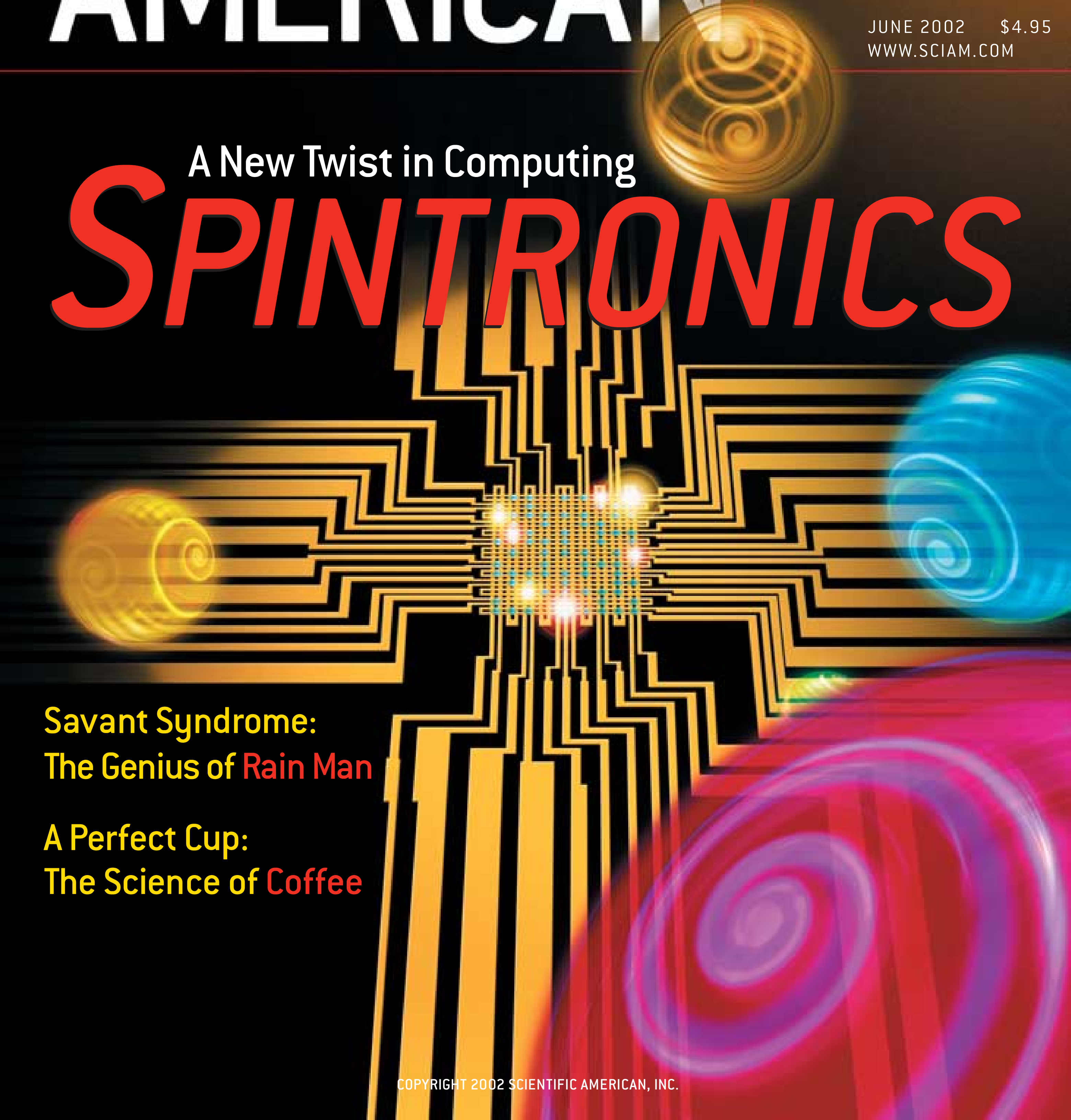
THE LIFE CYCLE OF GALAXIES ▪ VACCINES FOR AIDS ▪ CAN WE STOP AGING?

SCIENTIFIC AMERICAN

JUNE 2002 $4.95
WWW.SCIAM.COM



A New Twist in Computing SPINTRONICS

Savant Syndrome: The Genius of Rain Man

A Perfect Cup: The Science of Coffee

june 2002

contents

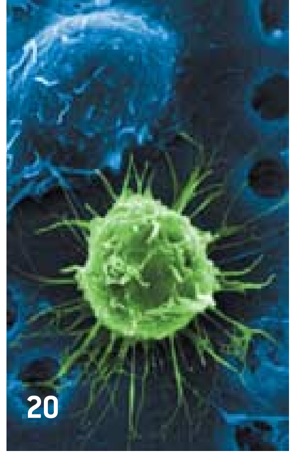
SCIENTIFIC AMERICAN Volume 286 Number 6

features

66 Computing with electron spins

SCIENTIFIC AMERICAN Volume 286 Number 6

departments



20

96

Eugene Chan and Ian Chan of U.S. Genomics 30

columns

Cover illustration 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 BN No. 127387652RT; QST No. Q1015332537. Subscription rates: one year $34.97, Canada $49, International $55. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. **Reprints available:** write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; (212) 451-8877; fax: (212) 355-0408 or send e-mail to sacust@sciam.com **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.** Printed in U.S.A.

A Death Every 30 Seconds

"**Where malaria prospers most,** human societies have prospered least," economist Jeffrey Sachs has observed of the world's preeminent tropical parasitic disease. In any year, 10 percent of the global population suffer its debilitating chills and fevers, and more than one million die. Ninety percent of these deaths occur in sub-Saharan Africa; most are children under the age of five. The disease is currently undergoing a resurgence because of resistance to drugs and insecticides; climate change may play a role as well.

ANOPHELES: malarial mosquito.

The link from malaria to underdevelopment is much more powerful than is generally appreciated. Well beyond medical costs and forgone income, the disease encumbers economic development indirectly. A high burden of malaria encourages a disproportionately high fertility rate—parents want additional children to replace the ones they are likely to lose. A high fertility rate, in turn, can lead to smaller investments in education and health for each child. And malaria can stifle foreign investment, depress tourism and hinder the movement of labor between regions.

Reducing the incidence of malaria would be an extremely cost-effective way to promote development and reduce poverty. So why isn't it happening? The review of *The Fever Trail: In Search of the Cure for Malaria* on page 102 of this issue traces the historical reason—the lack of a viable market for antimalarial pharmaceuticals. This situation is at least as pervasive today: drug companies are reluctant to fund research on vaccines and drugs for a disease that occurs mostly in countries unable to pay for treatment. A few commendable efforts in the public sector are taking up some of the slack, notably the Malaria Vaccine Initiative, supported by the Bill and Melinda Gates Foundation, and the Multilateral Initiative on Malaria, which coordinates research activities.

But developing new drugs is just part of the answer. We don't have to wait for a vaccine. The World Health Organization's Roll Back Malaria campaign, begun in 1998, aims to halve the burden of disease by 2010 through use of insecticide-treated bed nets and combinations of existing drugs, given in particular to pregnant women. And in 2001 the United Nations General Assembly established the Global Fund to Fight AIDS, Tuberculosis, and Malaria.

Sadly, the international efforts are unlikely to make great headway at their present, modest funding levels. Global spending to suppress malaria runs at less than $100 million a year. A basic control program in Africa alone would cost roughly $2 billion annually. Set against the $12 billion in lost GDP that economists estimate malaria costs Africa every year, the benefit clearly exceeds the cost, even when measured narrowly in dollars and cents, not in lives.

It is up to the governments and private institutions of the rich countries to make the required investment—by directly funding control, treatment and research programs and by committing to buy drugs and vaccines at a price sufficient to encourage R&D by pharmaceutical makers. Diseases such as malaria that afflict the poor affect the rich as well—through the spread of infections and the broader destabilization of society. Malaria is one disease we could control now using the technology we have in hand. As our book reviewer Claire Panosian Dunavan concludes, an all-out commitment to curing malaria is "an investment in humankind, global economic health and our own self-interest."

S. STAMMERS *Photo Researchers, Inc.*

THE EDITORS editors@sciam.com

How to Contact Us

EDITORIAL

For Letters to the Editors:
Letters to the Editors
Scientific American
415 Madison Ave.
New York, NY 10017-1111
or
editors@sciam.com
Please include your name and mailing address, and cite the article and the issue in which it appeared. Letters may be edited for length and clarity. We regret that we cannot answer all correspondence.

For general inquiries:
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-754-0550
fax: 212-755-1976
or
editors@sciam.com

SUBSCRIPTIONS

For new subscriptions, renewals, gifts, payments, and changes of address:
U.S. and Canada
800-333-1199
Outside North America
515-247-7631
or
www.sciam.com
or
Scientific American
Box 3187
Harlan, IA 51537

REPRINTS

To order reprints of articles:
Reprint Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8877
fax: 212-355-0408
reprints@sciam.com

PERMISSIONS

For permission to copy or reuse material from *SA*:
permissions@sciam.com
or
212-451-8546 for procedures
or
Permissions Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
Please allow three to six weeks for processing.

ADVERTISING

www.sciam.com has electronic contact information for sales representatives of Scientific American in all regions of the U.S. and in other countries.

New York
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8893
fax: 212-754-1138

Los Angeles
310-234-2699
fax: 310-234-2670

San Francisco
415-403-9030
fax: 415-403-9033

Midwest
Derr Media Group
847-615-1921
fax: 847-735-1457

Southeast/Southwest
MancheeMedia
972-662-2503
fax: 972-662-2577

Detroit
Karen Teegarden & Associates
248-642-1773
fax: 248-642-6138

Canada
Fenn Company, Inc.
905-833-6200
fax: 905-833-2116

U.K.
The Powers Turner Group
+44-207-592-8331
fax: +44-207-630-6999

France and Switzerland
PEM-PEMA
+33-1-4143-8300
fax: +33-1-4143-8330

Germany
Publicitas Germany GmbH
+49-69-71-91-49-0
fax: +49-69-71-91-49-30

Sweden
Andrew Karnig & Associates
+46-8-442-7050
fax: +49-8-442-7059

Belgium
Publicitas Media S.A.
+32-2-639-8445
fax: +32-2-639-8456

Middle East and India
Peter Smith Media & Marketing
+44-140-484-1321
fax: +44-140-484-1320

Japan
Pacific Business, Inc.
+813-3661-6138
fax: +813-3661-6139

Korea
Biscom, Inc.
+822-739-7840
fax: +822-732-3662

Hong Kong
Hutton Media Limited
+852-2528-9135
fax: +852-2528-9281

On the Web

WWW.SCIENTIFICAMERICAN.COM

FEATURED THIS MONTH

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

Position Statement on Human Aging

Anti-aging products are big business, but the marketing of these products often misrepresents the science. Rather than let their silence imply their support, S. Jay Olshansky and 50 other leading scientists in the field of aging research collaborated on a position paper that sets out the current state of the science and separates fact from fiction. An essay distilled from this paper appears on page 92 of this issue. Read the entire report on the *Scientific American* Web site.

Autonomic Computing

The latest catchphrase in computer science is autonomic computing. Researchers dream of creating computing systems that are capable of self-diagnosis, self-defense, self-repair and information sharing with unfamiliar systems. Indeed, IBM is so enamored of the idea that it has issued an eight-point manifesto on the topic. But does autonomic computing really represent a new mind-set in computer science, or is it just a lot of hand waving?

ASK THE EXPERTS

What exactly is déjà vu?
James M. Lampinen, assistant professor of psychology at the University of Arkansas, explains.
www.sciam.com/askexpert/

NEW TO THE SITE

TECHNOLOGY AND BUSINESS CHANNEL

www.sciam.com/techbiz/
From the Internet and microarrays to genetic engineering and robotics, this section provides current coverage of the science and technology influencing the business world.

SCIENTIFIC AMERICAN CAREERS

www.scientificamerican.com/careers
Looking to make a career change in 2002?
Visit Scientific American Careers for positions in the science and technology sectors.
Let the right employers find you.
POST YOUR RÉSUMÉ TODAY.

PLUS:

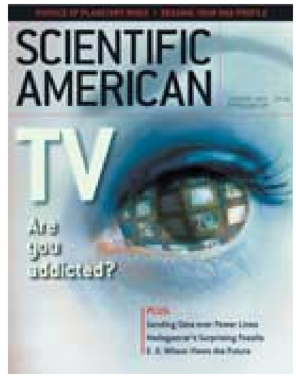
DAILY NEWS ■ DAILY TRIVIA ■ WEEKLY POLLS

Letters

EDITORS@SCIAM.COM

"IT WAS UNFORTUNATE that in their article 'Television Addiction' [February 2002], authors Robert Kubey and Mihaly Csikszentmihalyi did not mention that in the U.S. some 12,000 schools require students to watch *Channel One,* a 12-minute news program that is seen daily by more than 7.8 million students," writes Kristin L. Adolfson of Brooklyn, N.Y. "Each show has four 30-second ad spots. In some schools, children spend the equivalent of about one class week a year watching *Channel One,* including one full day just watching ads. How can we teach children to kick the habit of watching television when we require them to do so at school?"

Stay tuned for other comments on the February 2002 issue.



MORE BOOB-TUBE REFLECTIONS

I have long thought that television may have something to do with attention-deficit disorder (ADD). The story line of a TV movie, say, is interrupted every several minutes by commercials. This breaks the viewer's concentration on a single subject and, over hours of television viewing, instills a habit of jumping from idea to idea. How many people who are diagnosed with ADD merely have a habit caused by the on/off of television viewing? With this habit from television already formed before children ever go to school, is it any wonder they can't concentrate for any length of time? Study habits need to be learned.

Alice Ann Hiestand
Colorado Springs, Colo.

I wonder if Kubey and Csikszentmihalyi have pondered what I think is an important aspect of TV addiction—that is, when frequent viewers become dependent on the tube to fall sleep. This may sound like a joke, but I believe I inherited this trait from my father—not genetically, of course, but rather through the shared experience of nights spent up with the television on, in our most comfortable positions, perhaps even with a pillow, allowing the soothing changes of images and the endless monotone banter to lull us to sleep. But on a night when I did not turn the television on, I would be restless in bed, agitated and thinking I would never get to sleep.

Neil Raper
Flemington, N.J.

I wanted to let you know that I had every intention of reading the article about television addiction, but *Scientific American Frontiers* was on PBS, and I just *had* to watch it.

Todd Dart
Albuquerque, N.M.

Thank you for printing the excellent article "Television Addiction." Here at the Television Project, we have just launched a Web page of "testimonials," stories from parents about how they manage without television or with minimal television. The site is available on the Internet at www.thetelevisionproject.org

We invite your readers to send us their stories, and we will post them for others to read. In this way, we hope to emulate the curative method of Alcoholics Anonymous. Through the sharing of stories, made possible by the Internet, we hope that individuals will learn how they can be free of television addiction and will be inspired to take the first step and turn off the set.

Annamarie Pluhar
Executive Director, The Television Project
Silver Spring, Md.

PATENT PROTOCOLS

In "Intellectual Improprieties" [Staking Claims], Steve Ditlea perpetuates a common and flawed protocol for attacking granted patents. According to the protocol, a patent is read and the description contained therein is broadly generalized. A preexisting technology is then described as conforming to the generalization, thereby "proving" the impropriety of the granted patent.

Few existing patents are immune to this sort of attack. A patent's coverage can be properly evaluated only after careful review of the patent's full description; the patent's claims (a set of precisely worded paragraphs appearing in the patent); documents submitted to and received from the U.S. patent office during the patent application process; and a well-developed body of statutory, regulatory and case law.

After such review, the coverage can rarely be adequately described by a few words of prose and often extends to just part of the patent's description. Although I have no knowledge of the patents mentioned in the article, the respective patent holders deserve a more thorough analysis before their patents are disparaged in an authoritative public forum.

Nandu A. Talwalkar
Buckley, Maschoff, Talwalkar & Allison
New Canaan, Conn.

FIRST KNOCKOUT

In reference to "Count to 10," by Lisa Melton [News Scan], about the latest research on the mechanism of general anesthesia, I would like to point out an error of history. The first surgical general anesthetic, ether, was administered in March 1842 by Crawford W. Long, a doctor, in the rural hamlet of Jefferson, Ga. Several years later William Morton, a dentist, made the first public demonstration in Boston (shown in the article's accompanying photograph). Long did not publish details of his experiments until 1849—thus the continuing confusion.

Michael E. Maffett
Atlanta

WORK ON NETWORKS

In "The Network in Every Room," W. Wayt Gibbs writes that engineers "decided to use much higher frequencies than anyone had tried before, above four megahertz." That is inaccurate. Others have demonstrated spread-spectrum-based networks at those frequency levels as far back as the early 1980s. R. A. Piety characterized the power line up to 20 megahertz in 1983 and demonstrated a power-line network centered at seven megahertz and operating in the 3.5- to 10.5-megahertz range. The work was published in the May 1987 issue of the *Hewlett-Packard Journal.*

Bobbie Evelyn Piety
Palo Alto, Calif.

Gibbs states that the American power grid uses a "different design that makes it far too costly for utilities to compete with DSL and cable." He also quotes William E. Blair of the Electric Power Research Institute regarding the excessive cost of amplifiers needed at the much more numerous distribution transformers. Yet we believe that the special difficulty in the U.S. is in coupling signals onto and off the 2.4- to 33-kilovolt power distribution lines safely and economically, bypassing each of the many distribution transformers.

HOUSEHOLD DEVICES could eventually communicate over ubiquitous power lines.

Utilities have long used capacitors for such coupling on high-voltage lines, but they are indeed very expensive. Responding to that challenge, we have developed inductive couplers and low-cost network architecture, simplifying the high-voltage coupler insulation and dramatically reducing cost. These inductive couplers can transmit data over more than one mile of distribution lines with speeds reaching nearly 20 megabits a second, and we will be continuing initial network trials at major investor-owned utilities in the coming months. The longtime dream of exploiting the already built and maintained power grid at competitive costs may be closer than ever.

Yehuda Cern
Chief Engineer, Ambient Corporation
Brookline, Mass.

CONVERGENCE OF CALORIES

The February issue contained an astonishing convergence of evidence. Data Points [News Scan] noted the steadily rising incidence of obesity in the U.S., a condition that is believed to be preventable through proper diet and exercise.

The Innovations column reported on the invention of a vaccine meant to raise the level of beneficial HDL cholesterol in people who are at risk of atherosclerosis, which, in the majority of the population can also be treated with proper diet and sufficient exercise.

"Television Addiction" states that in the Western world, three hours a day is the average TV viewing time, during which, presumably, exercise is less important than channel surfing.

"The Bottleneck," by Edward O. Wilson, postulates that four additional planets would be needed to sustain the world's population if current Western consumption and lifestyle habits were practiced by every citizen on earth. Such a lifestyle evidently includes 1,000 hours of TV a year, an ample amount of junk food and precious little exercise.

I sense a disturbing pattern.

Will Breen
Kenora, Ontario

SCOTT GRIMANDO

Transistor Sales ■ Meat Business ■ Amazon Trade

JUNE 1952

TRANSISTORS FOR ALL—"Anyone who wants a junction transistor now can buy it. The arrival of this revolutionary substitute for the vacuum tube on the general commercial market was announced last month. The transistor has been extensively studied by Bell Telephone Laboratories, General Electric and the Radio Corporation of America, all of whom have made refinements in the device. The competitive rush to market it has now begun. One distributor quoted a price of $30 for a transistor."

DON'T WORRY—"Why does the same type of cancer grow rapidly in one patient but slowly in another? At the Veterans Hospital in Long Beach, Calif., researchers selected 25 patients whose cancers were growing rapidly and 25 in whom growth was slow. They examined all 50 with the Minnesota Multiphasic Personality Inventory, a standard psychological test which indicates the general type of personality. 'The findings suggest that the person with a rapidly growing tumor has a strong tendency to conceal his inner feelings and is less able to reduce tensions by doing something about them.' They say that measures to relieve the psychological tension may prolong the life of a patient."

MALARIA, ITALIAN-STYLE—"As recently as 1945 there were 411,600 malaria cases in Italy, though the death toll, thanks to atebrin, had been reduced to 386. Now, in six short years, Italy has utterly routed the pestilence. Not a single death from the disease has been reported in the past three years. At the end of the war Albert Missiroli, Italy's leading malariologist, formulated a five-year plan for eradicating malaria from the whole country. The ceilings and inside walls of every house and animal shelter in every malarious area of Italy were to be sprayed once a year, just before the malaria season [*see illustration*]. Italy is a model of what can be accomplished with mankind's new weapon against malaria: DDT and such related insecticides as benzene hexachloride."

DDT DELIVERY in the Italian antimalaria campaign, 1952

JUNE 1902

SUBMERGED HOPES—"The submarine is one of those devices which have suffered from the zeal of its friends. The naval world is now experiencing the first reaction of sentiment which was bound to follow the exaggerated praise of the submarine and the claims for unlimited powers of destruction which have been made for it. We would refer to the one important fact that all submarines are 'blind.' When at the surface, the craft can see; but when it is submerged to its working condition, it is as impossible for the craft to see as it is for it to be seen by the enemy."

CHICAGO MEATPACKING—"The industry of killing and packing beef, pork and mutton has reached such proportions at Chicago—the greatest center of this industry in the world—that the most modern processes have been introduced for the purpose of economizing both time and labor, as well as utilizing all of the products of the carcass. Yearly 3,000,000 cattle and 5,000,000 hogs are slaughtered and converted into packinghouse products in what is known as 'Packing Town.' As far as possible, machinery has been employed, with the result that one of the large companies treats 7,000 hogs in a day, where by hand less than 10 per cent of this number can be disposed of." [*Editors' note: The appalling conditions of this industry were exposed in Upton Sinclair's* The Jungle *in 1905.*]

JUNE 1852

GREEN ACRES—"Lieut. Matthew Fountaine Maury, in a singular memorial to the Senate and House of Representatives, says: 'Imagine an emigrant—a poor laboring man he may be—arriving from the interior of Europe, as a settler in the valley of the Amazon. Where he was, his labor could but support himself in the most frugal manner, and he was then no customer of the United States. But in his new home, where the labor of one day in seven is said to be enough to crown his board with plenty, he has enough to exchange with us for all the manufactured articles that he craves the most. It may be expected, whenever the tide of immigration shall begin to set into that valley, that New York and Boston will have to supply those people with every article of the loom or the shop, from the axe and the hoe up to gala dresses."

news SCAN

PROTEOMICS

Lifting the Screen

AN ACCURATE TEST IS NOT ALWAYS THE BEST WAY TO FIND CANCER **BY ALISON McCOOK**

Cancer screening is notoriously unreliable: a positive test often does not indicate disease, and a negative result does not always mean the patient can walk away with a handshake and a smile. In February many physicians and patients were encouraged by the results of a new test for ovarian cancer, hoping that it would be a noninvasive, cost-effective way to save thousands of lives. The findings offered proof of the enticing idea that within the thousands of proteins swimming in the blood lies a simple code that, if broken, will reveal whether cancer lurks in the body. But although the concept is promising, this technique is a long way from being useful within the general population.

TOO OFTEN, TOO LATE: Ovarian cancer cells, as seen by a scanning electron microscope. The image shows secretory cells with hairlike protrusions called microvilli (*pink*) as well as cilia (*green*) and mucus (*yellow*).

P. MOTTA AND S. MAKABE *Photo Researchers, Inc.*

News of this latest approach sparked widespread interest because none of today's diagnostic tests for ovarian cancer—including ultrasonography, pelvic exams and blood tests to detect levels of a protein called CA 125—can consistently detect the disease early, when the cure rate is around 90 percent. Instead most women are diagnosed once their cancer has progressed, when the chances of surviving five years drop to 35 percent.

In the recent paper, scientists led by Lance A. Liotta of the National Cancer Institute and Emanuel F. Petricoin of the U.S. Food and Drug Administration mapped, with the help of an artificial-intelligence algorithm, the particular blood proteins or protein fragments that differ in samples from women with ovarian cancer. Other researchers have published reports using proteomics to diagnose disease, but because Liotta and Petricoin's results appeared in a prestigious publication, the *Lancet,* they received additional attention. Indeed, they sound impressive: in 116 samples, that protein "fingerprint" picked out every woman with ovarian cancer, including 18 early cases, and designated 63 out of 66 healthy women as disease-free.

Within 48 hours of the study's publication, Carol L. Brown of Memorial Sloan-Kettering Cancer Center in New York City received calls from an estimated 75 percent of her patients who were in remission for ovarian cancer, asking about the test. But, as Brown told them, it is "not something that's going to be a commercially available test for, I think, many, many years—if at all," she says.

That's because, surprisingly, the ability to find all cases of cancer is not the best way to judge the value of a screening test. To calculate the likelihood that a positive test indicates cancer, epidemiologists use an equation that includes the test's sensitivity (how well it finds cancer when it is there), its specificity (its ability to diagnose healthy patients accurately) and the disease prevalence. The sensitivity of the new test is 100 percent, the specificity is around 95 percent (63 of 66 healthy patients found), and ovarian cancer occurs in only one in 2,500 women who are older than 35 years in the U.S. each year. Plugging those numbers into the equation shows that for every woman who gets a positive proteomics test result, there is a less than 1 percent chance she has the disease.

If a screened woman gets a positive result, her doctor conducts further analyses, such as a laparotomy, a surgery that opens the abdomen to explore for disease. In public health terms, subjecting 100 women to the anxiety, expense and risks of surgery to find cancer in just one patient is unacceptable. But the only value in the equation that can be improved is the specificity, which is already quite high. Ironically, increasing the test's specificity may mean lowering its overall accuracy, explains Sudhir Srivastava of the National Cancer Institute; in other words, the test would be capable of "finding" cancer in healthy people. But even if little tweaking of the numbers is possible, researchers may be able to give the test to women who are more likely to develop ovarian cancer, such as those with a family history of the disease. "It may be that in the high-risk population, these numbers are approaching acceptability," says Martee L. Hensley of Sloan-Kettering.

There is additional concern that other institutions may not be able to repeat the procedure using their own equipment and software. The unidentified proteins and protein fragments that make up the *Lancet* fingerprint are so small that any slight variations between machines, algorithms or the solutions used to prepare blood samples may skew the results. "So if you ran samples three months ago and got beautiful results, can you repeat that three months later, and can you repeat it on different instruments?" asks George L. Wright of Eastern Virginia Medical School.

Despite the reservations, these results may herald a future in which tests use multiple, not single, biomarkers to spot disease. Researchers are looking at patterns that may identify prostate and breast cancer, among others. Given the heterogeneity of cancer, this approach makes intuitive sense. Declares Wright: "One marker will not be found to improve the early detection, diagnosis, prognosis of any cancer or disease."

Alison McCook is a science writer based in New York City.

news SCAN

TO SCREEN OR NOT TO SCREEN?

Some screening techniques are facing increasing controversy. Experts debate whether mammography and PSA testing hurt more people than they help by detecting cancers at too early a stage, when it is unclear if the disease is benign or requires treatment. A study in the April 4 *New England Journal of Medicine* found that about two thirds of one-year-olds whose urine tests came back positive for neuroblastoma actually had completely harmless tumors.

But testing rates for most cancers remain high, says William C. Black of Dartmouth-Hitchcock Medical Center, because managed care physicians do not have the time to explain the nuances of screening and all are afraid of being sued by cancer patients who did not receive the test. And in the end, doctors can never be sure which patients treated for the disease could have postponed or even avoided the medical intervention. "Ironically, the people who are harmed by the overdiagnosis become the most vocal advocates for screening," Black remarks, "because they think, of course, they've been saved."

WASTE DISPOSAL

Divide and Vitrify

PARTITIONING NUCLEAR WASTE SAVES SPACE, BUT IT ISN'T EASY **BY STEVEN ASHLEY**

The tons of toxic waste left over from nuclear weapons production—including plutonium, uranium, cesium and strontium isotopes, as well as the now radioactive processing additives—sit unremediated in belowground storage tanks and bins at three U.S. Department of Energy sites. Even if the controversial "permanent disposal" effort at the Yucca Mountain repository in Nevada proceeds, there still will not be sufficient room to hold the entire mess.

To cram the waste into what space eventually opens up, nuclear scientists and engineers have been working on various methods to segregate the extremely dangerous wastes from the merely hazardous ones. The idea is to reduce the quantity of the most deadly high-level waste that must be buried, allowing the less threatening low-level waste to be consigned to cheaper belowground storage facilities nearer the surface. Separating out the highly radioactive materials also allows engineers to control the radioactivity and heat generated in the glass media that would store the waste, boosting the safe capacity of storage repositories.

TOXIC BREW of radioactive waste lies just 10 feet below technicians working to replace a pump in a million-gallon storage tank at the Hanford site in Washington State.

But dividing the bad from the not-as-bad has not proved simple. The DOE sites—namely, Savannah River in South Carolina, Idaho National Engineering and Environmental Laboratory (INEEL) and Hanford in Washington State—store various types of nuclear waste that require specially tailored separation technologies.

"When the Bush administration first arrived, it called for a review of the DOE's entire $300-billion Environmental Management program, which had been planned to run until 2070," explains Mark A. Gilbertson, director of the DOE's Office of Environmental Management. "As much as 50 percent of the cost of disposing of high-level waste is tied to pretreating it or the subsequent immobilization of it." Last year, Gilbertson says, his office spent about $16 million to find ways to hike the efficiency of nuclear waste handling and lower the environmental risks.

At the Savannah River site, where the bomb waste is highly alkaline, engineers had been removing cesium 137 from the soluble portion of the tank wastes through chemical precipitation. Unfortunately, that approach had to be halted because the process liberated flammable benzene gas. Bruce A. Moyer, group leader for chemical separations at Oak Ridge National Laboratory, and his team have developed a safer alternative that may soon be adopted. In their procedure, expensive "designer" solvent molecules called calixarenes selectively glom on to cesium, allowing it to be removed from the liquid. The cesium would then be stripped off chemically from the calixarene molecules, which could then be reused.

Meanwhile the nonsoluble sludge still in the tanks, which contains strontium 90 and transuranic elements (a mix of radioactive species heavier than uranium), would be washed with sodium hydroxide to remove bulky constituents such as aluminum, thus reducing the total mass. The sludge would at this point be added to the extracted cesium, and the mixture would be vitrified (turned into stable borosilicate glass logs) encased in stainless-steel canisters and then entombed.

Things are somewhat different at the Idaho facility, where the nuclear waste is stored in bins in the form of an acidic granular solid called calcine. Although technologies exist to separate the cesium, strontium and transuranics, each requires its own procedure, raising costs and slowing throughput, says R. Scott Herbst, consulting engineer at INEEL. Looking for a better option, INEEL scientists are studying a single-step chemical extraction process that is conceptually not unlike the Oak Ridge technique. Developed at the V. G. Khlopin Radium Institute in St. Petersburg, Russia, the procedure employs three compatible solvents that act simultaneously. "We still don't understand how this unitary process works, but it's worth following up since it would be substantially cheaper than the previous three-step procedure," Herbst states.

Hanford has the most complex hot refuse: it consists of a mix of wastes from many nuclear fuel reprocessing projects. Engineers are currently planning a two-stage ion exchange process to extract radioactive cesium and technetium from the soluble part of the alkaline tank waste. In this process, columns of polymer resin beads attract the harmful elements, which are later removed from the beads with acid.

A still speculative method may supplant that approach, however. Since 1998 Archimedes Technology Group in San Diego has been developing a filtering method that works via atomic mass rather than chemical properties. The technique, which borrows from fusion energy research, takes advantage of the fact that 99.9 percent of the radioisotopes in the waste are heavy elements, says company head John R. Gilleland. Radio waves would vaporize the waste, which would then be sent into a magnetic bottle containing a thin, trapped plasma. A radial electric field would then cause the plasma and most of the waste ions to "orbit" along a spiral path inside the

TAKING OUT THE NUCLEAR TRASH

The Department of Energy is trying to speed up the process of high-level waste disposal. In March the Bush administration committed an additional $450 million beyond the $2 billion already budgeted for 2003 as part of a scheme to halve the planned 70-year cleanup time at the Hanford site. Construction is slated to begin late this year on a giant vitrification plant to convert around 10 percent of Hanford's highly radioactive waste into borosilicate glass logs, which would be buried deep underground for 10,000 years or more.

DEPARTMENT OF ENERGY OFFICE OF RIVER PROTECTION cylindrical chamber. Ions below a certain mass would be confined to the magnetic field lines and travel to the ends of the chamber; the specially tuned magnetic field, however, could not hold the heavier ions (namely, the radioactive species), which would drop to the sidewalls for later removal. The Archimedes filter should be a high-throughput process, thereby saving time and money, but testing of a full-size prototype will not be done until 2003.

Splitting the atom has resulted in many long-term political, environmental and management headaches. Splitting the waste promises to offer a bit of relief.

PHYSICS

A Philosopher's Stone

COULD SUPERCONDUCTORS TRANSMUTE ELECTROMAGNETIC RADIATION INTO GRAVITATIONAL WAVES? BY GEORGE MUSSER

Raymond Chiao remembers the day, during his childhood in Shanghai, when his brother built a crystal radio set and invited him to try it. "When I put the earphones on, I heard voices," he says. "That experience had something to do with my going into physics." Chiao has since become well known for his work in quantum optics at the University of California at Berkeley. Now he is preparing an experiment that, if it works (a not insubstantial if), would be the biggest invention since radio.

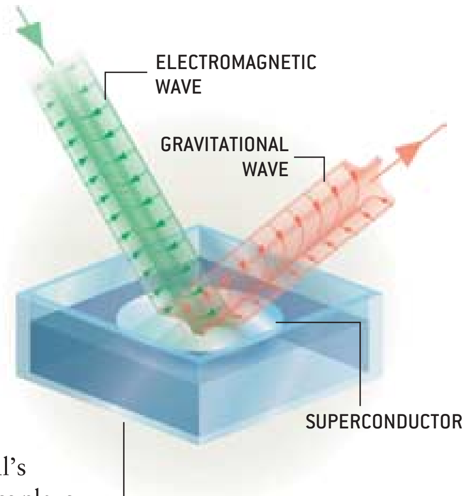
Chiao argues that a superconductor could transform radio waves, light or any other form of electromagnetic radiation into gravitational radiation, and vice versa, with near perfect efficiency. Such a feat sounds as amazing as transmuting lead into gold—and about as plausible. "It is fair to say that if Ray observes something with this experiment, he will win the Nobel Prize," says superconductivity expert John M. Goodkind of the University of California at San Diego. "It is probably also fair to say that the chances of his observing something may be close to zero."

Chiao presented his hypothesis at a March symposium celebrating the 90th birthday of Princeton University physicist John Archibald Wheeler (the paper is available at arXiv.org/abs/gr-qc/0204012).

His analysis, like most discussions of gravitational radiation, proceeds by analogy with electromagnetic radiation. Just as changes in an electric or magnetic field trigger electromagnetic waves, changes in a gravitational field trigger gravitational waves. The analogy is actually quite tight. To a first approximation, Einstein's equations for gravitation are a clone of Maxwell's equations for electromagnetism. Mass plays the role of electric charge, the only difference being that its value must be positive (at least in classical physics). Masses attract other masses via a "gravitoelectric" field. Moving masses exert forces on moving masses via a "gravitomagnetic" field. Gravitational radiation entwines gravitoelectric and gravitomagnetic fields.

Over the years a number of physicists have suggested that if a superconductor can block magnetic fields—giving rise to the famous Meissner effect, which is responsible for magnetic levitation over a superconductor—then it might block gravitomagnetic fields, too. When Chiao adds the gravitomagnetic field to the standard quantum equations for superconductivity, he confirms not only the gravitational Meissner-like effect but also a coupling between the two breeds of magnetic field. An ordinary magnetic field sets electrons in motion near the surface of a superconductor. Those electrons carry mass, and so their motion generates a gravitomagnetic field.

Thus, an incoming electromagnetic wave will be reflected partly as a gravitational wave, and vice versa. The same should occur in any



GRAVITY TRANSDUCER would reflect incoming electromagnetic radiation (*green*) as gravitational waves (*orange*). The radiation must be polarized in a so-called quadrupole pattern.

SAMUEL VELASCO

MAKING WAVES

Like an ordinary magnetic field, a gravitomagnetic field exerts a force on moving masses at right angles to their velocity. The rotating earth, for example, generates a gravitomagnetic field that torques satellite orbits, as observations over the past several years have confirmed. The Gravity Probe B satellite, scheduled for launch early next year, should precisely measure this effect, which is also known as the Lense-Thirring effect, or "frame dragging."

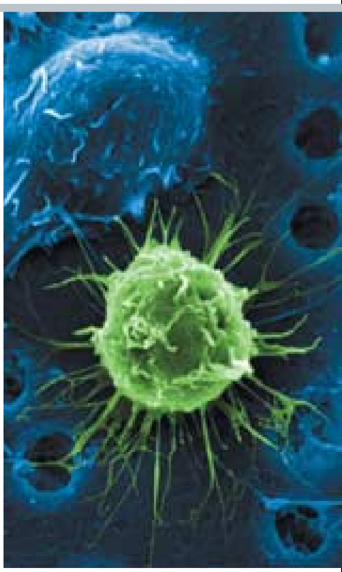
Even if Chiao's contraption works, it wouldn't allow the generation of antigravity fields, as Russian materials scientist Eugene Podkletnov, then at Tampere University of Technology in Finland, controversially claimed to have observed in 1992 (see www.sciam.com/askexpert/physics/physics29a.html). Antigravity requires canceling out a powerful, static gravitoelectric field, yet superconductors have no effect on such fields.

electrical conductor, but in a superconductor the electrons all move in unison, greatly amplifying the effect. In fact, Chiao ventures that the incoming energy will be divided evenly between the two types of radiation.

"His mathematical arguments seem to be correct," remarks Bryce DeWitt of the University of Texas at Austin, a pioneer of quantum gravity. But DeWitt, Goodkind and the half a dozen other leading lights of physics interviewed for this article have assorted ideas about where Chiao might have gone astray—pointing out, for instance, that he makes various simplifications and leaps of faith. And you have to wonder why this coupling, if it is really so strong, hasn't been noticed before.

By the time the theory is vetted, though, Chiao will probably have conducted his experiment and settled the question. Working with Berkeley electronics specialist Walter Fitelson, he plans to beam specially polarized microwaves onto one slab of superconductor and use a second slab to look for rebounding gravitational waves. The setup, which uses off-the-shelf parts, is not much more complicated than a crystal radio.

If it works, you could probably come up with 30 ideas for applications in as many seconds, from new gravitational-wave detectors for astronomy to graviton antennas for telecommunications, which could send signals through the solid earth. Chiao's idea is a reminder that for all the attention paid to cutting-edge research such as string theory, radical new physics may lie within the interstices of conventional theories.



BONE MARROW CELLS have been cited as a source for adult stem cells.

ANDREW LEONARD *Photo Researchers, Inc.*

CELL THERAPY

The Child Within

STEM CELLS FROM ADULTS MAY NOT BE SO USEFUL AFTER ALL **BY CAROL EZZELL**

Last August, when President George W. Bush banned the use of federal funds for any stem cell research that would require the creation and destruction of human embryos, one of the arguments his administration used was that embryonic stem cells might be unnecessary. Because such all-purpose cells could instead be isolated from adults and appeared to work in transplantation studies involving animals, administration officials alleged, why would anyone need stem cells derived from embryos, with their moral and ethical overtones?

The answer, possibly, is that what researchers once thought were stem cells from adults might not be. The premise of adult stem cell transplantation is that such cells are essentially undifferentiated and have therefore retained the capacity to become tissues as diverse as brain and liver. But two studies in the April 4 *Nature* suggest that transplanted adult stem cells merely fuse with a recipient's own cells without becoming a particular type of differentiated cell. If the results of the studies are upheld, it could bolster the case for using stem cells from embryos rather than adults.

The two groups of scientists that conducted the experiments—one led by Naohiro Terada of the University of Florida and the other by Austin G. Smith of the University of Edinburgh—initially set out to see whether adult stem cells from the bone marrow of mice would turn into embryolike cells when cultured with mouse embryonic stem cells. Instead they found that the adult cells simply fused with the embryonic cells to create giant cells with more than the normal number of chromosomes.

Ron Cohen, president and CEO of Acorda Therapeutics, a small biotech firm in Hawthorne, N.Y., worries that the transplants of adult stem cells might lead to cancer. Extra chromosomes, after all, are phenomena usually found in tumor cells. "Are you at risk for some genetic alteration in these cells?" he asks. "Could you be stimulating cancers somewhere down the road?"

Janis L. Abkowitz of the University of Washington agrees that the findings "are another reason for caution" in transplanting adult stem cells, "but there are so many others." In particular, she notes that it is very difficult to prove months later that a specific

CLONING ON THE HILL

Before its Memorial Day recess, the U.S. Senate was expected to vote on whether to join the House of Representatives in banning cloning—both for producing stem cells for transplantation and for generating babies. President Bush has indicated that he would sign such a ban into law. That would leave scientists to work with 64 existing cultures of embryonic stem cells—many of dubious quality—or with adult stem cells.

patch of tissue was once an injected adult stem cell.

Terada downplays the significance of the results for the safety of stem cell transplantation. "This is a cautionary tale for the plasticity of adult stem cells, not their safety," he asserts. But he acknowledges that the experiments were performed in laboratory culture dishes and that the results might not necessarily reflect what happens in a living organism. "We've never said this explains all the previous experiments in vivo," Terada says. But he adds that he and his colleagues have repeated their experiments using adult skin cells and found that adult stem cells also fuse with them. However one interprets the findings, they are sure to prompt a reevaluation of previous studies of adult stem cells and fuel the already raging debate over embryonic stem cells and cloning.

INFOTECH

URLs in Urdu?

INTERNATIONAL DOMAIN NAMES POSE A NEW SECURITY RISK BY WENDY M. GROSSMAN

Is this the Web address of tomorrow: http://ωωω.σχιαμ.χομ? At the moment, non-Latin alphabets and scripts are not compatible with ASCII, the lingua franca of the Internet also known as plain text. But as of March only 40 percent of the 561-million-strong global online population were native English speakers, according to online marketing firm Global Reach. Work has been proceeding for some time, therefore, to internationalize the system that assigns domain names (sciam.com, for example) to the dotted clumps of numbers that computers use (such as 192.1.1.0).

The technical side of things has been managed by the Internationalized Domain Name Working Group of the Internet Engineering Task Force (IETF). In April, VeriSign, the single largest registrar of domain names, claimed to have registered about a million international names. But turning Web addresses into a multilingual forum may open the door to a dangerous new hazard—hackers could set up fake sites whose domain names look just like the ASCII version.

One example is a homograph of microsoft.com incorporating the Russian Cyrillic letters "c" and "o," which are almost indistinguishable from their Latin alphabet counterparts. The two students who registered it, Evgeniy Gabrilovich and Alex Gontmakher of the Technion–Israel Institute of Technology in Haifa did so to make a point: they suggest that a hacker could register such a name and take advantage of users' propensity to click on, rather than type in, Web links. These fake domain names could lead to a spoof site that invisibly captures bank account information or other sensitive details.

In their paper, published in the *Communications of the ACM,* they paint scary, if not entirely probable, scenarios. For instance, a hacker would be able to put up an identical-looking page, hack several major portals to link to the homographed site instead of the real one, and keep it going unnoticed for perhaps years.

On a technical level, homograph URLs are not confusing. International domain names depend on Unicode, a standard that provides numeric codes for every letter in all scripts worldwide. And at its core, the internationalization of the domain name system is a veneer: the machines underneath can still only read ASCII.

According to the proposed standard, the international name will be machine-translated at registration into an ASCII string composed of an identifying prefix followed by two hyphens followed by a unique chunk of letters and numbers: "iesg--de-jg4avhby1noc0d," for example. This string would be translated back into Unicode and compared with the retranslation of the original. So right now anyone using a standard browser can easily see the difference between an internationalized domain name and an ordinary one.

This situation, however, is temporary. Technical drafts by the IETF state that users should not be exposed to the ugly ASCII strings, so increasingly users will have little way of identifying homographs. Computer scientist Markus G. Kuhn of the University of

NEED TO KNOW: ICANN CAN'T

The Internet Corporation for Assigned Names and Numbers (ICANN) was set up in 1998 to oversee several important technical functions that keep the Internet running. Ever since, it has been criticized for lack of accountability and openness. In February its current president, M. Stuart Lynn, issued a manifesto claiming that ICANN was seriously broken and proposing a complete reform. Although many concede that ICANN has failed, few agree with Lynn's specific proposals, which essentially call for a rebuilt organization with three to five times the budget, more than 50 percent additional staff and greater power. Critics argue that this plan will create a single point of failure, the very thing the Internet's design sought to avoid.

The upshot has been to reopen the intense debates that preceded ICANN's formation. Even former pacifists, including Peter G. Neumann, who moderates the online bulletin board RISKS Forum, and Lauren Weinstein of People for Internet Responsibility, are taking sides. They say an immediate handover to a less political, more strictly technical organization, such as the Internet Architecture Board, is necessary to avoid a meltdown.

Cambridge notes that for users to be sure they are connected to the desired site, they will have to rely on the secure version of the Web protocol (https) and check that the site has a matching so-called X.509 certificate. "That has been common recommended practice for electronic banking and commerce for years and is not affected by Unicode domain names," Kuhn observes. Certification agencies (which include VeriSign) ensure that encoded names are not misleading and that the registration corresponds with the correct real-world entity.

But experience shows that the Internet's majority of unsophisticated users "are vulnerable to all kinds of simple things because they have no concept of what's actually going on," explains Lauren Weinstein, co-founder of People for Internet Responsibility. Getting these users to inspect site certificates is nearly impossible. Weinstein therefore thinks that a regulatory approach will be necessary to prohibit confusing names. Such an approach could be based on the current uniform dispute resolution procedure of the Internet Corporation for Assigned Names and Numbers (ICANN), the organization that oversees the technical functions of handing out domain names. But it will require proactive policing on the part of the registrars, such as VeriSign, something they have typically resisted.

But are international domain names even necessary? Kuhn, who is German, doesn't think so: "Familiarity with the ASCII repertoire and basic proficiency in entering these ASCII characters on any keyboard are the very first steps in computer literacy worldwide." Internationalizing names might succeed only in turning the global network into a Tower of Babel.

Wendy M. Grossman, a frequent contributor on information technology, is based in London.

GEOPHYSICS

Scaling the Quakes

WHY AFTERSHOCKS MAY NOT REALLY BE AFTERSHOCKS AFTER ALL BY JR MINKEL

According to conventional earthquake wisdom, aftershocks represent the ground's "relaxing" after the main temblor has rattled the land. But researchers in Britain report that, statistically speaking, aftershocks are no different from main shocks.

Physicists Per Bak, Kim Christensen, Leon Danon and Tim Scanlon of Imperial College London mapped more than 330,000 earthquakes that struck California between 1984 and 2000. They found that all the quakes obeyed a single underlying scaling law, a mathematical relation that gives the statistical spread of events for a given area and magnitude. According to this law, earthquakes cluster in the same way at a range of timescales, from tens of seconds to tens of years. So from a wide enough perspective, an aftershock could come years after a primary event.

The scaling law supports the long-anticipated idea that earthquakes are self-organized critical phenomena, the investigators write in the April 29 *Physical Review Letters*. For such phenomena, a small change triggers a chain reaction of larger disturbances after some critical threshold is passed. A sandpile is the classic example of these systems: once it attains a certain slope, the addition of just a few extra grains will cause an avalanche. If real, the connection between earthquakes and self-organized critical phenomena suggests that one process is responsible for all quakes. "It shows that one cannot understand individual earthquakes independently," Christensen says.

Geophysicist Yan Y. Kagan of the University of California at Los Angeles agrees that "the distinction between aftershocks and main shocks is relative." Within slowly changing continental areas, he points out, aftershocks can rumble on for centuries.

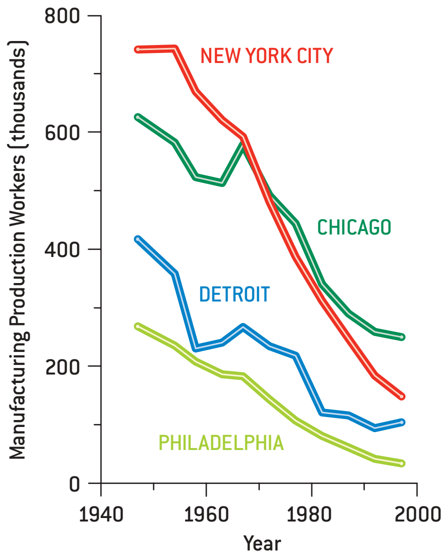
ROBYN BECK, ©AFP/CORBIS

AFTERMATH: The devastation in Puli in central Taiwan, after a 7.3-magnitude earthquake and aftershocks on September 22, 1999.

Bad Things Happen

HOW DEINDUSTRIALIZATION HAS AFFECTED COMMUNITIES **BY RODGER DOYLE**

The blue-collar middle class in the U.S. was built on manufacturing production jobs, but their number has dwindled. In major cities of the North, as the chart shows, the decline has been particularly steep. Furthermore, pay for these jobs, unlike that for highly skilled workers such as engineers, has declined relative to the national average. Several decades ago the typical production-line job did not require advanced skills but was unionized and so paid at or above the average. By 1997, however, production-line pay dropped below the average in most areas of the U.S., except where unions were still strong, such as in Detroit.



As manufacturing jobs dried up and older workers took early retirement, young people, instead of becoming assemblers or machine operators, became janitors and waiters. Such service-sector positions generally paid less than production work. The better-paying jobs were in hard-to-reach suburbs.

These disincentives left many young men unemployed. At about the same time, for reasons that are still not completely understood but that may include a dearth of eligible wage-earning men, the number of unmarried teenage mothers soared. Generally, these girls were not only economically insecure but lacked parenting skills, and so it is not surprising that their children tended to be disadvantaged. The children, moreover, grew up in neighborhoods that were coming apart. Churches, social clubs and unions—especially in black communities—were dissolving, in part because higher-income people fled, depriving the areas of key resources and role models for children. Black newspapers, once a vibrant force in many communities, all but disappeared.

These developments contributed to the surge in youth gangs and crime beginning in the 1960s. Other changes fed the crime wave, such as a large increase in the number of young men between the ages of 18 and 35, the most crime-prone age group, and the increasing availability of illegal drugs, particularly crack, which appeared in the 1980s.

Loss of jobs, together with a shortage of affordable housing that followed neighborhood gentrification and failure to maintain existing housing, added to the rising number of homeless people beginning in the 1970s. The legally mandated emptying of psychiatric hospitals was a factor in escalated homelessness, though apparently not a precipitating cause.

There are signs of improvement throughout the country as a whole. The number of babies born to teenage mothers has followed a downward trend since 1994; the poverty rate is below the level of a decade ago; drug use is down from the high levels of the 1980s; and most significantly, crime rates have plummeted since 1992.

But other signals suggest that the legacy of deindustrialization lingers. Wages of the bottom quarter of Americans have improved little in the past 25 years, and unions, which provided a measure of stability to working-class neighborhoods, have been severely weakened. According to the U.S. Conference of Mayors, homelessness and hunger went up sharply last year. Perhaps the most troubling news is that employment among young, undereducated black males fell from 62 percent in 1979 to 52 percent by the period 1999–2000, a development that probably traces in part to the decline of manufacturing production jobs.

Rodger Doyle can be reached at rdoyle2@adelphia.net

RODGER DOYLE

FAST FACTS: PAY BELOW PAR

Pay of manufacturing production workers in 1997 as a percent of average pay in:

Los Angeles	68
Chicago	82
San Jose, Calif.	73
Orange County, Calif.	76
New York City	46
Houston	88
Phoenix	85
Detroit	111
San Diego	84
Cleveland	105
Philadelphia	91
Pittsburgh	103
Boston	89
U.S. metropolitan areas	88

SOURCE: U.S. Census Bureau, Census of Manufactures. Data are for home counties of cities. In 1997 production workers accounted for 72 percent of all manufacturing employment.

news SCAN

PHOTONICS

Nice Threads

PHOTONIC-CRYSTAL THREADS are 0.2 millimeter wide.

Your clothing may someday reflect more than just your personality. Materials scientists at the Massachusetts Institute of Technology have made polymer threads coated with mirrors. They deposited a glassy substance, arsenic triselenide, onto a polymer and then rolled it up, creating a layered structure called a photonic crystal. Drawing the roll out produces long threads a few hundred microns thick that can be as reflective as gold. The fibers are more than high-tech sequins, though—the reflective properties can be adjusted by varying the diameter of the thread. Properly drawn and woven into normal fabric, the mirrored fibers could lead to wearable radiation barriers, optical bar-code tags for clothing and flexible filters for telecommunications. The April 19 *Science* contains the study. *—Philip Yam*

DATA POINTS: SHARK BITES MAN

Gruesome attacks provided for sensational news last summer, but 2001 actually saw a decline in the number of shark attacks worldwide compared with the number reported the previous year. Overall, the rate of attacks has risen during the past few decades because of increased human activity in the water, not because shark populations are growing.

Number of unprovoked shark attacks in:
2000: 85
2001: 76

Number of fatalities in:
2000: 12
2001: 5

Fatality rate in the 1990s: 12.7%

Favorite targets (percent of those attacked):
Surfers: 49%
Swimmers/waders: 29%
Divers/snorkelers: 15%
Kayakers: 6%

SOURCE: International Shark Attack File, Florida Museum of Natural History of the University of Florida

BIOLOGY

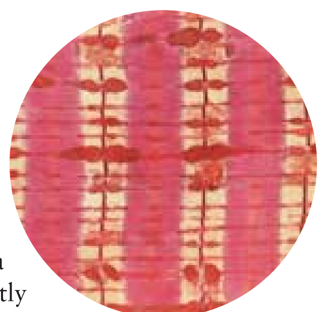
Battling Resistant Bacteria

Two recent results could help fight antibiotic-resistant bacteria. Netherlands researchers report a mathematical model for determining whether hospital patients' infections stem from bacteria they carried in with them or acquired from another patient—important knowledge for evaluating infection-control strategies. The existing method demands the expensive and time-consuming step of reading the bacterium's genome. In contrast, the model analyzes several months' worth of infection-prevalence data to give spontaneous infection and transmission rates. When fed numbers from two past studies, the new technique returned rates similar to those obtained with the genetic approach.

University of Rochester biologists have also developed a model that tracks antibiotic-resistant bacteria—by mimicking evolution. They generated many mutations in a 40-year-old version of a key bacterial gene and selected for variants that resisted antibiotics. The mutants they isolated were many of the same ones that emerged in people, suggesting that the model could predict how bacteria will respond to new drugs. The results already hint that resistance to the antibiotic cefepime may be forthcoming. The transmission model appears in the April 16 *Proceedings of the National Academy of Sciences*. The selection research is published in the March *Genetics*. *—JR Minkel*

CELL BIOLOGY

Gain without Pain



MUSCLE FIBRILS can be energized by more mitochondria (*brown spots along vertical structures*).

The microscopic powerhouses known as mitochondria energize all human activity—the more a cell possesses, the more stamina it has. Working out can pump up mitochondria numbers, but a study indicates that a protein apparently triggers the same effect, giving new meaning to the words "exercise supplement."

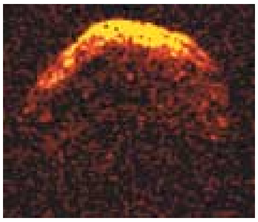
Researchers at the University of Texas Southwestern Medical Center in Dallas looked at easily fatigued muscles of sedentary mice and found that an enzyme known as CaMK can boost mitochondria levels in those muscles. A mitochondria-promoting drug could help bedridden patients or people with heart and lung problems enjoy the benefits of exercise. The scientists, who described their findings in the April 12 *Science*, also speculate that human performance could be enhanced by altering genetic activity to make more of the protein. *—Charles Choi*

FINK RESEARCH GROUP *M.I.T.* AND AAAS (*top*); QUEST *Photo Researchers, Inc.* (*bottom*); ILLUSTRATION BY MATT COLLINS

NEAR-EARTH OBJECTS

Hit or Miss

The bad news is that the kilometer-wide asteroid 1950 DA has up to a one-in-300 chance of striking the earth—the highest risk for any known asteroid, according to NASA physicists. The 100,000-megaton explosion resulting from a strike would cause global damage. The good news is that the impact wouldn't happen until March 16, 2880.



IT'S COMING: Radar image of asteroid 1950 DA.

The asteroid will more likely miss us by within a few days on either side of a 20-minute collision window. Many of the factors that affect the odds are uncertain, especially the rock's axis of spin. The orientation determines the direction of the push it gets after radiating absorbed sunlight back into space. We could exploit this source of drift, called the Yarkovsky effect, to nudge space rocks out of our way, suggests Joseph N. Spitale of the University of Arizona. Covering an asteroid in chalk powder or charcoal, painting it white or even wrapping it in Mylar could all subtly change its speed. Enacted decades or centuries in advance, such a scheme could divert rocks like 1950 DA. The April 5 *Science* has more details. *—JR Minkel*

MEDICAL TECH

Here's Magnet in Your Eye

Injecting a magnetic fluid into the eye could repair severely detached retinas. This light-sensitive layer of cells may tear away from the back of the eyeball because of disease or injury, potentially causing blindness. Doctors generally inject gas or silicone fluid to shove the retina back into place, but these methods don't always reach the bottom parts of the eye. Looking for something they could better control, chemist Judy S. Riffle of the Virginia Polytechnic Institute and her colleagues have combined tiny particles of cobalt or magnetite with a silicone-based fluid, they stated at an April meeting of the American Chemical Society. A magnetic band placed around the eye should hold the fluid against the retina at desired locations. Riffle says the group has also conducted the procedure in glass eyeballs and is set to begin in vitro toxicity testing. Animal studies could begin within a year. This approach might also work to deliver chemotherapy drugs or DNA for gene therapy. *—JR Minkel*

BRAIN AND BEHAVIOR

Double or Nothing

Gamblers often believe that after a string of losses they're due for a win. Scientists now think they have pinpointed areas in the brain that are partly behind this kind of false thinking. Using functional magnetic resonance imaging, investigators at Duke University found a brain region that automatically looks for patterns, real or imagined. When volunteers were shown random sequences of circles and squares, blood flow increased to the prefrontal cortex, which is located just behind the forehead and is involved in memorization during moment-to-moment activity. This brain layer reacted whenever there were violations to apparent short-term patterns in the sequences—even though subjects knew that they were random.

PLACE YOUR BETS: Your brain looks for patterns even when there aren't any.

Meanwhile researchers at the University of Michigan discovered that after losing a simple wager, volunteers were more likely to place larger, riskier bets if prompted to make another wager within a few seconds. Caps studded with electrodes revealed that when subjects learned they had won or lost wagers, electrical activity was highest in the medial frontal cortex, situated behind the prefrontal cortex. The Duke study appears online in the April 8 *Nature Neuroscience;* the Michigan work is in the March 22 *Science*. *—Charles Choi*

JPL/NASA (*top*); TIM FLACH *Stone* (*bottom*)

WWW.SCIAM.COM/NEWS

BRIEF BITS

- Two objects thought to be neutron stars might in fact be **strange quark matter stars**—denser, more exotic stellar objects consisting of strange quarks in addition to the usual up and down quarks. /041202/1.html
- The first drafts of **two rice genomes have been completed**, feats that should lead to hardier and more nutritious strains of one of the world's most important foods. /040502/1.html
- In a clinical trial of 340 patients, **St. John's wort proved to be ineffective** in alleviating moderately severe depression, working no better than a placebo. /041002/1.html
- Despite a 98.7 percent genetic similarity, **humans and chimpanzees are vastly different because of the rate of genetic activity** in the brain—gene expression evolved 5.5 times faster in humans than it did in chimps. /041502/1.html

Thinking Big

A Harvard Medical School dropout aims to usher in the personal-genomics era **By GARY STIX**

Science was king when Eugene Chan was growing up. His father, Ka-Kong Chan, an émigré from Hong Kong, would bring home ball-and-stick models that represented organic molecules—mementos from his job at Hoffmann–La Roche, where he received 40 U.S. patents as a chemist. Outside the home, however, rampant philistinism reigned. Chan's school environment in northwestern New Jersey had such slim science offerings that by the time he headed for the Ivy League, he had never even heard of the Westinghouse Science Talent Search (now the Intel Science Talent Search). Nevertheless, the boy propelled himself to become champion in a statewide physics contest in two separate years by grabbing physics and calculus books off library shelves. "I realized I had a lot of ability and didn't need formal training to compete with the best of the best," Chan remarks with characteristic bravado.

ONE-HOUR GENOME is the goal of U.S. Genomics, launched by Eugene Chan (*left*) and his brother Ian, shown here with a rapid sequencing machine.

JASON GROW

At Harvard his autodidactic skills served him well. He gained top honors, eventually graduating summa cum laude in 1996. But he still found enough time to contemplate the germ of an idea for a technology that would build on the scientific findings of the Human Genome Project, then in its middle phases. "Is it possible for us to gain complete sequence information from every single person on the planet?" he recalls wondering.

Later, at Harvard Medical School, he grew bored after a semester and returned to musing about a device that could read, within an hour or so, the variations in an individual's DNA that mark the essential genetic differences from person to person. During division each cell reads and replicates millions of DNA letters, or base pairs, in the course of a minute. Chan reasoned that a single blood test could be fashioned to achieve the more tractable task of rapidly discerning the variations in a genome, whereas the long unchanging segments of DNA would go unread. The probe would look for groupings of base pairs—several million in one genome—that correspond to a predisposition to disease or the ability to tolerate certain drugs.

Piles of books and journal articles on molecular biology, medical instrumentation, optics and physics covered much of Chan's dormitory room. Borrowing from semiconductor manufacturing and the nascent field of nanotechnology, Chan conceived of placing miniaturized channels on a quartz chip. The DNA, propelled along by a fluid flow, would stream down the channel as if it were a film running through a movie projector. As the DNA moved along, a laser, positioned about halfway down the channel, would illuminate groups of base pairs tagged with a fluorescent dye. Like a bar-code reader, an optoelectronic device would determine which groups lit up and would thus mark genetic variations. To make the test widely available, Chan estimated that it should cost no more than a few hundred dollars.

The concept became such an obsession that, after completing 18 months at Harvard Medical School, Chan left to found U.S. Genomics. His brother Ian, who worked at a lucrative investment-banking job with Morgan Stanley, decided to join him. Chan somehow convinced a prominent Boston intellectual-property law firm, Wolf, Greenfield and Sacks, to write a patent application for him on spec—the firm would be paid once Chan obtained financing. Then, to build credibility, he set about assembling a prominent panel of scientists, which grew to include a Nobel Prize winner. The scientific advisory board would help him gain entrance to the offices of venture capitalists.

The idea of a 23-year-old proposing a wholly new method of sequencing intrigued scientists and engineers on the Harvard–Massachusetts Institute of Technology axis. "I liked it that somebody his age was trying to tackle such a giant problem," says Robert S. Langer, a chemical engineering professor at M.I.T. and a member of the company's scientific advisory board. "If you could do the sequencing that rapidly, that would be a change-the-world kind of thing."

The first venture-capital infusion, a paltry $300,000, came from a Boston-area firm, the Still River Fund. The funding sufficed to rent space at a technology incubator at Boston University and served as an impetus to look for more money. To procure substantial backing, U.S. Genomics would have to show progress in its plan to create a personal-genomics sequencer. "The question people had for us was, 'Can you take that piece of DNA that looks like a big ball of spaghetti and unfurl this thing and move it past your reader device?'" Chan says. "In six months we demonstrated how we could do it."

With the help of five others who joined the newly formed company, Chan fabricated a series of upright posts, each spaced a few tens of nanometers apart, at the mouth of a channel down which the DNA was to travel. The posts snagged the ball of DNA, and the pressure of the molecule against the posts caused it to unravel and stream down the channel toward the optoelectronic detector.

A video that shows the DNA moving along the channel served as a proof-of-principle that allowed the company to return to the venture spigot in 1998 to raise $2 million. That led immediately to the next hurdle—the placement of fluorescent tags on the DNA and the detection of the base-pair groupings as they passed the detector at a rate of 30 million base pairs a second. The expanding U.S. Genomics team spent most of 2000 developing a technique that could train a laser on a two-nanometer spot on the elongated DNA and accurately detect whether the tags illuminate.

Chan claims that the Gene Engine, as the product is called, can spot variations on DNA segments of 200,000 base pairs in length, enough to make the technology commercially alluring. By year's end he wants to expand the readout capacity fivefold. Conventional sequencers evaluate about 1,000 base pairs at a time.

A 23-year-old proposing a wholly new method of DNA sequencing intrigued scientists on the Harvard-M.I.T. axis.

Until now, Chan has had to prove himself by convincing investors that it would be worth their while to lend $20 million to a 20-something medical school dropout, while persuading government and other funding sources to chip in $5 million more. With critical patents issued—and successes in meeting technology milestones—U.S. Genomics will now have to submit to the probing of prospective customers and the scientific community as well. It also faces competitors for rapid genome sequencing. The company, which is housed in virtually unmarked offices in an industrial park in Woburn, Mass., has yet to publish a paper in a scientific journal that details the Gene Engine's performance. But Chan and his brother have initiated a coming out. In January, U.S. Genomics announced that it would enter into a collaboration with a leading sequencing center, the Wellcome Trust Sanger Institute in Cambridge, England, and join in a separate endeavor with the Washington University School of Medicine to test the technology and to start publishing.

The road ahead is still long. Sequencing the variations in 200,000 base pairs is a far easier task than reading a full genome—more than three billion in all. In fact, M.I.T.'s Langer thinks that bioinformatics—milling through the wealth of data generated by reading the base pairs—remains a challenge. Chan is unconcerned. "Ninety percent of the major questions are answered," he says. And he predicts that the company will meet the goal of reading the variations in an entire genome by 2006. Even if that happens, Chan would not, at 32, be ready to rest on his laurels. Processing the information in whole genomes provides sufficient challenge, he contends, to last an entire career. SA

Wanted: More Mothers of Invention

Women patent holders are still a long way from parity with men By GARY STIX

C. D. Tuska, a patent director of RCA, tried to analyze the reasons for the dearth of women inventors in a 1957 book on inventors and invention: "Why is the percentage [of female inventors] so low? I am sure I don't know, unless the good Lord intended them to be mothers. I, being old-fashioned, hold that they are creative enough without also being 'inventive.' They produce the inventors and help rear them, and that should be sufficient."

The perception of the female inventor has changed a bit from the unabashed chauvinism of the Ozzie and Harriet era. In the past decade or so, a spate of books have feted women as something more than nursemaids for young Thomas Edisons-to-be. In the recent *Patently Female: From AZT to TV Dinners,* Ethlie Ann Vare and Greg Ptacek acknowledge the stereotyping by Tuska and others. Then they go to the opposite extreme by elevating women to an exalted status in the annals of human ingenuity: "Can there be any doubt that in the earliest civilizations the gatherers advanced agriculture through invention and innovation while the boys were out hunting? It was most likely a woman who first cultivated a crop, domesticated an animal and fashioned a plow."

One sex cannot claim sole responsibility for the origins of agriculture. But female inventors can point to concrete signs of progress. As recently as the late 1970s and early 1980s, less than 3 percent of patents issued to U.S. residents listed at least one woman's name, not a huge increase from the 1 percent or so that went to women in the period from 1790 to 1895. The ranks of female patentees expanded to 10.3 percent in 1998, however, the U.S. Patent and Trademark Office reported in a study called *Buttons to Biotech*.

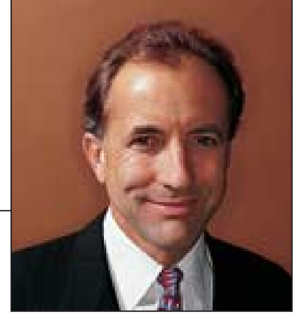
Women staged an especially good showing in obtaining patents for chemical technologies, garnering nearly 16 percent of those patents in 1996. In particular, they were well represented in chemical patents for biotechnology and pharmaceuticals. Vare and Ptacek document a number of prominent recent examples: Janet L. Rideout (AZT), M. Katharine Holloway and Chen Zhao (protease inhibitors), and Diane Pennica (tissue plasminogen activator).

Still, the number of female inventors falls woefully short when compared with other measures: women make up nearly half the workforce, and they play a larger role in science and engineering as a whole than they do at the patent office. The National Science Foundation reported that women represented 24 percent of the science and engineering workforce in 1999, more than twice the percentage of female patentees. "I think that women working in industry aren't in the same leadership positions where they get to do creative work; the leader of the team gets his name on the patent," notes Fred M. B. Amram, a professor emeritus at the University of Minnesota who has studied female patentees and is co-author of the book *From Indian Corn to Outer Space: Women Invent in America*.

An informal survey that Amram conducted of student inventor contests in Minnesota from 1989 to 1994 showed that the proportion of girls from later elementary to early high school who entered the competitions in a given year sometimes exceeded half of the participants. To Amram, this observation suggests that incentives to remain creative in this realm were somehow robbed from them in high school or college. And then the notion of what constitutes women's work became more narrowly cast. SA

Please let us know about interesting and unusual patents. Send suggestions to: patents@sciam.com

JOHN McFAUL



The Shamans of Scientism

On the occasion of Stephen W. Hawking's 60th trip around the sun, we consider a social phenomenon that reveals something deep about human nature By MICHAEL SHERMER

In 1998 God appeared at Caltech.

More precisely, the scientific equivalent of the deity, in the form of Stephen W. Hawking, delivered a public lecture via his now familiar voice synthesizer. The 1,100-seat auditorium was filled; an additional 400 viewed a video feed in another hall, and hundreds more squatted on the lawn and listened to theater speakers broadcasting this scientific saint's epistle to the apostles.

The lecture was slated for 8 P.M. By three o'clock a line began to snake around the grassy quad adjoining the hall. By five, hundreds of scientists flipped Frisbees and chatted with students from Caltech and other universities.

When Hawking rolled into the auditorium and down the aisle in his motorized wheelchair, everyone rose in applause—a "standing O" just for showing up! The sermon was his customary one on the big bang, black holes, time and the universe, with the theology coming in the question-and-answer period. Here was an opportunity to inquire of a transcendent mind the biggest question of all: "Is there a God?"

Asked this ultimately unanswerable question, Hawking sat rigidly in his chair, stone quiet, his eyes darting back and forth across the computer screen. A minute, maybe two, went by. Finally, a wry smile formed and the Delphic oracle spoke: "I do not answer God questions."

What is it about Hawking that draws us to him as a scientific saint? He is, I believe, the embodiment of a larger social phenomenon known as scientism. Scientism is a scientific worldview that encompasses natural explanations for all phenomena, eschews supernatural and paranormal speculations, and embraces empiricism and reason as the twin pillars of a philosophy of life appropriate for an Age of Science.

Scientism's voice can best be heard through a literary genre for both lay readers and professionals that includes the works of such scientists as Carl Sagan, E. O. Wilson, Stephen Jay Gould, Richard Dawkins and Jared Diamond. Scientism is a bridge spanning the abyss between what physicist C. P. Snow famously called the "two cultures" of science and the arts/humanities (neither encampment being able to communicate with the other). Scientism has generated a new literati and intelligentsia passionately concerned with the profound philosophical, ideological and theological implications of scientific discoveries.

Although the origins of the scientism genre can be traced to the writings of Galileo and Thomas Huxley in centuries past, its modern incarnation began in the early 1970s with mathematician Jacob Bronowski's *The Ascent of Man,* took off in the 1980s with Sagan's *Cosmos* and hit pay dirt in the 1990s with Hawking's *A Brief History of Time,* which spent a record 200 weeks on the *Sunday Times* of London's hardcover best-seller list and sold more than 10 million copies in 30-plus languages worldwide. Hawking's latest work, *The Universe in a Nutshell,* is already riding high on the best-seller list.

This being the Age of Science, it is scientism's shamans who command our veneration.

Hawking's towering fame is a result of a concatenation of variables that include the power of the scientism culture in which he writes, his creative insights into the ultimate nature of the cosmos, in which he dares to answer ersatz theological questions, and, perhaps most notably, his unmitigated heroism in the face of near-insurmountable physical obstacles that would have felled a lesser being. But his individual success in particular, and the rise of scientism in general, reveals something deeper still.

First, cosmology and evolutionary theory ask the ultimate origin questions that have traditionally been the province of religion and theology. Scientism is courageously proffering naturalistic answers that supplant supernaturalistic ones and in the process is providing spiritual sustenance for those whose needs are not being met by these ancient cultural traditions. Second, we are, at base, a socially hierarchical primate species. We show deference to our leaders, pay respect to our elders and follow the dictates of our shamans; this being the Age of Science, it is scientism's shamans who command our veneration. Third, because of language we are also storytelling, mythmaking primates, with scientism as the foundational stratum of our story and scientists as the premier mythmakers of our time. SA

Michael Shermer is founding publisher of Skeptic *magazine (www.skeptic.com) and author of* The Borderlands of Science.

BRAD HINES

Man of Two Cultures

As both scientist and administrator, John H. Marburger III tries to bring needed perspective into a White House not thought to be particularly interested in science By GARY STIX

A corner office on the fifth floor of a nondescript building a few blocks from the White House is adorned with large photographs of George W. Bush and Dick Cheney on one wall and an illustration of an American flag on another. Almost nothing decorative conveys the impression that this is the office of the president's science adviser—no scale models of space shuttles, no plastic double helices. Not even a plaster bust of Einstein.

TOM WOLFF

JOHN H. MARBURGER III: SCIENCE ADVISER

- Pioneered the mathematical and physical basis for self-focusing lasers, which are important in nonlinear optical devices and laser fusion.
- Built his own harpsichord and taught himself to play it and the piano.
- "He has such an outgoing and patient personality that graduate students wanted to work with him, even on difficult quantum electronic topics," says Larry G. DeShazer, who received tenure at the University of Southern California based on an experimental problem suggested by Marburger.

But on a small wooden table in the middle of the room sits an object that resembles a modernist sculpture—or the structural framework of a new Frank Gehry museum. Asked about the object, John H. ("Jack") Marburger III lights up. "The thing that's interesting about it is how nonintuitive the shapes are," he marvels. It is a collection of electromagnetic coils for a proposed fusion generator, and the twisted rings do not form the symmetric ovals expected in a series of coils. "No draftsman would ever come up with a design like that for an electromagnetic machine," he adds.

Marburger says he can't remember a time when he didn't hold such a passion for science and technology. The distinguished-looking 61-year-old can recall as a child during World War II how he stared in rapt fascination at a book that showed pictures of snowflakes and industrial processes. But for much of his working life, he has set aside his enthusiasm for physics to devote himself to a career in administration. "My interest is in science, but I do the other things because I feel obligated to do them because I know that I have a talent for getting people to work together and getting over obstacles to get things done," he says.

As a professor at the University of Southern California in the late 1960s and early 1970s, Marburger did research in theoretical physics on the study of quantum electronics and nonlinear optics and co-founded the university's Center for Laser Studies. His leadership abilities propelled him, by 1976, to become a U.S.C. dean and later to serve as president of the State University of New York at Stony Brook from 1980 to 1994. He then returned to teaching for several years. In 1998 he took over as director of Brookhaven National Laboratory on eastern Long Island. The previous management had been fired once it acknowledged belatedly that radioactive tritium had been seeping into groundwater. Marburger set about defraying tensions between the laboratory and local residents. "He built a culture that involved the community, and that just hadn't happened at that facility before," remembers Scott Cullen of the Standing for Truth about Radiation (STAR) Foundation, an Easthampton, N.Y., advocacy group that had fought the laboratory.

Marburger's deft handling of the crisis at Brookhaven gave him a visibility in Washington that made this registered Democrat the Bush administration's choice as science adviser. In its earliest months the administration had taken heat for failing to fill key science-related positions—a gap that became particularly obvious after September 11 and the anthrax incidents. "The connection with the science community had not been activated—that's the way I would phrase it. It was very passive," Marburger says. His attentions during recent months have been noticed. "He's been over here more than any other science adviser, maybe two or three times a week," says Bruce Alberts, president of the National Academy of Sciences, which collaborated with Marburger on a study on counterterrorism. "He's obviously very skilled at getting people to work together."

Marburger was nominated in June, and the Senate confirmed his appointment only in October. Within a few days of his arrival, Tom Ridge called on him to provide technical support to the Office of Homeland Security as it was trying to formulate a strategy to cope with contaminated mail. Marburger quickly brought the U.S. Postal Service together with several high-ranking science experts within the administration. By mid-November the technical team he had assembled had advised the postal service that existing irradiation technologies used for medical and food products would be capable of killing anthrax. Marburger also helped to defuse the overwrought atmosphere by quelling talk about the need for a "Manhattan Project" against terrorism. He believes that most of the basic technologies for detecting and analyzing pathogens, for instance, already exist and just need to be developed into working systems.

Some members of the science establishment fret that Marburger may not have much influence in the administration of a president who, unlike his opponent in the 2000 election, has not shown a great fascination for science and technology. The administration stripped the science adviser of the title "assistant to the president," fueling worries that Marburger, as science adviser and director of the Office of Science and Technology Policy, would have difficulty getting the president's ear. Constantly asked about this, Marburger wearily dismisses these concerns. "When the president needs science advice on a matter where science plays an important role in the decision, I'm present. I'm there. I'm part of the team that briefs him on the issues." Though not a science aficionado, Bush uses science "appropriately," Marburger says, weighing it as one of multiple factors in arriving at a decision. "Is President George W. Bush like Al Gore?" Marburger asks. "Definitely no. He is not, and I think that science has by no means suffered as a result."

"Is President George W. Bush like Al Gore? Definitely no. He is not, and I think that science has by no means suffered as a result."

Marburger's presence will not necessarily cause any fundamental shift in the administration's positions on controversial issues, such as limiting research on embryonic stem cells. "The president understands that he had to make a moral decision [on stem cells], not a science decision," Marburger notes. "Science doesn't tell you what you ought to do. What you ought to do depends on your moral principles, and I don't advise the president on moral principles." But Marburger does see his office serving as a faux pas detector, helping to avoid the awkward misstatements by administration officials that preceded his arrival. If he had been on board early in the administration, Marburger could have advised White House officials against making remarks contending that no scientific consensus had emerged on the contribution of human activity to global warming.

Marburger has provided intellectual firepower to defend the administration's position on issues such as nuclear waste storage at Yucca Mountain in Nevada, where the Energy Department wants to store spent radioactive materials. "I personally believe that the science is immensely strong and that the case for not moving forward with the Yucca Mountain program is weak," Marburger says.

He also has to explain the reasons for the haves and have-nots in the federal budget. The proposed 2003 research budget for the National Institutes of Health, for example, now totals almost as much as those of all the other civilian science agencies combined. Marburger has crafted an intricate rationale to justify the perceived lopsidedness, one that pays tribute to the physical sciences while still delivering the money to the NIH: extraordinary advances in instrumentation and information processing—hand-me-downs from physicists and chemists—will enable nanotechnological techniques that will yield large payoffs in medical research. "Given the new atomic-level capabilities, the life sciences may still be underfunded relative to the physical sciences," he said in February in a speech at the annual meeting of the American Association for the Advancement of Science.

When he completes his tenure with the administration, Marburger wants to return to teaching and studying physics. But being science adviser has allowed him to achieve a certain balance in his career. "I have more contact with scientists in this job than I did in previous jobs that I've held because I have fewer management responsibilities, and I have a greater responsibility to interpret science and to translate science into action." The position combines, better than any other administrative slot he has occupied, both his passion for science and his self-imposed obligation to engage in public service. SA

HOPE IN A

VIAL

Will there be an AIDS vaccine anytime soon?

By Carol Ezzell

HANK MORGAN *TimePix*

BOTTLES of a potential AIDS vaccine await use in human tests.

It wasn't supposed to be this hard. When HIV, the virus responsible for AIDS, was first identified in 1984, Margaret M. Heckler, then secretary of the U.S. Department of Health and Human Services, predicted that a vaccine to protect against the scourge would be available within two years. Would that it had been so straightforward.

Roughly 20 years into the pandemic, 40 million people on the planet are infected with HIV, and three million died from it last year (20,000 in North America). Although several potential AIDS vaccines are in clinical tests, so far none has lived up to its early promise. Time and again researchers have obtained tantalizing preliminary results only to run up against a brick wall later. As recently as two years ago, AIDS researchers were saying privately that they doubted whether even a partially protective vaccine would be available in their lifetime.

No stunning breakthroughs have occurred since that time, but a trickle of encouraging data is prompting hope to spring anew in the breasts of even jaded AIDS vaccine hunters. After traveling down blind alleys for more than a decade, they are emerging battered but not beaten, ready to strike out in new directions. "It's an interesting time for AIDS vaccine research," observes Gregg Gonsalves, director of treatment and prevention advocacy for Gay Men's Health Crisis in New York City. "I feel like it's Act Two now."

In the theater, Act One serves to introduce the characters and set the scene; in Act Two, conflict deepens and the real action begins. Act One of AIDS vaccine research debuted HIV, one of the first so-called retroviruses to cause a serious human disease. Unlike most other viruses, retroviruses insinuate their genetic material into that of the body cells they invade, causing the viral genes to become a permanent fixture in the infected cells and in the offspring of those cells. Retroviruses also reproduce rapidly and sloppily, providing ample opportunity for the emergence of mutations that allow HIV to shift its identity and thereby give the immune system or antiretroviral drugs the slip.

Act One also spotlighted HIV's opposition—the body's immune response—which consists of antibodies (Y-shaped molecules that stick to and tag invaders such as viruses for destruction) and cytotoxic, or killer, T cells (white blood cells charged with destroying virus-infected cells). For years after infection, the immune system battles mightily against HIV, pitting millions of new cytotoxic T cells against the billions of virus particles hatched from infected cells every day. In addition, the immune system deploys armies of antibodies targeted at HIV, at least early in the course of HIV infection, although the antibodies prove relatively ineffectual against this particular foe.

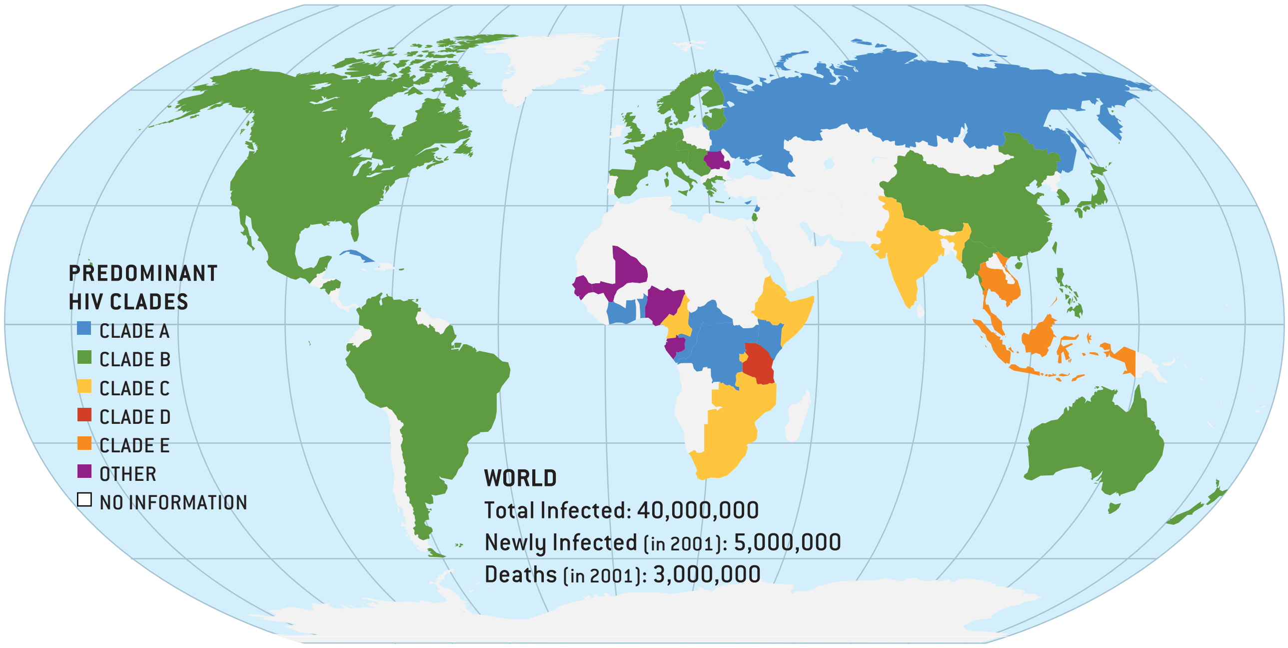
As the curtain rises for Act Two, HIV still has the stage. Results from the first large-scale trial of an AIDS vaccine should become available at the end of this year, but few scientists are optimistic about it: a preliminary analysis suggests that it works poorly. Meanwhile controversy surrounds a giant, U.S.-government-sponsored trial of another potential vaccine slated to begin this September in Thailand. But waiting in the wings are several approaches that are causing the AIDS research community to sit up and take notice. The strategies are reviving the debate about whether, to be useful, a vaccine must elicit immune responses that totally prevent HIV from colonizing a person's cells or whether a vaccine that falls somewhat short of that mark could be accept-

Overview/*AIDS Vaccines*

- Final results from the first large-scale test of a possible AIDS vaccine will be available at the end of this year, but few researchers are optimistic it will work.
- Scientists are now aiming to generate potential AIDS vaccines that stimulate both arms of the immune system: killer cells and antibodies.
- There are five main subtypes, or clades, of HIV. Researchers are debating whether it will be important to devise vaccines for a given area based on the predominant clade infecting that area.

WORLD AIDS SNAPSHOT

MOST OF THE GLOBE'S 40 million people infected with HIV live in sub-Saharan Africa and South and Southeast Asia, as reflected in the ranking below, which is based on 2001 data from the Joint United Nations Program on HIV/AIDS. There are five major strains of HIV, which are also called clades. Although more than one clade can usually be found in any given area, the map highlights the predominant clade affecting each region. The boundaries between prevailing clades are not exact; they change frequently.



1 SUB-SAHARAN AFRICA
Total Infected: 28,100,000
Newly Infected: 3,400,000
Deaths: 2,300,000

2 SOUTH/SOUTHEAST ASIA
Total Infected: 6,100,000
Newly Infected: 800,000
Deaths: 400,000

3 LATIN AMERICA
Total Infected: 1,400,000
Newly Infected: 130,000
Deaths: 80,000

4 EAST ASIA/PACIFIC IS.
Total Infected: 1,000,000
Newly Infected: 270,000
Deaths: 35,000

5 E. EUROPE/C. ASIA
Total Infected: 1,000,000
Newly Infected: 250,000
Deaths: 23,000

6 NORTH AMERICA
Total Infected: 940,000
Newly Infected: 45,000
Deaths: 20,000

7 WESTERN EUROPE
Total Infected: 560,000
Newly Infected: 30,000
Deaths: 6,800

8 N. AFRICA/MIDDLE EAST
Total Infected: 440,000
Newly Infected: 80,000
Deaths: 30,000

9 CARIBBEAN
Total Infected: 420,000
Newly Infected: 60,000
Deaths: 30,000

10 AUSTRALIA/NEW ZEALAND
Total Infected: 15,000
Newly Infected: 500
Deaths: 120

LAURIE GRACE; SOURCES: UNAIDS (*statistics*) AND VADIM ZALUNIN *Los Alamos National Laboratory* (*clade boundaries*)

able. Some scientists see potential value in vaccines that would elicit the kinds of immune responses that kick in soon after a virus establishes a foothold in cells. By constraining viral replication more effectively than the body's natural responses would, such vaccines, they argue, might at least help prolong the lives of HIV-infected people and delay the onset of the symptomatic, AIDS phase of the disease.

In the early 1990s scientists thought they could figure out the best vaccine strategy for preventing AIDS by studying long-term nonprogressors, people who appeared to have harbored HIV for a decade or more but who hadn't yet fallen ill with AIDS. Sadly, many of the nonprogressors have become ill after all. The key to their relative longevity seems to have been "a weakened virus and/or a strengthened immune system," says John P. Moore of Weill Medical College of Cornell University. In other words, they were lucky enough to have encountered a slow-growing form of HIV at a time when their bodies had the ammunition to keep it at bay.

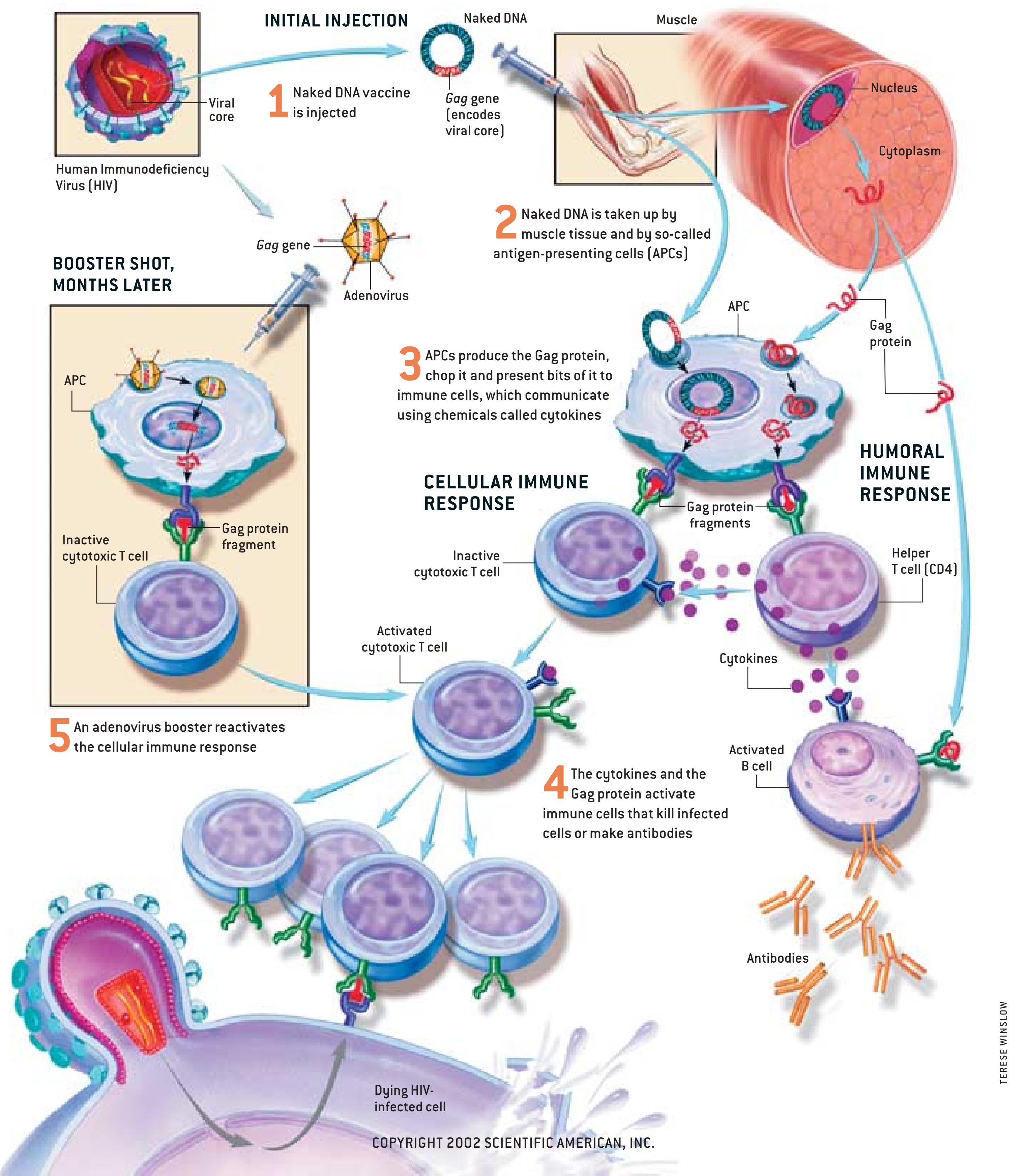
Not Found in Nature?

AIDS VACCINE developers have struggled for decades to find the "correlates of immunity" for HIV—the magic combination of immune responses that, once induced by a vaccine, would protect someone against infection. But they keep coming up empty-handed, which leaves them with no road map to guide them in the search for an AIDS vaccine. "We're trying to elicit an immune response not found in nature," admits Max Essex of the Harvard School of Public Health. As a result, the quest for an AIDS vaccine has been a bit scattershot.

To be proved useful, a candidate AIDS vaccine must successfully pass through three stages of human testing. In phase I, researchers administer the vaccine to dozens of people to assess its safety and to establish an appropriate dose. Phase II involves hundreds of people and looks more closely at the vaccine's im-

One AIDS Vaccine Strategy

A VACCINE APPROACH being pioneered by Merck involves an initial injection of a naked DNA vaccine followed months later by a booster shot of crippled, genetically altered adenovirus particles. Both are designed to elicit an immune response targeted to the HIV core protein, Gag, and to primarily arouse the cellular arm of the immune system—the one that uses cytotoxic T cells to destroy virus-infected cells. The naked DNA vaccine also results in the production of antibody molecules against Gag, but such antibodies are not very useful in fighting HIV.



munogenicity, its ability to prompt an immune response. In phase III, the potential vaccine is given to thousands of volunteers who are followed for a long time to see whether it protects them from infection. Phase III trials for any drug tend to be costly and difficult to administer. And the AIDS trials are especially challenging because of an ironic requirement: subjects who receive the vaccine must be counseled extensively on how to reduce their chances of infection. They are told, for instance, to use condoms or, in the case of intravenous drug users, clean needles because HIV is spread through sex or blood-to-blood contact. Yet the study will yield results only if some people don't heed the counseling and become exposed anyway.

The first potential vaccine to have reached phase III consists of gp120, a protein that studs the outer envelope of HIV and that the virus uses to latch onto and infect cells. In theory, at least, the presence of gp120 in the bloodstream should activate the recipient's immune system, causing it to quickly mount an attack targeted to gp120 if HIV later finds its way into the body.

This vaccine, which is produced by VaxGen in Brisbane, Calif.—a spin-off of biotech juggernaut Genentech in South San Francisco—is being tested in more than 5,400 people (mostly homosexual men) in North America and Europe and in roughly 2,500 intravenous drug users in Southeast Asia. The results from the North American/European trial, which began in 1998, are expected to be announced near the end of this year.

Many AIDS researchers are skeptical of VaxGen's approach because gp120 normally occurs in clumps of three on the surface of the virus, and the company's vaccine employs the molecule in its monomeric, or single-molecule, form. Moreover, vaccines made of just protein generally elicit only an antibody, or humoral, response, without greatly stimulating the cellular arm of the immune system, the part that includes activity by cytotoxic T cells. A growing contingent of investigators suspect that an antibody response alone is not sufficient; a strong cellular response must also be elicited to prevent AIDS.

Indeed, the early findings do not seem encouraging. Last October an independent data-monitoring panel did a preliminary analysis of the results of the North American/European data. Although the panel conducted the analysis primarily to ascertain that the vaccine was causing no dangerous side effects in the volunteers, the reviewers were empowered to recommend halting the trial early if the vaccine appeared to be working. They did not.

For its part, VaxGen asserts that it will seek U.S. Food and Drug Administration approval to sell the vaccine even if the phase III trials show that it reduces a person's likelihood of infection by as little as 30 percent. Company president and co-founder Donald P. Francis points out that the first polio vaccine, developed by Jonas Salk in 1954, was only 60 percent effective, yet it slashed the incidence of polio in the U.S. quickly and dramatically.

This approach could backfire, though, if people who receive a partially effective AIDS vaccine believe they are then protected from infection and can engage in risky behaviors. Karen M. Kuntz and Elizabeth Bogard of the Harvard School of Public Health have constructed a computer model simulating the effects of such a vaccine in a group of injection drug users in Thailand. According to their model, a 30 percent effective vaccine would not slow the spread of AIDS in a community if 90 percent of the people who received it went back to sharing needles or using dirty needles. They found that such reversion to risky behavior would not wash out the public health benefit if a vaccine were at least 75 percent effective.

The controversial study set to begin in Thailand is also a large-scale phase III trial, involving nearly 16,000 people. It combines the VaxGen vaccine with a canarypox virus into which scientists have stitched genes that encode gp120 as well as two other proteins—one that makes up the HIV core and one that allows it to reproduce. Because this genetically engineered canarypox virus (made by Aventis Pasteur, headquartered in Lyons, France) enters cells and causes them to display fragments of HIV on their surface, it stimulates the cellular arm of the immune system.

Political wrangling and questions over its scientific value have slowed widespread testing of the gp120/canarypox vaccine. Initially the National Institute of Allergy and Infectious Diseases (NIAID) and the U.S. Department of Defense were scheduled to conduct essentially duplicate trials of the vaccine. But NIAID pulled the plug on its trial after an examination of the data from a phase II study showed that fewer than 30 percent of the volunteers generated cytotoxic T cells against HIV. And in a bureaucratic twist, this past January the White House transferred the budget for the Defense Department trial over to NIAID as part of an effort to streamline AIDS research.

Peggy Johnston, assistant director of AIDS vaccines for NIAID, says she expects there will be a trial of the vaccine but emphasizes that "it will be a Thai trial; we won't have any [NIAID] people there on the ground running things."

Critics cite these machinations as a case study of politics getting in the way of progress against AIDS. "There's little science involved" in the trial, claims one skeptic, who wonders why the Thais aren't asking, "'If it's not good enough for America, how come it's good enough for us?'" Others point out that the trial, which was conceived by the Defense Department, will answer only the question of whether the vaccine works; it won't collect any data that scientists could use to explain its potential failure.

Results from the first large-scale trial of an AIDS vaccine should become available at the end of this year.

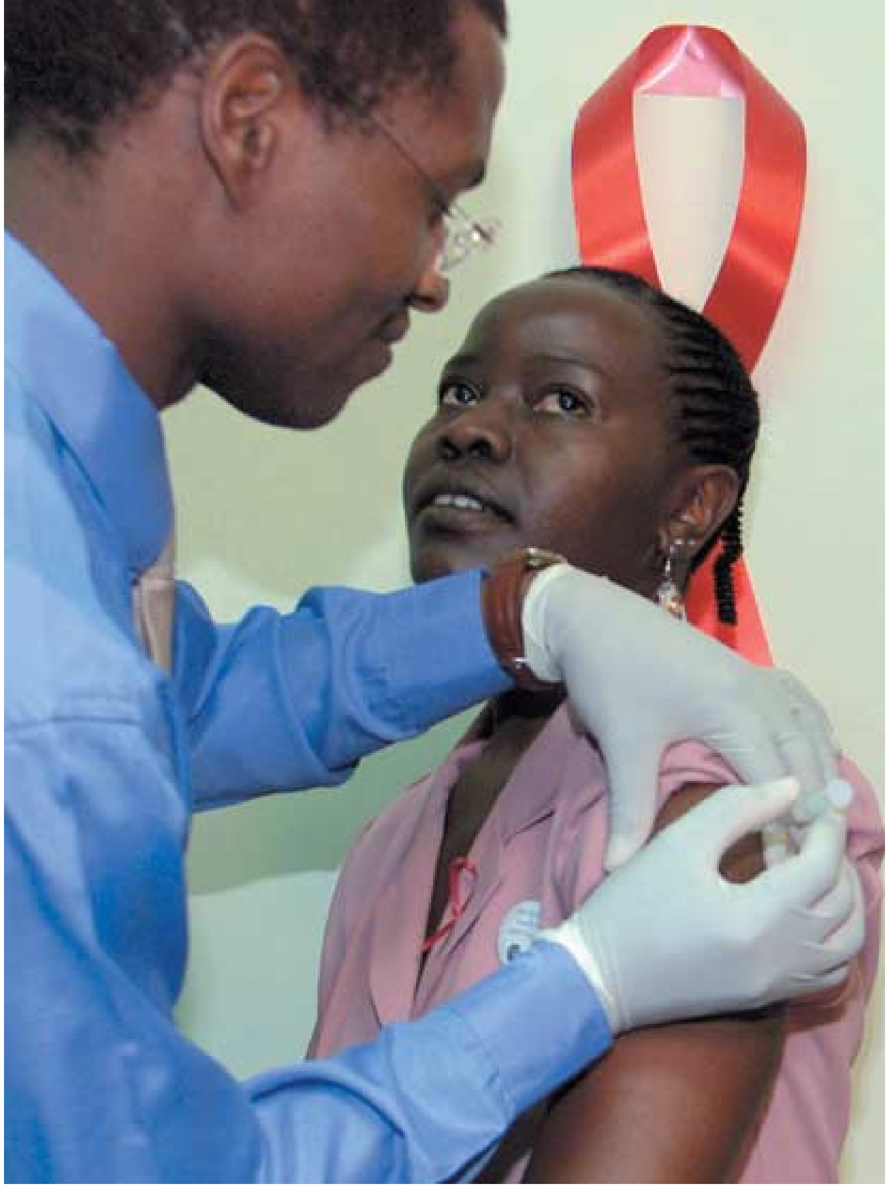
Partial Protection

INTO THIS SCENE comes Merck, which is completing separate phase I trials of two different vaccine candidates that it has begun to test together. In February, Emilio A. Emini, Merck's senior vice president for vaccine research, wowed scientists attending the Ninth Conference on Retroviruses and Opportunistic Infections in Seattle with the company's initial data from the two trials.

The first trial is investigating a potential vaccine composed of only the HIV *gag* gene, which encodes the virus's core protein. It is administered as a so-called naked DNA vaccine, consisting solely of DNA. Cells take up the gene and use it as a blueprint for making the viral protein, which in turn stimulates a mild (and probably unhelpful) humoral response and a more robust cellular response [*see illustration on page 42*]. Emini and his colleagues reported that 42 percent of volunteers who received the highest dose of the naked DNA vaccine raised cytotoxic T cells capable of attacking HIV-infected cells.

The second trial employs the HIV *gag* gene spliced into a crippled adenovirus, the class responsible for many common colds. This altered adenovirus ferries the *gag* gene into cells, which then make the HIV core protein and elicit an immune response targeted to that protein. Emini told the conference that between 44 and 67 percent of people who received injections of the adenovirus-based vaccine generated a cellular immune response that varied in intensity according to the size of the dose the subjects received and how long ago they got their shots.

Merck is now beginning to test a combination of the DNA and adenovirus approaches because Emini predicts that the vaccines will work best when administered as part of the same regimen. "The concept," he says, "is not that the DNA vaccine will be a good vaccine on its own, but that it may work as a primer of the immune system," to be followed months later by a booster shot of the adenovirus vaccine. A possible stumbling block is that most people have had colds caused by adenoviruses. Accordingly, the immune systems of such individuals would already have an arsenal in place that could wipe out the adenovirus vaccine before it had a chance to deliver its payload of HIV genes and stimulate AIDS immunity. Increasing the dose of the adenovirus vaccine could get around this obstacle.



VOLUNTEER in Kenya receives an injection as part of an AIDS vaccine study in that country.

ANTONY NJUGUNA *Reuters/TimePix*

Emini says he and his co-workers are emphasizing cellular immunity in part because of the disappointing results so far with vaccines designed to engender humoral responses. "Antibodies continue to be a problem," he admits. "There are a handful of reasonably potent antibodies isolated from HIV-infected people, but we haven't figured out how to raise those antibodies using a vaccine."

Lawrence Corey of the Fred Hutchinson Cancer Research Center in Seattle agrees: "You'd like to have both [a cellular and an antibody response], but the greatest progress has been in eliciting a cellular response," says Corey, who is also principal investigator of the federally funded HIV Vaccine Trials Network.

Antibodies are important, too, because they are the immune system's first line of defense and are thought to be the key to preventing viruses from ever contacting the cells they infect. Corey says that vaccines that are designed primarily to evoke cellular immunity (as are Merck's) are not likely to prevent infection but should give someone a head start in combating the virus if he or she does become infected. "Instead of progressing to AIDS in eight years, you progress in 25 years," he predicts. But, Corey adds, it is unclear whether a vaccine that only slowed disease progression would stem the AIDS pandemic, because people would still be able to spread the infection to others despite having less virus in their bloodstream.

Finding a way to induce the production of antibodies able to neutralize HIV has been hard slogging for several reasons. For one, the virus's shape-shifting ways allow it to stay one step ahead of the immune response. "The thing that distinguishes HIV from all other human viruses is its ability to mutate so fast," Essex says. "By the time you make a neutralizing antibody [against HIV], it is only against the virus that was in you a month ago."

According to many scientists, vaccines using a logical molecule, gp120—the protein the virus uses to invade immune cells, as discussed above—haven't worked, probably because the antibodies that such vaccines elicit bind to the wrong part of the molecule. Gp120 shields the precise binding site it uses to latch onto CD4, its docking site on immune cells, until the last nanosecond, when it snaps open like a jackknife. One way to get around this problem, suggested in a paper published in *Science* three years ago by Jack H. Nunberg of the University of Montana and his colleagues, would be to make vaccines of gp120 molecules that have previously been exposed to CD4 and therefore have already sprung open. But those results have been "difficult to replicate," according to Corey, making researchers pessimistic about the approach.

Another possible hurdle to getting an AIDS vaccine that elicits effective anti-HIV antibodies is the variety of HIV subtypes, or clades, that affect different areas of the world. There are five major clades, designated A through E [*see illustration on page 41*]. Although clade B is the predominant strain in North America and Europe, most of sub-Saharan Africa—the hardest-hit region of the globe—has clade C. The ones primarily responsible for AIDS in South and Southeast Asia—the second biggest AIDS hot spot—are clades B, C and E.

Several studies indicate that antibodies that recognize AIDS viruses from one clade might not bind to viruses from other clades, suggesting that a vaccine made from the strain found in the U.S. might not protect people in South Africa, for example. But scientists disagree about the significance of clade differences and whether only strains that match the most prevalent clade in a given area can be tested in countries there. Essex, who is gearing up to lead phase I tests of a clade C–based vaccine in Botswana later this year, argues that unless researchers are sure that a vaccine designed against one clade can cross-react with viruses from another, they must stick to testing vaccines that use the clade prevalent in the populations being studied. Cross-reactivity could occur under ideal circumstances, but, he says, "unless we know that, it's important for us to use subtype-specific vaccines."

Using the corresponding clade also avoids the appearance that people in developing countries are being used as guinea pigs for testing a vaccine that is designed to work only in the U.S. or Europe. VaxGen's tests in Thailand are based on a combination of clades B and E, and in April the International AIDS Vaccine Initiative expanded tests of a clade A–derived vaccine in Kenya, where clade A is found.

But in January, Malegapuru William Makgoba and Nandipha Solomon of the Medical Research Council of South Africa, together with Timothy Johan Paul Tucker of the South African AIDS Vaccine Initiative, wrote in the *British Medical Journal* that the relevance of HIV subtypes "remains unresolved." They assert that clades "have assumed a political and national importance, which could interfere with important international trials of efficacy."

Early data from the Merck vaccine trials suggest that clade differences blur when it comes to cellular immunity. At the retrovirus conference in February, Emini reported that killer cells from 10 of 13 people who received a vaccine based on clade B also reacted in laboratory tests to viral proteins from clade A or C viruses. "There is a potential for a substantial cross-clade response" in cellular immunity, he says, "but that's not going to hold true for antibodies." Corey concurs that clade variation "is likely to play much, much less of a role" for killer cells than for antibodies because most cytotoxic T cells recognize parts of HIV that are the same from clade to clade.

Johnston of NIAID theorizes that one answer would be to use all five major clades in every vaccine. Chiron in Emeryville, Calif., is developing a multiclade vaccine, which is in early clinical trials. Such an approach could be overkill, however, Johnston says. It could be that proteins from only one clade would be recognized "and the other proteins would be wasted," she warns.

Whatever the outcome on the clade question, Moore of Weill Medical College says he and fellow researchers are more hopeful than they were a few years ago about their eventual ability to devise an AIDS vaccine that would elicit both killer cells and antibodies. "The problem is not impossible," he says, "just extremely difficult." SA

"We're trying to elicit an immune response not found in nature." —*Max Essex, Harvard School of Public Health*

Carol Ezzell is a staff editor and writer.

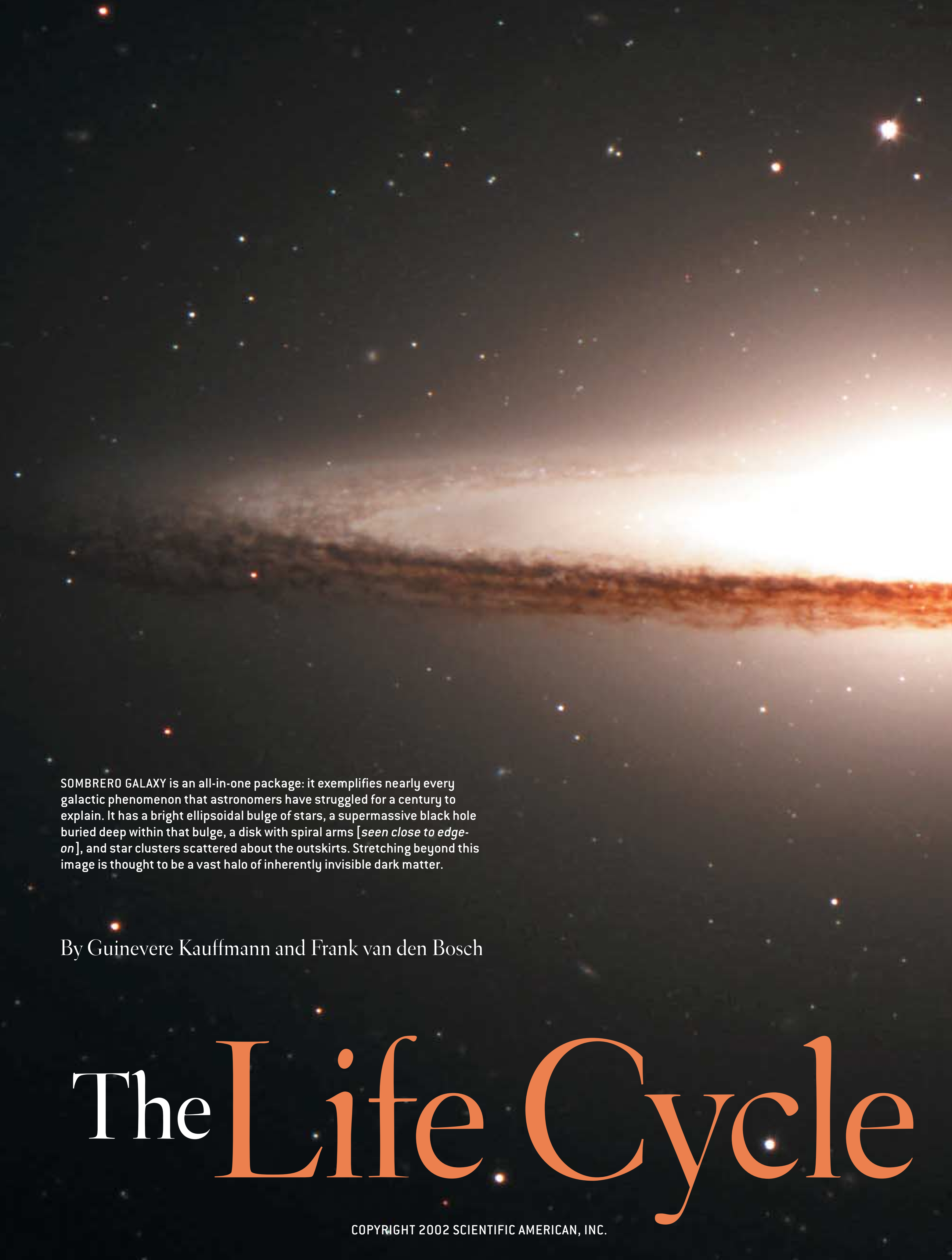
MORE TO EXPLORE

HIV Vaccine Efforts Inch Forward. Brian Vastag in *Journal of the American Medical Association,* Vol. 286, No. 15, pages 1826–1828; October 17, 2001.

For an overview of AIDS vaccine research, including the status of U.S.-funded AIDS clinical trials, visit **www.niaid.nih.gov/daids/vaccine/default.htm**

A global perspective on the AIDS pandemic and the need for a vaccine can be found at the International AIDS Vaccine Initiative Web site: **www.iavi.org**

Joint United Nations Program on HIV/AIDS: **www.unaids.org**



SOMBRERO GALAXY is an all-in-one package: it exemplifies nearly every galactic phenomenon that astronomers have struggled for a century to explain. It has a bright ellipsoidal bulge of stars, a supermassive black hole buried deep within that bulge, a disk with spiral arms [*seen close to edge-on*], and star clusters scattered about the outskirts. Stretching beyond this image is thought to be a vast halo of inherently invisible dark matter.

By Guinevere Kauffmann and Frank van den Bosch

The Life Cycle

Astronomers are on the verge of explaining the enigmatic variety of galaxies

# of Galaxies

In many science-fiction stories, a mighty empire dooms itself through its hubris: it presumes to conquer and rule an entire galaxy. That seems a lofty ambition indeed. To bring our Milky Way galaxy to heel, an empire would have to vanquish 100 billion stars. But cosmologists—those astronomers who study the universe as a whole—are unimpressed. The Milky Way is one of 50 billion or more galaxies within the observable reaches of space. To conquer it would be to conquer an insignificant speck.

A century ago nobody knew all those galaxies even existed. Most astronomers thought that the galaxy and the universe were synonymous. Space contained perhaps a billion stars, interspersed with fuzzy splotches that looked like stars in the process of forming or dying. Then, in the early decades of the 20th century, came the golden age of astronomy, when American astronomer Edwin Hubble and others determined that those fuzzy splotches were often entire galaxies in their own right.

Why do stars reside in gigantic agglomerations separated by vast voids, and how do galaxies take on their bewildering variety of shapes, sizes and masses? These questions have consumed astronomers for decades. It is not possible for us to observe a galaxy forming; the process is far too slow. Instead researchers have to piece the puzzle together by observing many different galaxies, each caught at a different phase in its evolutionary history. Such measurements did not become routine until about a decade ago, when astronomy entered a new golden age.

Spectacular advances in telescope and detector technology are now giving astronomers a view of how galaxies have changed over cosmic timescales. The Hubble Space Telescope has taken very deep snapshots of the sky, revealing galaxies down to unprecedentedly faint levels. Ground-based instruments such as the giant Keck telescopes have amassed statistics on distant (and therefore ancient) galaxies. It is as if evolutionary biologists had been handed a time machine, allowing them to travel back into prehistory and take pictures of the animals and plants inhabiting the earth at a series of different epochs. The challenge for astronomers, as it would be for the biologists, is to grasp how the species observed at the earliest times evolved into what we know today.

The task is of truly astronomical proportions. It involves physics on wildly disparate scales, from the cosmological evolution of the entire universe to the formation of a single star. That makes it difficult to build realistic models of galaxy formation, yet it brings the whole subject full circle. The discovery of all those billions of galaxies made stellar astronomy and cosmology seem mutually irrelevant. In the grand scheme of things, stars were just too small to matter; conversely, debates over the origin of the universe struck most stellar astronomers as hopelessly abstract. Now we know that a coherent picture of the universe must take in both the large and the small.

Overview/*Galaxy Evolution*

- One of the liveliest subfields of astrophysics right now is the study of how galaxies take shape. Telescopes are probing the very earliest galaxies, and computer simulations can track events in unprecedented detail.
- Researchers may soon do for galaxies what they did for stars in the early 20th century: provide a unified explanation, based on a few general processes, for a huge diversity of celestial bodies. For galaxies, those processes include gravitational instability, radiative cooling, relaxation (whereby galaxies reach internal equilibrium) and interactions among galaxies.
- Several vexing questions remain, however. A possible answer to these questions is that stars, seemingly insignificant to such large bodies as galaxies, actually have a profound and pervasive effect on their structure.

Galactic Species

TO UNDERSTAND HOW galaxies form, astronomers look for patterns and trends in their properties. According to the classification scheme developed by Hubble, galaxies may be broadly divided into three major types: elliptical, spiral and irregular [*see illustration on opposite page*]. The most massive ones are the ellipticals. These are smooth, featureless, almost spherical systems with little or no gas or dust. In them, stars buzz around the center like bees around a hive. Most of the stars are very old.

Spiral galaxies, such as our own Milky Way, are highly flattened and organized structures in which stars and gas move on circular or near-circular orbits around the center. In fact, they are also known as disk galaxies. The pinwheel-like spiral arms are filaments of hot young stars, gas and dust. At their centers, spiral galaxies contain bulges—spheroidal clumps of stars that are reminiscent of miniature elliptical galaxies. Roughly a third of spiral galaxies have a rectangular structure toward the cen-

EUROPEAN SOUTHERN OBSERVATORY/BARTHEL/NEESER (*preceding pages*)

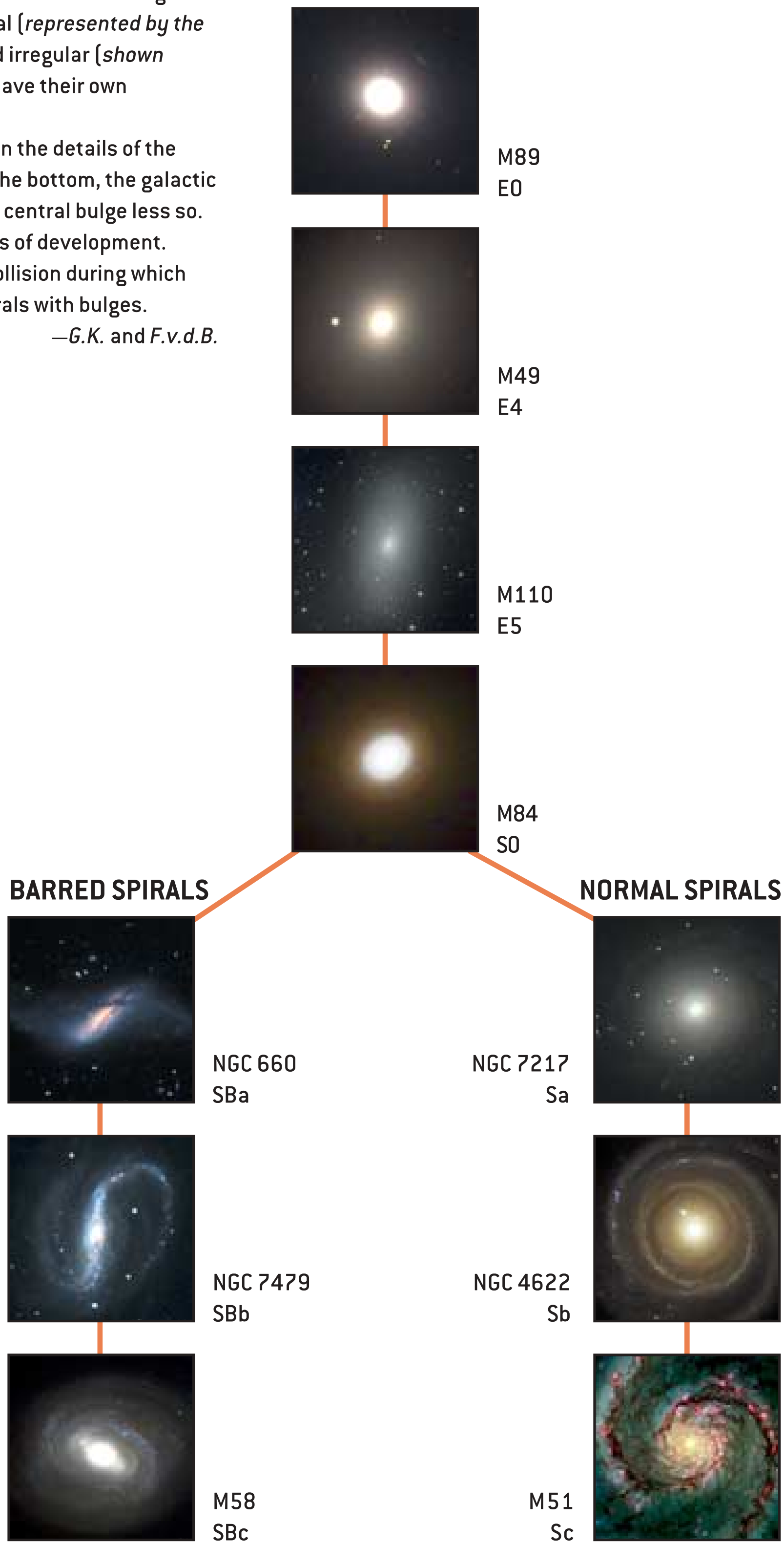
TYPES OF GALAXIES

ASTRONOMERS SORT GALAXIES using the "tuning fork" classification scheme developed by American astronomer Edwin Hubble in the 1920s. According to this system, galaxies come in three basic types: elliptical (*represented by the handle of the fork at right*), spiral (*shown as prongs*) and irregular (*shown below at left*). The smallest galaxies, known as dwarfs, have their own uncertain taxonomy.

Within each of the types are subtypes that depend on the details of the galaxy's shape. Going from the top of the tuning fork to the bottom, the galactic disk becomes more prominent in optical images and the central bulge less so. The different Hubble types may represent various stages of development. Galaxies start off as spirals without bulges, undergo a collision during which they appear irregular, and end up as ellipticals or as spirals with bulges.

—*G.K.* and *F.v.d.B.*

ELLIPTICALS

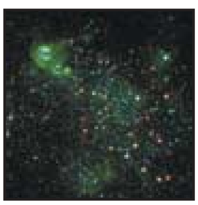


IRREGULARS

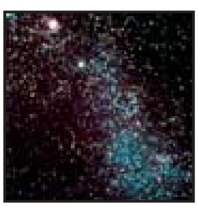
M82
Irregular

DWARF TYPES

M32
Elliptical



VII Zw 403
Blue Compact

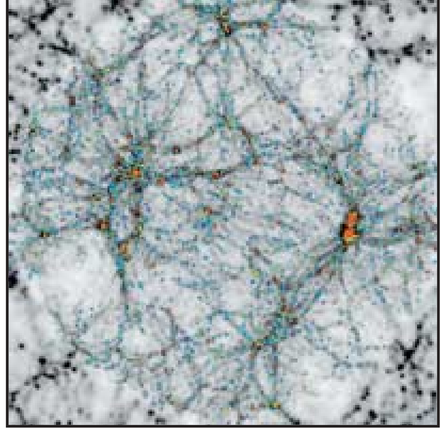


Small Magellanic Cloud
Irregular

Leo I
Spheroidal

N. A. SHARP/NOAO/AURA/NSF (*M82*); B. KEEL/HALL TELESCOPE/LOWELL OBSERVATORY (*M32*); R. SCHULTE-LADBECK/U. HOPP/M. CRONE/*ASTROPHYSICAL JOURNAL* (*blue compact dwarf*); NOAO/AURA/NSF (*Small Magellanic Cloud*); DAVID MALIN, © ANGLO-AMERICAN OBSERVATORY (*Leo I*); NOAO/AURA/NSF (*M89, M49, M110, M84*); R. BRANCH/R. MILNER/A. BLOCK/NOAO/AURA/NSF (*NGC 660*); A. BLOCK/NOAO/AURA/NSF (*NGC 7479*); F. CIESLAK/A. BLOCK/NOAO/AURA/NSF (*M58*); B. KEEL/R. BUTA/G. PURCELL/CERRO TOLOLO INTER-AMERICAN OBSERVATORY, CHILE (*NGC 7217*); G. BYRD/R. BUTA/T. FREEMAN/NASA (*NGC 4622*); NASA/STSCI/AURA (*M51*)

Supercomputer simulations of the spatial distribution of galaxies are in excellent agreement with observations.



H. MATHIS, V. SPRINGEL, G. KAUFFMANN AND S.D.M. WHITE *Max Planck Institute for Astrophysics, Garching, Germany,* AND G. LEMSON, A. ELDAR AND A. DEKEL *Hebrew University, Israel (simulation of galaxy formation in a region 900 million light-years across)*

ter. Such "bars" are thought to arise from instabilities in the disk.

Irregular galaxies are those that do not fit into the spiral or elliptical classifications. Some appear to be spirals or ellipticals that have been violently distorted by a recent encounter with a neighbor. Others are isolated systems that have an amorphous structure and exhibit no signs of any recent disturbance.

Each of these three classes covers galaxies with a wide range of luminosities. On average, however, ellipticals are brighter than spirals, and fainter galaxies are more likely than their luminous counterparts to be irregular. For the faintest galaxies, the classification scheme breaks down altogether. These dwarf galaxies are heterogeneous in nature, and attempts to pigeonhole them have proved controversial. Loosely speaking, they fall into two categories: gas-rich systems where stars are actively forming and gas-poor systems where no stars are forming.

An important clue to the origin of the galaxy types comes from the striking correlation between type and local galaxy density. Most galaxies are scattered through space far from their nearest neighbor, and of these only 10 to 20 percent are ellipticals; spirals dominate. The remaining galaxies, however, are packed into clusters, and for them the situation is reversed. Ellipticals are the majority, and the spirals that do exist are anemic systems depleted of gas and young stars. This so-called morphology-density relation has long puzzled astronomers.

Light and Dark

A SMALL PERCENTAGE of spirals and ellipticals are peculiar in that they contain an exceedingly luminous, pointlike core—an active galactic nucleus (AGN). The most extreme and rarest examples are the quasars, which are so bright that they completely outshine their host galaxies. Astronomers generally believe that AGNs are powered by black holes weighing millions to billions of solar masses. Theory predicts that gas falling into these monsters will radiate about 10 percent of its intrinsic energy, sufficient to generate a beacon that can be detected on the other side of the universe.

Once considered anomalies, AGNs have recently been shown to be integral to the process of galaxy formation. The peak of AGN activity occurred when the universe was approximately a fourth of its present age—the same time that most of the stars in ellipticals were being formed. Furthermore, supermassive black holes are now believed to reside in virtually every elliptical galaxy, as well as every spiral galaxy that has a bulge, regardless of whether those galaxies contain an AGN [see "The Hole Shebang," by George Musser; News and Analysis, SCIENTIFIC AMERICAN, October 2000]. The implication is that every galaxy may go through one or more episodes of AGN activity. As long as matter falls into the black hole, the nucleus is active. When no new material is supplied to the center, it lies dormant.

Most of the information we have about all these phenomena comes from photons: optical photons from stars, radio photons from neutral hydrogen gas, x-ray photons from ionized gas. But the vast majority of the matter in the universe may not emit photons of any wavelength. This is the infamous dark matter, whose existence is inferred solely from its gravitational effects. The visible parts of galaxies are believed to be enveloped in giant "halos" of dark matter. These halos, unlike those found above the heads of saints, have a spherical or ellipsoidal shape. On larger scales, analogous halos are thought to keep clusters of galaxies bound together.

Unfortunately, no one has ever detected dark matter directly, and its nature is still one of the biggest mysteries in science. Currently most astronomers favor the idea that dark matter consists mostly of hitherto unidentified particles that barely interact with ordinary particles or with one another. Astronomers typically refer to this class of particles as cold dark matter (CDM) and any cosmological model that postulates their existence as a CDM model.

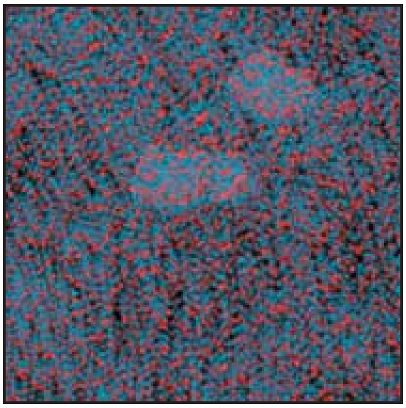
Over the past two decades, astronomers have painstakingly developed a model of galaxy formation based on CDM. The basic framework is the standard big bang theory for the expansion of the universe. Cosmologists continue to debate how the expansion got going and what transpired early on, but these uncertainties do not matter greatly for galaxy formation. We pick up the story about 100,000 years after the big bang, when the universe consisted of baryons (that is, ordinary matter, predominantly hydrogen and helium nuclei), electrons (bound to the nuclei), neutrinos, photons and CDM. Observations indicate that the matter and radiation were distributed smoothly:

THE AUTHORS

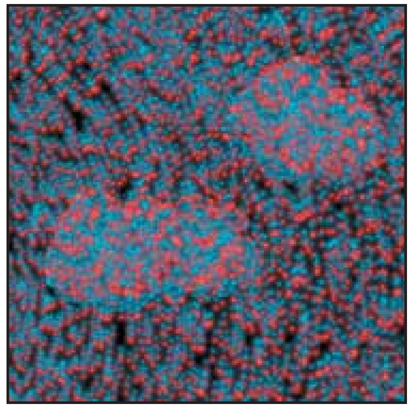
GUINEVERE KAUFFMANN and *FRANK VAN DEN BOSCH* are researchers at the Max Planck Institute for Astrophysics in Garching, Germany. They are among the world's experts on the theoretical modeling of galaxy formation. Kauffmann has recently turned her attention to analyzing data from the Sloan Digital Sky Survey, which she believes holds the answers to some of the mysteries highlighted in this article. In her spare time, she enjoys exploring Bavaria with her son, Jonathan. Van den Bosch is particularly intrigued by the formation of disk galaxies and of massive black holes in galactic centers. In his free time, he can often be found in a Munich beer garden.

COOKING UP A GALAXY

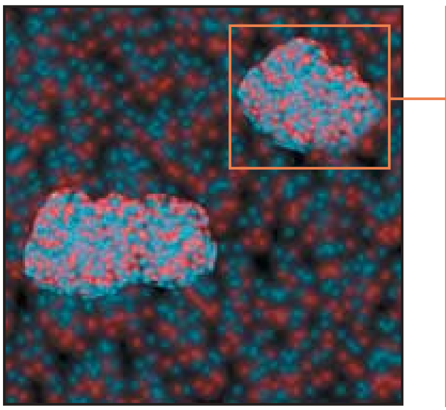
THREE BASIC PROCESSES dictated how the primordial soup congealed into galaxies: the overall expansion of the universe in the big bang, the force of gravity, and the motion of particles and larger constituents. The shifting balance among these processes can explain why galaxies became discrete, coherent bodies rather than a uniform gas or a horde of black holes. In this theory, small bodies coalesce first and then glom together to form larger objects. A crucial ingredient is dark matter, which reaches a different equilibrium than ordinary matter. —*G.K. and F.v.d.B.*



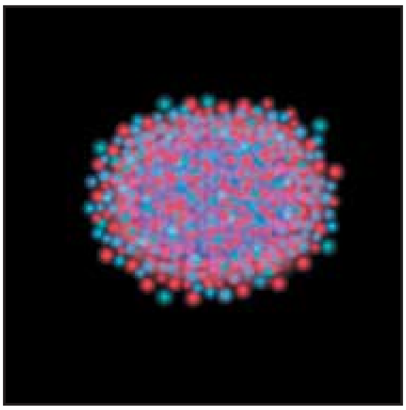
1 In the beginning, a primordial fluid—a mixture of ordinary matter (*blue*) and dark matter (*red*)—fills the universe. Its density varies subtly from place to place.



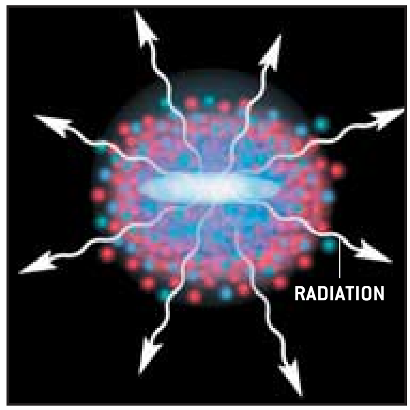
2 At first, cosmic expansion overpowers gravity. The fluid thins out. But patches of higher density thin out more slowly than other regions do.



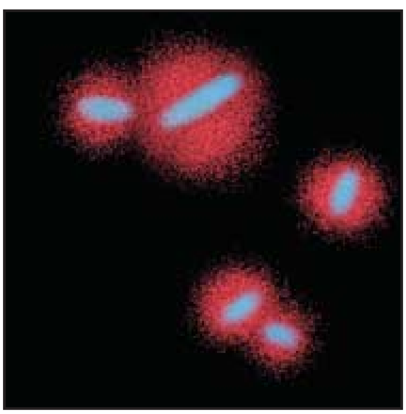
3 Eventually these patches become so dense, relative to their surroundings, that gravity takes over from expansion. The patches start to collapse.



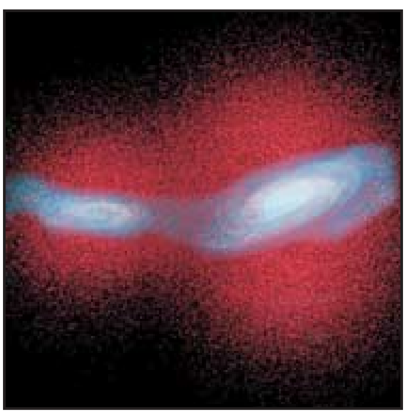
4 As each patch collapses, it attains equilibrium. The density, both of ordinary and of dark matter, peaks at the center and decreases toward the edge.



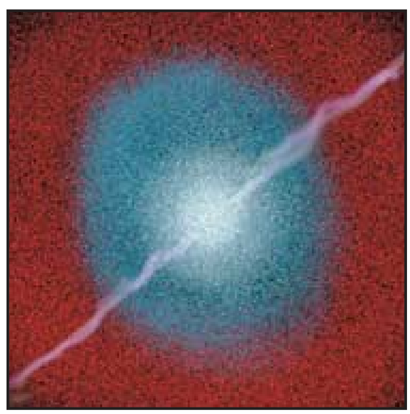
5 Dark matter, being unable to radiate, retains this shape. But ordinary matter emits radiation, collapses into a rotating disk and begins to condense into stars.



6 Protogalaxies interact, exerting torques on one another and merging to form larger and larger bodies. (This step overlaps with steps 4 and 5.)



7 When two disks of similar size merge, the stellar orbits become scrambled. An elliptical galaxy results. Later a disk may develop around the elliptical.



8 The merger triggers new star formation and feeds material into the central black hole, generating an active galactic nucleus, which can spew plasma jets.

DON DIXON

the density at different positions varied by only about one part in 100,000. The challenge is to trace how these simple ingredients could give rise to the dazzling variety of galaxies.

If one compares the conditions back then with the distribution of matter today, two important differences stand out. First, the present-day universe spans an enormous range of densities. The central regions of galaxies are more than 100 billion times as dense as the universe on average. The earth is another 10 billion billion times as dense as that. Second, whereas the baryons and CDM were initially well mixed, the baryons today form dense knots (the galaxies) inside gargantuan halos of dark matter. Somehow the baryons have decoupled from the CDM.

The first of these differences can be explained by the process of gravitational instability. If a region is even slightly more dense than average, the excess mass will exert a slightly stronger-than-average gravitational force, pulling extra matter toward itself. This creates an even stronger gravitational field, pulling in even more mass. This runaway process amplifies the initial density differences.

Sit Back and Relax

ALL THE WHILE, the gravity of the region must compete with the expansion of the universe, which pulls matter apart. Initially cosmic expansion wins and the density of the region decreases. The key is that it decreases more slowly than the density of its surroundings. At a certain point, the overdensity of the region compared with its surroundings becomes so pronounced that its gravitational attraction overcomes the cosmic expansion. The region starts to collapse.

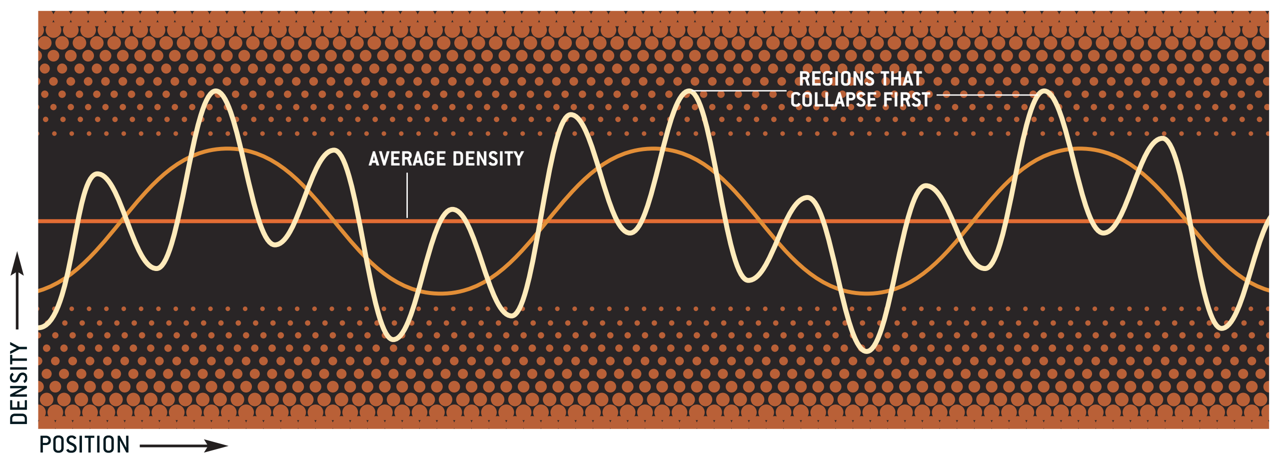
Up to this point, the region is not a coherent object but merely a random enhancement of density in the haze of matter that fills the universe. But once the region collapses, it starts to take on an internal life of its own. The system—which we shall call a protogalaxy from here on—seeks to establish some form of equilibrium. Astronomers refer to this process as relaxation. The baryons behave like the particles of any gas. Heated by shock waves that are triggered by the collapse, they exchange energy through direct collisions with one another, thus achieving hydrostatic equilibrium—a state of balance between pressure and gravity. The earth's atmosphere is also in hydrostatic equilibrium (or nearly so), which is why the pressure decreases exponentially with altitude.

For the dark matter, however, relaxation is distinctively different. CDM particles are, by definition, weakly interactive; they are not able to redistribute energy among themselves by direct collisions. A system of such particles cannot reach hydrostatic equilibrium. Instead it undergoes what is called, perhaps oxymoronically, violent relaxation. Each particle exchanges energy not with another individual particle but with the collective mass of particles, by way of the gravitational field.

Bodies traveling in a gravitational field are always undergoing an exchange of gravitational and kinetic energy. If you throw a ball into the air, it rises to a higher altitude but decelerates: it gains gravitational energy at the expense of kinetic energy. On the way down, the ball gains kinetic energy at the expense of gravitational energy. CDM particles in a protogalaxy behave much the same way. They move around and change speed as their balance of gravitational and kinetic energy shifts. But unlike balls near the earth's surface, CDM particles move in a gravitational field that is not constant. After all, the gravitational field is produced by all the particles together, which are undergoing collapse.

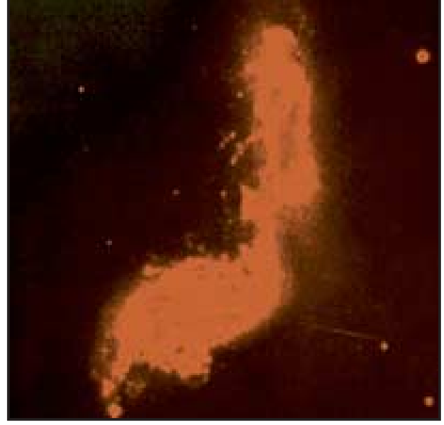
GALACTIC DENSITY VARIATIONS

DENSITY VARIATIONS in the pregalactic universe followed a pattern that facilitated the formation of protogalaxies. The variations were composed of waves of various wavelengths in a pattern that music connoisseurs will recognize as "pink noise." (Indeed, they originated as sound waves in the primordial plasma.) A small wave was superimposed on a slightly larger wave, which was superimposed on an even larger wave, and so on. Therefore, the highest density occurred over the smallest regions. These regions collapsed first and became the building blocks for larger structures. —*G.K. and F.v.d.B.*



SARA CHEN

Astronomers may be directly observing, for the first time, the formation of **elliptical galaxies.**



FROM *ATLAS OF PECULIAR GALAXIES*, BY HALTON ARP (*California Institute of Technology, 1966*); nedwww.ipac.caltech.edu/level5/Arp/frames.html (*merging galaxies known as Arp 270*)

Changes in the gravitational field cause some particles to gain energy and others to lose energy. Just as for the baryons, this redistribution of the energies of the particles allows the system to relax, forming a CDM halo that is said to be in virial equilibrium. The process is complicated and has never been worked out in great theoretical detail. Instead researchers track it using numerical simulations, which show that all CDM halos in virial equilibrium have similar density profiles.

The end point of the collapse and relaxation of a protogalaxy is a dark matter halo, inside of which the baryonic gas is in hydrostatic equilibrium at a temperature of typically a few million degrees. Whereas each CDM particle conserves its energy from then on, the baryonic gas is able to emit radiation. It cools, contracts and accumulates at the center of the dark matter halo. Cooling, therefore, is the process responsible for decoupling the baryons from the CDM.

So far we have focused on a single protogalaxy and ignored its surroundings. In reality, other protogalaxies will form nearby. Gravity will pull them together until they merge to form a grander structure. This structure will itself merge, and so on. Hierarchical buildup is a characteristic feature of CDM models. The reason is simple. Because small-scale fluctuations in density are superimposed on larger-scale fluctuations, the density reaches its highest value over the smallest regions. An analogy is the summit of a mountain. The exact position of the peak corresponds to a tiny structure: for example, a pebble on top of a rock on top of a hill on top of the summit. If a cloud bank descends on the mountain, the pebble vanishes first, followed by the rock, the hill and eventually the whole mountain.

Similarly, the densest regions of the early universe are the smallest protogalaxies. They are the first regions to collapse, followed by progressively larger structures. What distinguishes CDM from other possible types of dark matter is that it has density fluctuations on all scales. Neutrinos, for example, lack fluctuations on small scales. A neutrino-dominated universe would be like a mountain with an utterly smooth summit.

The hierarchical formation of dark matter halos cannot be described using simple mathematical relationships. It is best studied using numerical simulations. To emulate a representative part of the universe with enough resolution to see the formation of individual halos, researchers must use the latest supercomputers. The statistical properties and spatial distribution of the halos emerging from these simulations are in excellent agreement with those of observed galaxies, providing strong support for the hierarchical picture and hence for the existence of CDM.

Take a Spin

THE HIERARCHICAL PICTURE naturally explains the shapes of galaxies. In spiral galaxies, stars and gas move on circular orbits. The structure of these galaxies is therefore governed by angular momentum. Where does this angular momentum come from? According to the standard picture, when protogalaxies filled the universe, they exerted tidal forces on one another, causing them to spin. After the protogalaxies collapsed, each was left with a net amount of angular momentum.

When the gas in the protogalaxies then started to cool, it contracted and started to fall toward the center. Just as ice-skaters spin faster when they pull in their arms, the gas rotated faster and faster as it contracted. The gas thus flattened out, in the same way that the earth is slightly flatter than a perfect sphere because of its rotation. Eventually the gas was spinning so fast that the centrifugal force (directed outward) became equal to the gravitational pull (directed inward). By the time the gas attained centrifugal equilibrium, it had flattened into a thin disk. The disk was sufficiently dense that the gas started to clump into the clouds, out of which stars then formed. A spiral galaxy was born.

Because most dark matter halos end up with some angular momentum, one has to wonder why all galaxies aren't spirals. How did ellipticals come into being? Astronomers have long held two competing views. One is that most of the stars in present-day ellipticals and bulges formed during a monolithic collapse at early epochs. The other is that ellipticals are relative latecomers, having been produced as a result of the merging of spiral galaxies.

The second view has come to enjoy increasing popularity. Detailed computer simulations of the merger of two spirals show that the strongly fluctuating gravitational field destroys the two disks. The stars within the galaxies are too spread out to bang into one another, so the merging process is quite similar to the violent relaxation suffered by dark matter. If the galaxies are of comparable mass, the result is a smooth clump of stars with properties that strongly resemble an elliptical. Much of the gas in the two original disk galaxies loses its angular momentum and plummets toward the center. There the gas reaches high densities and starts to form stars at a frenzied rate. At later times, new gas may fall in, cool off and build up a new disk around the elliptical. The result will be a spiral galaxy with a bulge in the middle.

The high efficiency of star formation during mergers explains why ellipticals typically lack gas: they have used it up. The merger model also accounts for the morphology-density relation: a galaxy in a high-density environment will undergo more mergers and is thus more likely to become an elliptical.

Observational evidence confirms that mergers and interactions have been common in the universe, particularly early on. In Hubble Space Telescope images, many ancient galaxies have disturbed morphologies, a telltale sign of interaction. Moreover, the number of starburst galaxies—in which stars form at a frenetic pace—increases dramatically at earlier times. Astronomers may be directly observing, for the first time, the formation of elliptical galaxies.

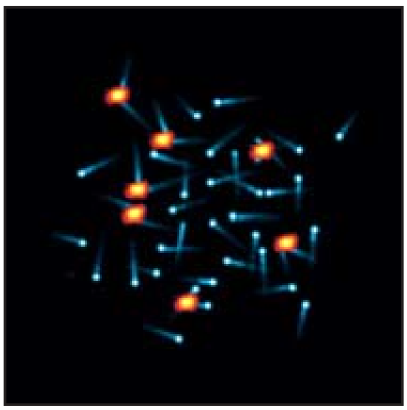
If elliptical galaxies and spiral bulges are linked to galaxy mergers, then it follows that supermassive black holes may be created in these events, too. Hole masses are strongly correlated with the mass of the surrounding elliptical galaxy or bulge; they are not correlated with the mass of the spiral disk. Merger models have been extended to incorporate supermassive holes and therefore AGNs. The abundant gas that is funneled toward the center during a merger could revive a dormant black hole. In other words, quasars were more common in the past because mergers were much more common then.

As for dwarf galaxies, in the hierarchical picture they are the leftovers—small clumps that have yet to merge. Recent observations show that star formation in dwarfs is particularly erratic, coming in short bursts separated by long quiescent periods [see "Dwarf Galaxies and Starbursts," by Sara C. Beck; Scientific American, June 2000]. In heftier galaxies such as the Milky Way, star formation occurs at a more constant rate. These results are intriguing because astronomers have often hypothesized that the mass of a galaxy determines its fertility. In lightweight galaxies, supernova explosions can easily disrupt or even rid the system of its gas, thus choking off star formation. Even the smallest perturbation can have a dramatic effect. It is this sensitivity to initial conditions and random events

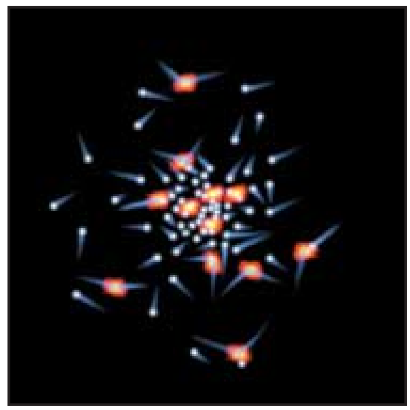
HOW RELAXING

AN INTERNAL STATE OF EQUILIBRIUM is what makes a galaxy a distinct object rather than merely an arbitrary patch of space. This equilibrium determines the overall properties of the galaxy, such as its shape and density profile. (An analogous equilibrium determines the size and temperature of stars.) The ordinary matter and dark matter attain equilibrium by different means.

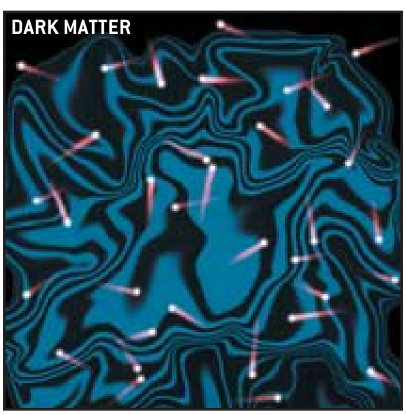
1 The ordinary matter—predominantly hydrogen gas—starts off moving every which way. Its density varies randomly.



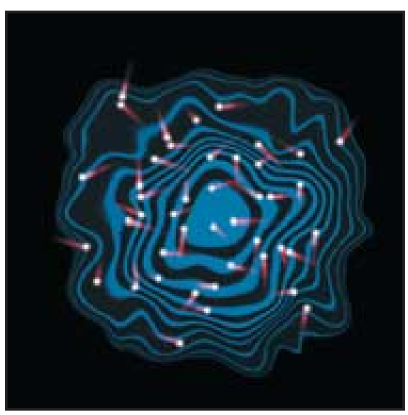
2 The gas particles bang into one another, redistributing energy and generating a pressure that resists gravity.



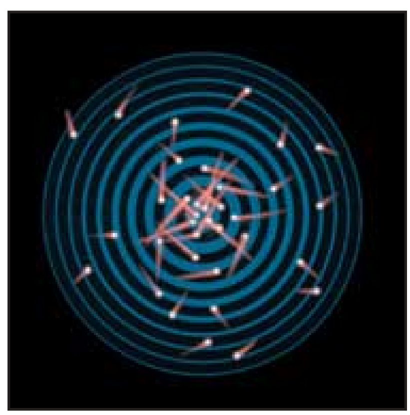
3 Eventually the gas settles down into hydrostatic equilibrium, with the density highest near the center of gravity.



1 Initially the dark matter has the same arrangement as ordinary matter. The difference is that the particles do not collide.



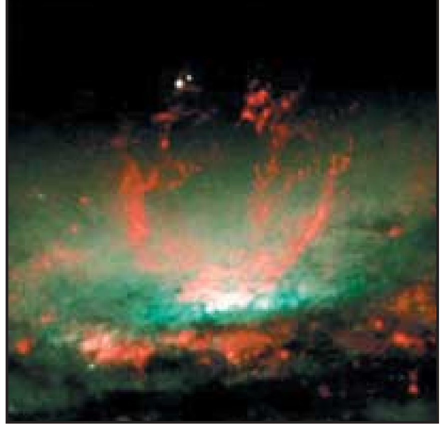
2 As the particles move around, the gravitational field changes, which causes particles to gain or lose energy.



3 Gradually the system settles down into virial equilibrium, in which the gravitational field no longer fluctuates.

DON DIXON

Supernova explosions could expel mass from low-mass galaxies so efficiently that hardly any stars would form.



that may account for the heterogeneity of the galactic dwarfs.

Although the standard picture of galaxy formation is remarkably successful, researchers are still far from working out all the processes involved. Moreover, they have yet to resolve some troubling inconsistencies. The simple picture of gas cooling inside dark matter halos faces an important problem known as the cooling catastrophe. Calculations of the cooling rates imply that the gas should have cooled briskly and pooled in the centers of halos, leaving intergalactic space virtually empty. Yet the space between galaxies is far from empty. Some extra input of energy must have prevented the gas from cooling down.

Some Feedback, Please

ANOTHER PROBLEM CONCERNS angular momentum. The amount of angular momentum imparted to protogalaxies in the models is comparable to the angular momentum that we actually see in spiral galaxies. So long as the gas retains its angular momentum, the CDM picture reproduces the observed sizes of spirals. Unfortunately, in the simulations the angular momentum leaks away. Much of it is transferred to the dark matter during galaxy mergers. As a result, the disks emerging from these simulations are a factor of 10 too small. Apparently the models are still missing an essential ingredient.

A third inconsistency has to do with the number of dwarf galaxies. Hierarchical theories predict a proliferation of low-mass dark matter halos and, by extension, dwarf galaxies. These are simply not seen. In the neighborhood of the Milky Way, the number of low-mass dwarfs is a factor of 10 to 100 lower than theories predict. Either these dark matter halos do not exist or they are present but have eluded detection because stars do not form within them.

Several solutions have been suggested for these problems. The proposals fall into two classes: either a fundamental change to the model, perhaps to the nature of dark matter [see "What's the Matter?" by George Musser; News and Analysis, SCIENTIFIC AMERICAN, May 2000], or a revision of our picture of how the cooling gas is transformed into stars. Because most astronomers are reluctant to abandon the CDM model, which works so well on scales larger than galaxies, they have concentrated on improving the treatment of star formation. Current models gloss over the process, which occurs on scales that are much smaller than a typical galaxy. Incorporating it in full is far beyond the capabilities of today's supercomputers.

Yet star formation can have profound effects on the structure of a galaxy [see "The Gas between the Stars," by Ronald J. Reynolds; SCIENTIFIC AMERICAN, January]. Some astronomers think that the action of stars might actually solve all three problems at once. The energy released by stars can heat the gas, obviating the cooling catastrophe. Heating also slows the descent of gas toward the center of the galaxy and thereby reduces its tendency to transfer angular momentum to the dark matter—alleviating the angular momentum problem. And supernova explosions could expel mass from the galaxies back into the intergalactic medium [see "Colossal Galactic Explosions," by Sylvain Veilleux, Gerard Cecil and Jonathan Bland-Hawthorn; SCIENTIFIC AMERICAN, February 1996]. For the lowest-mass halos, whose escape velocity is small, the process could be so efficient that hardly any stars form, which would explain why we observe fewer dwarf galaxies than predicted.

Because our understanding of these processes is poor, the models still have a lot of wiggle room. It remains to be seen whether the problems really can be fixed or whether they indicate a need for a completely new framework. Our theory of galaxy formation will surely continue to evolve. The observational surveys under way, such as the Sloan Digital Sky Survey, will enormously improve the data on both nearby and distant galaxies. Further advances in cosmology will help constrain the initial conditions for galaxy formation. Already, precise observations of the cosmic microwave background radiation have pinned down the values of the large-scale cosmological parameters, freeing galactic modelers to focus on the small-scale intricacy. Soon we may unite the large, the small and the medium into a seamless picture of cosmic evolution. SA

MORE TO EXPLORE

Cosmological Physics. John A. Peacock. Cambridge University Press, 1999.

The Formation of Ellipticals, Black Holes and Active Galactic Nuclei: A Theoretical Perspective. Guinevere Kauffmann, Stéphane Charlot and Martin G. Haehnelt in *Philosophical Transactions of the Royal Society of London*, Series A, Vol. 358, No. 1772, pages 2121–2132; July 15, 2000.

The Big Bang. Joseph Silk. W. H. Freeman and Company, 2001.

The Morphological Evolution of Galaxies. Roberto G. Abraham and Sidney van den Bergh in *Science*, Vol. 293, No. 5533, pages 1273–1278; August 17, 2001. Available at **astro-ph/0109358**

The Angular Momentum Content of Dwarf Galaxies: New Challenges for the Theory of Galaxy Formation. Frank C. van den Bosch, Andreas Burkert and Rob A. Swaters in *Monthly Notices of the Royal Astronomical Society*, Vol. 326, No. 3, pages 1205–1215; September 21, 2001. Available at **astro-ph/0105082**

New Perspectives in Astrophysical Cosmology. Martin Rees. Second edition. Cambridge University Press, 2002.

Galaxy Formation and Evolution: Recent Progress. Richard S. Ellis. Lecture given at the XIth Canary Islands Winter School of Astrophysics, "Galaxies at High Redshift" (in press). Available at **arxiv.org/abs/astro-ph/0102056**

GERALD CECIL ET AL. *University of North Carolina at Chapel Hill* AND NASA (*superbubble in NGC 3079*)

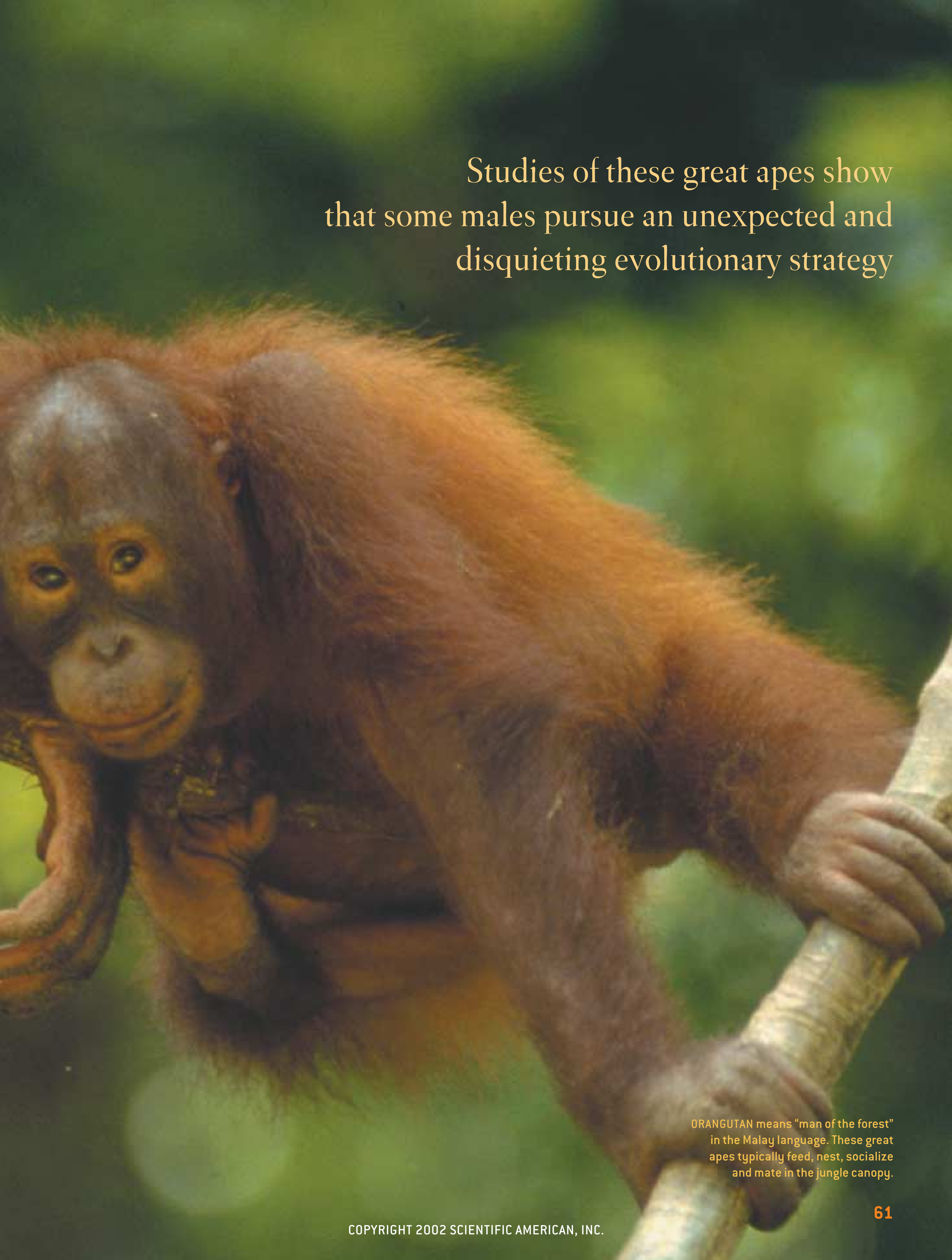
Disturbing Behaviors of the

ORANGUTAN

By Anne Nacey Maggioncalda and Robert M. Sapolsky

The orangutan is one of humankind's closest relatives. One of the four great apes (the other three are gorillas, chimpanzees and bonobos), *Pongo pygmaeus* is exquisitely adapted for life in the forest canopies of the Southeast Asian islands of Borneo and Sumatra. With their long arms and hooklike hands, orangutans are adept at swinging from tree to tree in search of tropical fruits. They are among the most solitary of large primates and the only great apes found outside Africa. Orangutans are also notable for the striking size difference between males and females: the average weight of an adult male (about 90 kilograms, or 198 pounds) is more than twice that of a female.

Studies of these great apes show that some males pursue an unexpected and disquieting evolutionary strategy



ORANGUTAN means "man of the forest" in the Malay language. These great apes typically feed, nest, socialize and mate in the jungle canopy.

An adult male orangutan is an impressive sight. The animal has a pair of wide cheek pads, called flanges, and a well-developed throat sac used for emitting loud cries known as long calls. The mature male also has long, brightly colored hair on its body and face. These are secondary sexual characteristics, the flamboyant signals that male orangutans flaunt to proclaim their fertility and fitness to the opposite sex. The features emerge during orangutan adolescence: males reach puberty at around seven to nine years of age, then spend a few years in a far-from-impressive "subadult" stage, during which they are about the same size as mature females. The males reach their adult size and develop secondary sexual traits by ages 12 to 14. Or at least that's what primate researchers used to think.

As stable social groups of orangutans were established in zoos, however, it became clear that an adolescent male could remain a subadult, in a state of arrested development, until his late teens. In the 1970s, studies of orangutans in the rain forests of Southeast Asia by Biruté M. F. Galdikas of Simon Fraser University in British Columbia and others produced the same finding: sometimes males were arrested adolescents for a decade or more, about half their potential reproductive lives. Variability of this magnitude is fascinating—it is like finding a species in which pregnancy could last anywhere from six months to five years.

Biologists are keenly interested in studying cases of arrested development because they often shed light on the processes of growth and maturation. In some instances, the cause of arrested development is a genetic disorder; for example, a mutation in the receptor for a growth factor in humans results in a form of dwarfism. Environmental factors can also slow or halt an organism's development. For instance, food shortages delay maturation in humans and many other animals. This response is logical from an evolutionary standpoint—if it is unclear whether you will survive another week, it makes no sense to waste calories by adding bone mass or developing secondary sexual characteristics. Gymnasts and ballet dancers who exercise to extremes and anorexics who starve themselves sometimes experience delayed onset of puberty.

Among male orangutans, though, the cause of arrested development seems to lie in the animals' *social* environment. The presence of dominant adult males appears to delay the maturation of adolescent males in the same vicinity. Until recently, researchers believed that they were observing a stress-induced pathology—that is, the adolescent orangutans stopped developing because the adult males bullied and frightened them. Over the past few years, however, we have conducted studies suggesting that arrested development among orangutans is not a pathology but an adaptive evolutionary strategy. The arrested adolescent males are capable of impregnating females, and by staying small and immature (in terms of secondary sexual features) they minimize the amount of food they need and lower the risk of serious conflict with adult males. But the strategy of these arrested adolescents has a disquieting aspect: they copulate forcibly with females. In other words, they rape.

THE AUTHORS

ANNE NACEY MAGGIONCALDA and *ROBERT M. SAPOLSKY* have been studying the evolution of male orangutan reproductive strategies for more than a decade. Maggioncalda, who received a Ph.D. in biological anthropology and anatomy from Duke University in 1995, is a lecturer in the department of anthropological sciences and the program in human biology at Stanford University and in the department of anatomy at the Stanford University School of Medicine. Sapolsky, who earned a Ph.D. in neuroendocrinology from the Rockefeller University in 1984, is professor of biological sciences and neurology at Stanford and a research associate at the National Museums of Kenya. His research interests include neuron death, gene therapy and the physiology of primates.

HUNDREDS OF THOUSANDS of orangutans roamed throughout Southeast Asia about 10,000 years ago, but their range is now limited to parts of Borneo and Sumatra. Their number has dwindled to fewer than 20,000. Given the current rate of hunting and habitat destruction, researchers say, orangutans could disappear from the wild within two decades.

Measuring Stress

THE FIRST INVESTIGATIONS into this subject focused on groups of captive orangutans. Terry L. Maple of Zoo Atlanta and other zoo biologists found that adolescent males remained developmentally arrested as long as there was a mature male in their enclosure. If the researchers removed that dominant male, the adolescents soon began to develop into adults. This kind of social regulation had been observed previously in other species. Among mandrill monkeys, for instance, socially dominant males develop dramatic secondary sexual characteristics, such as large testes and high testosterone levels, whereas subordinate males do not. In tree shrews and many rodent species, puberty is delayed in the subordinate animals. In another example, elephant poaching in certain areas of Africa has recently produced orphaned males that grew up in a fairly unsocialized manner. When in "musth"—a male elephant's mating period—these animals become quite aggressive and dangerous. Some zoologists have reported an effective solution: introducing older, more dominant males into the region, which results in social suppression of musth in the rogue males.

FRANS LANTING *Minden Pictures* (*preceding pages*); SARA CHEN (*above*)

In all these cases, researchers have generally agreed that the stress of being subordinate accounts for the developmental arrest. During a typical period of stress for a mammal—say, a sprint across the savanna to escape a predator—energy is mobilized to power the muscles. As part of this process, a variety of long-term building projects in the body are inhibited, including growth, tissue repair and reproductive functions. It is the logic of triage: the animal concentrates on survival during the emergency and resumes long-term tasks later, if there is a later. But when an animal undergoes chronic stress, such as that caused by social subordination, the triaging can have adverse consequences, such as decreased growth, lower levels of sex hormones, reduced fertility and delayed puberty. In humans, severe and prolonged psychological stress can cause growth to stop in children, a rare syndrome called psychogenic dwarfism.

At first glance, adolescent male orangutans also appear to be under chronic stress. Adult male orangutans are extremely aggressive to adolescents, particularly within the confines of a zoo. In the wild, orangutan males are dispersed and solitary, belligerently defending a large territory that encompasses several females' territories—sort of a scattered harem. But even there, adolescents are well aware of the threatening presence of a mature male. One signal is a musky odorant that adult males spread about their territories. In addition, mature male orangutans announce their presence by performing long calls; John C. Mitani of the University of Michigan has found that these resonant cries can travel for miles.

Researchers had made little effort, however, to test the hypothesis that the stress of being near a dominant male induces hormonal changes that arrest development in adolescents. In 1989 we began looking for a way to examine the hormones of arrested adolescent orangutans to determine whether these animals were indeed under chronic stress. Ideally one would want to measure the levels of relevant hormones in the orangutans' blood, but this was impossible to do, for ethical and practical reasons. So we took advantage of the fact that the average levels of various hormones in the animals' blood are reflected in a fairly parallel fashion in their urine. Getting urine from wild animals would be immensely difficult, so we studied captive populations. Thanks to the generous help of zookeepers, curators and veterinarians at 13 zoos, we obtained more than 1,000 urine samples from 28 male orangutans, along with information on their developmental status (juvenile, arrested adolescent, developing adolescent or adult), housing, diet, medical history and growth records. In collaboration with Nancy Czekala of the Center for Reproduction of Endangered Species at the San Diego Zoo, we measured the levels of nine hormones, comparing animals in different developmental stages.

First we focused on growth hormone, which is crucial for normal maturation. Among the juveniles, arrested adolescents and adults, growth hormone levels in the urine were low and extremely similar, within 15 percent of one another. In contrast, adolescent males that were maturing into adults had growth hormone levels approximately three times as high. This result basically served as an internal control, showing that the external assessments of an animal's development stage closely matched the hormonal profile relevant to growth. In other words, adolescent males going through a developmental spurt in terms of appearance—growing larger, increasing the size of their cheek flanges, and so on—were experiencing hormonal changes as well.

We then considered hormones that respond to stress. Probably the best known is adrenaline (also called epinephrine), which plays a central role in energy mobilization. Epinephrine, unfortunately, cannot be measured accurately in urine. We could, however, determine levels of another key class of stress hormones called the glucocorticoids, which can suppress growth, tissue repair and reproduction. In addition, we measured the levels of prolactin, a stress-indicative hormone that can inhibit reproduction.

This is where we got a surprise. Glucocorticoid levels did not differ among juveniles, arrested adolescents and adults. Prolactin levels did not differ either. But adolescents going through the developmental spurt had glucocorticoid and prolactin levels roughly double those of the other groups. It wasn't the devel-



FRANS LANTING *Minden Pictures* (*left*); THEO ALLOFS *Corbis* (*center*); W. PERRY CONWAY *Corbis* (*right*)

SECONDARY SEXUAL CHARACTERISTICS distinguish the adult male orangutan (*left*) from the arrested adolescent male (*center*) and the adult female (*right*). Perhaps the most prominent of these features are the adult male's wide cheek pads, called flanges, and the well-developed throat sac used for emitting loud cries known as long calls. Also, the average weight of adult males is more than twice that of arrested males and adult females.

opmentally arrested adolescents who seemed to be stressed—it was the *developing* adolescents.

We got another surprise when we examined reproductive hormones in these animals. As expected, adolescent males who were developing secondary sexual characteristics had hormonal profiles implying an active gonadal system. Developing males had higher levels of testosterone and luteinizing hormone (which stimulates the release of testosterone) than did the arrested adolescents. But the levels of these hormones in arrested adolescents were equivalent to those seen in adults. Moreover, arrested males had levels of follicle-stimulating hormone (FSH), which stimulates sperm maturation in males, equal to those of developing adolescents or adult males. And other investigators have found that arrested adolescents have mature functional sperm and that their testes are the same size as those of developing adolescents.

Evolutionary Strategies

THESE FINDINGS OVERTURNED some long-held assumptions about orangutans. Apparently, arrested adolescents are neither stressed nor reproductively suppressed. What is going on? It turns out that there is more than one way for a male orangutan to improve his chances of reproducing.

A cornerstone of modern evolutionary theory is that animal behavior has evolved not for the good of the species or the social group but to maximize the number of gene copies passed on by an individual and its close relatives. For a long time, the study of primates was dominated by simplistic models of how animals achieve this goal. According to these models, male behavior consists of virtually nothing but aggression and competition to gain access to females. If only one female is sexually receptive in a group with many males, this competition would result in the highest-ranking male mating with her; if two females are receptive, the males ranking first and second in the hierarchy would mate with them, and so on.

But this kind of behavior is rarely seen among social primates. Instead male primates can choose alternative strategies to maximize their reproductive success. Why should there be alternatives? Because the seemingly logical strategy—developing powerful muscles and dramatic secondary sexual characteristics to excel at male-male competition—has some serious drawbacks. In many species, maintaining those secondary characteristics requires elevated testosterone levels, which have a variety of adverse effects on health. The aggression that comes with such a strategy is not great for health either.

Furthermore, increased body mass means greater metabolic demands and more pressure for successful food acquisition. During famines, the bigger primates are less likely to survive. For an arboreal species such as the orangutan, the heavier body of the mature male also limits which trees and branches can be accessed for food. And the development of secondary sexual characteristics makes a male more conspicuous, both to predators and to other males that view those characteristics as a challenge.

The competition between adult males and developing adolescents probably explains the elevated levels of stress hormones in the latter. In the eyes of an adult male orangutan, a developing male is soon to be a challenger, so naturally he becomes a prime target for aggression. The same pattern is seen among horses and various other social ungulates: it is not until the young males start developing secondary sexual characteristics that the unrelated dominant males begin to harass them into leaving the group. Another example comes from work by one of us (Sapolsky) with wild baboons. Some socially subordinate male baboons have much higher glucocorticoid levels than do the dominant animals, primarily because these subordinates are actively challenging the high-ranking males.

In contrast, the key impression that a developmentally arrested male communicates to an adult male is a lack of threat or challenge, because the immature male looks like a kid. Arrested male orangutans are apparently inconspicuous enough to be spared a certain amount of social stress. What is more, the "low profile" of these animals may actually give them a competitive advantage when it comes to reproduction. In many primate species, the low-ranking males are actually doing a fair share of the mating. Genetic paternity testing of these primates has shown that the subordinate males are quite successful in passing on their genes. This finding extends to orangutans: studies of zoo populations have proved that arrested males mate and that these matings are fertile. More recently Sri Suci Utami of Utrecht University in the Netherlands has shown that arrested

CHARLES AND JOSETTE LENARS *Corbis*

adolescents fathered approximately half of the orangutan babies at her Sumatran study site.

Why are these low-ranking males taking part in so many matings? In some primate species, such as the savanna baboon, the females can decide with whom they will mate, and they frequently choose males who exhibit strong male-female affiliation and parental behavior rather than male-male competition. Even when dominant male baboons stand guard to prevent low-ranking males from mating, the females often initiate surreptitious matings—sometimes referred to as "stolen copulations"—with the subordinates. For low-ranking male baboons, the strategy of pursuing affiliative "friendships" with females is a viable one because it avoids the metabolic costs, injuries and stress of male-male competition.

THEO ALLOFS *Corbis*

ORANGUTANS ARE AMONG the most solitary of large primates, but the apes occasionally travel and feed in small bands. Three orangutans groom one another in a forest in Borneo (*left*). An adult female rests with her offspring in Borneo's Tanjung Puting National Park (*above*).

But arrested male orangutans do not engage in long-term affiliative relationships with females, although an arrested male may sometimes accompany a female for several days as she roams through the forest. Furthermore, the great majority of adult female orangutans are sexually receptive only to mature males. So how do the arrested males mate? Observations of orangutans both in the wild and in captive populations have indicated that the arrested males forcibly copulate with females. Rape is an apt term for these copulations: the adult females usually resist the arrested adolescents fiercely, biting the males whenever they can and emitting loud, guttural sounds (called rape grunts) that are heard only under these circumstances. Adult males sometimes rape, too, but not nearly as often as the arrested males. In a study conducted in Borneo during the early 1980s, Mitani and his field assistants observed 151 copulations by arrested males; 144 of the matings were forced.

Thus, two reproductive strategies appear to have evolved for adolescent male orangutans. If no fully mature males are nearby, the adolescent will most likely develop quickly in the hopes of attracting female attention. When adult males are present, however, a strategy of arrested development has its advantages. If the social environment changes—say, if the nearby adult males die off or migrate—the arrested males will rapidly develop secondary sexual features and change their behavior patterns. Researchers are now trying to determine exactly how the presence or absence of adult males triggers hormonal changes in the adolescents.

Unpleasant Findings

WHAT ARE THE LESSONS we can learn from the male orangutan? First, a situation that seems stressful from a human's perspective may not necessarily be so. Second, the existence of alternative reproductive strategies shows that the optimal approach can vary dramatically in different social and ecological settings. There is no single blueprint for understanding the evolution of behavior. Third, although the recognition of alternative strategies built around female choice has generally met with a receptive audience among scientists, the rape-oriented strategy of arrested male orangutans is not so pleasing. But the study of primates has demonstrated time and again that the behavior of these animals is far from Disney-esque. Just consider the strategic infanticide of langur monkeys or the organized aggression—sometimes called genocide—between groups of chimpanzee males.

One must be cautious, however, in trying to gain insights into human behavior by extrapolating from animal studies. There is a temptation to leap to a wrongheaded conclusion: because forcible copulation occurs in orangutans and something similar occurs in humans, rape has a natural basis and is therefore unstoppable. This argument ignores the fact that the orangutan is the only nonhuman primate to engage in forcible copulation as a routine means of siring offspring. Furthermore, close observations of orangutan rape show that it is very different from human rape: for example, researchers have never seen a male orangutan injure a female during copulation in an apparently intentional manner. Most important, the orangutan's physiology, life history and social structure are completely unlike those of any other primate. Orangutans have evolved a unique set of adaptations to survive in their environment, and hence it would be the height of absurdity to draw simpleminded parallels between their behaviors and those of humans. SA

MORE TO EXPLORE

Reflections of Eden: My Years with the Orangutans of Borneo. Biruté M. F. Galdikas. Back Bay Books, 1996.

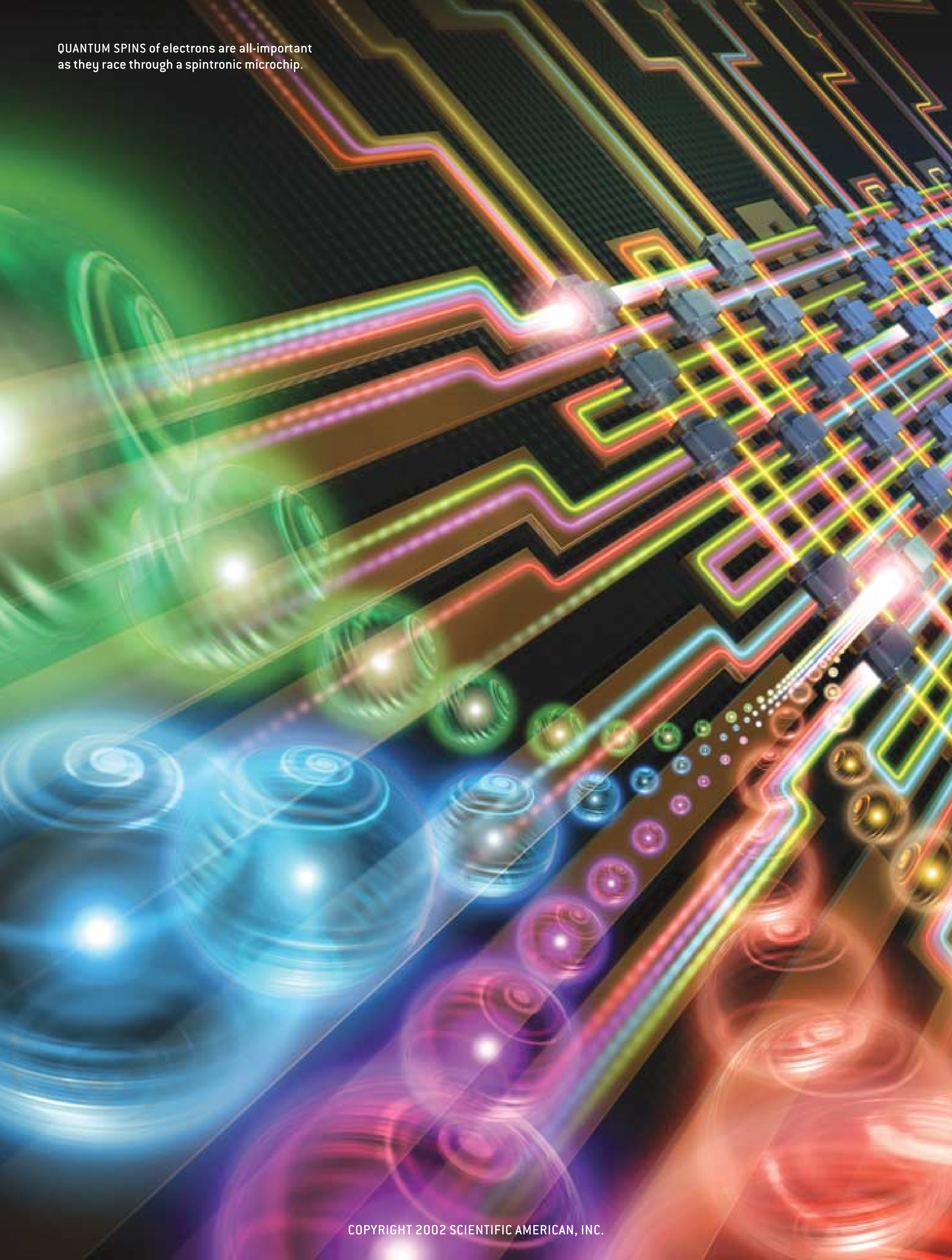
Reproductive Hormone Profiles in Captive Male Orangutans: Implications for Understanding Developmental Arrest. Anne N. Maggioncalda, Robert M. Sapolsky and Nancy M. Czekala in *American Journal of Physical Anthropology,* Vol. 109, No. 1, pages 19–32; May 1999.

Orangutans: Wizards of the Forest. Anne E. Russon. Firefly Books, 2000.

Male Orangutan Subadulthood: A New Twist on the Relationship between Chronic Stress and Developmental Arrest. Anne N. Maggioncalda, Nancy M. Czekala and Robert M. Sapolsky in *American Journal of Physical Anthropology,* Vol. 118, No. 1, pages 25–32; May 2002.

More information on orangutans is available at the Web sites of the Orangutan Foundation International (**www.orangutan.org**), the Balikpapan Orangutan Society (**www.orangutan.com**) and the Orangutan Network (**www.orangutannetwork.net**).

QUANTUM SPINS of electrons are all-important as they race through a spintronic microchip.



SPINTRONICS

Microelectronic devices that function by using the spin of the electron are a nascent multibillion-dollar industry—and may lead to quantum microchips

By David D. Awschalom, Michael E. Flatté and Nitin Samarth

Illustrations by Slim Films

As rapid progress in the miniaturization of semiconductor electronic devices leads toward chip features smaller than 100 nanometers in size, device engineers and physicists are inevitably faced with the looming presence of quantum mechanics—that counterintuitive and sometimes mysterious realm of physics wherein wavelike properties dominate the behavior of electrons. Pragmatists in the semiconductor device world are busy conjuring up ingenious ways to avoid the quantum world by redesigning the semiconductor chip within the context of "classical" electronics [see "A Vertical Leap for Microchips," by Thomas H. Lee; Scientific American, January]. Yet some of us believe that we are being offered an unprecedented opportunity to define a radically new class of device that would exploit the idiosyncrasies of the quantum world to provide unique advantages over existing information technologies.

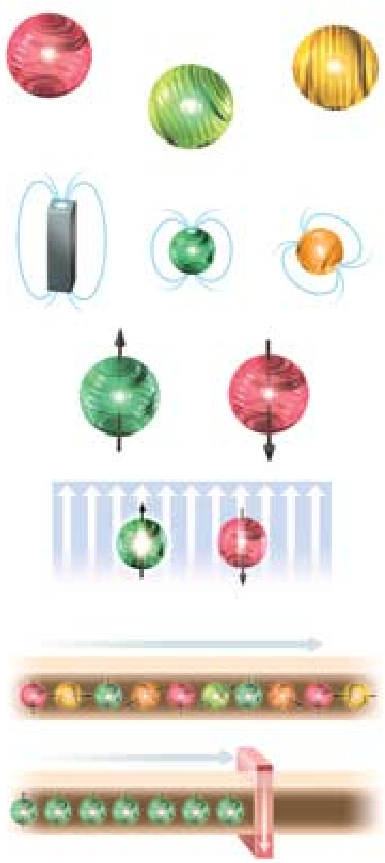
One such idiosyncrasy is a quantum property of the electron known as spin, which is closely related to magnetism. Devices that rely on an electron's spin to perform their functions form the foundation of spintronics (short for spin-based electronics), also known as magnetoelectronics. Information-processing technology has thus far relied on purely charge-based devices—ranging from the now quaint vacuum tube to today's million-transistor microchips. Those conventional electronic devices move electric charges around, ignoring the spin that tags along for the ride on each electron.

Magnetism (and hence electron spin) has nonetheless always been important for information storage. For instance, even the earliest computer hard drives used magnetoresistance—a change in electrical resistance caused by a magnetic field—to read data stored in magnetic domains. It is no surprise that the information storage industry has provided the initial successes in spintronics technology. Most laptop computers now come fitted with high-capacity hard drives that pack an unprecedented amount of data into each square millimeter. The drives rely on a spintronic effect, giant magnetoresistance (GMR), to read such dense data.

More sophisticated storage technologies based on spintronics are already at an advanced stage: in the next few years, MRAM (magnetic random-access memory), a new type of computer memory, will go on the market. MRAMs would retain their state even when the power was turned off, but unlike present forms of nonvolatile memory, they would have switching rates and rewritability challenging those of conventional RAM.

In today's read heads and MRAMs, key features are made of ferromagnetic metallic alloys. Such metal-based devices make up the first—and most mature—of three categories of spintronics. In the second category, the spin-polarized currents flow in semiconductors instead of metals. Achieving practical spintronics in semi-

FUNDAMENTALS OF SPIN



In addition to their mass and electric charge, electrons have an intrinsic quantity of angular momentum called spin, almost as if they were tiny spinning balls.

Associated with the spin is a magnetic field like that of a tiny bar magnet lined up with the spin axis.

Scientists represent the spin with a vector. For a sphere spinning "west to east" the vector points "north" or "up." It points "down" for the opposite spin.

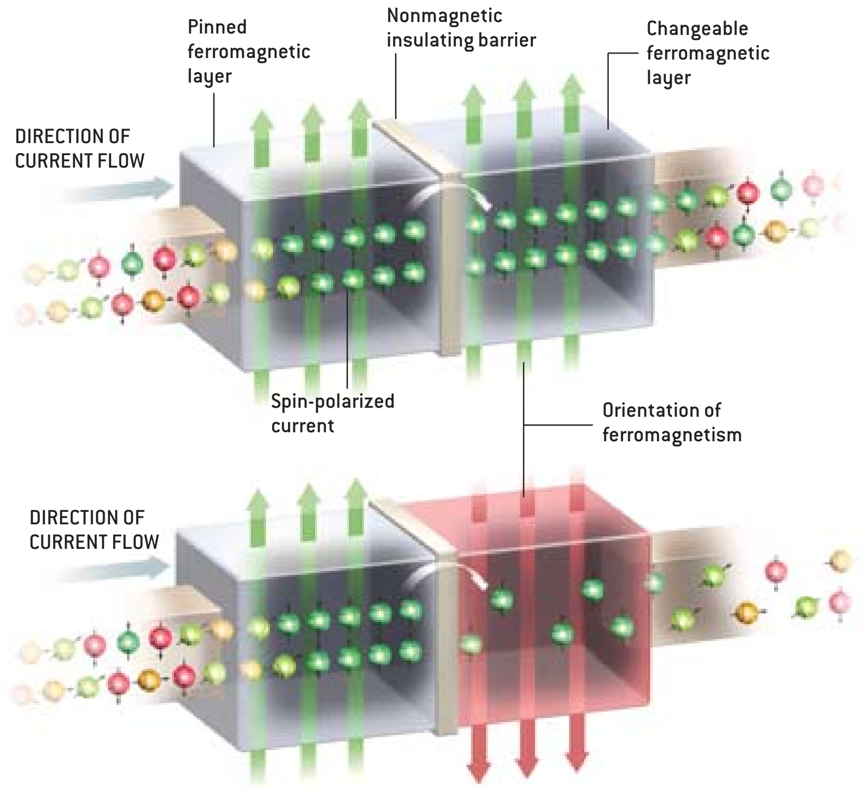
In a magnetic field, electrons with "spin up" and "spin down" have different energies.

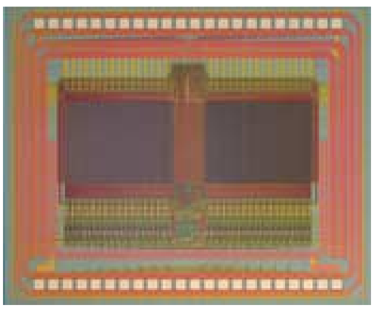
In an ordinary electric circuit the spins are oriented at random and have no effect on current flow.

Spintronic devices create spin-polarized currents and use the spin to control current flow.

MAGNETIC TUNNEL JUNCTION

MRAMS (magnetic random-access memories) store data in magnetic tunnel junctions, which retain their states even when the power is off. A 256-kilobyte MRAM chip is shown below. Magnetic tunnel junctions have two ferromagnetic layers separated by a thin insulating barrier. The first layer polarizes the spins of current-carrying electrons, which cross the barrier to the second layer by quantum tunneling when both layers are aligned (*"0," upper right*). When the magnetism of the second ferromagnetic layer is reversed, the tunneling is reduced (*"1," lower right*).





conductors would allow a wealth of existing microelectronics techniques to be co-opted and would also unleash many more types of devices made possible by semiconductors' high-quality optical properties and their ability to amplify both optical and electrical signals. Examples include ultrafast switches and fully programmable all-spintronics microprocessors. This avenue of research may lead to a new class of multifunctional electronics that combine logic, storage and communications on a single chip.

Researchers must answer several major questions before the second category of devices can take off as a viable industry: Can we devise economic ways to combine ferromagnetic metals and semiconductors in integrated circuits? Can we make semiconductors that are ferromagnetic at room temperature? What is an efficient way to inject spin-polarized currents, or spin currents, into a semiconductor? What happens to spin currents at boundaries between different semiconductors? How long can a spin current retain its polarization in a semiconductor?

Our own research groups are working on these questions but are keeping one eye also on the more distant and speculative quarry that is the third category of devices: ones that manipulate the quantum spin states of individual electrons. This category includes spintronic quantum logic gates that would enable construction of large-scale quantum computers, which would extravagantly surpass standard computers for certain tasks. A diverse assortment of exotic technologies is aimed toward that goal: ions in magnetic traps, "frozen" light, ultracold quantum gases called Bose-Einstein condensates and nuclear magnetic resonance of molecules in liquids—there are many ways to skin a quantum cat.

We believe it makes sense instead to build on the extensive foundations of conventional electronic semiconductor technology. Indeed, a recent series of unexpected discoveries appears to support our hunch that semiconductor spintronics provides a feasible path for developing quantum computers and other quantum information machines. Whether one looks at the near term for tomorrow's consumer electronics or at the more dis-

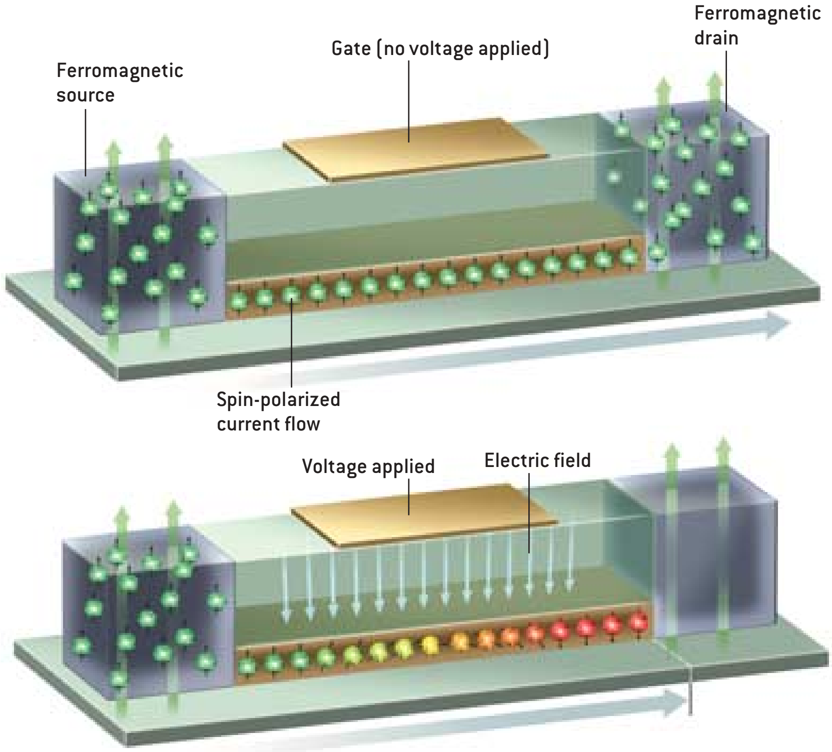
THE AUTHORS

DAVID D. AWSCHALOM, MICHAEL E. FLATTÉ and *NITIN SAMARTH* began working as a team when Flatté was recently on sabbatical at the University of California, Santa Barbara. Awschalom and Samarth had already been collaborating on experimental studies of semiconductor spintronics for more than a decade. Awschalom is professor of physics and director of the Center for Spintronics and Quantum Computation at U.C.S.B. Flatté works on condensed-matter theory and is associate professor at the University of Iowa. Samarth is professor of physics at Pennsylvania State University.

MOTOROLA (*chip*)

A SPIN TRANSISTOR CONCEPT

ONE PROPOSED DESIGN of a spin FET (spintronic field-effect transistor) has a source and a drain, separated by a narrow semiconducting channel, the same as in a conventional FET. In the spin FET, both the source and the drain are ferromagnetic. The source sends spin-polarized electrons into the channel, and this spin current flows easily if it reaches the drain unaltered (*top*). A voltage applied to the gate electrode produces an electric field in the channel, which causes the spins of fast-moving electrons to precess, or rotate (*bottom*). The drain impedes the spin current according to how far the spins have been rotated. Flipping spins in this way takes much less energy and is much faster than the conventional FET process of pushing charges out of the channel with a larger electric field.



tant prospect of quantum computing, spintronics promises to be revolutionary.

Exploiting Spin Currents

AN INTUITIVE NOTION of how an electron's spin works is suggested by the name itself. Imagine a small electrically charged sphere that is spinning rapidly. The circulating charges on the sphere amount to tiny loops of electric current, which create a magnetic field similar to the earth's magnetic field. Scientists traditionally depict the rotation by a vector, or arrow, that points along the sphere's axis of rotation. Immersing the spinning sphere in an external magnetic field changes its total energy according to how its spin vector is aligned with the field [*see box on page 68*].

In some ways, an electron is just like such a spinning sphere of charge—an electron has a quantity of angular momentum (its "spin") and an associated magnetism, and in an ambient magnetic field its energy depends on how its spin vector is oriented. But there the similarities end and the quantum peculiarities begin. Electrons seem to be ideal dimensionless points, not little spheres, so the simple picture of their spin arising from actual rotation doesn't work. In addition, every electron has exactly the same amount of spin, equal to one half the fundamental quantum unit of angular momentum. That property is hardwired into the mathematics that describes all the elementary particles of matter, a result whose significance and meaning are another story entirely. The bottom line is that the spin, along with a mass and a charge, is a defining characteristic of an electron.

In an ordinary electric current, the spins point at random and play no role in determining the resistance of a wire or the amplification of a transistor circuit. Spintronic devices, in contrast, rely on differences in the transport of "spin up" and "spin down" electrons. In a ferromagnet, such as iron or cobalt, the spins of certain electrons on neighboring atoms tend to line up. In a strongly magnetized piece of iron, this alignment extends throughout much of the metal. When a current passes through the ferromagnet, electrons of one spin direction tend to be obstructed. The result is a spin-polarized current in which all the electron spins point in the other direction.

A ferromagnet can even affect the flow of a current in a nearby nonmagnetic metal. For example, present-day read heads in computer hard drives use a device dubbed a spin valve, wherein a layer of a nonmagnetic metal is sandwiched between two ferromagnetic metallic layers. The magnetization of the first layer is fixed, or pinned, but the second ferromagnetic layer is not. As the read head travels along a track of data on a computer disk, the small magnetic fields of the recorded 1's and 0's change the second layer's magnetization back and forth, parallel or antiparallel to the magnetization of the pinned layer. In the parallel case, only electrons that are oriented in the favored direction flow through the conductor easily. In the antiparallel case, all electrons are impeded. The resulting changes in the current allow GMR read heads to detect weaker fields than their predecessors, so that data can be stored using more tightly packed magnetized spots on a disk, increasing storage densities by a factor of three.

Another three-layered device, the magnetic tunnel junction, has a thin insulating layer between two metallic ferromagnets [*see illustration on page 69*]. Current flows through the device by the process of quantum tunneling: a small number of electrons manage to jump through the barrier even though they are forbidden to be in the insulator. The tunneling current is obstructed when the two ferromagnetic layers have opposite orientations and is allowed when their orientations are the same.

Magnetic tunnel junctions form the basis of the MRAM chips mentioned earlier. Each junction can store one bit of data in the orientation of its unpinned ferromagnetic layer. That layer retains its magnetic state whether the power is on or off, at least until it is deliberately rewritten again.

Whereas the metallic spintronic devices just described provide new ways to store information, semiconductor spintronics may offer even more interesting possibilities. Because conventional semiconductors are not ferromagnetic, one might wonder how semiconductor spintronic devices can work at all. One solution employs a ferromagnetic metal to inject a spin-polarized current into a semiconductor.

In 1990 Supriyo Datta and Biswajit A. Das, then at Purdue University, proposed a design for a spin-polarized field-effect transistor, or spin FET [*see illustration on opposite page*]. In a conventional FET, a narrow semiconductor channel runs between two electrodes named the source and the drain. When voltage is applied to the gate electrode, which is above the channel, the resulting electric field drives electrons out of the channel (for instance), turning the channel into an insulator. The Datta-Das spin FET has a ferromagnetic source and drain so that the current flowing into the channel is spin-polarized. When a voltage is applied to the gate, the spins rotate as they pass through the channel and the drain rejects these antialigned electrons.

A spin FET would have several advantages over a conventional FET. Flipping an electron's spin takes much less energy and can be done much faster than pushing an electron out of the channel. One can also imagine changing the orientation of the source or drain with a magnetic field, introducing an additional type of control that is not possible with a conventional FET: logic gates whose functions can be changed on the fly.

As yet, however, no one has succeeded in making a working prototype of the Datta-Das spin FET because of difficulties in efficiently injecting spin currents from a ferromagnetic metal into a semiconductor. Although this remains a controversial subject, recent optical experiments carried out at various laboratories around the world indicate that efficient spin injection into semiconductors can indeed be achieved by using unconventional materials, called magnetic semiconductors, that incorporate magnetism by doping the semiconductor crystals with atoms such as manganese.

Some magnetic semiconductors have been engineered to show ferromagnetism, providing a spintronic component called a gateable ferromagnet, which may one day play an important role in spin transistors. In this device, a small voltage would switch the semiconductor between nonmagnetic and ferromagnetic states. A gateable ferromagnet could in turn be used as a spin filter—a device that, when switched on, passes one spin state but impedes the other.

The filtering effect could be amplified by placing the ferromagnet in a resonant tunnel diode. Conventional resonant tunnel diodes allow currents to flow at a specific voltage, one at which the electrons have an energy that is resonant with the tunneling barrier. The version incorporating a ferromagnet would have a barrier with different resonant voltages for up and down spins.

The most exciting developments in semiconductor spintronics will probably be devices we have not imagined yet. A key research question for this second category of spintronics is how well electrons can maintain a specific spin state when traveling through a semiconductor or crossing from one material to another. For instance, a spin FET will not work unless the electrons remain polarized on entering the channel and after traveling to its far end.

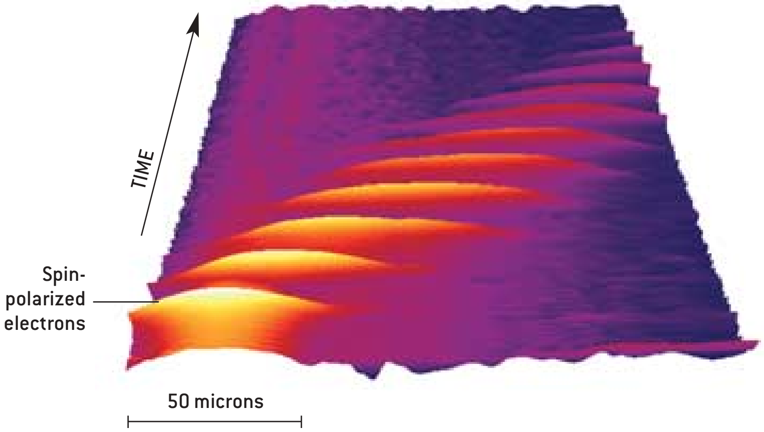
The question of how fast spin polarization decays becomes all the more acute if one is to build a quantum computer based on electron spins. That application requires control over a property known as quantum coherence, in essence the pure quantum nature of all the computer's data-carrying components. Quantum data in semiconductors based on the charges of electrons tend to lose coherence, or dissipate, in mere picoseconds, even at cryogenic temperatures. Quantum data based on spin should be intrinsically more sturdy. Curiously enough, our research groups stumbled on significant basic results regarding coherent electron spins while doing experiments aimed at developing practical magnetic semiconductors.

Semiconductor-based spintronics could combine logic, storage and communications on a single chip.

A Pleasant Surprise

IN 1997 AT the University of California at Santa Barbara, we were experimenting on zinc selenide (ZnSe), a long-studied conventional semiconductor. The ZnSe was meant to be a control for an ongoing project studying magnetic semiconductors. In our experiment we used pulses of circularly polarized light to excite pools of electrons in the ZnSe into identical spin states. In a circularly polarized light wave, instead of oscillating in intensity, the electric and magnetic fields rotate in a circle, transverse to the direction of the light.

SPIN DRAGGING IN SEMICONDUCTORS



POOLS OF ELECTRONS in spin-polarized states are dragged more than 100 microns by an electric field—a fundamental step toward technology relying on large-scale quantum coherence. The peaks diminish as quantum coherence of the states is lost.

We sent the ultrashort (100-femtosecond) pulses horizontally through the semiconductor, exciting electrons into horizontal spin states, initially aligned with the light beam. In a vertical ambient magnetic field the electron spins precess—the direction of each electron's spin vector rotates in the horizontal plane, similar to how a tilted gyroscope precesses in the earth's gravitational field. The precession enables us to monitor how long these states remain coherent, but the horizontal spin state has another, more important property.

For a baseball, say, horizontal spinning is nothing special and is quite distinct from the two vertical modes of spinning. For electrons, however, the horizontal quantum spin states are actually coherent superpositions of the spin-up and spin-down states. In effect, such electrons are in both the up and the down state at the same time. This is precisely the kind of coherent superposition of states employed by quantum computers [*see box on opposite page*].

Each electron spin can represent a bit; for instance, a 1 for spin up and a 0 for spin down. With conventional computers, engineers go to great lengths to ensure that bits remain in stable, well-defined states. A quantum computer, in contrast, relies on encoding information within quantum bits, or qubits, each of which can exist in a superposition of 0 and 1. By having a large number of qubits in superpositions of alternative states, a quantum computer intrinsically contains a massive parallelism so that quantum algorithms can operate on many different numbers simultaneously.

Unfortunately, in most physical systems, interactions with the surrounding environment rapidly disrupt these superposition states. A typical disruption would effectively change a superposition of 0 and 1 randomly into either a 0 or a 1, a process called decoherence. State-of-the-art qubits based on the charge of electrons in a semiconductor remain coherent for a few picoseconds at best—and only at temperatures too low for practical applications. The rapid decoherence occurs because the electric force between charges is strong and long range. In traditional semiconductor devices, this strong interaction is beneficial, permitting delicate control of current flow with small electric fields. To quantum coherent devices, however, it is anathema.

Electron-spin qubits interact only weakly with the environment surrounding them, principally through magnetic fields that are nonuniform in space or changing in time. Such fields can be effectively shielded. The goal of our experiment was to create some of these coherent spin states in a semiconductor to see how long they could survive. The results are also useful for understanding how to design devices such as spin transistors that do not depend on maintaining and detecting the quantum coherence of an individual electron's spin.

Our experiment measured the decoherence rate by monitoring the precession of the spins. Each electron would continue precessing as long as its superposition remained coherent. We used weak pulses of light to monitor the precession, in effect obtaining stroboscopic images of the spin dynamics. As the electrons precessed, the measured signal oscillated, in magnitude; as coherence was lost, the amplitude of the oscillations fell to zero.

Much to our surprise, the optically excited spin states in ZnSe remained coherent for several nanoseconds at low temperatures—1,000 times as long as charge-based qubits. The states even survived for a few nanoseconds at room temperature. Subsequent studies of electrons in gallium arsenide (GaAs, a high-quality semiconductor commonly used in everyday applications such as cellular phones and CD players) have shown that, under optimal conditions, spin coherence in a semiconductor could last *hundreds* of nanoseconds at low temperatures.

Hazards of Holes

THESE EXPERIMENTS also revealed characteristics that are crucial for attaining long spin coherence times. Of primary importance is the nature of the carriers of spin and charge. A semiconductor has two key bands of states that can be occupied by electrons: a valence band, which is usually full, and (at a slightly higher energy) a conduction band, which is usually empty. Charge carriers in semiconductors come in two flavors: conduction electrons, which are electrons in the conduction band, and valence holes, which are electrons missing from the valence band. The holes carry a spin because in a filled valence band all the spins cancel out: the removal of one electron leaves a net spin imbalance in the same way that it leaves behind a net positive charge.

Holes have dramatically shorter spin coherence times than electrons, and spin exchange between electrons and holes is very efficient, accelerating the decoherence of both. For these reasons, it pays to have no hole carriers, a condition that is achieved by using *n*-doped semiconductor crystals, which are doped to have some excess electrons in the conduction band without any corresponding valence holes.

When holes have been eliminated, the dominant remaining source of decoherence comes from a relativistic effect: a body moving at high speed through an electric field sees the field partially transformed into a magnetic field. For an electron moving in a semiconductor, the crystal structure of the material provides the electric field. The spin of a fast-moving electron precesses around the resulting local magnetic field that it sees. In each of our ensembles, the 10 billion or so excited electrons have a range of velocities and therefore precess in a variety of ways. Two electron spins that start off parallel can soon evolve to point in opposite directions. As this misalignment among the electrons grows, the average spin polarization of the population diminishes, which our experiment measures as loss of coherence. This population-based origin of decoherence holds forth the hope that the spin coherence times of individual electrons may turn out to greatly exceed even the remarkably long times seen in ensembles.

SPINTRONIC QUBITS

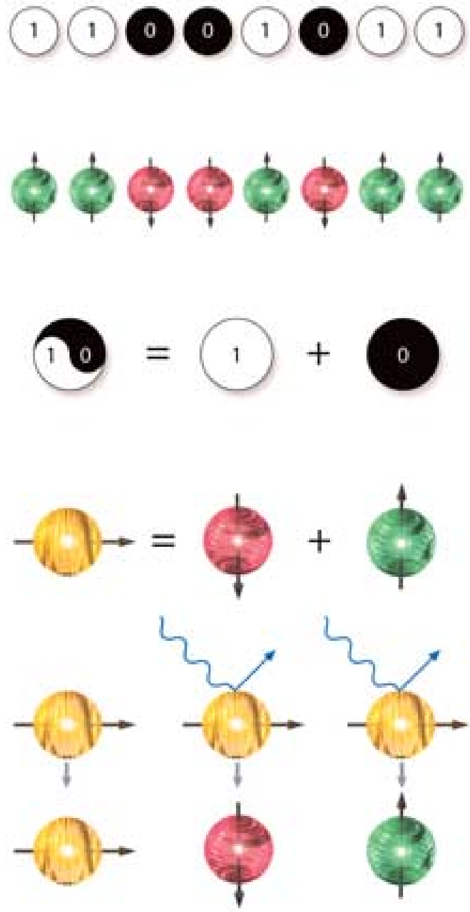
In a conventional computer every bit has a definite value of 0 or 1. A series of eight bits can represent any number from 0 to 255, but only one number at a time.

Electron spins restricted to spin up and spin down could be used as bits.

Quantum bits, or qubits, can also exist as superpositions of 0 and 1, in effect being both numbers at once. Eight qubits can represent every number from 0 to 255 simultaneously.

Electron spins are natural qubits: a tilted electron is a coherent superposition of spin up and spin down and is less fragile than other quantum electronic states.

Qubits are extremely delicate: stray interactions with their surroundings degrade the superpositions extremely quickly, typically converting them into random ordinary bits.



Spinning into the Future

IN CONJUNCTION WITH the carrier's lifetime, two other properties are crucial for semiconductor applications: how far excitations can be transported and how fast the state of a device can be manipulated. Macroscopic spin transport was first demonstrated in *n*-doped gallium arsenide. A laser pulse excited a "puddle" of coherently precessing electrons, much as in the lifetime experiments, but then a lateral electric field dragged the electrons through the crystal. Spin packets traveled more than 100 microns (a distance far exceeding the feature sizes in contemporary microelectronics) with only moderate loss of spin polarization [*see illustration on opposite page*]. Recent experiments have successfully driven coherent spins across complex interfaces between semiconductor crystals of different composition (for instance, from GaAs into ZnSe). A wealth of semiconductor applications, from lasers to transistors, are based on heterostructures, which combine disparate materials. The same design techniques can be brought to bear on spintronics.

Further advances toward quantum information processing have also taken place. For example, 150-femtosecond laser pulses have been used to tilt coherent electron spins, demonstrating that such spins can, in principle, be manipulated thousands of times before their coherence is lost. Meanwhile researchers with nearer-term goals have taken strides in making new magnetic semiconductors, which may at last open the door to practical spin transistors. On every front, the spintronics revolution is racing ahead and will continue to generate technologies that would be inconceivable in a nonquantum world. SA

MORE TO EXPLORE

Electrical Spin and Optical Coherence in Semiconductors. D. D. Awschalom and J. M. Kikkawa in *Physics Today,* Vol. 52, No. 6, pages 33–38; June 1999.

Meet the Spin Doctors. P. Ball in *Nature,* Vol. 404, pages 918–920; April 27, 2000.

Ultrafast Manipulation of Electron Spin Coherence. J. A. Gupta, R. Knobel, N. Samarth and D. D. Awschalom in *Science,* Vol. 292, pages 2458–2461; June 29, 2001.

Spintronics: A Spin-Based Electronics Vision for the Future. S. A. Wolf et al. in *Science,* Vol. 294, pages 1488–1495; November 16, 2001.

Microchips That Never Forget. A. Cho in *Science,* Vol. 296, pages 246–249; April 12, 2002.

Semiconductor Spintronics and Quantum Computation. Edited by D. D. Awschalom, D. Loss and N. Samarth. Springer Verlag, 2002.

For more about using spintronics for quantum computing, see **www.sciam.com/2002/602issue/0602awschalom.html**

KIM PEEK, who is developmentally disabled, knows more than 7,600 books by heart as well as every area code, highway, zip code and television station in the U.S. He provided the inspiration for the character Raymond Babbitt in the movie *Rain Man.*

Islands of Genius

Artistic brilliance and a dazzling memory can sometimes accompany autism and other developmental disorders

By Darold A. Treffert and Gregory L. Wallace

Photographs by Ethan Hill

LESLIE LEMKE is blind and has never studied piano. Although he suffers from cerebral palsy and is developmentally disabled, he composes music and is able to play thousands of pieces flawlessly, even when he has heard them only once.

Leslie Lemke is a musical virtuoso. At the age of 14 he played, flawlessly and without hesitation, Tchaikovsky's Piano Concerto No. 1 after hearing it for the first time while listening to a television movie several hours earlier. Lemke had never had a piano lesson—and he still has not had one. He is blind and developmentally disabled, and he has cerebral palsy. Lemke plays and sings thousands of pieces at concerts in the U.S. and abroad, and he improvises and composes as well.

Richard Wawro's artwork is internationally renowned, collected by Margaret Thatcher and Pope John Paul II, among others. A London art professor was "thunderstruck" by the oil crayon drawings that Wawro did as a child, describing them as an "incredible phenomenon rendered with the precision of a mechanic and the vision of a poet." Wawro, who lives in Scotland, is autistic.

Kim Peek is a walking encyclopedia. He has memorized more than 7,600 books. He can recite the highways that go to each American city, town or county, along with the area and zip codes, television stations and telephone networks that serve them. If you tell him your date of birth, he can tell you what day of the week it fell on and what day of the week it will be when you turn 65 "and can retire." Peek can identify most classical compositions and knows the date the music was published or first performed as well as the composer's birthplace and dates of birth and death. He is also developmentally disabled and depends on his father for many of his basic daily needs. His abilities provided the inspiration for the character Raymond Babbitt, whom Dustin Hoffman played in the 1988 movie *Rain Man*.

Lemke, Wawro and Peek all have savant syndrome, an uncommon but spectacular condition in which people with various developmental disabilities, including autism, possess astonishing islands of ability and brilliance that stand in jarring juxtaposition to their overall mental handicap. Savant syndrome is seen in about one in 10 people with autism and in approximately one in 2,000 people with brain damage or mental retardation. Of the known savants, at least half are autistic and the remainder have some other kind of developmental disorder.

Much remains mysterious about savant syndrome. Nevertheless, advances in brain imaging are permitting a more complete view of the condition, and a long-standing theory of left hemispheric damage has found support in these imaging studies. In addition, new reports of the sudden appearance of savant syndrome in people with certain forms of dementia have raised the intriguing possibility that some aspects of such genius lie dormant in all of us.

Down's Definition

DESCRIPTIONS of savant syndrome appear in the scientific literature as early as 1789. Benjamin Rush, the "father of American psychiatry," described the lightning-quick calculating ability of Thomas Fuller, who understood little math more complex than counting. When Fuller was asked how many seconds a man had lived by the time he was 70 years, 17 days and 12 hours old, he gave the correct answer of 2,210,500,800 a minute and a half later—and he had taken into account 17 leap years.

It was not until 1887, however, that the remarkable coexistence of deficiency and superiority was more completely laid out. That year J. Langdon Down, who is best known for having identified Down syndrome, described 10 people with savant syndrome. He had met these fascinating individuals during his 30 years as superintendent of the Earlswood Asylum in London. He coined the now discarded term "idiot savant," using the then accepted classification of an idiot as someone with an IQ of less than 25, combined with a derivative of the French word *savoir*, which means "to know."

More than a century has passed since Down's original description. Today we know a great deal more about this perplexing set of abilities from the 100 or so cases described in the scientific literature. It is now clear that savant syndrome generally occurs in people with IQs between 40 and 70—although it can occur in some with IQs as high as 114. It disproportionately affects males, with four to six male savants for every one female. And it can be congenital or acquired later in life following disease (such as encephalitis) or brain injury.

Narrow Repertoire

THE SKILLS THAT savant syndrome gives rise to are limited for the most part, and they tend to be based in the right hemi-

RICHARD WAWRO is an internationally renowned Scottish painter who has been exhibiting his work since he was 17 years old. He is autistic.

sphere. That is, they are predominantly nonsymbolic, artistic, visual and motor. They include music, art, mathematics, forms of calculating and an assortment of other abilities, such as mechanical aptitude or spatial skills. In contrast, left hemisphere skills are more sequential, logical and symbolic; they include language and speech specialization [see "The Split Brain Revisited," by Michael S. Gazzaniga; SCIENTIFIC AMERICAN, July 1998].

Most musical savants have perfect pitch and perform with amazing ease, most often on the piano. Some are able to create complex compositions. And for some reason, musical genius often seems to accompany blindness and mental retardation, as it does for Lemke. One of the most famous savants was "Blind Tom" Bethune, who lived from 1849 to 1908. In his time, he was referred to as "the eighth wonder of the world." Although he could speak fewer than 100 words, he could play beautifully more than 7,000 pieces on the piano, including many of his own works. (Some of his compositions were recently recorded by musician John Davis and released on CD.)

For their part, savant visual artists use a variety of media, although they most frequently express themselves through drawing and sculpture. Artistic savant Alonzo Clemons, for example, can see a fleeting image of an animal on a television screen and in less than 20 minutes sculpt a perfect replica of that animal. His wax model will be correct in every detail, every fiber and muscle and proportion.

Mathematical savants calculate incredibly rapidly and often have a particular facility with prime numbers. Curiously, the obscure skill of calendar calculating that Peek demonstrates is not confined to mathematical savants; it seems to coexist with many different skills.

Several other abilities appear less frequently. A rare savant may have extensive language ability—that is, the capacity to memorize many languages but not to understand them. Other unusual traits include heightened olfactory, tactile and visual sensitivity; outstanding knowledge in fields such as history, neurophysiology, statistics or navigation; and spatial ability. For instance, a musical and blind savant named Ellen can navigate in thick forests or other unfamiliar spaces without running into objects. Ellen also has a perfect appreciation of passing time despite the fact that she doesn't have access to a watch or clock, even in Braille. This ability was discovered one day when her mother let her listen to the "time lady" on the telephone. After listening for a short while to the recorded voice intone the hour and seconds, Ellen apparently set her own internal clock. Since then, she has been able to tell what time it is to the second, no matter the season.

Savant skills are always linked to a remarkable memory. This memory is deep, focused and based on habitual recitation. But it entails little understanding of what is being described. Some early observers aptly called this "memory without reckoning." Down himself used the phrase "verbal adhesion" to characterize it. One of his patients was a boy who had read the six-volume *History of the Decline and Fall of the Roman Empire,* by Edward Gibbon, and could recite it back word for word, although he did so without any comprehension.

Although they share many talents, including memory, savants vary enormously in their levels of ability. So-called splinter-skill savants have a preoccupation and mild expertise with, say, the memorization of sports trivia and license plate numbers. Talented savants have musical or artistic gifts that are conspicuously above what would be expected of someone with their handicaps. And prodigious savants are those very uncommon people whose abilities are so advanced that they would be distinctive even if they were to occur in a normal person. Probably fewer than 50 prodigious savants are alive at the moment.

Whatever their talents, savants usually maintain them over

THE AUTHORS

DAROLD A. TREFFERT and *GREGORY L. WALLACE* share a longstanding interest in savant syndrome. Treffert is a psychiatrist at St. Agnes Hospital in Fond du Lac, Wis., and has done research on autism and savant syndrome since 1962, the year he met his first savant. (His e-mail address is dtreffert@pol.net) Wallace is a visiting researcher at the Social, Genetic and Developmental Psychiatry Research Center at the Institute of Psychiatry in London. He is currently conducting research on why individuals with autism are more likely to develop savant skills. (His e-mail address is g.wallace@iop.kcl.ac.uk)

ALONZO CLEMONS can create perfect wax replicas of any animal he sees, no matter how briefly. His bronze statues are sold by a gallery in Aspen, Colo., and have earned him national repute. Clemons is developmentally disabled.

the course of their life. With continued use, the abilities are sustained and sometimes even improve. And in almost all cases, there is no dreaded trade-off of these wonderful abilities with the acquisition of language, socialization or daily living skills. Instead the talents often help savants to establish some kind of normal routine or way of life [*see box on page 85*].

Looking to the Left Hemisphere

ALTHOUGH SPECIALISTS TODAY are better able to characterize the talents of savants, no overarching theory can describe exactly how or why savants do what they do. The most powerful explanation suggests that some injury to the left brain causes the right brain to compensate for the loss. The evidence for this idea has been building for several decades. A 1975 pneumoencephalogram study found left hemispheric damage in 15 of 17 autistic patients; four of them had savant skills. (A pneumoencephalogram was an early and painful imaging technique during which a physician would inject air into a patient's spinal fluid and then x-ray the brain to determine where the air traveled. It is no longer used.)

A dramatic study published by T. L. Brink in 1980 lent further credence to the possibility that changes to the left hemisphere were important to savant syndrome. Brink, a psychologist at Crafton Hills College in California, described a normal nine-year-old boy who had become mute, deaf and paralyzed on the right side when a bullet damaged his left hemisphere. After the accident, unusual savant mechanical skills emerged. He was able to repair multigeared bicycles and to design contraptions, such as a punching bag that would weave and bob like a real opponent.

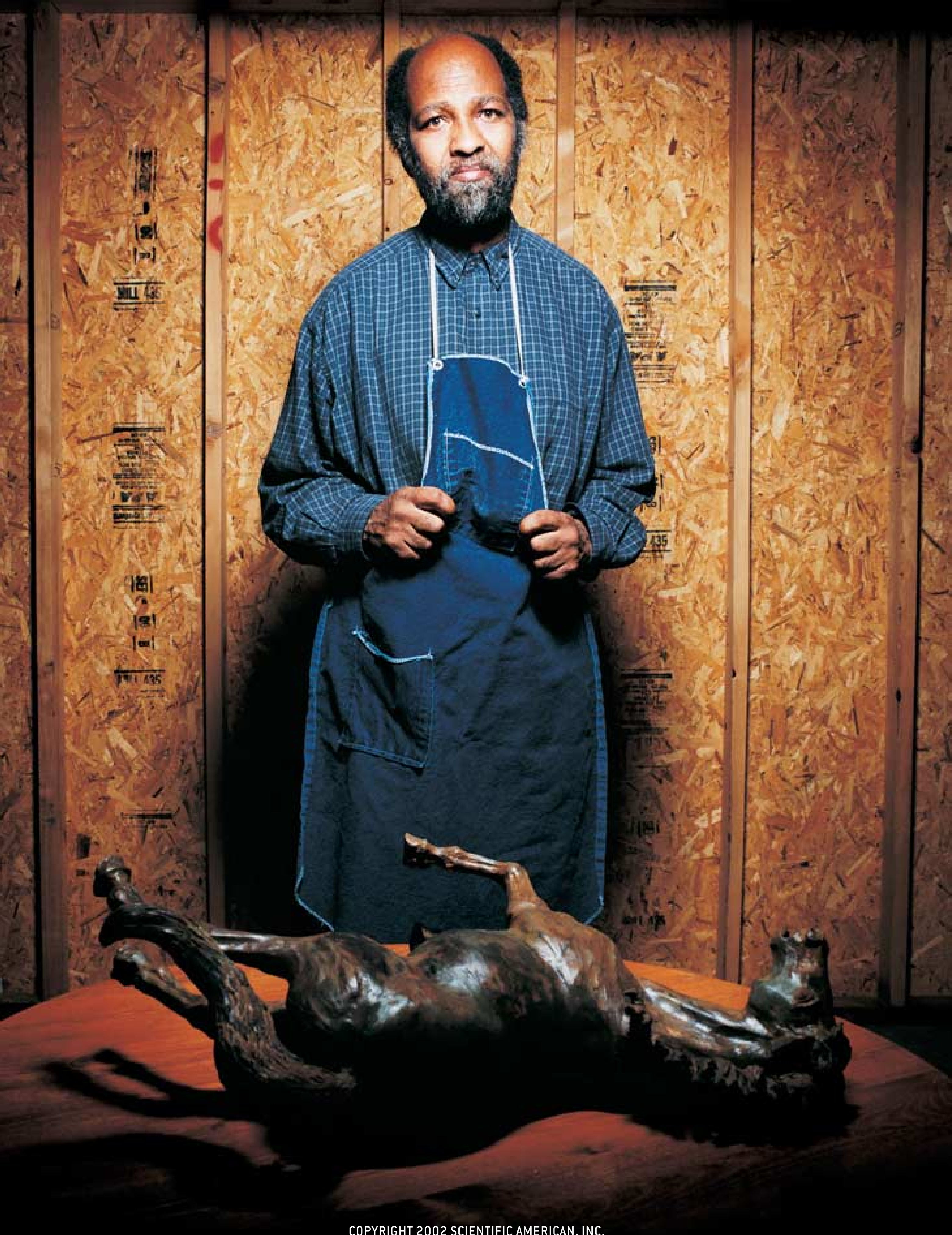
The findings of Bernard Rimland of the Autism Research Institute in San Diego support this idea as well. Rimland maintains the largest database in the world on people with autism; he has information on more than 34,000 individuals. He has observed that the savant skills most often present in autistic people are those associated with right hemisphere functions and the most deficient abilities are associated with left hemisphere functions.

In the late 1980s Norman Geschwind and Albert M. Galaburda of Harvard University offered an explanation for some causes of left hemispheric damage—and for the higher number of male savants. In their book *Cerebral Lateralization,* the two neurologists point out that the left hemisphere of the brain normally completes its development later than the right and is therefore subject to prenatal influences—some of them detrimental—for a longer period. In the male fetus, circulating testosterone can act as one of these detrimental influences by slowing growth and impairing neuronal function in the more vulnerable left hemisphere. As a result, the right brain often compensates, becoming larger and more dominant in males. The greater male-to-female ratio is seen not just in savant syndrome but in other forms of central nervous system dysfunction, such as dyslexia, delayed speech, stuttering, hyperactivity and autism.

Newly Savant

IN RECENT YEARS, more data have emerged to support the left hemisphere hypothesis. In 1998 Bruce L. Miller of the University of California at San Francisco examined five elderly patients with frontotemporal dementia (FTD), one form of presenile dementia. These patients had developed artistic skills with the onset and progression of their dementia. They were able to make meticulous copies of artworks and to paint beautifully. Consistent with that in savants, the creativity in these five individuals was visual, not verbal. Single-photon-emission computed tomography (SPECT) showed that injury was predominantly on the left side of the brain. Miller examined seven other patients who had developed musical or artistic ability after the appearance of FTD. He found damage on the left as well.

Miller, Craig Hou of Washington University and others then compared these images with those of a nine-year-old artistic autistic savant named DB. SPECT scans of DB revealed a higher-than-normal blood flow in part of his neocortex but decreased flow in his left temporal lobe. (The neocortex is involved with high-level cognitive function; the temporal lobe is responsible for some aspects of memory and emotion.) Miller is hoping to study other artistic savants to see if the findings hold true for them as well. But the fact that DB and older FTD patients with newfound savant skills have the same pathology is quite striking and suggests that researchers will soon be able to identify pre-



cisely the neurological features associated with savant syndrome.

The seemingly limitless memory of savants will mostly likely be harder to pinpoint physiologically. Mortimer Mishkin of the National Institute of Mental Health has proposed different neural circuits for memory, including a higher-level corticolimbic circuit for what is generally referred to as semantic or cognitive memory, and a lower-level corticostriatal circuit for the more primitive habit memory that is most often referred to as procedural memory. The memory of savants seems to be the noncognitive habit form.

The same factors that produce left hemispheric damage may be instrumental in producing damage to higher-level memory circuits. As a result, savants may be forced to rely on more primitive, but spared, habit memory circuits. Perhaps brain injuries—whether they result from hormones, disease, or prenatal or subsequent injury—produce in some instances certain right brain skills linked with habit memory function. In those situations, savant syndrome may appear.

Rain Man in Us All?

THE EMERGENCE of savantlike skills in people with dementia raises profound questions about the buried potential in all of us. Accordingly, several researchers are seeking to unlock what has been called the "little Rain Man in each of us." One group has used a technique called repetitive transcranial magnetic stimulation (rTMS) in 17 normal individuals, eight male and nine female. Tracy Morrell of the University of South Australia, Robyn L. Young of Flinders University in Adelaide and Michael C. Ridding of Adelaide University applied magnetic stimulation to the area in the left temporal lobe that Miller identified as damaged in his FTD patients.

In their as yet unpublished study, the team reports that only two participants experienced a series of short-lived skills, such as calendar calculating, artistic ability and enhanced habit memory. Others discovered a new skill here and there, also lasting just a few hours. The researchers suggest that savant skills may be limited to a small percentage of the normal population in the same way that they are limited to a small percentage of the disabled population.

Nevertheless, many experts believe that real potential exists to tap into islands of savant intelligence. Allan Snyder and John Mitchell of the Centre for the Mind in Canberra, Australia, argue that savant brain processes occur in each of us but are overwhelmed by more sophisticated conceptual cognition. Autistic savants, they conclude, "have privileged access to lower levels of information not normally available through introspection."

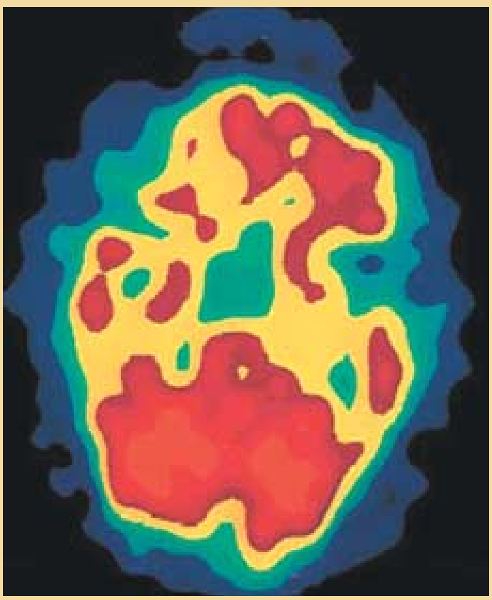
Our view is also that all of us have some of the same circuitry and pathways intrinsic to savant functioning but that these are

Emerging Savants

IN RECENT YEARS, researchers have discovered that certain patients who have frontotemporal dementia (FTD) can paint beautifully when they previously had no such talent. In short, they have become savantlike as dementia has taken hold.

This painting of horses (*right*) was made by one of these patients, a 64-year-old woman with FTD. Bruce L. Miller of the University of California at San Francisco has examined many FTD patients and has documented damage to the left side of their brain. One theory suggests that savant skills may emerge in the more artistic right hemisphere as a way of compensating for damage in the left. In this SPECT image (*left*) of an FTD patient, enhanced blood flow can be seen in a part of the right hemisphere (*red*).

—D.A.T. and G.L.W.



COURTESY OF BRUCE L. MILLER *U.C.S.F.*

Living with Savant Syndrome

A FEW REPORTS in the literature suggest that when savants are encouraged to acquire better language skills they lose their special artistic talents. Perhaps the most famous of these cases is that of Nadia, a girl with autism who by the age of three was producing astounding drawings. When she turned seven, Nadia entered a school for autistic children that focused on verbal abilities; by the time she was a teenager, Nadia was more verbal but could no longer create brilliant and intricate drawings.

This trade-off between talent and language or socialization is not something we have witnessed. Instead the exceptional abilities of savants have proved to be strengths that are built on and used as a conduit toward normalization; these skills have helped individuals develop improved social skills, better language acquisition and greater independence. Savants gain a sense of accomplishment because of their talent; that sense, in turn, allows them to participate more fully in the world. Musical prodigy Leslie Lemke has become more animated, performing concerts and interacting with audiences. Painter Richard Wawro feels delight and excitement when he finishes a work, and he seeks out celebration. And memory wizard Kim Peek has emerged from the social isolation that characterized him before the movie *Rain Man* was made; he now travels the country talking to hundreds of school groups.

Fortunately, simultaneously encouraging savant abilities and normalization is proving to be the accepted approach to such individuals' care. Savants are being placed in some classes for the gifted and talented, an opportunity that promotes social growth for both them and their classmates. Some new programs, such as the one at Hope University in Anaheim, Calif., cater entirely to these exceptional individuals. Others include people with similar disorders as well; for example, music and art camps have been established for those with Williams syndrome, many of whom have savantlike musical skills [see "Williams Syndrome and the Brain," by Howard M. Lenhoff, Paul P. Wang, Frank Greenberg and Ursula Bellugi; SCIENTIFIC AMERICAN, December 1997]. Nurturing the talent is the most fulfilling approach. *—D.A.T. and G.L.W.*

FERRY BOAT TO TIREÉ, SCOTLAND, was painted in 1978 by Richard Wawro. Wawro is cared for by his father, who enthusiastically supports his painting. Most researchers believe that such artistic skill may help savants develop better social skills.

less accessible—in part because we tend to be a left-brain society. Sometimes, though, we can find elements of the savant in ourselves. At certain moments, we just "get" something or discover a new ability. And some procedures—including hypnosis; interviews of subjects under the influence of the barbiturate sodium amytal, which induces relaxation; and brain stimulation during neurosurgery—provide evidence that a huge reservoir of memories lies dormant in every individual. Dreams can also revive those memories or trigger new abilities.

A Window into the Brain

NO MODEL OF BRAIN FUNCTION will be complete until it can explain this rare condition. Now that we have the tools to examine brain structure and function, such studies can be correlated with detailed neuropsychological testing of savants. We hope the anecdotal case reports that have characterized the literature on this topic for the past century will soon be replaced by data comparing and contrasting groups of normal and disabled people, including prodigies, geniuses and savants.

Savant syndrome provides a unique window into the brain with regard to questions of general intelligence versus multiple forms of intelligence. It may also shed light on brain plasticity and central nervous system compensation, recruitment and repair—areas of research that are vital in understanding and treating such diverse conditions as stroke, paralysis and Alzheimer's disease.

But savant syndrome has relevance outside the scientific realm. Many lessons can be learned from these remarkable people and their equally remarkable families, caretakers, therapists and teachers. One of the greatest lessons is that they have been shaped by far more than neural circuitry. The savants thrive because of the reinforcement provided by the unconditional love, belief and determination of those who care for them. Savant syndrome promises to take us further than we have ever been toward understanding both the brain and human potential. SA

MORE TO EXPLORE

Emergence of Artistic Talent in Frontotemporal Dementia. B. Miller, J. Cummings and F. Mishkin et al. in *Neurology,* Vol. 51, No. 4, pages 978–982; October 1, 1998.

Extraordinary People: Understanding Savant Syndrome. Darold A. Treffert. iUniverse.com, Inc., 2000.

www.wisconsinmedicalsociety.org/savant

CREATIVE LEARNING ENVIRONMENTS, AUSTIN, TEX.

One of life's simple pleasures is really quite complicated

The Complexity of Coffee

By Ernesto Illy

Photographs by Tina West

For sheer sensory enjoyment, few everyday experiences can compete with a good cup of coffee. The alluring aroma of steaming hot coffee just brewed from freshly roasted beans can drag sleepers from bed and pedestrians into cafés. And many millions worldwide would find getting through the day difficult without the jolt of mental clarity imparted by the caffeine in coffee. But underlying this seemingly commonplace beverage is a profound chemical complexity. Without a deep understanding of how the vagaries of bean production, roasting and preparation minutely affect the hundreds of compounds that define coffee's flavor, aroma and body, a quality cup would be an infrequent and random occurrence.

Connoisseurs agree that the quintessential expression of coffee is espresso: that diminutive heavy china cup half-filled with a dark, opaque brew topped by a velvety thick, reddish-brown froth called *crema*. Composed of tiny gas bubbles encased in thin films, the surprisingly persistent *crema* locks in the coffee's distinctive flavors and aromas and much of its heat as well. Espresso—the word refers to a serving made on request expressly for the occasion—is brewed by rapidly percolating a small quantity of pressurized, heated water through a compressed cake of finely ground roasted coffee. The resulting concentrated liquor contains not only soluble solids but also a diverse array of aromatic substances in a dispersed emulsion of tiny oil droplets, which together give espresso its uniquely rich taste, smell and "mouthfeel."

Aficionados consider perfectly brewed espresso to be the ultimate in coffee because its special preparation amplifies and exhibits the inherent characteristics of the beans. Espresso is useful for our purposes as it is in effect a distillation of all the numerous techniques by which coffee can be made, including the Turkish method and various infusion and filter drip processes [*see box on page 91 for descriptions of alternative coffee-preparation methods*]. To know espresso is to know coffee in all its forms.

High-quality coffee arises from maintaining close control over a multitude of factors in the field, in the plant and in the cup. Coffee cultivation entails myriad variables that must be monitored and regulated. Once a coffee bean is grown, nothing can be added or removed: the quality must already be present. For a single portion of espresso, 50 to 55 roasted coffee beans are required; a single imperfect bean will taint the whole sufficiently to be noticeable. This is because human olfaction and taste senses originated as defense mechanisms that protected our ancestors from rotten—hence, unhealthy—foods. Only through modern technology can one economically and consistently identify 50 nearly perfect beans.

Growing Coffee

RAW COFFEE BEANS are the seeds of plants belonging to the Rubiaceae family, which comprises at least 66 species of the genus *Coffea*. The two species that are commercially exploited are *Coffea arabica,* which accounts for two thirds of world production, and *C. canephora,* often called robusta coffee, with one third of global output. Robusta coffee plants and all wild coffee species have 22 chromosomes, whereas arabica has 44. Therefore, arabica and other coffee species cannot be crossed to produce a hybrid plant.

Robusta is a high-yielding and disease-resistant tree standing up to 12 meters tall that grows best in warm, humid climes. It produces a cup featuring substantial body, a relatively harsh, earthy aroma, and an elevated caffeine content that ranges from 2.4 to 2.8 percent by weight. Although robusta is sold by many purveyors, it does not give rise to the highest-quality coffee.

Arabica, which originated in the Ethiopian highlands, is a medium- to low-yielding, rather delicate tree from five to six meters tall that requires a temperate climate and considerable growing care. Commercially grown coffee bushes are pruned to a height of 1.5 to 2.0 meters. Coffee made from arabica beans has an intense, intricate aroma that can be reminiscent of flowers, fruit, honey, chocolate, caramel or toasted bread. Its caffeine content never exceeds 1.5 percent by weight. Because of its superior quality and taste, arabica sells for a higher price than its hardy, rougher cousin.

A good rainfall induces arabica coffee plants to blossom, and some 210 days afterward red or yellow fruit called cherries appear. Each cherry contains two oblong seeds—the coffee beans. Because both flower and fruit can be present simultaneously on the same branch, the picker's forefinger and thumb are the best tools to gather just the ripe cherries. Stripping entire branches by hand or using automated harvesting machines does not discriminate between the ripe and the unripe cherries.

The ultimate quality of the resulting coffee beans depends on the genetics of the plant, the soil in which it grows and the microclimate, which encompasses factors such as altitude, the amount of rainfall and sunlight, and daily temperature fluctuations. Along with the roasting processes that are applied, these agricultural and geographical considerations are responsible for the taste differences among the many varieties of coffee beans that suppliers combine to produce the various distinctive blends one can purchase.

Processing Coffee

COFFEE CHERRIES must be processed immediately after harvest to prevent spoilage. Producers employ two processing methods: sun-drying and washing. Effective sun-drying is accomplished by spreading the cherries out on a patio and stirring the desiccating fruit frequently to evenly heat and aerate it. The dried cherries are run through a machine that crushes the hulls and then removes both the hulls and the surrounding parchment membrane layer, thus freeing the beans for sorting and bagging. In the alternative approach, the fruit is mechanically pulped, washed, and finally dried and liberated from the parchment covering. The goal of either route is the same: the 65 percent water content of the coffee cherry is reduced to the 10 to 12 percent moisture level of a prime raw, or green, coffee bean.

One of the greatest challenges in producing superior coffee is ensuring that one starts with exceptional green beans. Premium producers, such as illycaffè, based in Trieste, Italy, use many sophisticated process-control techniques to minimize the percentage of defective coffee beans, including ultraviolet fluorescence analysis to spot moldy beans and trichromatic mapping to generate a color fingerprint (yellow-green, red and infrared) of each lot of beans. At illycaffè, a dichromatic sorting system developed in collaboration with the English company Sortex is applied as a final control right before roasting. As beans fall into bins, photoelectric cells detect duds, signaling for them to be rejected individually with a puff from an air nozzle. The sorting operation is accomplished at a speed that no human hand can match (400 beans a second) and with a precision that even the most highly trained eye is incapable of.

A perfect mature green coffee bean is composed of cells with uncommonly thick walls: as much as five to seven microns, an exception in the vegetal kingdom. During roasting, these 30- to 40-micron-diameter cells serve as tiny reactors in which all the key heat-driven chemical reactions occur that generate coffee's seductive taste and fragrance. The cells of immature beans feature thinner walls. Unripe beans also lack the important aromatic precursor proteins that develop in the last stages of the ripening process.

THE AUTHOR

ERNESTO ILLY is chairman of illycaffè, a family business based in Trieste, Italy, that his father founded in 1933. Over two million illy espressos are served every day in Italy alone. Illy holds a doctorate in chemistry and has completed advanced studies in molecular biology. His goal is to harness science to create the truly perfect cup of espresso.

Fermented beans are composed of cells that have been emptied of these crucial constituents by molds or bacteria.

Roasting Coffee

ROASTING is a pyrolytic (heat-driven) process that greatly increases the chemical complexity of coffee. The aroma of green coffee contains some 250 different volatile molecular species, whereas roasted coffee gives rise to more than 800.

When subjected to the staged heating of a roasting machine (basically, a huge, hot rotating cylinder), residual water inside each cell is converted to steam, which promotes diverse, complicated chemical reactions among the cornucopia of sugars, proteins, lipids and minerals within [*see box on next page*]. At high heat, from 185 to 240 degrees Celsius, sugars combine with amino acids, peptides and proteins according to a well-known caramelization process called Maillard's reaction. The end products are brownish, bittersweet glycosylamine and melanoidins—which give rise to coffee's dominant taste—along with carbon dioxide (up to 12 liters per kilogram of roasted coffee).

Simultaneously, a wide variety of lower-mass aromatic molecules emerge; these volatile compounds give coffee its familiar fragrance. Pressure inside each cell increases to as much as 20 to 25 atmospheres as the steam and carbon dioxide try to escape but are sealed in by the thick, low-porosity cell walls and a coating of oil. Some cells eventually burst, creating the characteristic popping sound of roasting coffee. During roasting, coffee bean volume increases by half or more; bean mass decreases by a fifth.

Depending on the temperatures and procedures applied, the roasting process can last from 90 seconds to as long as 40 minutes. Twelve minutes, however, is the traditional duration. The thermodynamics of the intracellular reactions differ according to roasting time, and so does the final result. A short roasting time, which requires a great deal of thermal energy, minimizes weight loss but imparts to the cup a metallic bitterness stemming from the presence of polyphenols that do not have enough time to react properly. Long roasting periods, frequently used in poorer countries in which many consumers can afford only low-priced, defective beans, forces all the off-flavors and fragrances to leave the beans. Sadly, the desirable tastes and aromas flee as well, yielding a rather bitter cup.

The higher the final temperature of the roasting, the less desirable the aroma will be and the stronger the bitterness. Conversely, low roasting temperatures fail to develop fully the welcome aromas, and acidity tends to come to the fore.

Smelling Coffee

AROMA SCIENCE is highly complex. Researchers typically analyze the fragrances evolved during coffee bean roasting by gas chromatography coupled with olfactometry, in which skilled testers sniff and define the smell of each recognizable element. Mass spectrometry is frequently then applied to identify the chemical composition of each odor. Sniffing roasted coffee aromas that have been fractionated by a gas chromatograph is an enlightening experience: one may recognize the aromas of roses, Darjeeling tea, chocolate, vanilla and violets, as well as truffles, soup, cheese, sweat and even what is called cat scent, which, if diluted, smells like sauvignon blanc wine but in a concentrated sample is disgusting.

In the laboratories of illycaffè, technicians focus on the strongest odorants. Imagine listening to a recording of a choir of 800 singers that includes the strong solo voices of Jessye Norman, Luciano Pavarotti and several other virtuosi who tend to dominate the ensemble. If the volume of the playback is reduced, the stronger voices will still be recognizable even as the choir's sound fades away. Diluting the aroma of coffee is analogous; beyond a certain point, only the strongest compounds are perceived. Unfortunately, the most powerful molecules in the smell of a coffee sample are those originating from defective beans.

Molecules such as ethylbutanoate and ethylglycolate, which are responsible for the unpleasant aroma of immature beans, ruin a cup of coffee by their very presence. Likewise, methylisoborneol and trichloroanisole (TCA) molecules produce the characteristic earthy, chemical smell of ro-

With more than 400 BILLION CUPS consumed every year, coffee is the world's most popular beverage (besides water, of course).

✷

Coffee ranks SECOND ONLY TO PETROLEUM in terms of dollars traded worldwide (about $10.4 billion in exports in 2000).

✷

ACCORDING TO A SURVEY OF 3,000 AMERICANS CONDUCTED BY THE NATIONAL COFFEE ASSOCIATION IN 2001:

Fifty-two percent of adults (older than 18 years) in the U.S. drink coffee every day, representing 107 MILLION daily drinkers. Another 28 percent of adults (57 MILLION) drink coffee occasionally.

American coffee drinkers consume, on average, 3.3 (NINE-OUNCE) CUPS OF COFFEE a day.

✷

Controversy continues over whether SHADE-GROWN coffee should be promoted to enhance bird habitats.

✷

Coffee and caffeine have been the subject of extensive scientific study during the past quarter of a century, with 1,500 TO 2,000 PAPERS PUBLISHED EVERY YEAR on the topic. Despite this close scrutiny, few negative health effects have been definitively linked to the moderate consumption (TWO CUPS A DAY) of caffeinated coffee. In fact, recent work indicates that roasted coffee can be a good source of antioxidants.

✷

1000 B.C. TO A.D. 500 *The nomadic Oromos tribe, living in the kingdom of Kefa (in modern-day Ethiopia), eat coffee, crushed, mixed with fat and shaped into golf ball–size portions, as a pick-me-up.*

CIRCA 600 *Coffee is brought by traders across the Red Sea into Arabia (modern-day Yemen).*

LATE 1400s TO EARLY 1500s *Coffee beans, heretofore an Arabian monopoly, are brought to Turkey, Egypt and Syria by Muslim pilgrims returning from Mecca. Arabian-influenced coffeehouses open in Constantinople, Damascus and other Near Eastern cities, where European traders, particularly Venetians, are introduced to the beverage.*

CIRCA 1600 *Calling it "the bitter invention of Satan," advisers urge Pope Clement VIII to reject the favorite drink of the infidel Ottoman Turks. Instead he decides to give papal authority to coffee, making it acceptable for Roman Catholics.*

1616 *Dutch entrepreneurs start to cultivate coffee commercially, beginning with a coffee plant obtained from Yemen. By 1658 (according to some sources, the 1690s), the Dutch are growing coffee in Ceylon and their East Indian colony of Java.*

1714 *The mayor of Amsterdam presents the French king, Louis XIV, with a coffee plant from Java.*

1723 *Gabriel Mathieu de Clieu, a French naval officer, carries three coffee seedlings obtained under questionable circumstances from the Royal Botanical Gardens on a perilous voyage to the Caribbean island of Martinique, where one of the plants thrives.*

1727 *After having been called to arbitrate a border dispute between two coffee-growing colonies, Dutch Guiana and French Guiana, Francisco de Melo Palheta, a Portuguese Brazilian official, smuggles several coffee seedlings to his home estates.*

1903 *German coffee importer Ludwig Roselius perfects decaffeinated coffee.*

1933 *Francesco Illy patents the first automatic espresso machine.*

1961 *Ernesto Valente of Faema, an Italian coffee machine maker, designs the archetype of the modern espresso machine.*

busta coffees. TCA, which is also called Rio taste because it was first discovered in coffees grown around Rio de Janeiro, can be found in corked wines as well. Its perception threshold to the human olfactory system is shockingly low—six millionths of a billionth of a gram per milliliter.

Preparing Coffee

THE NEXT MAJOR STEP in the transformation of roasted beans into a cup of espresso is the extraction of the active components in the roasted and ground coffee by heated water. The interaction of hot water and coffee grounds is, however, subtly different when making common drip coffee than when making espresso.

When filter drip coffee is prepared, hot water passes through a loose aggregation of medium-size coffee grounds. During the four to six minutes of contact with the boiling water, most of the soluble substances present in the roasted coffee pass into solution. Thus, large quantities of highly soluble acids and caffeine dissolve into the cup. In contrast, the much shorter percolation time of espresso allows less acid and only 60 to 70 percent of the caffeine to dissolve into the brew.

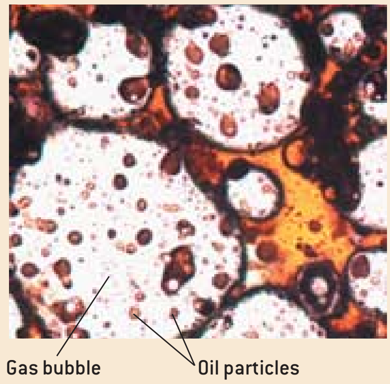
Brewing espresso requires specialized equipment that can heat water to a temperature of between 92 and 94 degrees C and pressurize it to nine atmospheres. Coffee, ground to a fine to medium consistency, is placed in a perforated basket and firmly tamped down to create a compacted bed of particles. The compressed grounds adhere to one another thanks to a thin coating of oil, which is as viscous as honey. The oil binds the particles together into a condensed maze of minuscule air passages. Experimentation has shown that the hydraulic resistance of this bed of coffee grounds must be slightly less than the pressure of the steaming-hot extraction water, allowing it to flow through at a rate of around a milliliter a second.

Using the recommended 30 seconds of percolation, a skilled *barista* (coffee bar technician) produces about 30 milliliters of dense coffee liquor covered by the all-important *crema*. If the color of the foam topping is light, it means that the espresso has been underextracted, probably because the grind was too coarse, the water temperature too low or the time too short. If the *crema* is very dark in hue and has a "hole" in the middle, it is likely that the consistency of the coffee grounds was too fine or the quantity of grounds was too large. An overextracted espresso exhibits either a white froth with large bubbles if the water was too hot or just a white spot in the center of the cup if the brewing time was too long.

The percolation process also washes out components present on the surface of the coffee grounds, including aroma-filled oil and bits of the cellular structure. The high pressure generated by the espresso

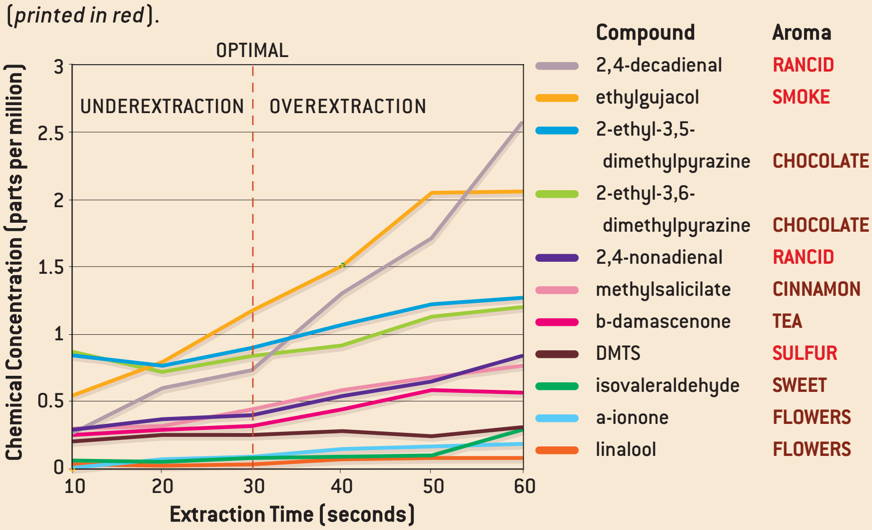
THE CHEMISTRY OF COFFEE

CREMA, the dense, reddish-brown foam that tops an espresso, is shown in an enlarged cross section. Composed mainly of tiny carbon dioxide and water vapor bubbles (*large circles*) surrounded by surfactant films, the *crema* also includes emulsified oils containing key aromatic compounds (*particles with red borders*) and dark fragments of the coffee bean cell structure.

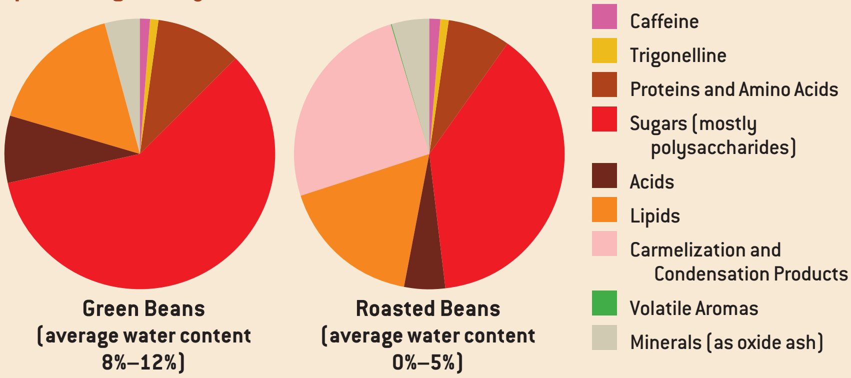


Cumulative Chemical Composition of Espresso with Increasing Extraction Time

THE OVEREXTRACTION of espresso (beyond the recommended 30 seconds) leads to the incorporation of undesirable and less soluble aromatic compounds into the drink (*printed in red*).



Chemical Composition of Raw and Roasted Arabica Coffee (percentage of dry matter)



ILLYCAFÉ (*micrograph*); NINA FINKEL (*illustrations*)

OTHER COFFEE-BREWING TECHNIQUES

FILTER DRIP METHODS (automatic drip, Melitta, Chemex pots). These popular techniques employ finely ground coffee in receptacles lined with filter paper. A medium grind should be used with a reusable gold filter. There are two keys to making superior coffee using these processes: first, rinse the paper filter with boiling water to remove the papery smell; second, ensure that the near-boiling-hot brewing water takes no more than four to six minutes to pass through the grounds, thereby producing optimal extraction levels. The brewing time of an automatic drip machine can be controlled by tailoring the quantity of water so that it flows for the recommended four to six minutes.

French press or plunger pot. This apparatus steeps the coffee in the hot water before the grounds are filtered out. Combine heated water and coarsely ground coffee in the pot and allow it to infuse from two to five minutes, depending on the desired strength. Then press the wire-mesh filter/plunger slowly through the infusion, segregating the grounds at the bottom of the pot.

Turkish method. Unlike other brewing processes, gentle boiling of the coffee is desirable when using this method. Mix equal amounts of pulverized coffee, water and sugar in a special pot called an *ibrik,* which sits directly over the heat. Stir the mixture as it comes to a slow boil. Stop stirring when the powdered coffee no longer sticks to a spoon. As the brew just begins to boil and foam up, remove the *ibrik* from the heat. Tap the *ibrik* to reduce the foam somewhat. Repeat the process at least two additional times. The result is a uniquely thick, sweet brew.

Adapted from The Great Coffee Book, *by Timothy J. Castle and Joan Nielsen (Ten Speed Press, Berkeley, Calif., 1999).*

machine emulsifies a small amount of the oils, about 0.1 gram a cup. Intact cells in the grounds create a fine effervescence, which is derived from gases (especially carbon dioxide) passing through tiny pores in the cell walls. Some very fine grounds can also find their way into the beverage, along with cell wall fragments, which endow the foamy *crema* with what is called the tiger-skin look.

The final result is a polyphasic colloidal system, in which water molecules are bound to the dispersed gas bubbles, oil droplets and solid fragments, all of which are less than five microns in size. The colloidal character of the dispersion gives the drink high body, high viscosity and low surface tension. Espresso thus visibly coats our tongues and continues to release the aromatic volatiles dissolved in the emulsified oils as long as it remains there. These oily flavor/fragrance carriers mean that the great taste and aroma of a good espresso can be savored for as long as 20 minutes after it has been drunk. Luckily, the drinker need not know anything about the complex chemistry of coffee to enjoy it. SA

A broadcast version of this article will air May 28 on *National Geographic Today,* a program on the National Geographic Channel. Please check your local listings.

MORE TO EXPLORE

Coffee, Vols. 1–6. R. J. Clarke and R. Macrae. Elsevier Applied Science, 1985.

Coffee: Botany, Biochemistry and Production of Beans and Beverage. M. N. Clifford and K. C. Willson. Croom Helm, London, 1985.

Caffeine, Coffee and Health. Edited by S. Garattini. Raven Press, 1993.

Espresso Coffee: The Chemistry of Quality. A. Illy and R. Viani. Academic Press, 1995.

Coffee: Recent Developments. R. J. Clarke and O. Vitzthum. Blackwell Science, 2001.

International Coffee Organization: **www.ico.org**

International Scientific Association of Coffee: **www.asic-cafe.org**

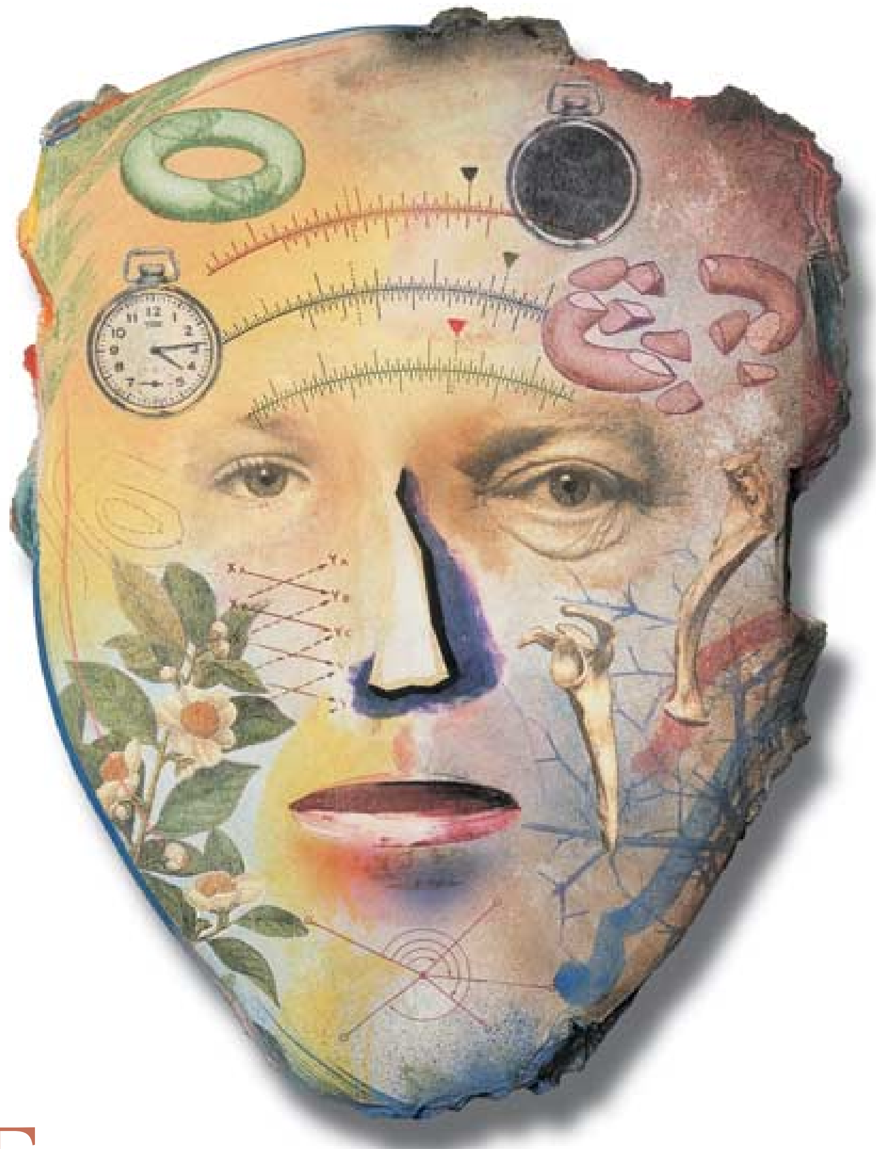
Coffee industry–supplied information on coffee, caffeine and health: **www.coffeescience.org**

No Truth to the Fountain of Youth

Fifty-one scientists who study aging have issued a warning to the public: no anti-aging remedy on the market today has been proved effective. Here's why they are speaking up

By S. Jay Olshansky, Leonard Hayflick and Bruce A. Carnes

Illustrations by J. W. Stewart



Efforts to combat aging and extend human life date at least as far back as 3500 B.C., and self-proclaimed experts have touted anti-aging elixirs ever since. Indeed, the prospect of immortality has always had universal appeal, spurring Alexander the Great and Ponce de León to search for the legendary Fountain of Youth and feeding alchemists' desire to manufacture gold (once believed to be the most potent anti-aging substance in existence). But the hawking of anti-aging "therapies" has taken a particularly troubling turn of late. Disturbingly large numbers of entrepreneurs are luring gullible and frequently desperate customers of all ages to "longevity" clinics, claiming a scientific basis for the anti-aging products they recommend and, often, sell. At the same time, the Internet has enabled those who seek lucre from supposed anti-aging products to reach new consumers with ease.

Alarmed by these trends, scientists who study aging, including the three of us, have issued a position statement containing this warning: no currently marketed intervention—none—has yet been proved to slow, stop or reverse human aging, and some can be downright dangerous. While the public is bombarded by hype and lies, many biologists are intensively studying the underlying nature of aging in the belief that their research will eventually suggest ways to slow its progression and to thereby postpone infirmity and improve quality of life. But anyone purporting to offer an anti-aging product today is either mistaken or lying. The full position statement, drafted and endorsed by 51 internationally recognized investigators, can be found on the *Scientific American* Web site [*see bottom of page 95*]. Here we state the case as we see it, speaking for ourselves.

What Aging Is ... and Isn't

ANY DISCUSSION OF AGING should first clarify its terms. Various definitions have been proposed, but we think of aging as the accumulation of random damage to the building blocks of life—especially to DNA, certain proteins, carbohydrates and lipids (fats)—that begins early in life and eventually exceeds the body's self-repair capabilities. This damage gradually impairs the functioning of cells, tissues, organs and organ systems, thereby increasing vulnerability to disease and giving rise to the characteristic manifestations of aging, such as a loss of muscle and bone mass, a decline in reaction time, compromised hearing and vision, and reduced elasticity of the skin.

This accretion of molecular damage comes from many sources, including, ironically, the life-sustaining processes involved in converting the food we eat into usable energy. As the energy generators of cells (mitochondria) operate, they emit destructive, oxidizing molecules known as free radicals. Most of the damage caused by these reactive molecules gets repaired, but not all. Biologists suspect that the oxidative assaults ultimately cause irreparable injury to the mitochondria, thereby impeding the cell's ability to maintain the integrity of the countless molecules needed to keep the body operating properly. The free radicals may also disrupt other parts of cells directly.

Aging, in our view, makes us ever more susceptible to such ills as heart disease, Alzheimer's disease, stroke and cancer, but these age-related conditions are superimposed on aging, not equivalent to it. Therefore, even if science could eliminate today's leading killers of older individuals, aging would continue to occur, ensuring that different maladies would take their place. In addition, it would guarantee that one crucial body component or another—say, the cardiovascular system—would eventually experience a catastrophic failure. It is an inescapable biological reality that once the engine of life switches on, the body inevitably sows the seeds of its own destruction.

Men and women in the developed world typically live longer now (75 and 80 years, respectively) than they did throughout much of history (about 25 years) because human ingenuity—which brought us sanitation systems, vaccines, antibiotics and so on—has had phenomenal success in thwarting the infectious and parasitic diseases responsible for a great deal of premature death. We live longer now not because we have altered the way we age but because we have altered the way we live.

Though inevitable, aging is not, as some might think, a genetically programmed process, playing itself out on a rigidly predetermined time schedule. The way evolution works makes it impossible for us to possess genes that are specifically designed to cause physiological decline with age or to control how long we live. Just as an automobile does not have a built-in plan for decline written in its blueprints, we do not possess genetic instructions that tell our bodies how to age or when to die.

The logic behind this assertion goes basically like this: Genes perpetuate themselves by orchestrating the transformation of a fertilized egg into a sexually mature adult that produces offspring. Clearly, any genetic variant that compromises this developmental process would be self-eliminating. Conversely, evolution is totally blind to the consequences of gene action (whether good, bad or indifferent) after reproduction is achieved. Genes or genetic variants that prove detrimental in the postreproductive part of the life span can become commonplace, but only if they participate in important processes early on. For

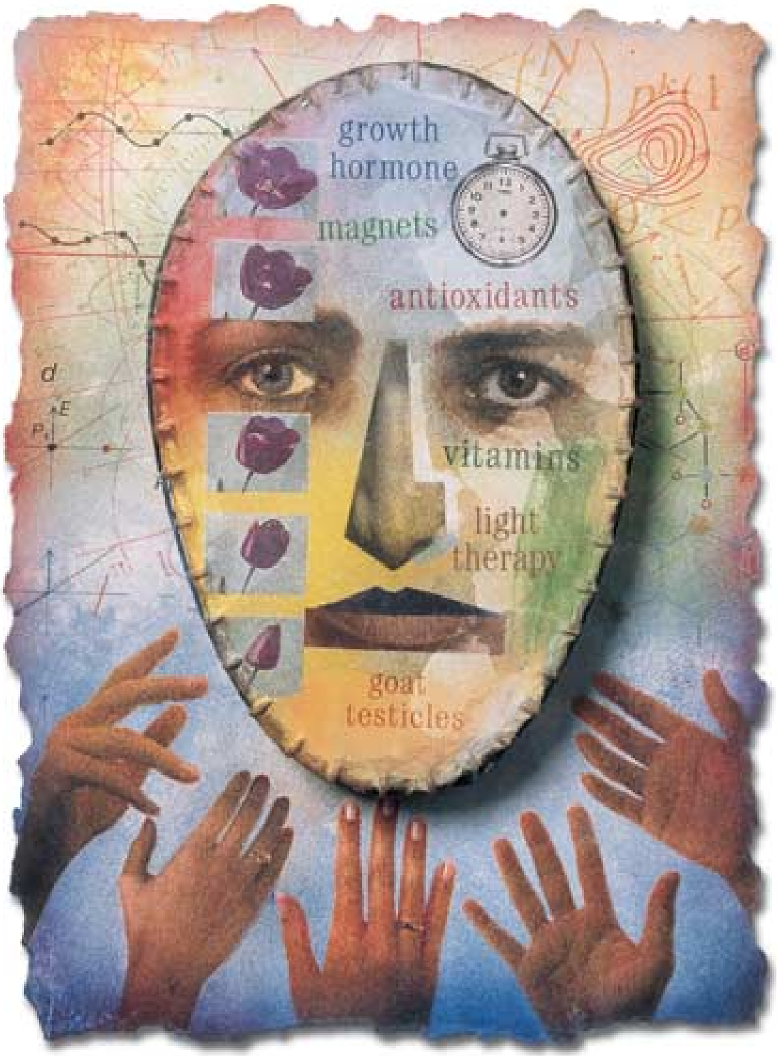
The primary goal of biomedical research and efforts to slow aging should not be the mere extension of life. It should be to prolong the duration of healthy *life.*

THE AUTHORS

S. JAY OLSHANSKY, LEONARD HAYFLICK and *BRUCE A. CARNES* have all studied aging for many years and spearheaded the drafting of the position statement on aging discussed in this essay. Olshansky is professor of public health at the University of Illinois at Chicago. Hayflick is professor of anatomy at the University of California, San Francisco. Carnes is senior research scientist at the National Opinion Research Center/Center on Aging at the University of Chicago, where Olshansky works as well.

example, several genes that contribute to cancer in the later years are known to participate in growth and development early in life.

Without a doubt, a host of our genes influence aging, but they do so indirectly, as an inadvertent by-product of processes involved in growth, development, and the maintenance of health and vigor. The lack of a specific genetic program for aging and death means that there are no quick fixes that will permit us to treat aging as if it were a disease. A single genetic intervention in an organism as complex as a human being would have little chance of combating the probably vast array of genes and biological activities that play subtle, unpredictable parts in the timing of our ultimate demise.



False Claims

DESPITE THIS COMPLEXITY, some researchers believe that they may manage to find ways to slow the rate of human aging. If they succeed, many people will live longer than would otherwise be expected, and a few people might even surpass the modern longevity record of 122 years. But the primary goal of biomedical research and efforts to slow aging should not be the mere extension of life. It should be to prolong the duration of *healthy* life. Slowing the rate of aging could help postpone the onset of age-related diseases and infirmities, essentially enabling people to stay younger longer.

On what grounds do we assert so vehemently that no purported anti-aging intervention has been proved to modify aging? To assess whether an intervention has affected a biological process, researchers need a yardstick for measuring that process. In this case, no single or aggregate age-related phenomenon has proved to be a reliable indicator of the rate of aging in humans or other species. Without a yardstick, there can be no measurements, and without measurements there can be no assurance that an intervention was successful.

People eager to retain or restore their youthful biology might well recognize the paucity of proof but decide to try a putative anti-aging intervention anyway, thinking they have little to lose. They should think again. For instance, the U.S. Food and Drug Administration does not require products sold as dietary supplements to undergo the rigorous tests of safety and effectiveness that medicines must pass before they can be sold to the public. Consequently, these supplements come with no guarantees of purity or potency, no established guidelines on dosage, and often no warnings about side effects that may result when the products are taken along with approved medications.

Antioxidants constitute one popular class of supplements touted to have anti-aging powers. Such chemicals occur naturally in the body and in fruits and vegetables and are believed to neutralize free radicals. Proponents claim that if taken in sufficient quantities, antioxidant supplements will sop up the radicals and slow down or stop the processes responsible for aging. But eliminating all free radicals would kill us, because they perform certain necessary intermediary steps in biochemical reactions. Further, although epidemiological studies have demonstrated that the antioxidant vitamins E and C contained within the foods we eat may reduce the risk of cancer, macular degeneration and other disorders, no one has established that vitamin supplements containing antioxidants limit oxidative damage in the body or influence aging.

Like antioxidants, another fashionable anti-aging intervention, hormone replacement, has a plausible rationale. This strategy was first popularized early in the 20th century, when older men occasionally submitted to the grafting of testicles from goats or monkeys or received injections of macerated testicles. Today pure forms of hormones can be administered. The replacement strategy seems logical in principle because the blood levels of most hormones—among them melatonin, growth hormone, testosterone and dehydroepiandrosterone (DHEA)—commonly decrease with age. Also, experiments on older men have demonstrated that some physical and physiological attributes that show declines over time, notably muscle mass and skin elasticity, respond favorably in the short term to growth hormone replacement.

On the other hand, hormones can cause worrisome side effects. In mice, for instance, delivery of melatonin increases the risk of tumor development, and the overproduction of growth hormone leads to kidney problems, premature heart and lung failure, and an increased probability of early death. Human adults given growth hormone have suffered from acromegaly (excess bone growth) and carpal tunnel syndrome. Estrogen replacement therapy may offer health benefits to some postmenopausal women; however, this form of therapy has recently been challenged and has risks of its own, such as breast cancer and blood clots. In short, hormone replacement therapy has a place in the treatment of specific age-associated disorders, but evidence that it affects the rate of aging is lacking.

Some people might wonder whether following today's public health recommendations for diet and exercise can serve as a more natural Fountain of Youth. Good nutrition and regular exercise do reduce the risk of various diseases and, in that way, may extend the duration of life for many people—thereby serving as the best current prescription for a long and healthy life. As is true of other interventions, though, no one has shown that diet or exercise, or both, directly influences aging.

What Science Says

WE FIND IT IRONIC that a phony anti-aging industry is proliferating today, because serious efforts to understand aging have advanced greatly in recent years. Biologists who work with yeast, roundworms, fruit flies and mice have extended life by manipulating the genes of those species. These genetic alterations did not affect what is believed to be an important hallmark of aging in a population (an exponential increase in the risk of dying with time after puberty), which means that the longevity extensions in those experiments cannot safely be interpreted as resulting from an intervention in the aging process. Nevertheless, further study of those genes could offer clues to the influences on longevity and to approaches that might postpone infirmity and age-related disorders.

Another avenue of research may also lead to true aging interventions. Investigators have known for decades that caloric restriction extends life and the duration of good health in all species in which it has been studied, as long as the diet includes enough nutrition for routine maintenance of the body. These findings suggest that caloric restriction might have similar effects in humans. Given that few people would ever reduce their food intake enough to lengthen their lives, biologists are now trying to discover the mechanism that underlies the benefits of caloric restriction and to find agents that might mimic those helpful effects in people without forcing them to go hungry.

A number of scientists look at current research trends and feel hopeful. They can envision a time when treatments based on an understanding of aging can help slow its progression and when not yet specialized (stem) cells can be coaxed to repair and rejuvenate damaged tissues, enabling people to remain vigorous longer than they would without medical assistance. Not all researchers share that optimism, though. Some assert that aging's complexity will forever militate against the development of anti-aging therapies.

One thing is indisputable: the number of elderly people is growing worldwide, and opportunists stand ready to cash in on the burgeoning market for anti-aging products. The researchers who wrote and endorsed the position paper appearing on *Scientific American*'s Web site do not necessarily agree on every word written there, but everyone realized that we had to set aside our minor differences to raise awareness of the growing scam. The public needs to know that the products sold as anti-aging remedies at longevity clinics and elsewhere have no scientifically proven efficacy and may at times be harmful. Systematic investigations into aging and its modification are in progress and could one day provide methods to slow our inevitable decline and extend health and longevity. That day, however, has not dawned yet. SA

People might well recognize the paucity of proof but decide to try a putative anti-aging intervention anyway, thinking they have little to lose. They should think again.

On the Web...

To read the full position statement on aging and its extensive list of references, follow this link to the "Explorations" part of the *Scientific American* Web site: **www.sciam.com/explorations/2002/051302aging/**

GYROSCOPE GUIDANCE

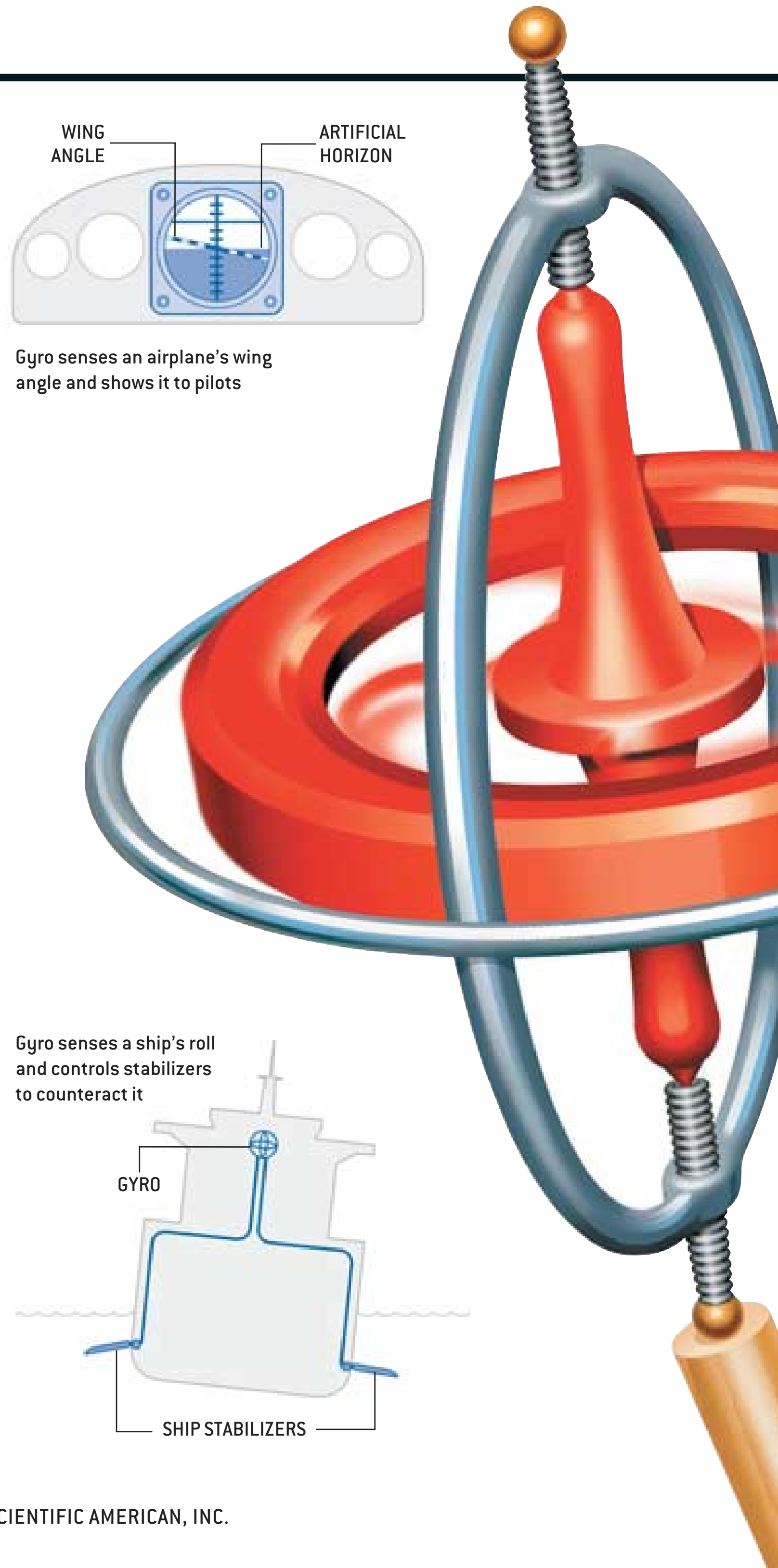
Hidden Guides

Gyroscopes are the silent brains that keep planes in the air, satellites in orbit and ships on course. American Elmer Sperry, who in 1910 founded the Sperry Gyroscope Company, invented the first practical gyroscopic instruments. The company built ship stabilizers and the "artificial horizon" that shows pilots the attitude of their plane. Since Sperry's death in 1930, a variety of Sperry companies have converted his nearly 400 patents into automatic navigation and guidance systems for military planes, rockets, bombs, satellites and spacecraft. Today gyros guide robots, antiskid systems and dashboard navigators in cars, the space shuttle, the Hubble Space Telescope, the Mars Sojourner rover, and astronaut propulsion backpacks.

Two principles make gyros handy. In accordance with gyroscopic inertia, a rotating mass will maintain a fixed position in space. A spinning gyro keeps a satellite always pointing toward the earth, making it stable and easy to communicate with. A rotating mass will also resist a force that tries to tip it. In the mid-1900s huge gyroscopes weighing tons were bolted into ship hulls and spun by motors; they "pushed back" against waves to keep ships upright.

According to the precession principle, a spinning gyro not bolted down will move at right angles to an external force. A spinning top on a table does not topple over when pushed but moves sideways to the push. A gyro in a level-flying airplane wing will precess when the wings begin to tip; by sensing the precession, instruments can show a pilot the wing angle—crucial when he can't see anything outside the windshield. A second gyro can indicate the pitch of the plane, nose to tail. When linked with accelerometers that measure speed, the gyros can fly the plane: autopilot.

Fewer and fewer gyros are mechanical these days. In the 1980s ring laser gyros and fiber-optic gyros that monitor changes in light patterns proved more accurate, smaller and lighter. Now they sell for $3,000 to $4,000. Companies are also making tiny gyros by micromachining quartz or silicon, which sense changing vibrations in the materials. They are less accurate but can be mass-produced like integrated circuits and cost as little as $20, opening up consumer applications such as those in automobiles. —*Mark Fischetti*



Gyro senses an airplane's wing angle and shows it to pilots

Gyro senses a ship's roll and controls stabilizers to counteract it

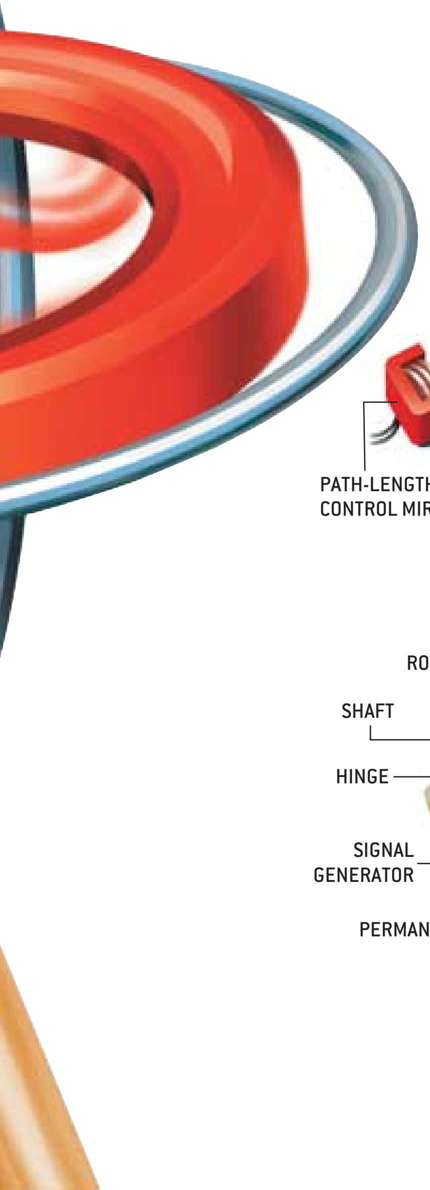
ILLUSTRATIONS BY GEORGE RETSECK

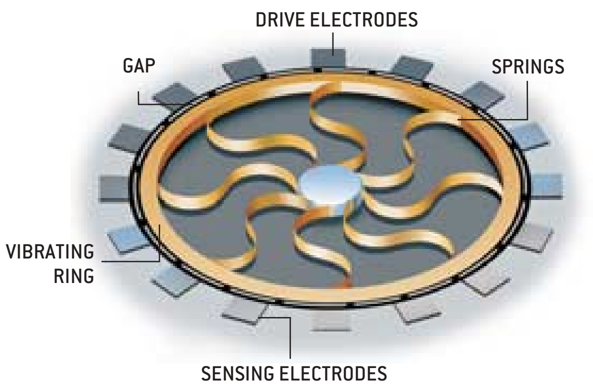
DID YOU KNOW . . .

➤ MATTER OF DEGREE Gyroscopes vary in accuracy. “Bias,” which arises from friction and temperature changes, is critical. Neil Barbour, gyro guru at Charles Stark Draper Laboratory in Cambridge, Mass., says the best gyros keep bias uncertainty below 0.01 degree an hour, meaning they can guide a projectile to within one nautical mile of a target after an hour’s flight. The guided bombs dropped over Afghanistan have a bias of about 1 degree an hour. The gyros in antiskid systems have biases as high as 3,600 degrees an hour but are useful because they operate for only a few seconds to correct the vehicle.

➤ BUZZED At a 1914 Paris air competition, inventor Elmer Sperry’s son Lawrence demonstrated his gyroscopic airplane stabilizer. Flying a biplane, he buzzed the judge’s stand after taking his hands off the flight controls, while a mechanic walked on the plane’s wing. Father and son later devised an “automatic pilot,” which helped Wiley Post make the first solo flight around the world in 1933. The autopilot on ships is often called Metal Mike, as if it were an invisible crew member.

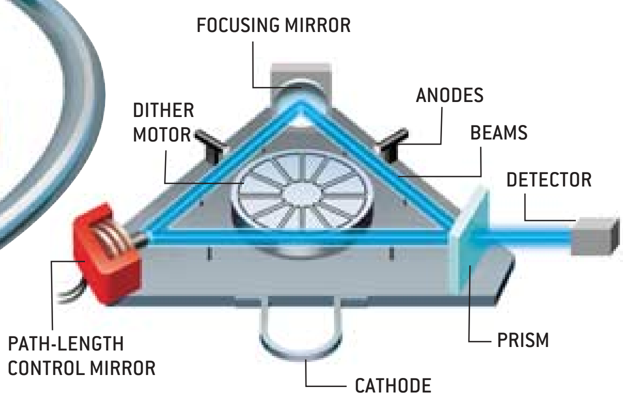
➤ OUT OF CONTROL U.S. law prohibits export of highly accurate gyroscopes. In 1999 American authorities arrested a Chinese businessman they said was trying to buy fiber-optic gyros for targeting “smart” bombs. In 1995 United Nations divers found gyros in the Tigris River bed near Baghdad and traced them to a closed Soviet military institute that had made guidance systems for ballistic missiles.





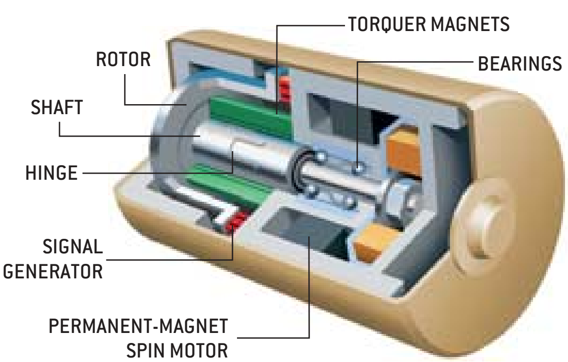
MICRO GYRO

In a silicon micromechanical vibrating ring gyroscope made by Farrokh Ayazi of the Georgia Institute of Technology, a ring is suspended by free-floating, curved support springs attached to a central, fixed post. Drive electrodes apply an electrostatic force to the ring, which sets up a standing vibration pattern monitored by sensing electrodes. If the ring is rotated around its axis by an outside force, the vibration pattern distorts, indicating the direction of rotation. The amplitude of the distortion indicates the rate of rotation.



RING LASER GYRO

Two anodes and a cathode in a ring laser gyroscope excite a gas, sending two light waves of the same frequency in opposite directions. If the ring is rotated by an outside force, one wave will travel slightly farther. A detector senses the resulting frequency mismatch, which indicates the rate of rotation. The light waves entering the detector create a fringe pattern; a change in the pattern shows the direction of rotation. To prevent the beams from locking up in one frequency—a natural tendency—a dither motor vibrates the unit to introduce a slight phase change.



DYNAMICALLY TUNED GYRO

A motor in a dynamically tuned mechanical gyroscope keeps an iron rotor on bearings spinning at a constant rate. If the gyro is rotated by an outside force, the rotor will begin to precess, causing magnetic field changes at a signal generator, which indicate the direction and rate of rotation. The generator also instructs the torquer magnets to counteract the precession so the rotor doesn’t collide with its housing.

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

Whatever You Say

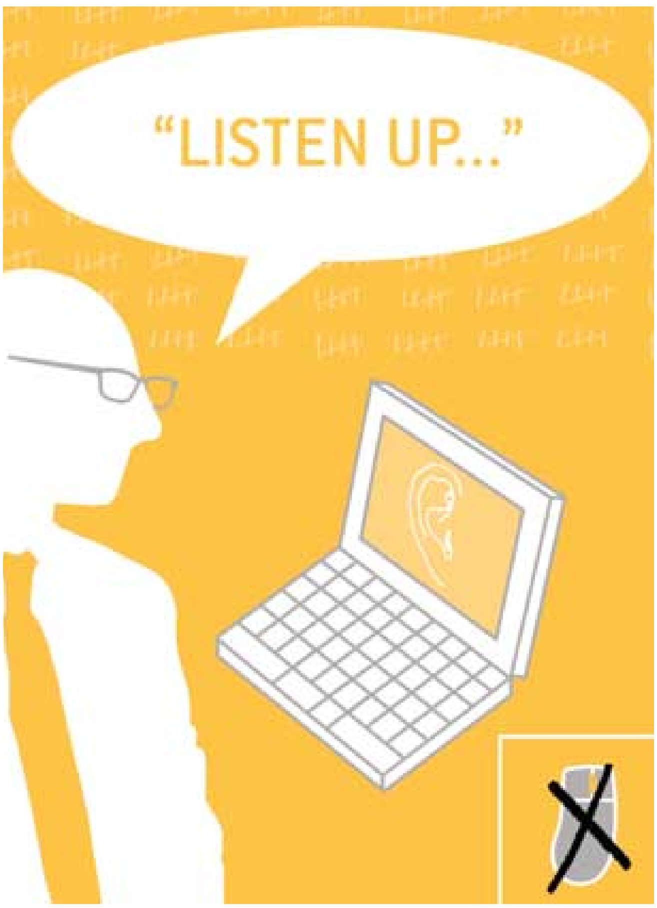
WITH SPEECH-RECOGNITION SOFTWARE, YOUR VOICE IS THE COMPUTER'S COMMAND **BY W. WAYT GIBBS**

Last month I installed an odometer on my mouse. Actually, it's not a mouse; I switched to a trackball a year ago when the pinky on my mousing hand started hurting so much that I couldn't bend it. And it's not a real odometer that I installed but a free program (available for Windows at www.modometer.com) that keeps track of cursor movements, button clicks and keystrokes. I was astounded when I checked the results. In five days my fingertips had skittered that cursed little arrow 2,440 feet around the screen, and my thumb had made 21,719 clicks. No wonder the pain had spread into the rest of my hand.

I knew where this was going. My wife had already developed a repetitive stress injury from massaging work out of her computer. The injury degenerated for six months until just picking up a magazine made her wince. It took a season of nightly icing, physical therapy, sleeping with a wrist brace and using a funky joystick-shaped pointing device called a 3M Renaissance mouse before she regained full use of her forearm. I wasn't anxious to follow her lead. It was time to try a new way of controlling my PC: by voice.

I was curious anyway to assess the current state of this particularly challenging art. I first tried voice control nearly 10 years ago, when Apple included it in the Macintosh operating system. The software then could recognize only a few dozen spoken commands, and it did so at such a slow pace that it was no more than a toy.

But then the Defense Advanced Research Projects Agency (DARPA) started injecting millions of dollars a year into automated speech transcription, and the field took off. Error rates began dropping by 10 to 15 percent a year. By 1996 Dragon Systems's NaturallySpeaking package, running on a high-end PC, could take dictation about as well as a sleep-deprived undergraduate. But for tasks beyond text entry, it was all but helpless.



MUTTER AT YOUR COMPUTER, as we all do from time to time, and it will probably ignore you. But install one of the new voice-control programs, and the PC will heed your spoken commands.

JOHN McFAUL

With the latest versions of NaturallySpeaking and its chief competitor, IBM's ViaVoice, that has changed. Both are still optimized for dictation. But both now enable you to select menu options, push buttons, check e-mail, browse the Web and even nudge the cursor around the screen without stressing a finger.

I elected to try ViaVoice, for three reasons. ViaVoice Pro USB Release 9 is, at $219, less than a third the cost of NaturallySpeaking Professional Version 6. You get a headset microphone with either package, but the one bundled with ViaVoice Pro has a custom signal-processing chip built in to filter out the background noise that confuses the computer no end.

And I cannot help feeling uncertain about the future of NaturallySpeaking. Dragon Systems was bought out by Lernout & Hauspie, which last year imploded in an accounting scandal. Although the company's speech products were sold to ScanSoft in December, it seems unlikely that this relatively small firm will be able to keep up with IBM's massive R&D effort.

> **Within a week, ViaVoice reduced my mousing by roughly a third in e-mail, by more than half in file management and by two thirds in Web surfing.**

Indeed, IBM announced in March that it is increasing the number of researchers working on speech technologies. Its ambitious, decade-long goal is to build systems that can reliably transcribe (and act on) normal conversations taking place in noisy rooms among people whose voices the computer has never been exposed to before. "We now have more than 100 researchers working on speech technologies," says David Nahamoo, who manages that group at IBM Research, "and a similar number working on natural-language understanding."

Understanding the spoken word is, for better or worse, what we have been trained to expect from listening computers. In films and on TV, we've seen a thousand times how we *should* be able to talk to our machines. "Computer, open a secure channel to Admiral Nechayev," Captain Picard says to the screen. "Open the pod bay doors, please," astronaut Dave implores HAL 9000. We command, the machines execute. If they fail, they do so politely: "I'm sorry, Dave. I'm afraid I can't do that."

ViaVoice has no understanding of English. It uses what is called a "context-free grammar"—essentially, a list of phrases it listens for, along with some primitive rules for combining phrases. This allows it to fake comprehension.

I installed ViaVoice on my computer at the office, then trained the software to recognize my style of speaking. If I say "check my messages," the software launches my e-mail program, Microsoft Outlook. ViaVoice was built to work with Outlook and the other programs in the Microsoft Office suite. So I can also say "check e-mail" or "view e-mail" or even "send an e-mail message to John Rennie" or to anyone else listed in my address book. ViaVoice recognizes such an extensive list of synonyms that in Outlook it seems almost intelligent.

Prepare to be frustrated, however, if you try to talk to other programs with such casual familiarity. "Open 'My Documents' folder," I say—the third variation I have tried in an attempt to get ViaVoice to have Windows Explorer open the folder in which I keep all my work. To the first two attempts, ViaVoice played deaf. This time it does display what it thinks I said: "Open new document and its alter." It then tries to execute this baffling command. Evidently, ViaVoice's programmers didn't code in as many phrases for the Windows file manager as they did for Outlook.

But in every application I opened, ViaVoice reliably responded to the names of menu options and the labels on buttons in dialogue boxes (those annoying gray rectangles that pop up to display errors and ask questions). Once I learned to state commands in the precise form that ViaVoice expects, I could scroll up and down and resize windows, open and close programs, and switch among various tasks without reaching for the trackball. Voice control was most useful for Web surfing; read the first few words of a link, and that page opens. Utter the name of any site you have bookmarked, and it appears. Say "show commands," and ViaVoice labels each graphic hyperlink on the current page with numbers, allowing you to "click" on images as well as words by simply saying the appropriate numbers.

ViaVoice soon grew quite adept at understanding my commands regardless of how softly or quickly I spoke. If anything, its accuracy improved when I ran my words together. After using the software for a week, I looked at my mouse odometer, which showed 1,564 feet. Even while I was climbing the learning curve, ViaVoice had reduced my mousing by roughly a third in e-mail, by more than half in file management and by two thirds in Web surfing. Pain no longer shot up my forearm, and my hand also felt much better. For all the flaws and immaturity of voice control, its benefits now outweigh the hassles and the cost. I think I'll install the software at home as well. It can save multiple voice models, so my wife and I can both mutter at the PC and rub our hands in satisfaction. SA

Men, Money and Malaria

IN THE QUEST FOR A CURE, IDEALISM FELL VICTIM TO GREED **BY CLAIRE PANOSIAN DUNAVAN**

THE FEVER TRAIL: IN SEARCH OF THE CURE FOR MALARIA
by Mark Honigsbaum
Farrar, Straus and Giroux
New York, 2002 ($25)

Guadalcanal, November 1942. Another night of war. Aboveground, the air buzzed with mosquitoes and enemy ordnance. In foxholes, weary men in boots, fatigues and steel helmets shifted restlessly, counting the hours until dawn.

If a soldier on Guadalcanal wanted to stay in one piece, a hole in the ground was the safest place to spend the night. But nocturnal refuge had its price: disease-bearing *Anopheles* mosquitoes. For proof, one need only to have visited the hospital field tent. There, by flickering electric lanterns, medics tended delirious patients whose blood swarmed with the delicate rings, crescents and clusters we know as malaria.

During World War II, nearly half a million American servicemen came down with malaria. Atabrine, a drug that stained skin yellow and could also trigger psychosis, was the main defense. Why was real quinine, still the world's leading cure for malaria in the 1940s, so precious in the Pacific theater? Because 80 years earlier, cinchona—the natural source of quinine—was abducted from its native habitat in the Andean cloud forest. By the 20th century, virtually all the world's cinchona bark came from the island of Java. And by 1942, Java lay in Japanese hands.

How the elusive cinchona, renowned for "febrifugal" powers since the 1600s, journeyed from South America to Dutch plantations in Java is a major thread of *The Fever Trail*. Lovingly researched and written by Mark Honigsbaum, former chief reporter for the *Observer* in London, this book is an adventure story cum historical account of dreams and schemes to steal the valuable tree with the crimson-lined leaf. In its closing chapters, *The Fever Trail* fast-forwards to malaria today: the stalled race to create a vaccine, the politics of new drug development and distribution, and, most troubling of all, the continued toll of *Plasmodium falciparum*. *Falciparum* is the malarial strain that killed in World War II and still kills one million to two million of the world's poor every year, especially in Africa.

But first the back story, starring three leading cinchona hunters of the 19th century. Honigsbaum, who retraced their steps, lists them as "Richard Spruce, a hypochondriac Yorkshireman and moss collector who, despite his fear of disease, spent fifteen years wandering the Amazon and the Andes on behalf of the Royal Botanic Gardens at Kew; Charles Ledger, a cockney trader who came to Peru to make his fortune at the age of eighteen and left South America broken and penniless after a series of incredible expeditions through Bolivia and Chile; and Sir Clements Markham, the patron of Robert Scott's two expeditions to the Antarctic and the so-called father of polar exploration."

The story of Ledger's loyal guide, Manuel Mamani, is one of *The Fever Trail*'s poignant subplots. According to Honigsbaum, Mamani was the unsung hero who collected the seeds that later spawned Java's quinine-rich bark. At Ledger's behest, Mamani returned to the forest for more seeds, only to suffer a brutal imprisonment for threatening Bolivia's cinchona monopoly. Within days of his release, he was dead.

Clements Markham was another idealist whose quest for cinchona ended on a melancholy note. Son of a canon to the royal court at Windsor, himself a noted historian and explorer of Peru, Markham was the most philanthropic of the Andean botanical explorers. Along with Florence Nightingale, he decried malaria's toll of death and disability on Britain's colonial subjects and shared her dream of saving them through better sanitation and treatment.

Years later Markham's dream was dashed. Yes, the worldwide price of quinine had plummeted, thanks to Java's prolific *Cinchona ledgeriana* groves. But lives lost to human greed far outweighed lives saved by less costly malaria pills. One villain was Britain's East India Company, whose punitive taxes and neglect of

CRESCENT-SHAPED malarial parasites, *Plasmodium falciparum*, and red blood cells.

OLIVER MECKES AND NICOLE OTTAWA *Photo Researchers, Inc.*

public works fueled outbreaks of malaria, cholera and other infectious diseases. The other culprit? Cold commerce. In Honigsbaum's words, "There was a limit to philanthropy and in India the British had reached it. If any profits were to be made from the bark in the future, they would come from selling quinine to those who could afford it—in other words, 'rich' Europeans and Americans."

This north-south disparity continues today. It's reflected in brisk sales of expensive new malaria drugs for safari-goers while an African child dies of malaria every 30 seconds following second-rate treatment or no treatment at all. To be sure, decades of worsening resistance by *P. falciparum,* skyrocketing pharmaceutical costs and changing patterns of transmission have upped the ante of conquering the ancient blood-borne parasite. But the sad truth is, we don't have to wait for a malaria vaccine. Were there the international will and infrastructure, currently available drugs could yet achieve Markham's dream.

Meanwhile the U.S. military has recently beefed up its ongoing malaria research, both for the protection of troops and, perhaps, for a larger purpose Honigsbaum never mentions. Could it be, post–September 11, that global security might hinge not just on strategic accords but on more muscled control of the worldwide plagues of the poor? Read between the lines of *The Fever Trail.* An all-out commitment to curing malaria in this century is no longer foreign aid in the grand noblesse oblige tradition. It's an investment in humankind, global economic health and our own self-interest. SA

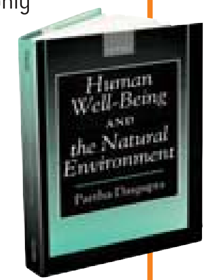
Claire Panosian Dunavan is professor of medicine at the University of California at Los Angeles School of Medicine, a specialist in infectious and tropical diseases, and a medical journalist. Her father served nine months on Guadalcanal in 1942–43, where he, like many others, suffered repeated attacks of malaria.

THE EDITORS **RECOMMEND**

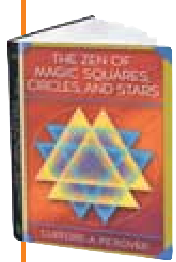
HUMAN WELL-BEING AND THE NATURAL ENVIRONMENT

by Partha Dasgupta. Oxford University Press, New York, 2001 ($35)

Current measures of the quality of life are, by and large, insensitive to our dependence on the natural environment. Dasgupta, a distinguished professor of economics at the University of Cambridge, aims to remedy that. In a style that is both engaging and rational, he argues that the most valid measure of human well-being encompasses not only manufactured assets but also human capital (skills), knowledge (ideas) and the natural environment, which includes "minerals and fossil fuels, soils, fisheries, sources of water, forests and woodlands, watersheds, the oceans, places of beauty and tranquility, and the atmosphere." The sobering picture that emerges from this important book contrasts sharply with the one portrayed in most literature on economic development. *Human Well-Being* is intended both for scholars and for "the general citizen interested in what are among the deepest and most urgent social problems we face today."



THE ZEN OF MAGIC SQUARES, CIRCLES, AND STARS: AN EXHIBITION OF SURPRISING STRUCTURES ACROSS DIMENSIONS



by Clifford A. Pickover. Princeton University Press, Princeton, N.J., 2002 ($29.95)

"In this book," Pickover writes, "we will go far beyond ordinary magic squares and consider many unusual variations, some in higher dimensions, all with mind-boggling patterns." You do not have to reach the "miniature epiphany" he says you might have while contemplating the intriguing structures he describes, but you should get instruction and pleasure from them. Pickover, a research staff member at the IBM Thomas J. Watson Research Center, is the author of many other books on mathematical subjects.

THE STRUCTURE OF EVOLUTIONARY THEORY

by Stephen Jay Gould. Belknap Press of the Harvard University Press, Cambridge, Mass., 2002 ($39.95)

When America's best-known scientist publishes his major work, it is an event. And *The Structure of Evolutionary Theory* is indeed an event, but not one aimed at the general reader. It is a monumental work, both in size (1,400-plus pages) and in scope—it sets out to do nothing less than reformulate Darwin's theory of evolution. "Nothing of Darwin's central logic has faded or fully capsized," Gould writes, "but his theory has been transformed, along his original lines, into something far different, far richer, and far more adequate to guide our understanding of nature." Among the many threads that weave through the book runs a major Gouldian theme, the idea of species selection—that species, not just individuals, compete against one another for survival.

The style is vintage Gould. He opens with an extended exposition on a metaphor comparing evolutionary theory to the Duomo (Cathedral) of Milan: Darwin laid the foundation of a great edifice, but in the process of erection, the superstructure has been altered by his successors, just as construction of the Duomo occupied several centuries and incorporated an amalgam of changing styles, from Gothic to Baroque. Fascinating historical digressions, even asides about "Mr. Berra," pepper the book, so that scholarly though the work is, the nonexpert, dipping into the deep waters here and there or reading just the opening chapter, will find much to delight and to inform.

The books reviewed are available for purchase through www.sciam.com

Privacy Taboos

BY DENNIS E. SHASHA

In a god-fearing culture called the Paranoimos, each adult woman has a personal temple where she prays. The temples differ in that each woman puts symbols of her devotion in her temple hut. If woman A enters woman B's hut, then the Paranoimos believe that B's hut is "weakened" and B's prayers will lose their force, whereas A's will gain power.

For this reason, if a woman suspects that another woman has entered her hut, then she is duty-bound to strike the intruder with a ritual stick. Suspicion comes from accusations by third parties, never from direct observation. Note that several women might accuse someone of intruding.

Woman C may tell B that A entered B's hut. At that point B must strike A. When striking A, B tells A that it was C who accused her. (C, of course, might be one of several accusers.) B must be honest about this. If the accusation is false, A is duty-bound to strike C. If the accusation is true, A does not strike C by tradition and for fear that C may come up with proof that would cause A's banishment. Because these intrusions lead to bad feelings, most women avoid them.

An anthropologist interested in privacy taboos visited the Paranoimos recently but did not know the language. Our job is to interpret her journal to determine who accused whom and, if possible, who was guilty of intrusion. Here are the journal entries:

C speaks to F
B speaks to E
A speaks to F
F speaks to A
E speaks to B
A speaks to B
E strikes A
A strikes C
A strikes B
B strikes A
F strikes D
A strikes E
B strikes C
C strikes A

A warm-up problem is illustrated below; the answer is at the bottom of this page.

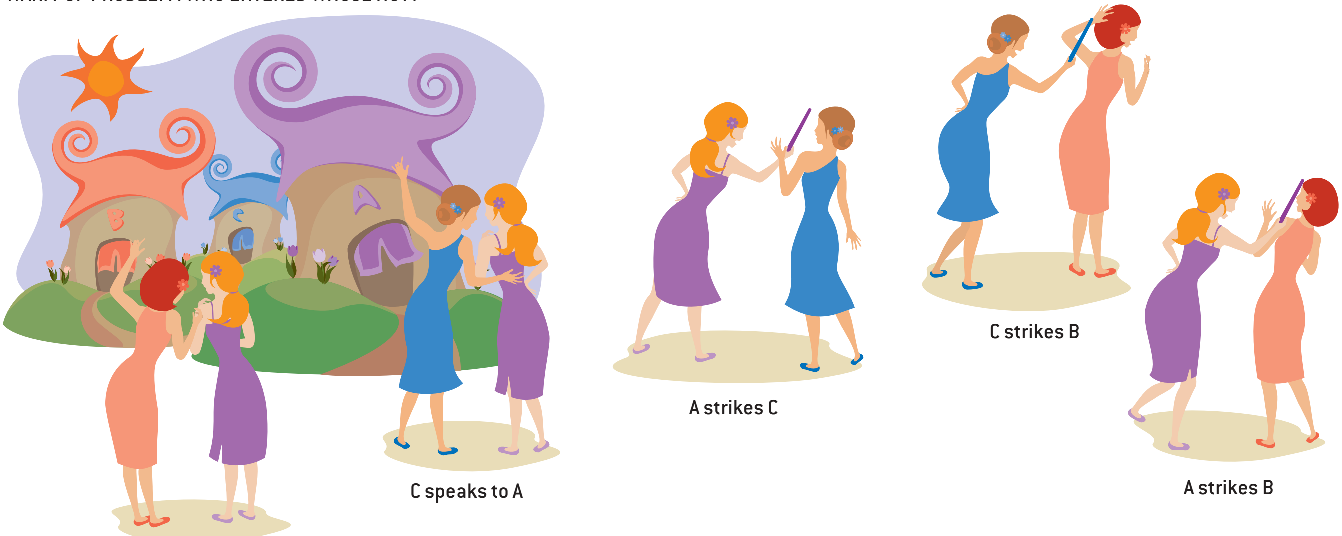
Answer to Last Month's Puzzle

If the tackles can start the football game at least three spaces from the runner, they can win every time. But if the tackles must start in the southernmost row, the runner can win every time. For a full explanation, go to www.sciam.com

Web Solution

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

WARM-UP PROBLEM : WHO ENTERED WHOSE HUT?



SARA CHEN

Answer to Warm-up Problem: B accused C of entering A's hut, and C accused B of the same. But in fact, only B did so.

About the Size of It

A CONTEMPLATION ON THE STATES OF OUR PERCEPTION **BY STEVE MIRSKY**

In February and March an Antarctic ice shelf fell apart, leaving bobbing evidence of the end of its shelf life. The disintegrated Larsen B ice shelf covered approximately 1,250 square miles, which, numerous media outlets noted, was about the size of Rhode Island. Some scientists who study our planet, which is about the size of Venus, immediately blamed global warming. Whether the cause was indeed widespread warming or merely a more geographically isolated heat-up will no doubt be a subject of discussion elsewhere. And despite the fact that the Larsen B shelf weighed about 720 billion tons, which is equivalent to the 720 billion cigarettes that the U.S. produced in 2000 if each cigarette weighed one ton, I've been thinking about less weighty stuff: namely, the practice of describing the sizes of things in terms of other things.

My interest in this obscure area of journalism began a few years ago, when I was reading an article about the environmental impact of golf courses. Par for the discourse, the writer observed that all the world's golf courses combined would cover an area one third the size of Belgium. He then committed what I have since thought of as the Belgian waffle: he pointed out that Belgium was about the size of Maryland. Why bother with the Belgian middleman? I wondered. Why not just say that the world's golf courses cover one third the area of Maryland? And if Maryland's surface area is about 10,461 square miles, one third of which is 3,487 square miles, one could just as easily say that the world's golf courses cover an area about one half the size of Hawaii (6,473 square miles). Which makes more sense anyway, because Hawaii eventually *will* be completely covered with golf courses.

Even more to the point, does Belgium or Maryland immediately conjure up a useful impression of size? The average American would be as likely to look for Belgium next to francium on the periodic table as on a map of Europe. By the way, Europe, at about four million square miles, covers an area a little bit bigger than the U.S., including Alaska and Hawaii, or including just Alaska without Hawaii. As for Maryland, it only looks big because it's all over Delaware.

Speaking of Delaware, at roughly the same time that Larsen B hit the sea came the story of an iceberg called B-22. (Coincidentally, a Mrs. Goldberg called B-22 once in a bingo game in Florida, which is about the size of Arkansas.) News accounts of B-22 busting loose from Antarctica remarked that the iceberg was about the size of Delaware. Such a description probably means that B-22 is a really big iceberg, although Delaware is a really small state. Which brings us back to Rhode Island and Larsen B.

Considering that the single most identifiable attribute of Rhode Island is that it's the *smallest* state in the union, tying the Larsen B shelf to Rhode Island may have inadvertently sent the message that gigantic portions of Antarctic ice falling away isn't such a big deal. But, as one scientist pointed out, the Larsen B collapse might signal the instability of the entire Antarctic Ross ice shelf, which, as some news reports felt compelled to reveal, is about the size of France. Changes in the integrity of the Ross ice shelf might indicate future worldwide climate changes and a rise in sea level, which in turn means that Rhode Island might someday be even smaller. And as Rhode Island goes, so goes Delaware. Not to mention Hawaii. SA

ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

ASK THE EXPERTS

Q Do people lose their senses of smell and taste as they age?

—N. Sly, Windsor, Australia

Charles J. Wysocki, a neuroscientist at the Monell Chemical Senses Center in Philadelphia who studies variation among individuals in the perception of odors and the response of the human nose to chemical irritation, offers this answer:

It's true that as people age they often complain about a decrease in—or even the loss of—their ability to taste a superb meal or appreciate a fine beverage. When people eat, however, they tend to confuse or combine information from the tongue and mouth (the sense of taste, which uses three nerves to send information to the brain) with what is happening in the nose (the sense of smell, which utilizes a different nerve input).

It's easy to demonstrate this confusion. Grab a handful of jellybeans of different flavors with one hand and close your eyes. With your other hand, pinch your nose closed. Now pop one of the jellybeans into your mouth and chew, without letting go of your nose. Can you tell what flavor went into your mouth? Probably not, but you most likely experienced the sweetness of the jellybean. Now let go of your nose. Voilà—the flavor makes its appearance.

SARA CHEN

This phenomenon occurs because smell provides most of the information about the flavor. Chemicals from the jellybean, called odorants, are inhaled through the mouth and exhaled through the nose, where they interact with special receptor cells that transmit information about smell. (It's the reverse process that one experiences downwind from a pig farm or chocolate factory.) These odorants then interact with the receptor cells and initiate a series of events that are interpreted by the brain as a smell.

Estimates for the number of odorant molecules vary, but there are probably tens of thousands of them. Taste, in contrast, is limited to sweet, sour, bitter, salty and umami (the taste of monosodium glutamate, or MSG).

The sense of smell diminishes with advancing age—much more so than the sensitivity to taste. This decrease may result from an accumulated loss of sensory cells in the nose. The loss may be perhaps as much as two thirds of the original population of 10 million. Although the elderly are in general less sensitive than young people to the overall perception of the food they eat, there are exceptions: some 90-year-olds may be more sensitive to smells than some 20-year-olds.

Q What happens when an aircraft breaks the sound barrier?

—M. Kerr, Marlow, England

Tobias Rossmann, a research engineer with Advanced Projects Research, Inc., and a visiting researcher at the California Institute of Technology, provides this explanation:

A discussion of what happens when an object breaks the sound barrier must begin with the physical description of sound as a wave with a finite propagation speed. Anyone who has been far enough away from an event to see it first and then hear it is familiar with the relatively slow propagation of sound waves. At sea level and a temperature of 22 degrees Celsius, sound waves travel at 345 meters per second (770 mph). As the local temperature decreases, the sound speed also drops, so that for a plane flying at 35,000 feet—where the ambient temperature is, say, –54 degrees C—the local speed of sound is 295 meters per second (660 mph).

Because the propagation speed of sound waves is finite, sources of sound that are moving can begin to catch up with the sound waves they emit. As the speed of the object increases to the sonic velocity, sound waves begin to pile up in front of the object. If the object has sufficient acceleration, it can burst through this barrier of unsteady sound waves and jump ahead of the radiated sound, thus breaking the sound barrier.

An object traveling at supersonic speeds generates steady pressure waves that are attached to the front of the object (a bow shock). An observer hears no sound as an object approaches. After the object has passed, these generated waves (Mach waves) radiate toward the ground, and the pressure difference across them causes an audible effect, known as a sonic boom. SA

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

FUZZY LOGIC

BETWEEN NOTHING AND SOMETHINGNESS