunyi, J. M. Dickins and Yin Hongfu. Cambridge University Press, 1992.

THE GREAT PALEOZOIC CRISIS: LIFE AND DEATH IN THE PERMIAN. Douglas H. Erwin. Columbia University Press, 1993.

THE PERMO-TRIASSIC EXTINCTION. Douglas H. Erwin in *Nature,* Vol. 367, pages 231–236; January 20, 1994.

SAM OGDEN

Who Owns Digital Works?

Computer networks challenge copyright law, but some proposed cures may be as bad as the disease

by Ann Okerson

Millions of readers since 1926 have found A. A. Milne's stories of Pooh and Piglet and their friends Eeyore and Tigger delightfully simple and yet profound. So it is not surprising that James Milne (no relation) of Iowa State University thought that it would be a wonderful idea to put *Winnie-the-Pooh* on the World Wide Web. A computer attached to the Internet could take a few files containing linked text and pictures from the books and make them available to children of all ages around the world. In April 1995, shortly after he created the Web site, Milne received a *very* polite letter (as have other Pooh fans) from E. P. Dutton, the company that holds the rights to the text and classic Pooh illustrations, telling him in the nicest way imaginable to cease and desist. His other choice was to sequester a substantial part of his life's savings for the coming legal bills.

About the same time, a scandalous new book about the private life of former French president François Mitterand was banned from distribution in print in France. It turned up anonymously on the Internet days later. There was little anyone could do to prevent its rapid digital dissemination.

Some network enthusiasts assert that "information wants to be free," but an equally vociferous band of digital pioneers contend that the real future of the global Internet lies in metering every drop of knowledge and charging for every sip. How will society's legal and cultural institutions react? Will tomorrow's readers be able to browse electronic works as easily as they have been able to peruse books at their favorite bookstore? Will they be able to borrow from virtual libraries? Authors, publishers, librarians and top-level government officials are debating these questions.

No More Yawns

Even five years ago few people would have thought of electronic copyright as an issue for heated national controversy. But today there are vast sums to be gained—or lost—as a result of the inevitable legal decisions to be made regarding ownership of "intellectual property" transmitted via electronic media. By the early 1990s the core copyright industries of the U.S. (which include publishing, film and music) accounted for more than $200 billion in business annually, or about 3.6 percent of the gross domestic product. In 1993, when QVC and Viacom battled for control of Paramount and its archive of classic films, it became clear that both companies believe the future lies in ownership of "content." Since 1981 the National Writers Union has sued large publishing organizations, including the New York Times Company and Mead Data Central, for allegedly selling unauthorized digital copies of its members' works. Even universities now think about how to maximize the return on the intellectual property they produce, rather than simply assigning full rights to publishers.

For the most part, the copyright in-

BROWSING is a time-honored tradition at newsstands and in bookstores and libraries. Proposed changes to copyright law could regulate electronic browsing and make it illegal unless expressly permitted. Although it is unlikely that electronic pub-

dustries create mass-market products such as trade books, films and related items. (The novel *Jurassic Park*, for example, spawned a major movie, videotapes, audiotapes, T-shirts, toy dinosaurs and other derivatives, all protected by various rights.) Scholarly and literary publishing—the scientific, critical and artistic record of human knowledge, culture and experience—accounts for only about half a percent of the total, or $1 billion a year. (Publicly funded government information, freely distributed, plays no important part in that market.)

Most scientists and scholars are far more interested in the widest possible distribution of their work to their professional colleagues than in capturing every possible royalty dollar. The Internet can deliver information more quickly and cheaply than traditional print formats can, which makes it an appealing vehicle for publishing. An electronic copy of a document or program will also usually be identical to the original and exactly as functional.

Yet such authors are merely passengers on a mass-media ocean liner, required to abide by the same copyright laws as the makers of action-figure toys based on Saturday-morning cartoons. And publishers' exhilaration about new products and markets is offset by fear that a single sale to a library or an individual could result in the endless reproduction of a document over the global Internet, eliminating hopes of further revenue.

Questions about how to apply current copyright law to new formats and media abound: To what extent are works on the newer—let alone not yet created—electronic media protected by law? Is cyberspace a virtual Wild West, where anyone can lay claim to anyone else's creations by scanning and uploading them or simply copying a few files? Many works are being created through extensive electronic communities or collaborations—who owns and benefits from these? How can we track who owns what, assuming that ownership makes sense at all? How do we efficiently compensate information owners when their works can be sold by the word, phrase or even musical note? What are the liabilities of Internet access providers, who may be unaware of copyright violation over their facilities? Should we dispense with copyright as we have known it entirely and seek new paradigms, as the Office of Technology Assessment advocated in 1986?

Where the Law Stands

The roots of copyright are old, and the lines along which it has grown are complex. One of the earliest copyright disputes, from sixth-century Ireland, sets the tone: St. Columba had copied out for himself a manuscript of the Latin Psalter, and the owner of the original, Finnian of Druim Finn, objected. The king ruled: "As the calf belongs to the cow, so the copy belongs to its book." A war ensued; the "copyright violator" prevailed and held on to the book. (The manuscript has had a long history as a good luck charm for the Columba clan's military adventures and survives to this day in the library of the Royal Irish Academy in Dublin.)

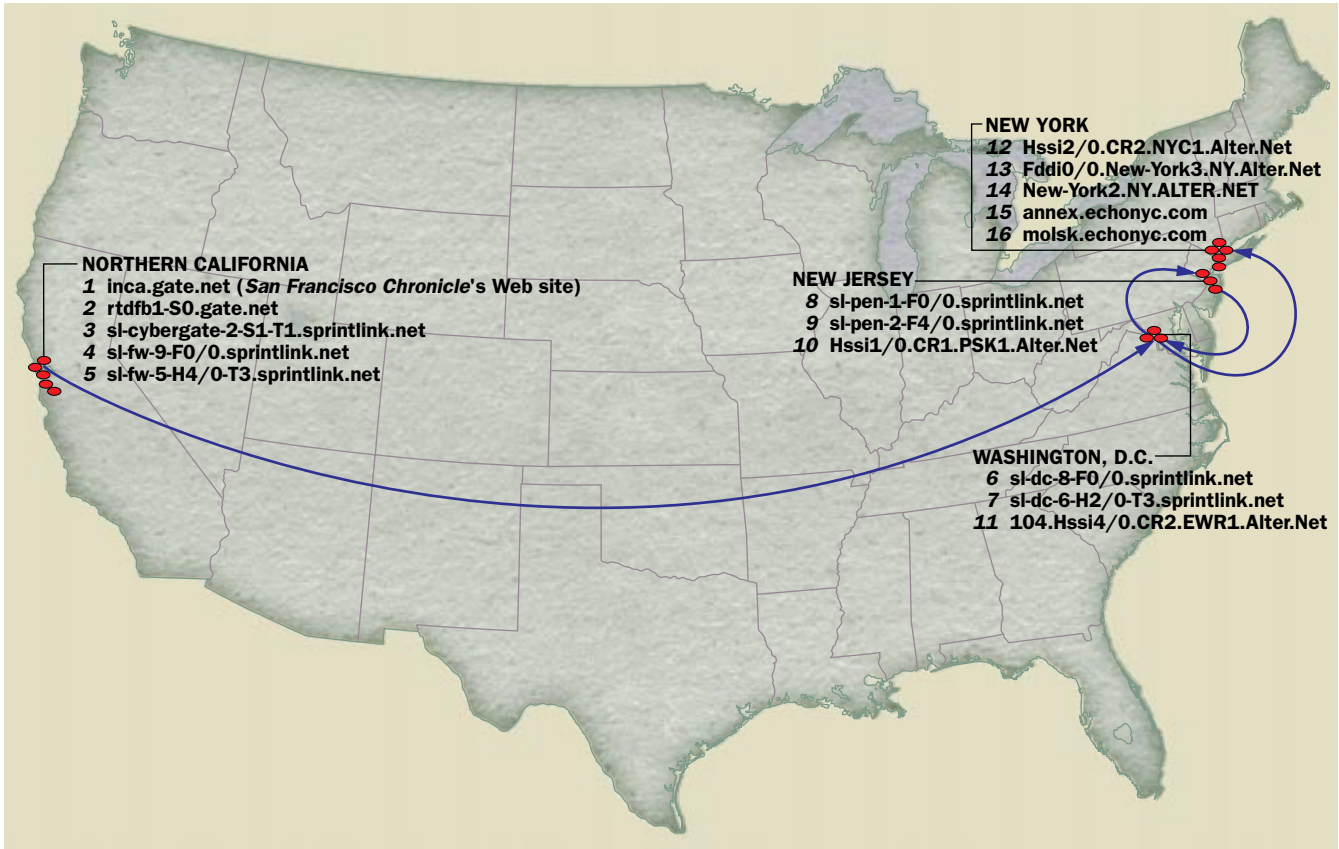
The Statute of Anne, enacted in England in 1710, was the first national copyright law. It gave authors rights to their work and limited the duration of those rights; it served as a model for the first statute governing copyrights in the New World, enacted in 1790. On both sides of the Atlantic, copyright in its nascent stages balanced neatly the interests of private property and public use. Indeed, the constitutional authority for U.S. copyright is based on its potential to "promote the progress of science and useful arts."

In successive revisions, Congress has extended the period of copyright, expanded the types of works that are protected, and joined in global copyright agreements, such as the Berne Convention. Berne signers agree to give copyrighted works from other countries the same protection they would have if they had been produced in the home nation.

American publishers have not always been so scrupulous in observing foreign copyrights. Pirate editions were common during the 19th and early 20th centuries (when, according to an argument made by Paul Goldstein of Stanford University and others, the U.S. was a net importer of intellectual property). British artists then, from Gilbert and Sullivan to J.R.R. Tolkien, were acutely

SAM OGDEN

lishers would utterly prohibit people from leafing through their publications, the author asserts that even a partial ban would upset the current balance between the rights of readers and copyright holders.



LAURIE GRACE

MULTIPLE COPIES OF A DIGITAL TEXT may be made as it travels across the Internet from source to destination. In the example above, information passes through 14 different computers (*shown by name and approximate location*) on its way from the *San Francisco Chronicle*'s World Wide Web site to a browser in New York City. Some copyright experts assert that each of these copies must be sanctioned by the copyright owner. As the proposed law now stands, the owners of all the intermediate machines might be found liable if it is determined that infringement has occurred.

aware of such trespasses. Today, however, Americans look at such lax-copyright countries such as China as disapprovingly as Britain looked westward 100 years ago.

Some observers have argued that cyberspace is a similarly underdeveloped territory with respect to intellectual property—many words and images from other media find their way there, but relatively few cyberworks have crossed in the opposite direction. That asymmetry is changing rapidly, though: the land rush of media companies to the Internet in the mid-1990s may already have put an end to the frontier era. And, of course, because cyberspace has no physical territory, its citizens are subject to the laws of whatever jurisdiction they live in.

A Tilted Playing Field

The most recent revision of the U.S. copyright law, made in 1978, is far more thorough than its predecessors. It protects creative works in general, including literature, music, drama, pantomime, choreography, pictorial, graphical and sculptural works, motion pictures and other audiovisual creations, sound recordings and architecture. (Patents and trademarks are governed by their own laws, as are trade secrets.) Copyright explicitly grants the owners of the expression of an idea the right to prevent anyone from making copies of it, preparing derivative works, distributing the work, performing it or displaying it without permission.

At the same time, the law limits the exclusive rights of owners in various ways. The most important of these exceptions is fair use, which allows copies to be made without either payment or permission under certain conditions. Fair use includes copying for purposes of research, teaching, journalism, criticism, parody and library activities.

Much of the current debate about electronic copyright stems from questions about the future of fair use raised by the Lehman Commission, known more officially as the National Information Infrastructure Task Force's Working Group on Intellectual Property Rights, chaired by Bruce A. Lehman, the U.S. commissioner of patents and trademarks. In mid-1994 the 25-member group released a first draft of its report for comment. Hearings were held in Washington, D.C., Chicago and Los Angeles, and the group took comments by post, fax and e-mail. Individual readers and copyright market participants offered well over 1,000 pages of opinion. In September 1995 the group released its final draft—a white paper—containing a legislative package intended to update the current Copyright Act.

In general, the information-producing industries have greeted the white paper's recommendations with relief and acclaim. It forestalls publishers' and authors' worst-case scenario, which could have reduced income to the point where there would be no incentives to produce new works and market them on-line. The tighter controls over digital reproduction proposed in the white paper appear to secure the industry's financial well-being in the on-line environment.

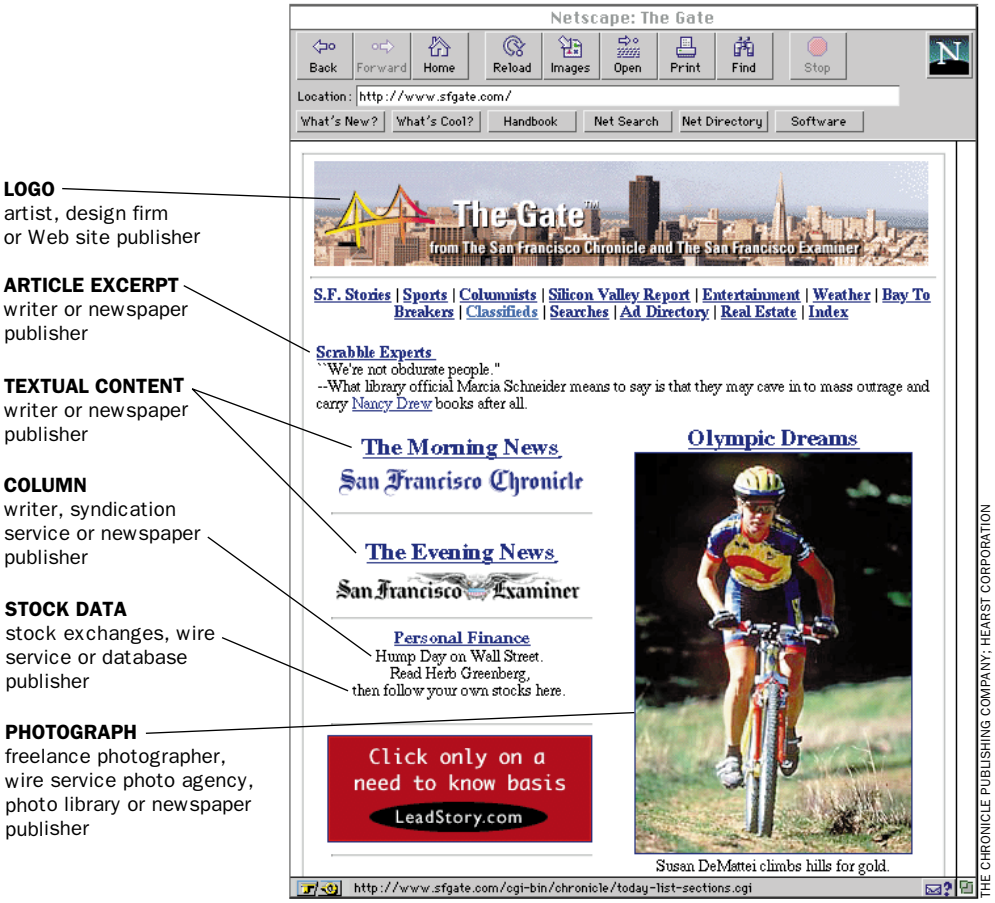
In contrast, library and education groups, on-line services and private citizens have been mostly negative—and very voluble—in their responses to the Lehman proposals. Their nightmare future is one in which nothing can be

looked at, read, used or copied without permission or payment. Many libraries are already feeling pinched as costs for information, particularly scientific books and journals, increase by 10 percent or more annually.

Fees charged for electronic information licenses (which give libraries or schools permission to use material that they do not own) are generally even higher than prices for the equivalent books or periodicals. Thus, the working group's suggestion that the use of licensing should be greatly expanded has an ominous ring for librarians in most American institutions. Under the typical license, such terms as price, permission for users to download sections of a database, liability and long-term ownership favor the information provider in significant ways. If license terms continue to make electronic information more expensive than its print counterparts, and the digital domain continues to grow, libraries will eventually be unable to afford access. At that point, of course, this imbalance will have to change, because there can be no marketplace without a ready supply of customers to buy new products.

Who Owns the Pieces?

Multimedia creation involves taking text, pictures and sound from many different sources. Each item (*as shown below*) may belong to a different entity. The resulting issues of ownership and compensation are complex and can hamper the creative process.



Furthermore, in the eyes of many citizens and legal scholars, the Lehman commission's suggested changes upset the balance that the current law maintains between the rights of copyright owners and those of users. For example, the commission affirms that any information alighting in a computer's memory for any length of time—however fleeting—is "fixed" for purposes of copyright. The Copyright Act governs only ideas "fixed in a tangible medium of expression, when its embodiment… is sufficiently permanent or stable to permit it to be perceived, reproduced, or otherwise communicated for a period of more than transitory duration." This distinction is crucial. The white paper implies that anyone who for any reason transfers a sequence of bits representing copyrighted information between computers without permission of the copyright owner breaks the law. Indeed, the working group recommends that the Copyright Act be amended expressly to recognize that transmissions fall within the exclusive right of the copyright owner. Even the act of viewing a Web page, which involves transmitting it from a server to a user's computer, could be interpreted as illegal without specific prior authorization.

In addition, the group refuses to extend to electronic copies the so-called doctrine of first sale. Someone who buys a book or magazine can sell or give away that copy without paying additional royalties, but this would not be true in cyberspace. This apparently illogical recommendation follows the argument that during an electronic transfer, a work is "fixed" in at least two computers, even if only for a few milliseconds—and hence is duplicated rather than being transferred the way that a book might be. Legitimate ways for a lawful owner of a copy of an electronic work to sell the copy or to give it to a friend, an act that is perfectly legal in the world of print-on-paper, are left unexplored. As a result, in the electronic information omniverse, the ability even to glance at materials, an act we take for granted in libraries and bookstores, could vanish. Browsing works on-line without permission could be considered a violation of the law.

Universities and other organizations that supply access to the Internet are particularly concerned by the commission's assertion that they should be liable for any copyright violations committed by their users. Such a situation would force them into the role of unpaid digital police, checking on every piece of data that students, staff or subscribers read or published.

Although the white paper proposes a future inconsistent with the grand tradition of public access, one way around these controversies might be disarming-

ly simple. The commission emphasizes technological aspects of "transmission" and "fixation," but many critics have found those discussions imperfect precisely on technological grounds. A more thorough analysis of the range of technological possibilities for transferring files—including cryptographic methods that effectively limit the number of permanent copies produced—might make the Lehman approach more useful than it now seems likely to be.

Fair Use—The Balancing Act

If access to electronic materials without payment for every use is to be recognized, then fair use is the area in which the bridges can be built between the rights of copyright owners and those of information users. At least that is where they have been built in the medium of the printed word.

Just what fair use means in the electronic environment is unclear. Other than stating that fair use should continue in the electronic realm and that the need for it will diminish as licenses and other automatic accounting techniques become more widespread, the white paper says little about it. Advocates of readers' and users' rights find this imprecision particularly troubling. Although the Lehman commission made a clear statement in favor of owners' rights, it balked at the chance to clarify users' rights similarly.

Lehman's office has, however, continued to foster an unofficial series of meetings, at which between 50 and 70 users, authors, librarians, lawyers and publishers' representatives meet every month in Washington, D.C., in an attempt to evolve guidelines for electronic fair use. This Conference on Fair Use is affectionately called CONFU, a reference to the CONTU (Commission on New Technological Uses) group that drafted the guidelines that have helped to clarify the 1978 copyright revisions. From the start of CONFU, it quickly became clear that little agreement would be reached by the time the white paper was due or copyright legislation would be introduced and debated on Capitol Hill.

What policymakers may not appreciate is that the inability of the CONFU participants to agree is by no means bad. On the one hand are fears that without a set of electronic fair use guidelines, confusion about the law and the likelihood of litigation will increase—particularly damaging for elementary schools and others who can least afford the risk. Yet on the other hand are the advantages of proceeding slowly with legislation. Because there are not a lot of rules about the new media, publishers, librarians and scholars are free to conduct electronic experiments—many of them governed by written agreements between commercial publishers and educational or library organizations.

Progress in the creation and distribution of electronic information is being made nicely, though not rapidly. Commercial copyright owners seem a long way from suing libraries or elementary schools. The individual scientist or teacher preparing a Web page may be in technical violation of one or another owner's rights but seems similarly immune, at least for now, from legal action. (If this truce breaks down, of course, the consequences for electronic distribution of information could be grim.)

In the view of many participants, the disagreements at CONFU meetings deserve to be cherished. Many believe the technology is not mature enough for agreement about fair use guidelines. They shy away from making legal commitments before they really understand the implications of what they agree to, and at this writing it appears that the process of reaching adequate voluntary electronic fair use agreements will take a long time.

Coming to Terms with the Future

For all the criticism that some aspects of the Lehman commission's report have generated, there is substantial consensus on many others. Many recommendations have generated dissension not so much about their general appropriateness as about the degree to which they should be codified in law. For example, few take issue with the notion that malicious tampering with encryption methods intended to secure copyright should be illegal. Questions arise only about how draconian the punishment for such an offense should be and whether investigators should be able to presume guilt.

Similarly, everyone, except for the small minority who believe copyright protection has no future on the Internet, agrees that there is an urgent need to educate citizens about copyright. Now that everyone with a computer and modem is a publisher, rules that once applied to only a few companies bind millions of people.

How society ultimately changes the Copyright Act will largely determine the nation's information future. The power of new technologies already transforms the way creators work and how authors and publishers deliver information. Is it too much to hope that widely and cheaply accessible public and academic information will coexist with information sold by publishers at prices that earn profits and foster the copyright industries? We do have the potential, if we act wisely and well, to arrange matters so that most participants in the new technologies will be winners. SA

The Author

ANN OKERSON is associate university librarian at Yale University, where her work includes making digital materials available to the library's users. Before going to Yale in September 1995, she was director of the Office of Scientific and Academic Publishing for the Association of Research Libraries in Washington, D.C. Okerson received her M.L.S. in library and information science from the University of California, Berkeley, in 1967. Since January 1993 she has been co-moderator of *NewJour,* an electronic mailing list that announces new electronic journal, magazine and newsletter start-ups. The list's World Wide Web site (http://gort.ucsd.edu/newjour/) has links to the journals described in the announcements and handles about 1,000 searches a day.

Further Reading

THE NATURE OF COPYRIGHT: A LAW OF USERS' RIGHTS. L. Ray Patterson and Stanley W. Lindberg. University of Georgia Press, 1991.

COPYRIGHT'S HIGHWAY: FROM GUTENBERG TO THE CELESTIAL JUKEBOX. Paul Goldstein. Hill and Wang, 1994.

COPYRIGHT, PUBLIC POLICY, AND THE SCHOLARLY COMMUNITY. Association of Research Libraries, Washington, D.C., July 1995.

COPYRIGHT LAW OF THE UNITED STATES OF AMERICA. Contained in Title 17 of the United States Code. Obtain copies from U.S. Government Printing Office or the Copyright Office at the Library of Congress.

For access to the World Wide Web site maintained by the creator of the Pooh FAQ: http://www.clark.net/pub/rinzel/muppooh/

For current information on copyright law: http://www.library.yale.edu/~okerson/copyproj.html

Exoskeletal Sensors for Walking

To move their limbs, cockroaches, crabs and spiders rely on organs in their exoskeletons that act as strain gauges. Their method of locomotion could facilitate the design of multilegged robots

by Sasha N. Zill and Ernst-August Seyfarth

When spiders search for prey, or crabs charge into the surf, or cockroaches flee into a crevice, they are relying on natural sensors that are the envy of many human engineers. These arthropods have specialized organs in their legs that detect small deformations, or strains, in their exoskeleton. These so-called biological strain gauges are exquisitely sensitive, rivaling those that engineers can currently make. In fact, they go beyond human-made gauges in that they both measure strain and exert control: they regulate the creatures' walking movements and can even be used to "memorize" stepping patterns.

Studying the strain gauges of arthropods not only enhances understanding of these invertebrates, it also provides important clues for the science of robotics. To build machines that could navigate rough terrain on Earth—or even on the moon or Mars—roboticists have leaned toward crafting insectlike automatons, which have the best balance (rarely does a crab trip and fall). Perhaps the most famous such robot is Dante, the eight-legged mobile mechanism that descended into volcanic craters in Antarctica and Alaska in the early 1990s. Yet that robot, which tipped over and suffered other mechanical troubles, never came close to matching the grace and agility of a spider. The natural strain gauges of arthropods, coupled with other aspects of their locomotion, such as gait and body posture, could be a key to building more nimble automatons.

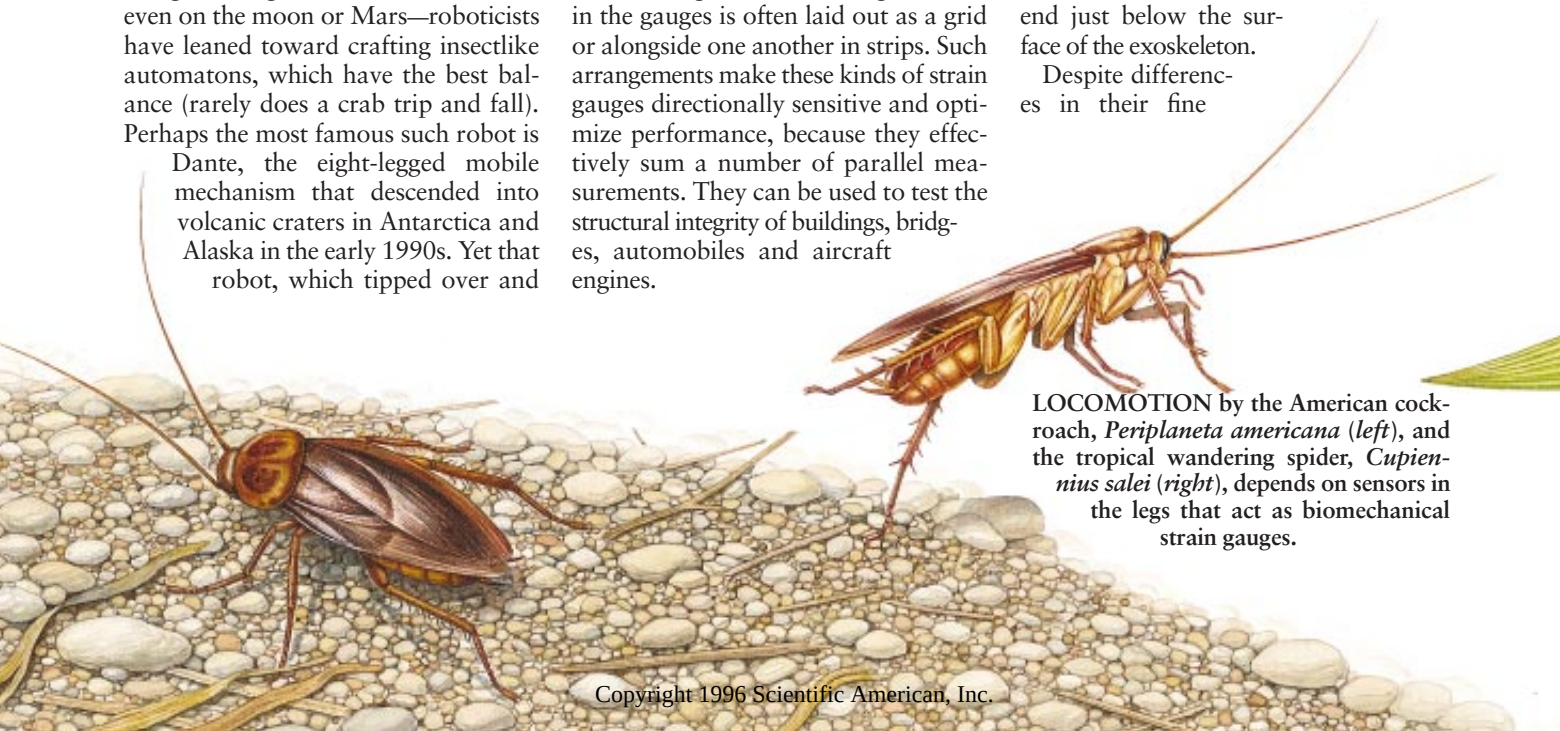
In engineering terms, strain is the ratio of the change in the length of an object to its original length. Strain gauges consist of a metal foil or flexible semiconductor, the electrical resistance of which varies with the amount of strain. The electronic bathroom scale is a familiar application of strain gauges. A person stepping on the scale distorts a mechanical beam, and an attached gauge transforms the beam's deflection into an electrical signal that is displayed as a readout of pounds or kilograms.

More sophisticated strain gauges have widespread industrial applications. In these settings the conducting material in the gauges is often laid out as a grid or alongside one another in strips. Such arrangements make these kinds of strain gauges directionally sensitive and optimize performance, because they effectively sum a number of parallel measurements. They can be used to test the structural integrity of buildings, bridges, automobiles and aircraft engines.

The strain gauge is a relatively new tool in science and engineering. But arthropods have used comparable devices for more than 300 million years. Each of the three main arthropod groups—insects, arachnids and crustaceans—has its own version of a sense organ for detecting deformations or strains in its cuticle, or exoskeletal material.

In insects this organ is known as a campaniform sensillum (from the Latin *campana,* meaning "bell," which roughly describes the appearance of the organ). The sense organ of spiders is called a slit sensillum. It consists of a fine canal, or slit, in the cuticle and is covered by an extremely thin membrane on its outer surface; on the legs these sensilla are found in tightly organized groups that usually appear harp-shaped (lyriform). The sensors of crabs, collectively known as, simply, force-sensitive organs, tend to look like canals that end just below the surface of the exoskeleton.

Despite differences in their fine



LOCOMOTION by the American cockroach, *Periplaneta americana* (*left*), and the tropical wandering spider, *Cupiennius salei* (*right*), depends on sensors in the legs that act as biomechanical strain gauges.

structure, all three types of sense organ are linked to the central nervous system in similar ways. Each is attached to a part of a sensory nerve cell (specifically, the dendrite) hidden underneath the exoskeleton. When a force deforms the exoskeleton around a sensillum, the tip of the dendrite deforms with it. In response, the nerve cell generates bursts of electrical activity (action potentials), which are conducted to the central nervous system via the axon of the neuron.

Although the basic anatomy of arthropod cuticle sensors had been known for some time, their strain-gauge ability did not become evident until the advent of modern neurophysiological recording techniques in the 1930s. Such instrumentation relies on electronic amplifiers to detect nerve cell activity. The late John W. S. Pringle of the University of Oxford was the first to explore arthropods in this way. During his research, Pringle noted a consistent orientation of these receptors, which hinted at the directional sensitivity of a strain gauge. Moreover, the sense organs are concentrated in critical strain regions of the exoskeleton—areas near the joints and tips of the legs. Furthermore, Pringle observed that many groups of sense organs are clustered near exoskeletal sites attached to muscle tendons, suggesting that arthropods' sensilla respond to the strains of internal muscle contractions.

More recently, a series of remarkable experiments by Reinhard Blickhan and Friedrich G. Barth of the University of Frankfurt revealed the extreme sensitivity of these arthropod strain gauges. The investigators directly measured the strains in the leg of the tropical wandering spider, *Cupiennius salei,* by attaching miniature gauges to segments of the leg. They concluded that, kilogram for kilogram, arthropod cuticle is nearly as stiff as human bone. (Technically speaking, cuticle and bone have an elastic modulus—the ability to withstand an amount of force per area—on the order of 10 billion newtons per square meter, or one million pounds per square inch.) That the sensors can detect changes in such a stiff structure indicates high performance: they are as sensitive to displacement as the receptors in the human ear are to sound.

Subsequent physiological studies have probed the pronounced directional sensitivity of these arthropod sensors. Stanley M. Spinola and Kent M. Chapman of Brown University showed that in insects, the campaniform sensillum responds best to forces that compress the organ perpendicular to its long axis. Barth (now at the University of Vienna) and his associates discovered a similar property in the sensory slits of spiders. They also found that spiders have a wide range of slit lengths within a single lyriform organ. Each slit is angled slightly differently from another, a layout that enhances the spider's ability to resolve the direction of incoming forces. Recordings of crustacean receptors have revealed a similar capacity, although the mechanical basis for their directionality has not been established.

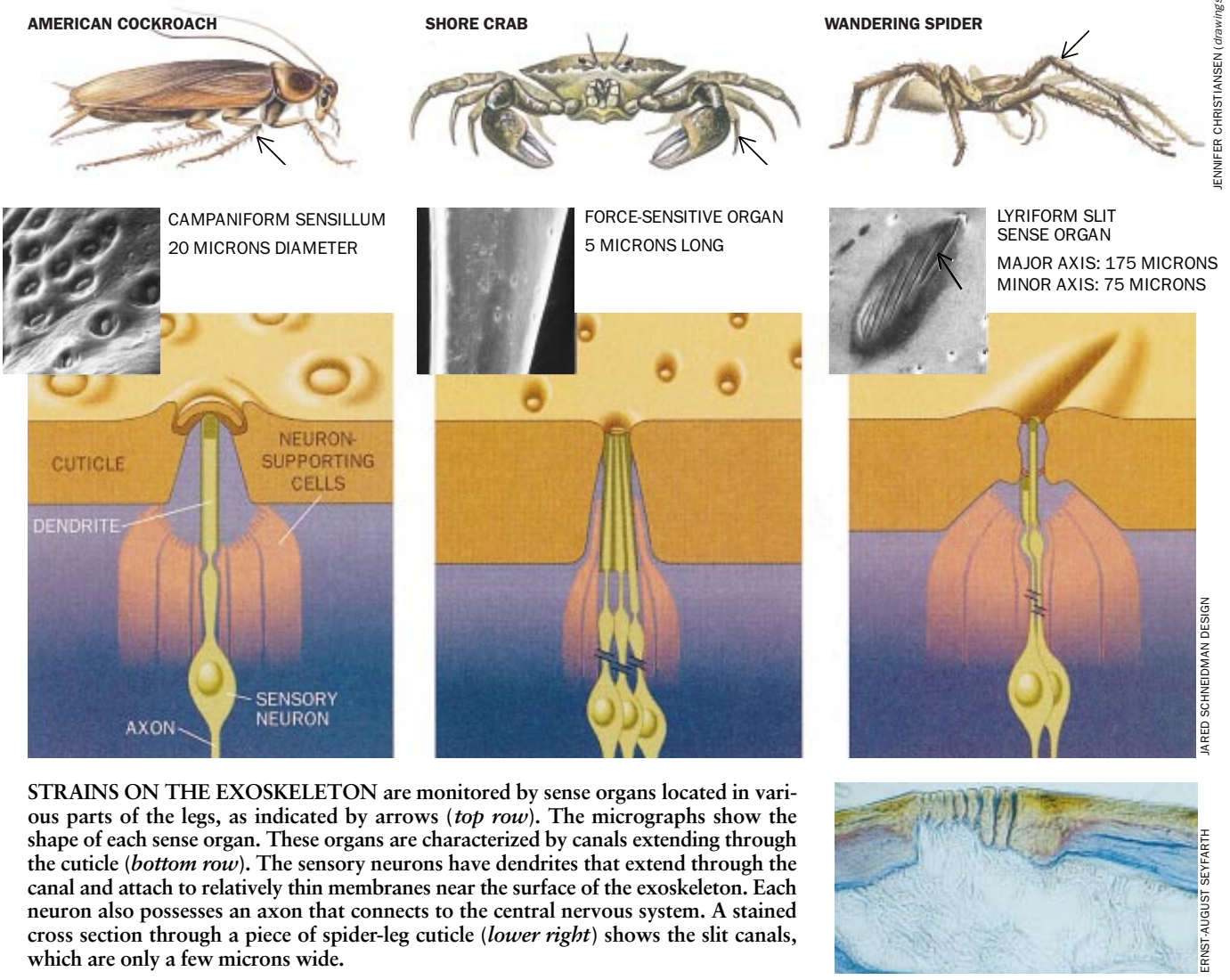
A Walking Machine

Whereas earlier studies demonstrated the nature of strain gauges in arthropod organs, they did not indicate their role in the creatures' actual movements. So several researchers, including the two of us, began to investigate how arthropods use such gauges in locomotion. For this purpose, the American cockroach makes an excellent subject. Although roaches are a pest in the home, in the biomechanics laboratory they make ideal walking models. Lacking any other defense mechanism, they depend on fleetness to escape.

Cockroach campaniform sensilla are linked to the exoskeleton by a hard cap made from the exoskeletal cuticle. The sensilla located on the leg's tibial segment—the shin or the forearm, so to speak—form a precise array. They are divided into two subgroups: proximally located receptors, which have caps oriented perpendicular to the long axis of the leg, and distal receptors, which have caps pointing parallel to the leg axis. (Structural engineers often situate strain gauges in the same way to monitor the full range of forces acting on an object, arranging them in mutually perpendicular orientations described as two-axial, 90-degree rosettes.)

To look at how these sensors affect the muscles, we first restrained the animals by trapping their legs under staples. We then stimulated the caps of individual organs with a fine tungsten wire, only microns in diameter, and monitored the tibial muscles. Receptors with perpendicular caps re-

ROBERTO OSTI



MICROGRAPHS BY ULRIKE GRUENERT (*left*), MANFRED SCHMIDT (*center*), ERNST-AUGUST SEYFARTH (*right*) *University of Frankfurt*

JENNIFER CHRISTIANSEN (*drawings*)

JARED SCHNEIDMAN DESIGN

ERNST-AUGUST SEYFARTH

STRAINS ON THE EXOSKELETON are monitored by sense organs located in various parts of the legs, as indicated by arrows (*top row*). The micrographs show the shape of each sense organ. These organs are characterized by canals extending through the cuticle (*bottom row*). The sensory neurons have dendrites that extend through the canal and attach to relatively thin membranes near the surface of the exoskeleton. Each neuron also possesses an axon that connects to the central nervous system. A stained cross section through a piece of spider-leg cuticle (*lower right*) shows the slit canals, which are only a few microns wide.

sponded to dorsal, or upward, bending of the tibia. That stimulus in turn excited the extensor muscle (equivalent in function to the human quadriceps muscle). Sensilla with parallel caps fired in response to ventral, or downward, bending and inhibited extensor activity.

Moreover, the receptors responded to strains from contractions of the animals' own leg muscles, confirming previous suspicions. Sensilla with perpendicular caps fired only during the bending of flexor muscles (similar in function to human hamstrings), and those with parallel caps responded only to contraction of extensor muscles. Thus, the discrete functions of the insect's sensilla are quite similar to those of simple strain gauges used in mechanical engineering.

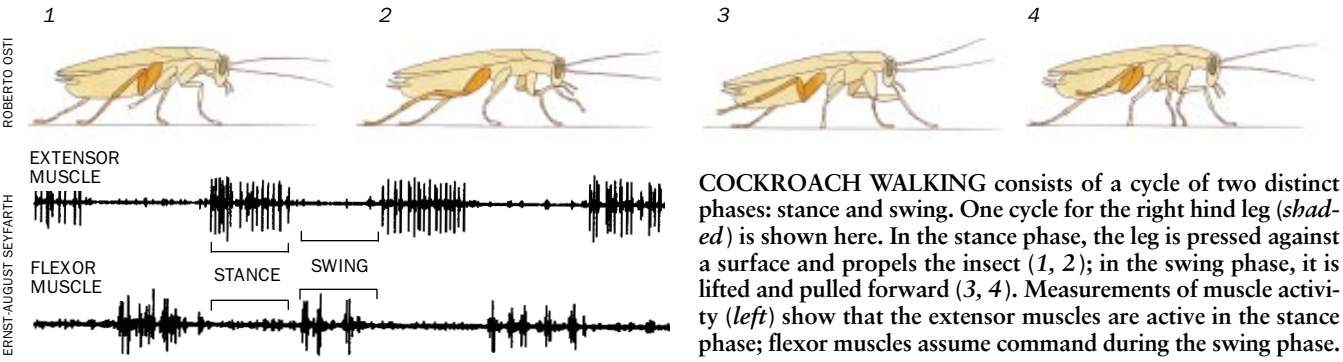
Testing stapled cockroaches, of course, does not indicate how the sensillum controls actual walking. To answer that question, we recorded the receptor activities in freely moving insects and correlated them with the workings of the tibial muscles. The hind legs of cockroaches follow the relatively simple pattern that most animals use in walking: a leg is first lifted and pulled forward in what is known as the swing phase. It is then placed down and pushed back during the stance phase; it is here that the leg actively supports and propels the animal. The flexor muscles are active during the swing phase, whereas the extensors contract during the stance phase.

To record the electrical activities of the sensilla, we implanted fine wires in the leg, close to the nerve containing the receptor axons. The tibial sensilla showed consistent patterns of activity during walking. At the onset of the stance phase, the animal places a leg on the walking surface and then applies its weight to it. This force bends the upper surface (dorsal side) of the tibia and excites the sensilla with perpendicular caps. The sensors cause the extensor muscle to contract, propelling the animal.

The muscle keeps working until the end of the stance phase, when continued contractions eventually bend the lower (ventral) part of the tibia. That bending in turn activates the second kind of gauge: the receptors with caps parallel to the leg axis. The firing of these receptors essentially reverses the process: it inhibits the extensor muscle, limiting the muscle contractions and the resulting cuticular strains. It also activates the flexor muscle, which swings the cockroach's leg back to start another step. Thus, by utilizing directional information from biological strain gauges, the animal monitors the alternate dorsal and ventral bending forces acting on the leg and regulates its walking movements accordingly.

Shoes for Crabs

Strain-gauge organs can also help solve more complex coordination tasks. Consider the crab, which confronts a set of problems different from that faced



ROBERTO OSTI

ERNST-AUGUST SEYFARTH

COCKROACH WALKING consists of a cycle of two distinct phases: stance and swing. One cycle for the right hind leg (*shaded*) is shown here. In the stance phase, the leg is pressed against a surface and propels the insect (*1, 2*); in the swing phase, it is lifted and pulled forward (*3, 4*). Measurements of muscle activity (*left*) show that the extensor muscles are active in the stance phase; flexor muscles assume command during the swing phase.

by cockroaches. First, crabs have 10 legs, which demands more coordination. Second, many crustaceans, such as the shore crab, *Carcinus maenas,* walk sideways. Therefore, crab legs are used to push and pull the animal, unlike the appendages of cockroaches, which are used mainly in forward walking.

The location of the crab's force-sensitive organs enables it to take full advantage of its 10 legs and lateral mobility. They are clustered in dense concentrations on the dactyl, or foot, of the leg, rather than on the tibia, as in the cockroach. For the crab, this positioning seems to offer an ideal advantage, allowing the organs to keep track of the forces throughout the leg: the sensilla respond to the bending of the dactyl as well as to strains produced by muscle contractions. The force-sensitive organs also modulate the activity of the crab's main walking muscle, called the depressor muscle, which presses the leg against the surface on which the crab walks.

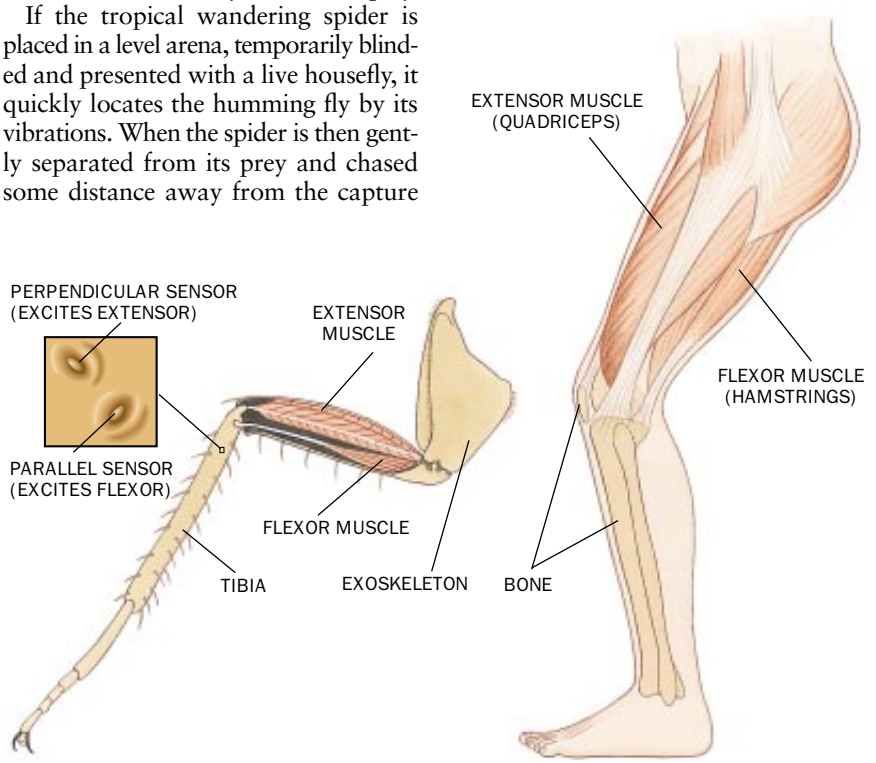
These sensors, though, seem to be involved in much more complex behaviors than simple reflexes. Fred Libersat and François Clarac, working at the CNRS laboratory in Arcachon, France, discovered this possibility when they tested the coordination of leg movements after fitting the crab with, in effect, a very tight shoe, a miniature clamp that is tightened by turning a screw. It would thereby deform the dactyl as the crab moved about. As a result, it continuously stimulated the force-sensitive receptors on that leg segment.

Squeezing the crab's dactyl in this way produced several effects. First, it substantially decreased the intensity and duration of activity in the depressor muscle of the stimulated leg. Second, it altered the activity in the depressor muscles for the adjacent legs by greatly prolonging their bursts of activity and forcing them to fire in an alternating pattern. In other words, when weight is applied on a leg, the sensor in that leg tells the depressor muscle to ease up and instructs the adjacent legs to pick up the weight in a coordinated fashion. That an organ on one leg helps to control the reflex activity of several other legs indicates that the arthropod strain gauge creates behavior that is more complex than simple reflex action.

In fact, the role of sense organs can be as subtle as it is complex. One illustration is the "kinesthetic orientation" of spiders. This process enables the animals to memorize information about their own previous walking movements in order to find the way back to lost prey.

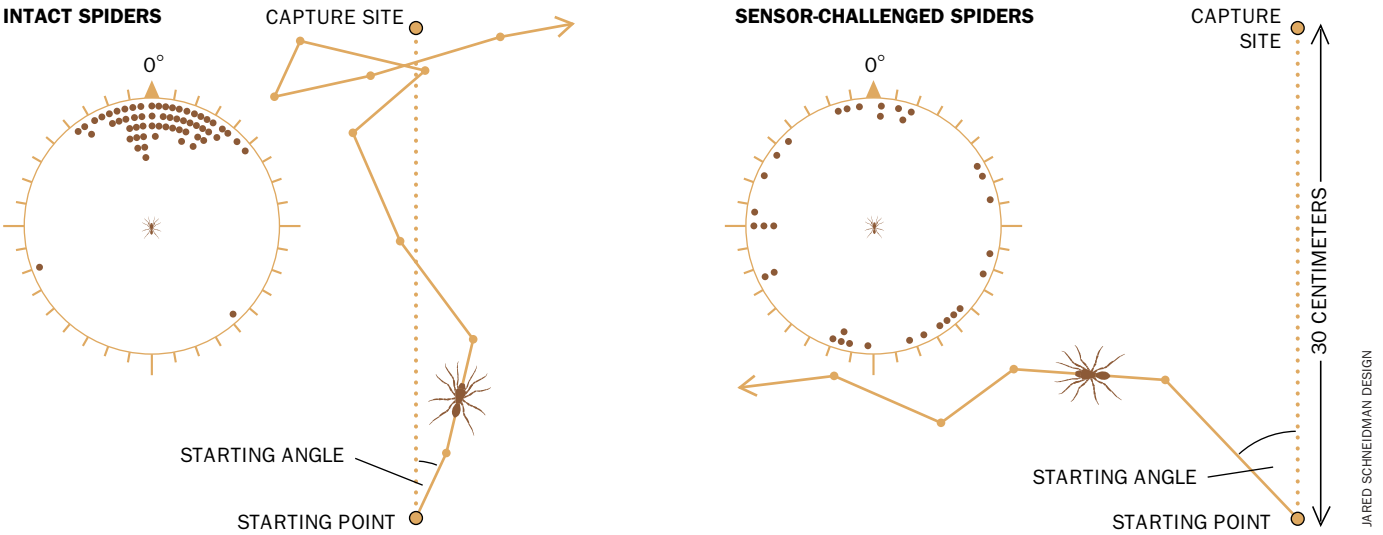
If the tropical wandering spider is placed in a level arena, temporarily blinded and presented with a live housefly, it quickly locates the humming fly by its vibrations. When the spider is then gently separated from its prey and chased some distance away from the capture site, it turns around and returns to the original spot, even though the fly has been removed and placed elsewhere in the arena. If guided away from the capture site along a curving detour path, the spider does not simply retrace its curved path. It actually cuts corners, thus indicating that it remembered the position of the fly.

This type of ability does not depend on external cues from vision or smell or gravity. Instead it relies on strain gauges. If the sense organs near the leg joints are destroyed, the spiders become disoriented. These sensor-challenged spiders start out in the wrong direction and appear unable to correct their erroneous choice along the way, eventually missing their target. The spiders show no other apparent deficits in their ability to catch prey. Hence, when external cues are not available for orientation, these nocturnal animals rely on information



ROBERTO OSTI; JARED SCHNEIDMAN DESIGN (*detail at far left*)

COCKROACH AND HUMAN LEGS rely on similar types of muscles. The hamstring muscles correspond to the insect flexor muscles, and the quadriceps to the extensors. The insect sensors are oriented to detect strain in all directions and to control particular muscles.



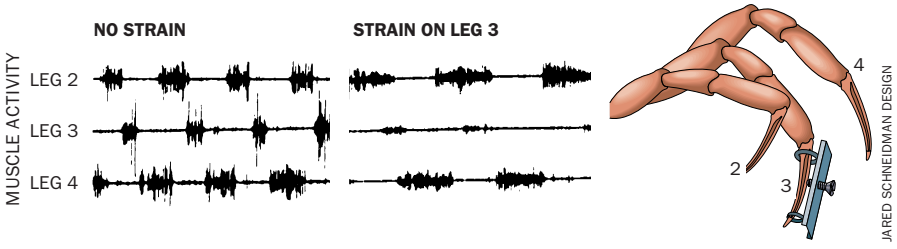
MEMORY OF PREVIOUS STEPS by the wandering spider *Cupiennius* depends on its slit sense organs. When temporarily blinded, most spiders managed to start back in the direction where they had left a previously captured fly (*left*), often heading back in a zigzag fashion. Their starting angles are shown as dots in the circular diagram. But spiders that had damaged slit sense organs went off in random directions (*right*) and often came nowhere close to the capture site.

about their own movements provided by strain receptors.

Several groups of engineers are now trying to use arrays of simple sensors to coax behavior as complex as that of arthropods out of their mechanical constructions. Ken S. Espenschied, working with Roger D. Quinn, Randall D. Beer and their colleagues at Case Western Reserve University, has constructed six-legged robots based on the insect nervous system, in which the control of movement is not localized in just one central processor but instead is distributed in various segments. These robots can walk on a grid of widely spaced slats at a moderately fast pace. Such a design overcomes, at least in part, some of the problems plaguing other kinds of multilegged robots, many of which must laboriously calculate data obtained from many sensors in order to take a step.

A robot as nimble and well balanced as an arthropod would have many uses. It would be ideally suited to move on dangerous and treacherous terrain where wheeled vehicles would not stand a chance. Investigations of toxic-waste sites and of craters on Earth, the moon and Mars are among the most frequently mentioned possibilities. Others are even more complex. For instance, the U.S. Navy is seeking to design crablike robots that could jump out of ships, scramble to shore and hunt for mines.

It will be a while before a robot arthropod will be as quick-footed as the real thing. These invertebrates have evolved sensors that take advantage of the laws of mechanics in ways we are only beginning to understand. SA



COORDINATION OF CRAB LEGS is affected by input from force-sensitive organs, as revealed when a small bending device is placed on the foot of one leg (here, leg number 3). The continuous stimulation of the sense organ not only prevents the depressor muscles in leg 3 from working but also excites the muscles of adjacent legs.

The Authors

SASHA N. ZILL and ERNST-AUGUST SEYFARTH met when Zill went to Frankfurt on an academic exchange program in 1982. Zill received his Ph.D. in anatomy from the University of Colorado at Denver and now teaches at the Marshall University School of Medicine in Huntington, W. Va. For his strain-gauge research, he primarily relies on cockroaches and grasshoppers. Seyfarth has concentrated his research on sensory physiology and the behavior of spiders since his studies at the University of Munich, where he received a Ph.D. in zoology. He currently teaches zoology and neuroscience as a privatdozent at the University of Frankfurt.

Further Reading

THE EXOSKELETON AND INSECT PROPRIOCEPTION. Sasha N. Zill and David T. Moran in *Journal of Experimental Biology,* Vol. 91, pages 1–24 and 57–75; 1981.

NEUROBIOLOGY OF ARACHNIDS. Edited by F. G. Barth. Springer-Verlag, 1985.

FORCE-SENSITIVE MECHANORECEPTORS OF THE DACTYL OF THE CRAB: SINGLE-UNIT RESPONSES DURING WALKING AND EVALUATION OF FUNCTION. F. Libersat, F. Clarac and S. Zill in *Journal of Neurophysiology,* Vol. 57, No. 5, pages 1618–1637; May 1987.

BIOLOGICAL NEURAL NETWORKS IN INVERTEBRATE NEUROETHOLOGY AND ROBOTICS. Edited by R. D. Beer, R. E. Ritzmann and T. McKenna. Academic Press, 1993.

INTRACELLULAR CHARACTERIZATION OF IDENTIFIED SENSORY CELLS IN A NEW SPIDER MECHANORECEPTOR PREPARATION. Ernst-August Seyfarth and Andrew S. French in *Journal of Neurophysiology,* Vol. 71, No. 4, pages 1422–1427; April 1994.

by Shawn Carlson

Covert Observations of Nesting Sparrows

Naturalists are among the most passionate of scientists. Their deepest wish is to get close enough to record a species' most intimate behaviors but not so close as to disturb the animals. This desire for intimacy is indelibly etched into a naturalist's soul and is so strong that many spend years in the wild seeking new scientific understanding of their subjects.

Sadly, few amateurs can devote years of their lives to fieldwork. But that doesn't mean enthusiasts cannot make important contributions. Thousands have already joined the ranks of urban naturalists, studying species that live quite literally in their own backyards. A few have even distinguished themselves by developing new methods to uncover nature's secrets.

To this short list of distinguished amateur naturalists, we can now add an ophthalmologist from Omaha. Robert Slaughter has developed an ingenious method for spying on house sparrows. He combined a birdhouse, a camera and a blind in one unit to get close—within a foot or so—to a brood of nesting birds. The method opens up new vistas for investigating animal behavior.

The idea is simple. Slaughter modified a standard birdhouse, removing the top and extending the sides. He next cut an observation port in his elevated wooden patio deck and mounted the birdhouse underneath the deck so he could peer down into the chamber. (Deckless amateurs take heart—you can adapt this project with a video camera, as I'll explain later.) An old magician's trick enables Slaughter to view the birds surreptitiously. The entire arrangement lets him make detailed observations of nesting birds, day or night.

The magician's trick requires a black veil thin enough to see through easily. It is stretched across the open top of the birdhouse. By lighting the interior of the birdhouse, a naturalist can see the goings-on inside the nest. In Slaughter's device, the fabric is mounted on a sliding frame to allow a fresh section to be moved into place whenever material from exuberant nest building adheres to the fabric and obstructs the view.

The birds will not sense that they are being watched as long as the observer remains in darkness. To accomplish that, Slaughter built a blind (basically, a small tent) that completely enclosed the observation area. The frame, which was made from plastic polyvinyl chloride (PVC) pipe fittings, can house two people. He draped the frame with "blackout" cloth purchased from a neighborhood fabric store and covered the cracks between the floorboards of the deck with opaque tape. You can also place aluminum foil over the cracks and seal it in place with duct tape.

ROBERT SLAUGHTER

HOUSE SPARROW CHICKS (**Passer domesticus**) *were surreptitiously viewed through an ingenious birdhouse designed by Robert Slaughter of Omaha, Neb.*

Although the blind technique works well, you may want to consider less ambitious (and less costly) strategies. A hood similar to those used to cut down glare on oscilloscope displays should do nicely; you will also need some additional black cloth to keep out the light when you are not observing. Alternatively, you can permanently affix a 35-millimeter, single-lens reflex camera just above the hole. Amateur ornithologists may then watch the birds directly through the viewfinder of the camera without having to conceal themselves. Whatever you do, make sure that rainwater cannot leak into the nest and, more important, that a cat or other predator cannot get in. When not observing, Slaughter protected his subjects by turning a plastic washbasin upside down over the observation port and securing it with bungee cord.

The interior lighting is critical. A bulb inside the birdhouse would overheat it, so Slaughter illuminated the nest from the outside. He first painted the nest's interior white to minimize shadows. Then he cut into the back of the box a rectangular slot one and a half inches wide, three and a half inches long and six inches above the bottom of the nest chamber. To diffuse the light and illuminate the nest evenly, Slaughter covered the slot with a piece of opalescent glass. A thin sheet of white plastic would serve the same purpose.

Almost any light source is suitable so long as it is weatherproof. Slaughter used low-voltage yard lights available from local hardware stores. The brightness of the lights can be controlled by wiring in a dimmer switch. To reflect more light into the box, he placed a white board behind the lights.

Slaughter sent me some wonderful photographs taken from his ornithological observatory. He gets good depth of field using a flash and a 100-millimeter macro lens with its f-stop set at 32. The problem was directing enough light from the flash into the birdhouse. Slaughter removed the flash from his camera and positioned it over a hole he cut in the side of the birdhouse. He triggered the flash with a short cable purchased at a camera store. Slaughter reports that, oddly enough, the flash does not seem to bother the birds.

Nifty as this setup is, the project could be developed into an even better scientific tool. Consider, for example, how to automate the observing process, quantify the activity in the nest, record the weights of the birds throughout the nest-

ing season or monitor the temperature of the eggs during incubation. Any of these additional endeavors would make a great amateur project.

To get you started, let me suggest a way to record your observations of the nest continuously. Try replacing the still-photography camera with a small video camera. It can continuously record the movements in the birdhouse, and the signal can be piped directly into your television or videocassette recorder. That way, you can observe your subjects comfortably in your own home. It's also a dandy solution for amateurs who do not have a wooden patio deck: you can mount the birdhouse wherever your subject roosts. Also, many camcorders can detect near-infrared light. A bright infrared-light-emitting diode may be able to illuminate the inside of your birdhouse and allow clear videos to be shot at night without disturbing the birds.

It may be possible to quantify the activity inside the nest using an ultrasonic motion sensor. Such devices can be cannibalized from old burglar alarms. Position the sensor next to the camera and point it down toward the birds. If they do not flee when the sensor is turned on, they probably can't hear the ultrasound.

The fraction of time the sensor spends in its triggered mode is a measure of the birds' activity. You'll need to experiment with both the intensity and sensitivity levels of the sensor. You will also need a chart recorder or a computer to monitor the signal. Data recorded 24 hours a day throughout the nesting season could reveal fascinating tidbits about avian behavior. The sensor can also alert you when the birds enter or leave the nest or set off your photographic equipment.

Keep in mind that improving these techniques could allow you to answer several open scientific mysteries. For instance, a house sparrow typically lays a clutch of four eggs, one each day for four days, and so the first egg is usually three days older than the last. Yet all the eggs hatch within about 24 hours of one another. Perhaps the nesting bird positions its body to warm new eggs more efficiently than the older ones, so the warmer eggs develop faster. Measuring the eggs' temperatures when the birds leave the nest may answer this question. Let me know what you discover with this or any other variation.

For more information about this project, including suggestions for fascinating ornithology experiments, send $5 to the Society for Amateur Scientists, 4951 D Clairemont Square, Suite 179, San Diego, CA 92117 or download the information for free from the SAS Web site at http://www.thesphere.com/SAS/ You can also post messages about this or any project on the SAS Web site.

BRYAN CHRISTIE (*drawing*); ROBERT SLAUGHTER (*photographs*)

BIRDHOUSE FOR SECRET VIEWING *consists of a see-through black veil mounted on a removable top panel. Interior lighting comes from a bulb that shines through a slot in the back of the birdhouse* (not shown). *For photographs, a camera is placed over the observation port.*

by Ian Stewart

Arithmetic and Old Lace

Who is a mathematician? Some years ago, in a rare flash of insight, it dawned on me that a mathematician is somebody who sees an opportunity for doing mathematics where others might not. Consider shoelaces. The potential for extracting significant mathematics from shoelaces is not widely recognized. That it exists was made clear to me by the article "The Shoelace Problem," by John H. Halton, who is in the computer science department of the University of North Carolina at Chapel Hill. It appeared in the Fall 1995 issue of the *Mathematical Intelligencer.*

There are at least three common ways to lace shoes [*see illustration below*]: American zigzag, European straight (from which the term "straitlaced" is derived, though perhaps by way of garments rather than shoes) and quick-action shoe store. To the purchaser, styles of lacing can differ in their aesthetic appeal and in the time required to tie them. To the shoe manufacturer, a more pertinent concern is which type of lacing requires the shortest—and therefore cheapest—laces. In this month's column I shall ask the shoe manufacturer's question.

To find the amount of lace, I will focus only on the length represented by straight line segments. The amount of extra lace required to tie an effective bow is the same for all methods of lacing, so it can be ignored.

My terminology will refer to the lacing as seen by the wearer, so that the "top" row of eyelets in the figure lies near the ankle. I will also idealize the lace to be a mathematical line of zero thickness and the eyelets to be points. Using a brute-force attack, one can then calculate the length of the lace in terms of three parameters:

- The number n of pairs of eyelets
- The distance d between successive eyelets
- The gap g between corresponding left and right eyelets

With the aid of Pythagoras's theorem (one wonders what the great man would have made of this particular application), it is not too hard to show that the lengths for the lacings are as follows:

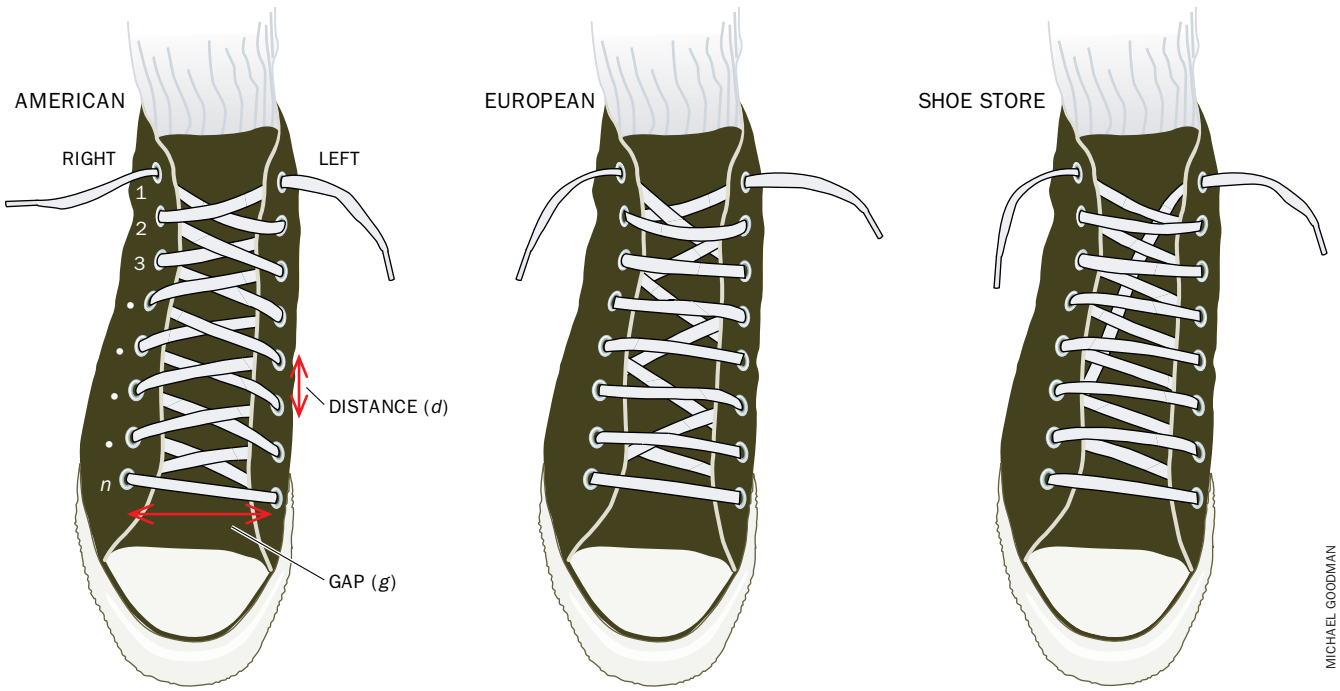
- American: $g + 2(n-1)\sqrt{d^2+g^2}$
- European: $(n-1)g + 2\sqrt{d^2+g^2} + (n-2)\sqrt{4d^2+g^2}$
- Shoe store: $(n-1)g + (n-1)\times \sqrt{d^2+g^2} + \sqrt{(n-1)^2d^2+g^2}$

Which is the smallest? Suppose, for the sake of argument, that $n = 8$, as in the figure, $d = 1$ and $g = 2$. Then the lengths are:

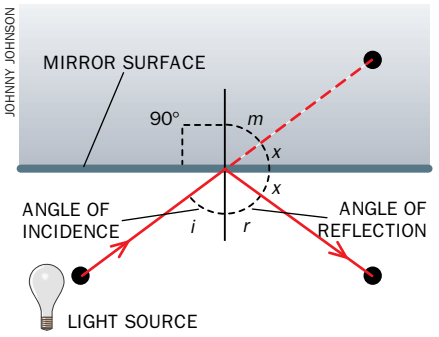
- American: $2 + 14\sqrt{5} = 33.305$
- European: $14 + 2\sqrt{5} + 6\sqrt{8} = 35.443$
- Shoe store: $14 + 7\sqrt{5} + \sqrt{53} = 36.933$

But can we be certain that American is always the shortest? Some careful high school algebra shows that if d and g are nonzero and n is at least 4, then the shortest lacing is always American, followed by European, followed by shoe store. If $n = 3$, American is still shortest, but European and shoe-store lacings are of equal length. If $n = 2$, then all three lacings are equally long, but only a mathematician would worry about such cases!

Still, this algebraic approach is com-



LACING PATTERNS on a sneaker can have different complexities and lengths. Which of these patterns requires the least lace?



FERMAT'S PRINCIPLE
for deriving the path of a light ray reveals that the angle of reflection equals the angle of incidence.

plicated and offers little insight into what makes different lacings more or less efficient. Halton instead observes that a clever geometric trick makes it completely obvious that American lacing is the shortest of the three. The idea owes its inspiration to optics, the study of paths traced by rays of light.

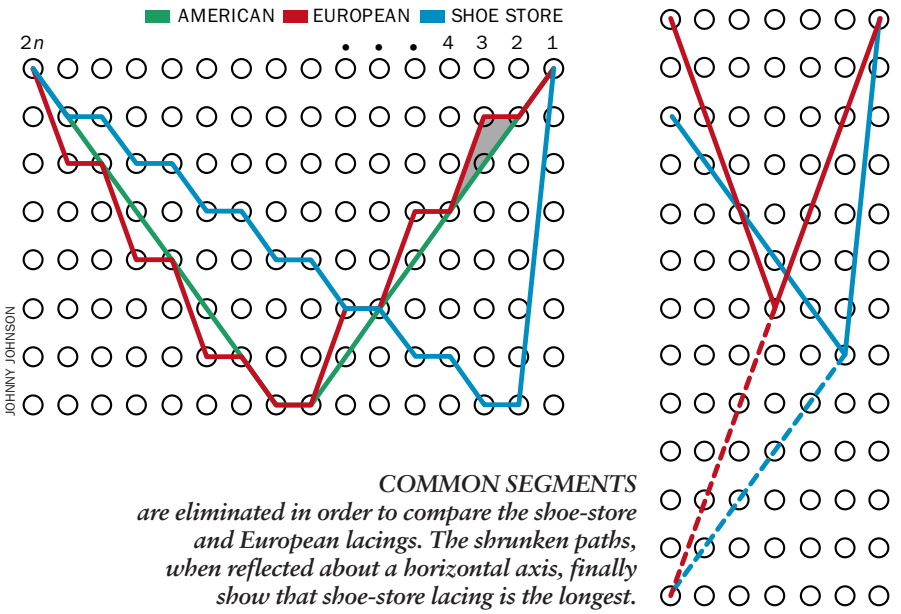
Mathematicians discovered long ago that the geometry of light rays can be made more transparent (so to speak) by applying carefully chosen reflections to straighten out a bent light path. For example, to derive the classical law of reflection—"angle of incidence equals angle of reflection"—consider a light ray that hits a mirror and bounces off. If you reflect the second half of the path about the plane of the mirror [*see illustration at top left*], the result is a path that passes through the mirror and enters Alice's world behind the looking glass.

According to the principle of least time, a general property of light rays enunciated by Pierre de Fermat, such a path must reach its destination in the shortest time—which in this case implies that it is a straight line. Draw a line perpendicular to the mirror at the point of incidence. Then the angle marked m in the figure is equal to the angle of incidence i; further, the two angles marked x are equal. But $m + x = 90°$, and if r is the angle of reflection, $r + x = 90°$ as well. So $m = r = i$.

Halton derives geometric representations of all three types of lacing by an extension of this reflection trick. He draws a diagram [*see illustration at far left*] made of $2n$ columns of eyelets spaced distance d apart in the vertical direction. Successive rows are spaced distance g apart horizontally. (To shrink the diagram, we have reduced g; the method works for any values of d and g.) The last column of the diagram represents the left-hand column of eyelets; the second-last column represents the right-hand column of eyelets. Overall the odd-numbered columns represent left-hand eyelets, and even-numbered columns represent right-hand eyelets.

The polygonal paths that zigzag across this diagram correspond to the lacings, but with an extra "twist." Start at the top left-hand eyelet of a lacing pattern and draw the first segment of lace, running from left to right of the shoe, be-

REFLECTION TRICK
allows the shortest lacing to be derived geometrically. Instead of zigzagging, the lacing path is reflected at each eyelet, so that it is straightened out as much as possible. American lacing becomes the straightest, because it runs along one side, rather than two, of each small triangle (gray). *It is therefore the shortest.*



COMMON SEGMENTS
are eliminated in order to compare the shoe-store and European lacings. The shrunken paths, when reflected about a horizontal axis, finally show that shoe-store lacing is the longest.

FEEDBACK

Tom Sales of Somerset, N.J., sent me a fascinating letter inspired by the February column on zero-knowledge protocols—means of proving that you know the rules without revealing them. Years ago in this journal, Martin Gardner introduced a card game called Eleusis, in which one player invented rules and the others had to guess them by being told whether a given play was legal or illegal. At that time, Sales invented a similar game, involving a mouse, Alpha, that inhabits a triangular room.

In each of the three corners is an array of colored lightbulbs. Alpha is frightened by the lights and scurries from corner to corner according to rules such as "If the light in my corner is red and the light clockwise from me is green, then I will run to the clockwise corner." One player sets up the rules, secretly, and the other(s) try to deduce them by setting combinations of lights and watching where the mouse moves. Note that the rules depend only on the state of the lights relative to Alpha's current position.

Now eliminate the mouse! If you can't see Alpha, then there is no way to deduce the rules. But at any random instant the mouse can be rendered visible, so that an observer can check that the rules are indeed being followed. Mouse movements thus form the basis of a zero-knowledge protocol. Now let Alpha's movements represent a message, so that the rules for moving the mouse act as an enciphering algorithm, and you have a very interesting system, with a zero-knowledge flavor, for transmitting code messages.

And it's a fun game, too. —*I.S.*

tween columns 1 and 2 of the diagram. Draw the next segment of lace between columns 2 and 3 instead of going back from column 2 to column 1 as for a real shoe. In effect, the segment is reflected as though the columns of eyelets were replaced by mirrors. Continue in this manner, reflecting the position of each successive segment whenever it encounters an eyelet. Instead of zigzagging between the two columns of eyelets, the path now moves steadily to the left of the figure.

Because reflection of a segment does not alter its length, this representation leads to a path that has exactly the same length as the corresponding lacing pattern. The added advantage, however, is that it is now easy to compare the American and European patterns. In a few places they coincide, but everywhere else the American pattern runs along one edge of a thin triangle, whereas the European one runs along the other two edges. Because any two sides of a triangle exceed the third side in length (that is, a straight line is the shortest path between two given points), the American lacing is clearly shorter.

It is not quite so obvious that the shoe-store lacing is longer than the European. The simplest way to see this is to eliminate from both paths all horizontal segments (each path has $n - 1$ horizontal segments, which contribute the same amount to both lengths) and also the two sloping segments that match up. The elimination results in two V-shaped paths; shift the truncated European path so that it again shares the first eyelet with the shoe-store path. If each V is now straightened out by reflection about a horizontal axis placed at the tip, it finally becomes easy to see that the shoe-store path is longer, again because two sides of a triangle exceed the third side.

These cunning reflection tricks can do more than compare particular lacing patterns. Halton uses them to demonstrate that the American zigzag lacing is the shortest among all possible lacings. More generally, shoelaces and Fermat-style optics become united in the mathematical theory of geodesics, the shortest paths in various geometries. There, mirror reflections answer fundamental questions in physics, as well as confirming the superiority of the American way of lacing shoes. SA

PITT RIVERS MUSEUM, UNIVERSITY OF OXFORD

AN ANTHROPOLOGIST'S ATTIC

Review by Meredith F. Small

Pitt Rivers Museum, University of Oxford
South Parks Road, Oxford, England OX1 3PP
Telephone: 011-44-1865-270927

Anthropology is the granddaddy of multiculturalism. Long before anyone even considered putting the word "correct" next to "politically," stalwart ethnographers were taking notes by firelight in isolated camps, documenting the fact that there are many ways to create a life. And it is the beads, baskets and spears hauled back from points previously unknown that first demonstrated how every group is sophisticated in its own right, every society deserves reverence and every culture is in its own way "correct."

I have found that it is all too easy to lose touch with such insights. For me, over the years, the feeling of cultural magic that usually clings to anthropology had been buried under the daily grind of academic life. But last winter a most unusual place reminded me why I love my work. I was spending a few days in Oxford, visiting a former teaching assistant, Nick Fowler. This trip was for pleasure, not business, but Nick thought that nonetheless we should honor our anthropological roots by touring the Pitt Rivers Museum, the oldest anthropology museum in the world. "Oh great," I muttered. "Just what I need, a dusty, badly lit collection of bowls, baskets and spears." On the other hand, it might be amusing. We could laugh at what the world looked like before deconstruction and semiotics became as important as vaccinations and see displays in which certain ethnic groups were still called "Eskimos" and "American Indians." We decided to go primarily because of the museum's kitsch appeal.

From its inception, the Pitt Rivers was never intended to be a standard museum. General Augustus H. Pitt-Rivers (who lived from 1827 to 1900) was originally an avid collector of firearms and other weaponry but on reading Charles Darwin's *On the Origin of Species* began collecting other artifacts instead. Pitt-Rivers believed that material culture could be arranged on a progressive evolutionary scale, much as his contemporaries were organizing animal taxonomy. He wanted to demonstrate, through objects, that there existed a relation between the simple technology of aboriginal peoples and the more "advanced" artifacts of our more recent past. And he felt the keys to understanding this historical continuum were the everyday objects used by regular folks. After retiring from the military in 1882, Pitt-Rivers focused full-time on archaeological pursuits, overseeing meticulous excavations of prehistoric, as well as Roman and Saxon sites on his Wiltshire estate.

Pitt-Rivers's ideas fit nicely with the times; in the mid-1800s British naturalists and explorers were traveling the globe, collecting all sorts of plants, animals and human doodads. Through a network of friends and contacts, Pitt-Rivers was able to amass a huge personal collection of artifacts, and he soon began lobbying the government for a national anthropological collection. In 1874 he obtained exhibit space at what is now the Victoria and Albert Museum in London. The collection continued to grow and found a permanent home at Oxford in 1884.

So, on a rainy weekday afternoon, Nick and I entered the Oxford University Museum of Natural History on Parks Road. Dinosaur fossils collected in Darwin's day sat forlornly, illuminated by the meager sun filtering through a lattice of Victorian iron grillwork and antique glass panes. This room recalled an era of men in pith helmets on expeditions into remote places where tea was served at 4 P.M., no matter what. I assumed the Pitt Rivers would be just as nostalgic.

We passed into the forecourt of the Pitt Rivers, which housed a temporary display of textiles, and headed around the corner into the main museum. And then we stopped. The glass case in front of us mirrored our matching expressions: we were dumbstruck.

Ahead stretched a central court, three stories high, and each floor was stacked to the rafters with more junk than I had ever seen in my life. We had walked into a wonderland of stuff, a jumble of human stuff, stuff so strange and fasci-

nating that I wanted to shout, "Look at this!" and run from case to case like a three-year-old in a toy store. There were shrunken heads here, boats there, magic beads over there, cases of cloth and costumes, drawers of shells, samples of yarn, hundreds of musical instruments. At every turn we saw piles of junk thrown together like the world's richest yard sale. I felt a familiar joy rise in my soul; I felt at home.

This bounty, I realized, is what anthropology is all about. Not dissertations, not books on cultural theory, not lectures and slides, but *stuff*—human culture at its most expressive. The Pitt Rivers, in all its junky glory, captures the basic impulse of our species to create—and to solve a problem by inventing a tool to fix it. Nick and I did not just want to stand there and ponder the meaning of all this *stuff*—we wanted to play with it.

The Pitt Rivers is unlike any other museum for one basic reason: it is a mess. The museum does not try to impart a full understanding of another way of life. Instead the artifacts are grouped by mind-set. A case labeled "Measures of Time" shows how people from any number of cultures solved the problem of marking moments; the boats from every sea, river and ocean stacked together demonstrate how humans from all places and all generations figured out how to get from place A to place B when faced with large intervening bodies of water; endless drawers of amulets and potions bring home the ubiquitous fear of the unknown.

People, the Pitt Rivers reveals, have common problems but find local solutions. Visual excess turns out to be a brilliant curatorial technique for emphasizing the universality of problem solving and creativity. A paean to global humanity sings from every case.

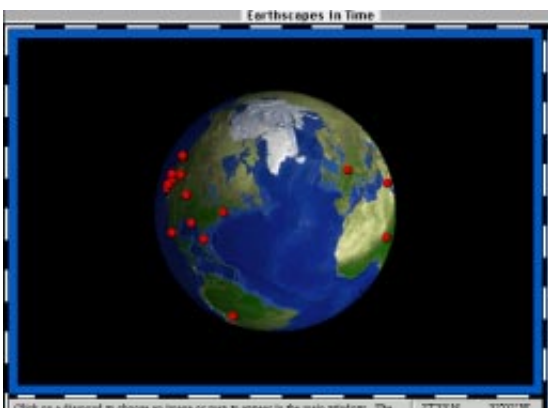
The museum owes much of its style to Henry Balfour, a zoologist by training who was its first curator. Balfour traveled extensively with his wife, sometimes commissioning pieces for the mu-

THE CD EXAMINED

Seeing Earth in a New Way: The SIR-C Missions of Endeavour
National Aeronautics and Space Administration/Jet Propulsion Laboratory, 1995 (CD-ROM for Macintosh, gratis; ordering information is available by calling (605) 594-6116 or by visiting the World Wide Web site at http://southport.jpl.nasa.gov/education.html)

Earthscapes in Time: The See-Through Satellite Atlas
Now What Software, 1995 (Hybrid CD-ROM for Windows or Macintosh, $49.95)

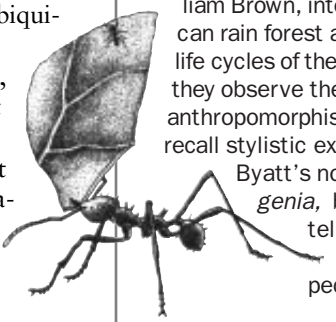
The NASA CD-ROM presents a no-frills introduction to the stunning Earth images produced by the Spaceborne Imaging Radar, which flew onboard the space shuttle in 1994. An included copy of Netscape's Web browser serves both as the CD-ROM interface and as a gateway to extensive on-line resources. Although aimed at students, the disc has much to offer anyone interested in earth science. *Earthscapes in Time* (*below*) contains an attractively packaged collection of satellite images that highlight natural and human-generated environmental changes. Sequences showing the shrinking of the Aral Sea and deforestation of the Amazon grimly illustrate our power to despoil our planet; linked (nonsearchable) articles provide helpful context.



—Corey S. Powell

BRIEFLY NOTED

THE EARTH DWELLERS: ADVENTURES IN THE LAND OF ANTS,
by Erich Hoyt. Simon & Schuster, 1996 ($24).
For as long as there have been myrmecologists, ant colonies have served as metaphors for human society. The author follows two giants of the field, Edmund O. Wilson and William Brown, into the Costa Rican rain forest and describes the life cycles of the superorganisms they observe there. Erich Hoyt's anthropomorphisms sometimes recall stylistic excesses of A. S. Byatt's novella *Morpho Eugenia,* but the tales he tells are interesting and full of unexpected detail.



THE CIGARETTE PAPERS,
by Stanton A. Glantz, John Slade, Lisa A. Bero, Peter Hanauer and Deborah E. Barnes. University of California Press, 1996 ($29.95).
Based on several thousand pages of legal documents that mysteriously found their way from a tobacco-company law firm to the public domain (the originals are on-line at http://www.library.ucsf.edu/tobacco), *The Cigarette Papers* chronicles the response of BAT Industries and its American subsidiary, Brown and Williamson, to the growing evidence that their product kills. The authors relate that efforts in the 1950s and 1960s to develop less harmful cigarettes gave way to a multimillion-dollar campaign of obfuscation driven by fear of litigation.

THE WING OF MADNESS: THE LIFE AND WORK OF R. D. LAING,
by Daniel Burston. Harvard University Press, 1996 ($35).
R. D. Laing was for a time among the most influential of psychologists, but he is now almost forgotten. His work with schizophrenics helped form the notion of a "therapeutic community" to supersede the near barbaric conditions prevalent in mental institutions during the 1960s and 1970s. Yet Laing's own life was marked by alcoholism, erratic behavior and personal strife. This tentative biography offers at least a measure of insight into Laing's often contradictory life and work.

Continued on page 102

seum from artisans along the way. More important, Balfour (like Pitt-Rivers) insisted on extensive documentation; each object in the museum is painstakingly labeled—some have yellowed tags handwritten by Balfour himself.

The Pitt Rivers Museum has a long history of influencing both budding anthropologists and the direction of British anthropology. It was at Pitt-Rivers's insistence in 1883 that Sir Edward B. Tylor was appointed to the university's faculty in the first full-time position in anthropology in Britain. And Beatrice Blackwood, a museum employee, was one of the first female anthropologists to travel to remote places in the 1930s, living among headhunters in New Guinea and collecting artifacts for the museum.

Today anthropologists returning from the field continue to donate objects. These newer artifacts—the African toy trucks made from flattened oilcans or the miniature clay cars made by native South Americans and sold to tourists—demonstrate ongoing changes in material culture. And I discovered that the Pitt Rivers remains as vital an academic resource today as it was at the turn of the century. It is closed in the mornings so that teaching and research can go on without interruption; during our visit, a class of young artists was busy drawing masks from the collections.

The museum is also completely unpretentious. Yes, the 40-foot Haida totem pole with its carved bear, raven and frog faces is imposing, and it certainly merits the status of "art." But just as impressive in its own way is the case packed full of hair combs: combs from Africa and Asia, combs made of shell and of bone. If these objects were spotlighted and displayed on velvet, they, too, could surely be called art. Instead they are presented as practical devices, things that humans make to accomplish a task.

By casually stuffing the combs together in one glass case, the museum seems to say that we are a species that just cannot let well enough alone. We don't just make a tool and be done with it. No, we have to shrink it, elongate it, paint it or carve it. We have to add beads, give it spiritual significance and pass it down through generations. In a sense, human

THE ILLUSTRATED PAGE

"Americans Waiting to Enter the Olmec World" (New York City, 1989–93)

Truths & Fictions
BY PEDRO MEYER
Aperture, 1995 ($40)

Digital photography is a doubly difficult art; one must possess a talent not only with the camera but also with the computer. In this collection, Pedro Meyer demonstrates considerable aptitude on both scores. He manipulates pixels and juxtaposes images to create absurdist situations, snippets of social commentary or startling mystical visions. On one page, Mexican migrant workers labor beneath a gaudy casino billboard; on another, weary museum goers remain oblivious to the—literally—uplifting experience taking place beneath their feet (*above*). These images are fake (in the sense that they do not represent the literal world) and yet simultaneously real (in the sense that, like all photographs, they capture the essence of the photographer's vision). The book's final section unveils some of Meyer's compositional techniques—an interesting lesson in imaging technology and an apt punch line to his visual jokes. *—Corey S. Powell*

artfulness is documented more effectively in this cluttered sanctuary than in museums, where every painting cries for reverence. Here the creativity that underlies art appears plainly as part of human nature rather than as a diversion for an elite few.

The collection is also a testimony to common concerns that cut across ethnic lines. Technology often passes from one group to another in a process that anthropologists call cultural diffusion. But when faced with mummies from ancient Egypt and mummies from rather recent Peru, the viewer is struck by the way that people in two distinct cultures and times, with absolutely no contact with one another, were compelled (presumably for similar emotional reasons) to preserve the things they loved.

As the cases and drawers of the Pitt Rivers Museum suggest, humans all over the world, and throughout history, have been motivated independently to make music, smoke weeds, defile their enemies, keep away evil and gamble, among other activities. The Pitt Rivers confirms the oneness of our species while presenting the myriad ways in which different cultures solve the same problems. When we left the museum several hours later and headed for the Kings Arms for a pint of beer, Nick and I crowded close under an umbrella, chatting about why we had become anthropologists in the first place. As I told Nick, who couldn't love these strange creatures who make rain gear from walrus intestines and shrink the heads of their enemies?

I often think about all that clutter in the Pitt Rivers Museum, especially when I am in a shopping mall staring at all the new stuff my own culture is trying to make me buy. I wonder about those shrunken heads, I recall the various hair combs and I ponder the diverse measures of time. And I silently thank those fellow anthropologists who still go far away and come back loaded down with the detritus of other cultures. They are the guardians of the true purpose of our field—to make multiculturalism come alive.

MEREDITH F. SMALL is an associate professor of anthropology at Cornell University.

A FORCEFUL LIFE

Review by Istvan Hargittai

Force of Nature: The Life of Linus Pauling
BY THOMAS HAGER
Simon & Schuster, 1995 ($35)

Thomas Hager first met Linus C. Pauling in 1984 at a presentation on vitamin C by the then 83-year-old scientist. Both men arrived early and found themselves alone in the seminar room. Pauling introduced himself and, wasting no time, proceeded to deliver an enthusiastic "minilecture on the chemical binding properties of tin."

Hager's description of this encounter reminded me of my own meeting with Pauling, only a couple of years before Hager's, at the University of Oslo. Pauling lectured a packed auditorium about structural chemistry. He was deriving complicated expressions without using so much as a scrap of paper, marching back and forth in front of the long blackboard, which he covered with formulas. He kept his enthusiastic Norwegian

BRIEFLY NOTED

Continued from page 99

THE SIMPLE SCIENCE OF FLIGHT: FROM INSECTS TO JUMBO JETS, *by Henk Tennekes. MIT Press, 1996 ($20).*

Frequent flyers will find a good companion in this paean to the creatures and devices that take to the air. The book begins with casual ruminations on the energy consumed in airplane travel and culminates in a discussion of the Boeing 747, "one of the great engineering wonders of the world." Along the way, Henk Tennekes painlessly carries the reader through some of the fundamental ideas and equations of aerodynamics. He makes persuasive connections between science and everyday life—explaining, for instance, why a 737 has stubby wings or how birds adapt to a wide range of airspeeds.

AIR & SPACE SMITHSONIAN DREAMS OF FLIGHT. *Creative Multimedia, 1995 (CD-ROM for Windows, $29.99).*

This ambitious disc includes a multimedia encyclopedia of aviation, a collection of aircraft clip art and a flight-simulator program. The result is decidedly mixed. The interface is attractive but cumbersome, the encyclopedia entries interesting but often simplistic. Enthusiasts may enjoy the numerous images and the charming—albeit difficult to install—software that permits stunt flying on the PC. Those seeking a definitive reference, however, should look elsewhere.

THE ENCHANTED WORLD OF SLEEP, *by Peretz Lavie. Translated by Anthony Berris. Yale University Press, 1996 ($27.50).*

We spend approximately a third of our lives asleep; Peretz Lavie explores what occurs during those lost hours. The book rocks intriguingly between well-told clinical anecdotes and scientific analysis: the dreaming behavior of Holocaust survivors, the biochemical function of melatonin, and how to handle children who refuse to go to sleep. If only Lavie delved more boldly into the most basic question—*Why* do we sleep?

audience in awe and only gradually did it dawn on me that the sophisticated derivations were superfluous to an understanding of the subject matter. During the luncheon after the talk, he stayed fresher and more alert than any of us.

Hager faced a daunting task in trying to document the energy and diversity that marked Pauling's long scientific career (spanning almost 70 years until his death in 1994). To the public, Pauling is probably best known for his championing of the health benefits of vitamin C. In scientific circles, however, he is most renowned as the principal architect of structural chemistry, the fundamental science of the spatial arrangements of atoms in molecules and crystals and the interactions that bond substances.

This work, crowned by his book *The Nature of the Chemical Bond* (first published in 1939), earned him the 1954 Nobel Prize for Chemistry. Although Pauling possessed only a tiny fraction of what we know today about structural chemistry, his observations have withstood the test of time. His achievements have also demonstrated that a method of collecting information and arriving at a discovery may have as lasting an impact on the development of science as a discovery itself.

Hager, a science journalist, describes Pauling's science well—not a simple task, considering its breadth. He explains complicated concepts easily yet correctly and fixes ideas in the reader's mind with succinct descriptions. For example, the ability of hydrogen to bond simultaneously to two atoms instead of the usual one—a structural feature of vital importance to chemistry—becomes "hydrogen bigamy." Just as Hager uses human terms in explaining science, he also speaks about the "chemical bond" between Pauling and a fellow scientist.

The author has exercised much restraint in condensing Pauling's exceptionally productive and inspiring life into a manageable book. Those aspects of the scientist's work that are left out of *Force of Nature* might suffice as life achievements for lesser researchers. For example, Pauling and his graduate student assistant Lawrence O. Brockway (whose name is misspelled in the book) used gas-phase electron diffraction to determine the structure of volatile molecules during the early 1930s; they introduced a technique called Fourier transformation with which the distances between atoms can be determined directly. This approach is now applied daily in electron diffraction laboratories.

Hager also omits Pauling's role in the development of Corey-Pauling-Koltun (CPK) space-filling models, Tinkertoy-like objects whose relative sizes and connection points are based on those of actual atoms and molecules. They facilitate hands-on testing of proposed molecular structures and are still in widespread use. They were instrumental in launching the theory of host-guest chemistry, which earned Donald J. Cram a Nobel Prize in 1987, and have aided in many a chemist's education.

Force of Nature presents Pauling not only as a great scientist but also as an exceptional human being. Hager reaches back to Pauling's ancestors from Germany and Ireland, tracing his childhood in Oregon and his youthful travels to his happy and productive decades at the California Institute of Technology. (Hager also tells the love story of Pauling and Ava Helen Miller. We learn about their dogged pursuit of their goals; a pursuit they sometimes engaged in at the expense of their four children.)

Pauling played a major role in making Caltech a world center of scientific research. Yet the institute attempted to ease him out in the 1950s, when his leftist political activism began to embarrass its mostly conservative administration. His resistance to the actions of the House Un-American Activities Committee, the Federal Bureau of Investigation and the Passport Office of the State Department showed him to be a true champion of the spirit of American independent thinking. Even so, Pauling eventually felt compelled to choose be-

CALIFORNIA INSTITUTE OF TECHNOLOGY

Linus Pauling in the early 1960s

tween conspicuous political resistance and research opportunities, so he scaled down his political activism. There are lessons in this chapter of Pauling's story that remain important today.

Hager's book gives great emphasis to the other arena of Pauling's political work: his fight against nuclear weapons testing, for which he received the 1962 Nobel Peace Prize. In these activities he appears to have been somewhat one-sided, trying to pressure the U.S. (and Great Britain) more than the Soviet Union. Pauling explained that it was more natural for him to criticize his own country's government than that of the U.S.S.R.; nevertheless, he apparently was fooled by Soviet propaganda and did not see the Soviet Union for what it was.

It is ironic, then, that just as Pauling was facing political problems at home, he was declared a public enemy by the Soviet chemistry establishment. Some mediocre but influential professors considered his resonance theory to be ideological heresy and managed to terrorize the entire Soviet chemistry community into reviling it. Pauling thought Soviet chemists merely needed more time to appreciate his theory. In fact, generations of talented young Russians considered theoretical chemistry hazardous and continued to shy away from it long after the resonance theory had become a nonissue.

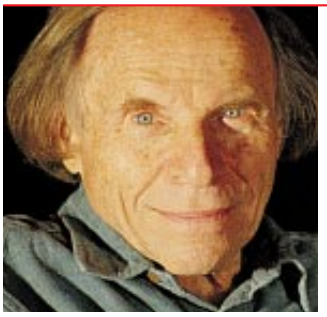
At times, Hager appears to succumb to the temptation to make his subject larger than life. Something of this bias may be seen when Hager details Pauling's unsuccessful attempts to help the son of a German crystallographer escape Nazi Germany. The relatively large weight given to this one episode—and the lack of similar ones—suggests that Pauling's aid to victims of German National Socialism was limited. This early stance is in pointed contrast to his work on behalf of Japanese-Americans interned in the U.S. during World War II and his later dedication to other causes involving the persecuted and oppressed.

Hager does not flinch, however, from recounting some of Pauling's personal and professional relationships that became very close, only to break apart, sometimes ending in lawsuits. Pauling's attitude toward the mathematician Dorothy Wrinch and her original (albeit probably erroneous) protein model, for instance, appears anything but magnanimous.

Another such story, which Hager does not mention, involves Pauling's unbending hostility toward quasicrystals after their discovery in 1984 by Dan Shechtman. Quasicrystals are regular but nonperiodic structures that scientists once considered to be a physical impossibility; the evidence for their existence necessitated a change in the very definition of what a crystal is. Pauling never did believe in quasicrystals, and his immense influence may have hindered the broadening of crystallographic concepts.

Despite the 627 pages of text and more than 50 pages of notes, it is inevitable that *Force of Nature* omits chunks of Pauling's life. Nevertheless, Pauling comes alive on the page—forceful, creative and unyielding. Hager has produced a book worthy of its subject.

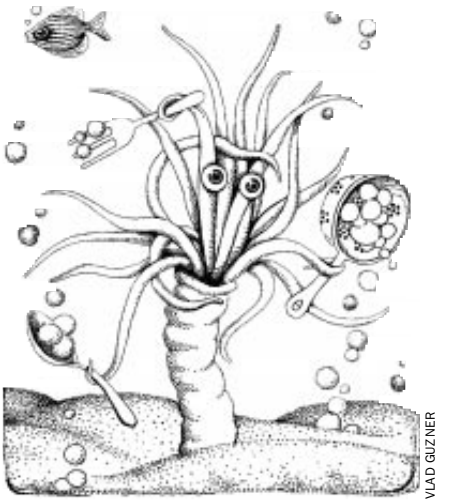
ISTVAN HARGITTAI is professor of chemistry at Budapest Technical University.



WONDERS

by Philip Morrison

Gutless



VLAD GUZNER

Decades ago a brilliant physiologist showed a crowd of abiological number crunchers like myself how to model a generalized animal. Easy: suppress all finer points—limbs, wings, fur, fins—and consider a simplified, living cylinder, pierced rather like a bead but doubly, by two passages, one for the breath of life, the other for food and water.

The reason lies deep in chemistry: the cells of any familiar animal require a regular double uptake. Each day we nourish ourselves on close to two pounds of airy oxygen and another pound, by dry weight, of denser foodstuffs. We routinely breathe through one compound passage, while we eat and drink through quite another. Almost every second, we draw a new breath, although the other foods, no less essential, can enter in timed batches. Even aquatic vertebrates use two passageways: Who has not watched tank fish steadily and visibly gulp watery "breaths"? Their gills extract from the flow the dissolved oxygen the fish must have, but fish digest what they eat via quite another elaborate tube, which empties into their wake.

Yet there are certain wondrous exceptions. Twenty years ago the little submersible *Alvin* first found them about two miles down in the darkness of the Galápagos Spreading Center, where new ocean floor issues as quick-hardening lava, and the seawater leaks into a field of volcanic fissures. Once the water reaches the molten magma below, it is superheated and soon rises out of many ridge vents, which nurture vigorous oases of remarkable animal life. By now such hydrothermal vents are known to exist on the seafloor from the Sea of Japan to the Mid-Atlantic Ridge.

Several hundred new species have been found thriving at these undersea sites. Among the most conspicuous are the giant red-tipped tube worms. These creatures are leathery white cylinders, longer than a tall pitcher's arm. They stand upright in showy clusters, each tube topped by its soft, crimson gill plume: quite literally, these are "gutless wonders." They have no guts at all: no jaws, no mouthparts or teeth, no stomach, no intestinal tract, no anus.

Tube worms breathe water but eat no organic foodstuffs in all the decades of their fast-growing, if sedentary, adult lives. All the energy feedstocks they need flow in as seawater streams, some magma-heated, some cold as iced beer. The seawater provides life-supporting gases in solution: hydrogen sulfide, carbon dioxide and the same dissolved oxygen that fish enjoy.

Is there any life now on Earth that owes no obligation to sunlight at all?

Vent life cannot engage in photosynthesis, for no sunlight at all reaches those murky depths. Instead a large, odd-looking blood-rich organ occupies most of the lower part of each tube worm. Those tissues are crammed with microscopic bacteria clustered among the tube worm's own cells, all embedded in innumerable flecks of yellow sulfur, the solid wastes of the organism's lifetime. (The aerobic zoologist wielding the scalpel will smell brimstone.)

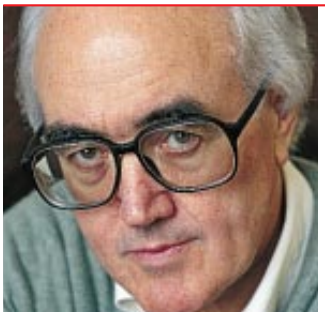
Whereas green plants use photons from the generous sun, this symbiotic bacterial plantation exploits chemically stored energy, the product of inorganic reaction between seawater gases and plentiful sulfates in the very hot rock below the vent. The product is the dissolved H_2S, a hydrogenous gas toxic to us. That fuel is gathered freely by the tube worm from the hot, outpouring water. The gas circulates past the resident microbes, which split and oxidize it, using the oxygen in cold seawater. Coupling to that energy source, the microorganisms can spin as well as any leaf can the molecule glucose, simplest of sugars, out of water itself and dissolved CO_2. Thus do the bacterial partners nourish their big, sheltering host lifelong. The solid sulfur so split off remains behind as waste in the tube worm's body.

That is most unlike the recycling of the analogous waste, oxygen, from sun-split H_2O. In our sunnier world, oxygen replaces sulfur as the waste atom. Thus atmospheric oxygen gas diffuses from every leaf in support of the animal kingdom. Of course, the intricate reactions of solar photosynthesis store energy much more effectively than any simple water-mineral reaction generated by magma. For each atom of oxygen freed by sunlight, a cell gains an order of magnitude more useful energy than by freeing one atom of sulfur. Vent life is elegantly adapted to exceptional existence in that hot and sulfurous darkness.

But the blood-red plumes of the tube worms bear witness to the vital importance of oxygen even in their deep-water ecosystem. Without that critical co-fuel oxygen, the creative sulfur biochemistry of the tube worm and its bacterial partners would come to naught even where geothermal heat is plentiful.

It is natural to imagine that the life of the vents is independent of the sun. But it is just not so. The luxuriant sun-fed

Continued on page 107



CONNECTIONS

by James Burke

Folies de Grandeur

Sitting here by the Thames at my word-processing machine, I look out on a beautiful Isambard Kingdom Brunel railway bridge, so I'm constantly reminded of the way 19th-century iron and steel technology produced so many machine-assisted *folies de grandeur.* And as I dredge my mental silt for a start to this column, one comes floating past. A dredger, that is.

Which brings to mind the Suez Canal, *folie de grandeur* of them all. Everybody, from the Romans on, had a go at linking the Mediterranean and the Red Sea. Even Napoleon, after he invaded Egypt in 1798, tried (and gave up when his commission of scientists told him the 30-inch difference in water level between the seas made it inadvisable). But in 1869, 25,000 *fellahin* backed by a consortium of Switzerland, Italy, Spain, Holland and Denmark succeeded. In the last stages of construction, suction dredgers were employed.

Both the canal and the pneumatic sand removal had been French ideas. The canal itself was masterminded by a think-big entrepreneur named Ferdinand-Marie de Lesseps (who went on to bankruptcy over a similar job, in the Isthmus of Panama). Industrial-scale pneumatics were introduced while the French were digging the first railway tunnel through the Alps under Mont Cenis. This project was intended to unite Italian Savoy (north of the mountains) through Switzerland with the rest of Italy to the south. Also, people coming home from India and the East could then pick up a train somewhere like Brindisi, instead of having to sail all the way around Spain. But before the tunnel was complete, there was a war, and the French got Savoy. Still, it would be good for tourism.

In 1861, after three years of tedious hand-boring into the Alpine rock face (advancing all of eight inches a day), the chief engineer, Germain Sommeiller, decided to try to finish the job in less than his lifetime. To get through faster, he built a special reservoir, high above the tunnel entrance, to produce a head of water that compressed air for pneumatic drills, which sped things up 20 times.

The Mont Cenis Tunnel amazed everybody almost as much as the Suez Canal did, and the new wonder drills were showcased in a magazine perused by an American whiz kid, George Westinghouse. In 1869 he turned the pneumatic concept into an air brake for use on trains. Compressed air, in pipes running under the train, held back pistons. In the event of a release of air pressure, the pistons slammed forward, driving brake shoes against wheels. Thus a 103-foot train going 30 miles per hour could be stopped in 500 feet.

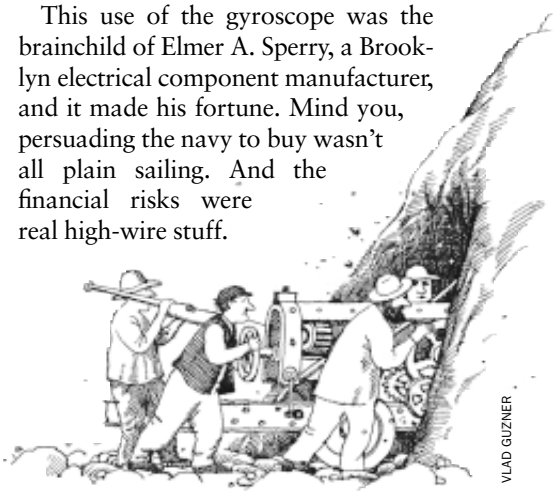
Everybody, from the Romans on, had a go at linking the Mediterranean and the Red Sea.

Such a feat encouraged the idea of scheduling more trains, more closely spaced than had previously been wise, which in turn required better signaling. Which is why, in 1888, Westinghouse fell in with an inventive Croat who wore a new red-and-black tie every week and lived in a hotel room full of pigeons. Name of Nicola Tesla, he figured out how to send electrical power long distances over train tracks to operate railway signals. And then he invented a device so fundamental to the modern world that most of the time you don't know it's there. Tesla sent alternating current into two sets of coils wound on iron, setting up currents that were 90 degrees out of phase with each other. These generated a magnetic field that rotated with each successive burst of current, causing a copper disk to spin. When you put a belt on the disk, you had an electric motor. By World War I, this trick was just what the captains of the newest monster-size battleships were desperate for.

Desperate, first of all, because metal ships with onboard electrical power made things difficult for a magnetic compass, so you could easily get lost. And second of all, the giant 14-inch guns, capable of firing 850-pound shells nearly 10 miles, were unlikely to do any harm if the ship were rolling so much you couldn't hit an enemy barn door.

Tesla's little motor helped to solve both these problems because it could spin gyroscopes of different sizes. There were tiny gyros, for true-north pointing (once you spin the gyro, it stays pointing the way you set it spinning, come hell or, more appositely, high water). Then there were humongous, 4,000-ton gyros, spinning in the center of a ship and compensating for the swell of the sea. And, finally, midsize gyros, doing the same favor for all the gun platforms. Now dreadnoughts could live up to their name. In her first wartime encounter, the newly gyrostabilized USS *Delaware* shot every attacking plane out of the sky. During a storm.

This use of the gyroscope was the brainchild of Elmer A. Sperry, a Brooklyn electrical component manufacturer, and it made his fortune. Mind you, persuading the navy to buy wasn't all plain sailing. And the financial risks were real high-wire stuff.



VLAD GUZNER

Which, as it happened, was where Sperry had originally (and only once) failed. Early on, he'd tried to talk P. T. Barnum into featuring a gyrostabilized wheelbarrow in one of his circus trapeze acts.

The likely reason for Barnum's refusal was that he didn't have much truck with technology, except briefly in the 1840s, when he first set out to be a showman and went looking for curiosities to exhibit. His wish list included "industrious fleas … fat boys … rope dancers and … knitting machines." Besides, by the time Sperry was pitching the gyro idea to him, Barnum was well beyond wheelbarrows (or fleas), touring "The Greatest Show on Earth" with 800 people, 10,000 miles a year by special train.

From time to time Barnum would give it all up for temperance work, or, from 1851 to 1852, to manage a U.S.-Cuba tour for the greatest soprano in the world, Jenny Lind. In 1844 Lind had given her first performance outside Sweden (in Berlin) and was so extravagantly successful that she became an instant diva at 24. One fan wanted only to touch her shoulder "to see where the wings began"(she was known as the "Swedish Nightingale"). In the street she caused scenes that wouldn't be witnessed again until the Beatles. In 1845 Her Majesty's Theater in London commissioned an opera for Lind and asked the other operatic superstar of the day, Giuseppe Verdi, to write it. Two years later Verdi obliged with *I masnadieri,* starring Lind in the role of Amalia. Boffo.

This might be why Verdi got the chance, in the late 1860s, to write what became the most popular opera ever: *Aïda.* Running Egypt at the time was a khedive called Ismail Pasha, whose local engineering efforts had cost so much he was severely short of funds and had to sell his shares in a potentially major moneymaker (the project for which *Aïda* was originally commissioned). The opera, being set in Pharaonic Egypt, was supposed to glorify the country's ancient past and cock a snook at Ismail's Turkish overlords. It didn't do much in that direction, however, and perhaps because it took so long to strike a deal Verdi would accept, delivery of the score was two years overdue. Too late for the occasion it was supposed to celebrate: the opening of the Suez Canal.

And that's my machine-assisted *folie de grandeur* for this month. SA

Wonders, *continued from page 104*

life on land and sea generates reactive oxygen in the atmosphere so fast that it recycles through the biosphere in some 10,000 years. Some of it dissolves in the sea. Should the sun die, the sea, too, would run out of oxygen after a delay of a few thousand years, and the vent fauna would suffocate in a geologic instant.

Indeed, anaerobic bacteria that are biochemical virtuosi abound in such airless environments as soil, marshes and hot springs. Most of them use organic breakdown products of oxygen-dependent life; some may use mineral sources alone. But they (and we) depend on the sun for the whole unfrozen ambiance of surface life, winds and currents, day and night, the climate itself. The mobile diversity of the biosphere above the bedrock is a debt owed the sun.

Is there any life now on Earth that owes no obligation to sunlight at all? In 1992 an unexpected regime of fully sunless life was compellingly suggested by Thomas Gold of Cornell University, based on his own experiences while drilling for natural gas through five miles of granite in Sweden. Much microbial life now does appear to dwell in the wetted pores of deep crustal rock, at least within a couple of miles of the surface, where temperatures are not too high.

Since Gold's proposal, a number of finds of such rock microbes have been reported, although neither the full extent nor the diversity of this newly entered domain of life is known. Presumably, this subterranean life, too, draws on gases vented from the geothermal interior to make a living, for example, by oxidizing abundant iron in the surrounding rock while transforming it from a rust-red state to a dark one. So modest a chemical change offers only a hardscrabble living compared to tapping solar photons, but it can arise without any exposure to perils from the open sky.

We humans and all surface life owe our energy to the enduring thermonuclear gift of bright Apollo. The life we are finding in the warm, rocky strata survives on the much less abundant energy (chemical) secreted below by the god Pluto since the assembly of Earth. And the denizens of the vents belong to Poseidon, ruler of the seas that lie between; they mainly make use of gases empowered by glowing lava and of oxygen set free by the sunlight of years past. SA

HALOGEN LIGHTS

by Terry McGowan

During the 1950s, researchers at General Electric wanted to build tiny, powerful lights that could fit within the razorlike wing tips of supersonic jet aircraft. They first tried increasing the temperature of the tungsten filament to boost its light output. But, as expected, the tungsten metal quickly evaporated, blackening the bulb and causing the filament to break.

Then they hit on a clever idea. Rather than filling the bulb with an inert gas, as in standard bulbs, they used a highly reactive element, iodine. This experiment spawned a variety of lamps now known as halogens for the type of gas that fills the interior (usually iodine or bromine).

DON WILSON *West Stock*

AIRCRAFT LIGHTS have to be small and bright. Such demanding requirements first spurred scientists and engineers at General Electric to develop the halogen lamp.

LAURIE GRACE

TUNGSTEN FILAMENT, when heated by an electric current, gives off light. The immensely hot temperatures that result cause the metal filament to evaporate rapidly, sending out a constant effusion of tungsten vapor (*purple*).

COOLING TUNGSTEN in the vicinity of the glass wall can combine with the halogen gas (*blue*) that fills the bulb. The tungsten halide gas created in this way can migrate freely back to the hot filament.

DECOMPOSITION of the tungsten halide occurs when these molecules are heated by the filament. This process deposits tungsten metal on the filament and releases halogen gas to the bulb, completing the ongoing "halogen cycle."

CONVENTIONAL BULBS contain a chemically inert gas, and so the tungsten vapor given off by the hot filament is simply deposited on the glass, darkening it. Standard bulbs thus have to be made large enough so that the tungsten coating does not become opaque.

Halogen lights take advantage of the peculiar chemistry of tungsten. At the extreme temperatures around the hot filament—about 3,000 degrees Celsius—tungsten vapor does not chemically combine with halogen gas. As the tungsten atoms approach the glass, however, they cool to below 800 degrees C and spontaneously react with the halogen, creating gaseous tungsten halide.

Molecules formed in this way migrate toward the center of the bulb and reach the eroding filament. Because the tungsten halide is unstable there, it decomposes, releasing halogen gas and depositing solid tungsten back on the glowing filament.

Halogen lamps thus shine brighter for longer because tungsten metal does not darken the glass but instead cycles back to help "heal" the eroding filament. Halogen lights are widely employed as floodlights, for automobile headlights and as desk lamps for the home or office. Some still find their way into airplane wings.

TERRY MCGOWAN is manager of worldwide application development for GE Lighting in Cleveland, Ohio.