HOW TO MOVE ASTEROIDS • AIRCRAFT WITH MORPHING WINGS

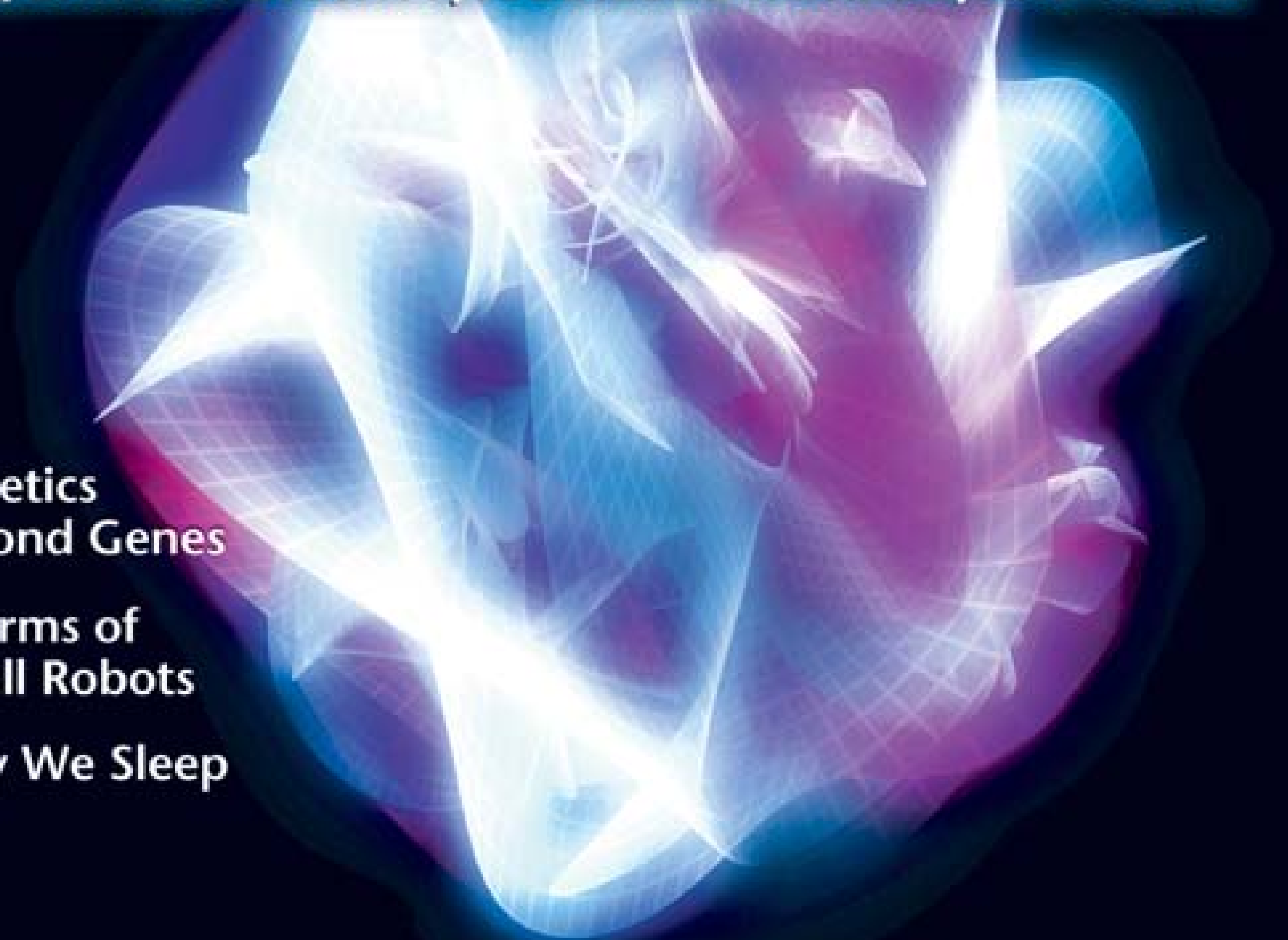
SCIENTIFIC AMERICAN

Explorers from 1,750,000 B.C.

NOVEMBER 2003 $4.95
WWW.SCIAM.COM

STRINGS & SPACETIME WITH 11 DIMENSIONS

PHYSICIST BRIAN GREENE TALKS ABOUT STRING THEORY, QUANTUM GRAVITY, PARALLEL WORLDS, AND MORE



Genetics beyond Genes

Swarms of Small Robots

Why We Sleep

november 2003 contents

SCIENTIFIC AMERICAN Volume 289 Number 5

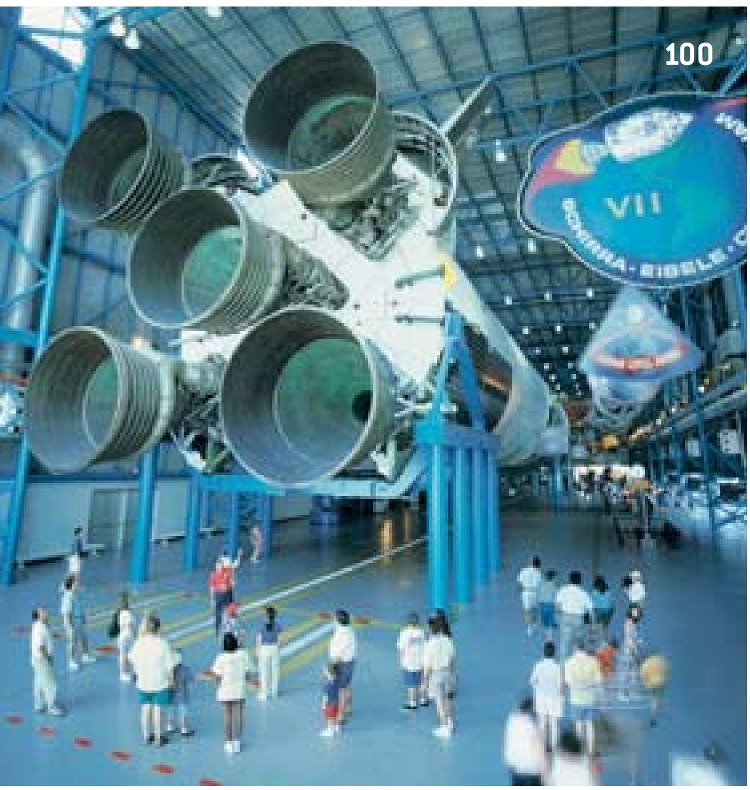
features

74 Dmanisi skull

departments


32


105



100

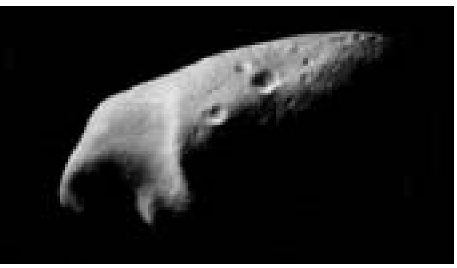
columns

Cover image: courtesy of *NOVA*, with special thanks to Andrew J. Hanson of Indiana University; preceding page: Gouram Tsibakhashvili

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian BN No. 127387652RT; QST No. Q1015332537. Subscription rates: one year $34.97, Canada $49 USD, International $55 USD. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. **Reprints available:** write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; (212) 451-8877; fax: (212) 355-0408 or send e-mail to sacust@sciam.com **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.** Printed in U.S.A.

Penny-Wise, Planet-Foolish

Somewhere in the inner solar system, there's a rock with our name on it. Literally. In March the International Astronomical Union named a newly discovered asteroid 14145 SciAm, on the recommendation of its discoverer, Edward Bowell of Lowell Observatory. Fortunately for the magazine's public relations image, the asteroid does not cross paths with Earth. Others after whom asteroids are named may not be so lucky. As most people now recognize, killer rocks are a fact of life on our planet. Doubters can ask the dinosaurs for their opinion.



ASTEROID 433 Eros, as seen by the NEAR Shoemaker probe.

Is the world doing enough to cope with the threat of impacts? In this issue, a team of scientists and astronauts argues for going beyond the current telescope surveys to begin developing a rocket that could land on an asteroid and push it out of the danger zone [see "The Asteroid Tugboat," on page 54]. The project could cost $1 billion, spread out over a decade or so. Is it worth it?

Some question whether we should spend even a penny on distant threats when we face so many immediate ones. One counterargument is that the world doesn't have the luxury of tackling its problems one by one. It needs to cope with many at once by allocating resources among them. Certain problems deserve more, others less—but all need something.

Actuarial calculations can help us perform this juggling act. By the latest estimate, every year Earth has a one-in-600,000 chance of getting whacked by an asteroid wider than one kilometer—big enough to wreak global havoc and kill billions of people. Averaged out over time, several thousand people a year will die from such impacts, which is greater than the toll from plane crashes or international terrorism. If you value their lives at $1 million apiece (a common ballpark figure used by insurers), you could justify putting several billion dollars each year into anti-asteroid efforts. This calculation is crude, but the conclusion is clear: the roughly $10 million a year that the world pays to scan for big asteroids is money well spent.

What about extending the search to smaller ones? Because they are harder to find and would do less damage, the cost goes up and the benefit goes down. But recent studies, most notably a NASA report released in September, suggested that looking for the small guys still makes economic sense. Every year they have a roughly one-in-5,000 chance of taking out a city or triggering the mother of all tsunamis. On average, it works out to a couple hundred million dollars of damage a year. The search would cost a tenth of that.

When it comes to making active preparations, however, the balance of cost and benefit is unclear. Should we get a jump on deflection technologies, evacuation plans and the like, or can we prudently wait until we're sure that an asteroid is headed our way? To answer that, the world needs a high-level, high-profile study conducted not just by astronomers and geologists but also by economists and disaster planners. One of the authors of the article in this issue, asteroidologist Clark Chapman, has called for the National Academy of Sciences to weigh in. We agree.

Human beings are notoriously inconsistent about evaluating risks. Even by that low standard, though, we are ill prepared for threats of the asteroidal kind—so devastating that our existence could be at stake yet so infrequent that they sound practically like fairy tales. The difficulty of comprehending the threat makes a sober, comprehensive and authoritative analysis all the more urgent.

NASA

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*:
www.sciam.com/permissions
or
212-451-8546 for procedures
or
Permissions Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
Please allow three to six weeks for processing.

ADVERTISING

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

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

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

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

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

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

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

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

U.K.
The Powers Turner Group
+44-207-592-8331
fax: +44-207-630-9922

France and Switzerland
PEM-PEMA
+33-1-46-37-2117
fax: +33-1-47-38-6329

Germany
Publicitas Germany GmbH
+49-211-862-092-0
fax: +49-211-862-092-21

Sweden
Publicitas Nordic AB
+46-8-442-7050
fax: +46-8-442-7059

Belgium
Publicitas Media S.A.
+32-(0)2-639-8420
fax: +32-(0)2-639-8430

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

India
Yogesh Rao
Convergence Media
+91-22-2414-4808
fax: +91-22-2414-5594

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

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

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

On the Web

WWW.SCIENTIFICAMERICAN.COM

FEATURED THIS MONTH

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

Bacterial Battery Converts Sugar into Electricity

A tiny bacterium recovered from sediment may power batteries of the future. Researchers report that a primitive microbial fuel cell can convert simple sugars into electricity with 81 percent efficiency. Unlike previous attempts at creating such batteries, the novel design does not require unstable intermediaries to shuttle electrons. It thus holds promise for producing energy from waste materials containing sugar.

ENVIRONMENTAL BIOTECHNOLOGY CENTER/UNIVERSITY OF MASSACHUSETTS, AMHERST

Silkworm's Secret Unraveled

Scientists have long envied the lowly silkworm's ability to spin the strongest natural fiber known. Now they are one step closer to comprehending just how the creature manages the feat. According to the results of a new study, the key lies in the animal's ability to carefully control the water content of its silk glands. The findings should help improve artificial silk-making techniques.

Ask the Experts

What is the cosmic microwave background radiation?

Astronomer Erik M. Leitch of the University of Chicago enlightens.

Exclusive online issue: Forces of Nature

(On sale now for only $5)

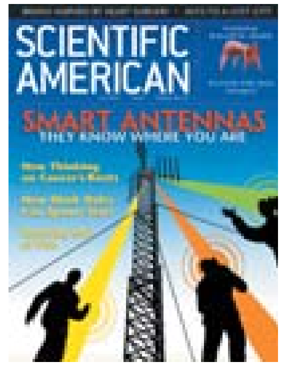
Earthquakes, volcanoes, tornadoes, hurricanes—for all the control humankind holds over its environment, sometimes nature just can't be contained. Scientists may never be able to tame these thrilling and terrifying forces, but advances in understanding them are leading to ways to save lives. In this exclusive online issue, experts share their insights into asteroid impacts, tornado and hurricane formation, and earthquake prediction. Other articles probe the mysteries of lightning and contemplate the future of an increasingly menacing volcano.

Find out more at www.sciam.com/special/

Letters

EDITORS@SCIAM.COM

FROM CELL PHONES to stem cells, cell-related technologies inspired many responses to the July issue. In a month marked by its celebration of independence, readers wrote about the liberties these various systems allow and reflected on how to keep business and law current with available technology. Some addressed the complexities of the freedom granted by wireless communications. Several pursued the issue of self-imposed limits on independent research and applications of cloning. American statesman and science enthusiast Thomas Jefferson once pondered this theme himself, postulating in 1810 that "laws and institutions must go hand in hand with the progress of the human mind." Feel free to read more about the July issue on the following pages.



WIRELESS IS MORE

Martin Cooper's article "Antennas Get Smart," on adaptive antenna arrays, trivializes some difficult technical and business problems. For example, the text includes only a short segment on multipath, but the vast majority of mobile calls are connected by multiple reflected signals (not direct line of sight) for at least part of the call. Multipath is the heart of the difficulty of achieving the full potential of smart antenna technology, but the mathematics underlying the processing for a phased array in a dynamic multipath environment with moving users and moving reflectors (like the bus going by your window) is daunting. Another concern is multicarrier performance. Network operators are building base stations operating at multiple carrier frequencies, so single frequency adaptive arrays are out of step with the market. But multicarrier adaptive arrays are harder to design, and more expensive to produce, than the single carrier type.

Steve Roemerman
CEO, Incucomm, Inc.
Richardson, Tex.

COOPER REPLIES: We certainly did not intend to trivialize either the technical or business challenges facing adaptive array technology. Both areas are indeed complex; ArrayComm has spent about $250 million over the past 11 years working toward a solution. At least a dozen other companies are currently in the smart antenna business as well. As the article states, the "personal cells" that characterize the most advanced adaptive array antennas are created by processing multipath data. Almost all cellular telephone calls involve multipath, and that is one of the reasons adaptive arrays are so effective.

Although multicarrier operation is complicated, AirNet Corporation is nonetheless demonstrating adaptive arrays in adaptive-array-equipped base stations for widely used standards. A European manufacturer is producing a third-generation cellular station, similarly equipped. ArrayComm's iBurst high-speed wireless Internet system is now operating in Australia with multiple carriers, lots of users and performance 40 times as great as systems without smart antennas.

The success of smart antenna technology is directly correlated with, among other things, the availability of cheap computing power. When we started attacking the task more than a decade ago, few computers existed at any price that were powerful enough to solve the complexities Roemerman mentions. A $100 chipset now does the huge computational job effectively for certain cellular standards. Of course, the technology of smart antennas is a challenge, but less daunting problems rarely yield such powerful results.

FISH GUARDS

"Counting the Last Fish," by Daniel Pauly and Reg Watson, stated that no nation had stepped up to its duties with regard to managing marine fisheries. Coincidentally, the truncated map adjacent to this misinformation omitted the single nation that has: New Zealand.

Colin MacGillivray
Auckland, New Zealand

PAULY AND WATSON REPLY: Our maps were intended to show the scope and intensity of changes in the global marine environment, and we regret that New Zealand was omitted. Fisheries management in that country is regarded by many as exemplary for its early establishment of (unfortunately small) marine protected areas and its efforts to limit fishing by privatizing fisheries through individual transferable quotas. These measures did not, however, prevent the crash of the country's valuable orange roughy stock in the late 1990s. Some experts, including Bjørn Hersoug in his book Unfinished Business, *have questioned whether New Zealand's quotas alone are adequate for ecosystem management, especially when only 9 percent of the nation's fish stocks can be evaluated in detail.*

ANCIENT IDENTITY ISSUES

Jonathan Mark Kenoyer's article "Uncovering the Keys to the Lost Indus Cities" refers to the animal shown on page 68 as a unicorn. I believe this is actually a bull seen in profile. Viewed from the side, curved horns seem to straighten, and the horn in the background becomes obliterated by the one in the foreground. It is not surprising that 65 percent of the images on the seals Kenoyer discovered depicted "unicorns." The bull was a widespread religious symbol throughout the ancient Middle East.

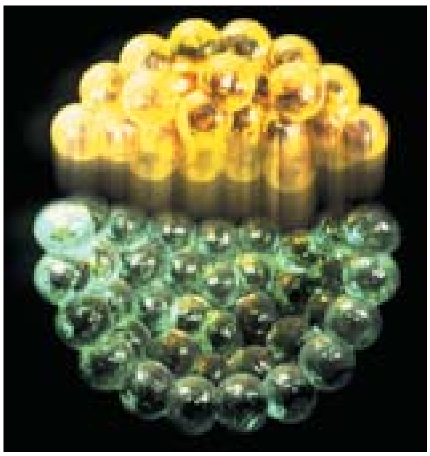
David M. Lank
Dobson Center for Entrepreneurial Studies
McGill University

KENOYER REPLIES: Available evidence indicates that Indus seal artists depicted side views of two-horned animals with two horns visible. Numerous seals show humped zebu and some nonhumped cattle with two horns. Furthermore, the discovery of one-horned "unicorn" terra-cotta figurines at the ancient sites of Mohenjo Daro, Chanhu Daro, Harappa and Dholavira confirms that the Indus people believed in a mythical animal with one horn, which we refer to as a unicorn.

THE ETHICS OF CELLING

Regarding "Terms of Engagement," by Sally Lehrman [Insights], I write to correct the implication that I, or other members of the President's Council on Bioethics, have acted on the basis of sectarian beliefs, rather than publicly accessible reasons, in reaching our judgments about the ethics of human cloning.

One need not be religious to have ethical concerns about the production, use and destruction of cloned human embryos—even in the service of the noble cause of science and medicine. In keeping with our mandate, the council has sought to "articulate fully the complex and competing moral positions" in terms that would help to educate and inform the national dialogue. By joining with the majority of the council in calling for a four-year moratorium on cloning for biomedical research, I sought time to deepen and extend the scientific and ethical understanding essential for discussion of a subject of such significance for the character of our society as a whole. And, as with all of the council's deliberations and recommendations, my own positions were formulated and expressed drawing on scientific evidence and reasoned moral argument. I would direct the reader to the council's report "Human Cloning and Human Dignity" (see bioethics.gov/reports/).



BIOLOGISTS' VIEWS on human cloning are as divided as the cells represented in this artwork.

Irving Weissman makes a comment to the effect that there is no "assay for a human soul." But if by "soul" we mean the principle of the dignity and moral nature of a human life, then we must seek something beyond empirical evidence to guide our scientific project. Here the enduring religious and moral traditions that have always been part of the practice of medicine can inform our moral reflection and moral reasoning. Although I agree with Weissman that the Hippocratic oath can help serve as a moral guide, his paraphrasing of the oath was inaccurate. Far from a repudiation of "personal ethical, religious [and] moral concerns," it advocates the alignment of medical practice with strict moral principles demanding respect for human life. For example, as originally formulated, it directly prohibits both euthanasia and abortion. Anthropologist Margaret Mead aptly described the Hippocratic tradition of "separation between killing and curing" to be a "priceless possession which we cannot afford to tarnish."

In keeping with the principles of the democratic process, I hope we will stop misrepresenting and dismissing the views of those with whom we disagree. We can then engage in genuine and productive dialogue to open scientific progress within a wider moral consensus.

William B. Hurlbut
Program in Human Biology
Stanford University

GUN SAFETY?

I very much enjoyed Steve Mirsky's description of the Stupid Security Awards in "The Yanked Clippers" [Anti Gravity]. I recently attended a gun show here in Albany, N.Y. Security was tight as people entered the parking garage under the Empire State Plaza, backing up traffic for half a mile. Visitors brought in guns and ammunition with no problem. What was security keeping out?

Warren Redlich
Republican candidate for Congress
New York State, 21st Congressional District

ERRATUM In "Brief Points" [News Scan], the full name of the publication listed as *Psycopharmacology* should be *Psychopharmacology Bulletin.*

50, 100 & 150 Years Ago
FROM SCIENTIFIC AMERICAN

Mathematics of Children ▪ Culture of Crete ▪ Philosophy of Heat

NOVEMBER 1953

CHILD LEARNING—"It is interesting to study how children spontaneously learn to measure. One of my collaborators, Dr. Bärbel Inhelder, and I have made the following experiment: we show the child a tower of blocks on a table and ask him to build a second tower of the same height on another table (lower or higher than the first) with blocks of a different size. He begins to look around for a measuring standard. Interestingly enough, the first measuring tool that comes to his mind is his own body. He puts one hand on top of this tower and the other at its base, and then, trying to keep his hands the same distance apart, he moves over to the other tower to compare it. Children of about the age of six often carry out this work in a most assured manner, as if their hand could not change position on the way! —Jean Piaget"

COMPACT POWER—"The gas turbine, today popularly known as the jet engine, born barely a dozen years ago, has come forward with enormous speed, not only in aircraft but also in a range of other applications. By 1965, if not sooner, it will be indisputably the engine of the age. It is likely to reshape all surface transportation and revolutionize the stationary generation of power. The gas turbine, indeed, is the most versatile prime mover that man has yet built. The two big U.S. steam-turbine builders, General Electric and Westinghouse, put their first stationary gas-turbine power units into operation almost simultaneously in 1949, and there are now 20 in the U.S."

NOVEMBER 1903

PRINTING REVOLUTION—"Some ten years ago aluminum began to be manufactured in a sufficient quantity to make it commercially useful, and it was soon discovered that this light, white metal could be treated to give it the property of printing like lithographic stone. As long as stone was the only surface printing material, only one form of press, the flatbed, was practical. With a metallic plate it was possible to bend the metal to a cylinder. With the rotary press it was simple to pass the paper sheets between two cylinders, as clothes are passed through a laundry wringer, and get twice as many impressions as from the slow-moving flatbed. There has been indeed a revolution in lithographic establishments, until some of the larger shops now print 90 per cent of their work from rotary presses."

HOW CHILDREN LEARN, as studied by Jean Piaget, 1953

ANTIQUITIES OF CRETE—"Dr. Arthur Evans has ceased, for the time being, his great labors in Crete. Where are his treasures to be stored? Many have hoped that some of them might find their way, considering Dr. Evans's nationality, to the British Museum. It is now reported from Munich, however, that the foundation stone of a Cretan museum has been laid in Candia, wherein will be stored the priceless antiquities which have rewarded Dr. Evans for his spadework in Knossos. Remembering the shame of the Elgin marbles, we can only say that this is well. Crete, to which we owe an inestimable debt, is surely entitled to the possession of those great beginnings of fine art and those significant clay tablets with which she initiated European history three thousand five hundred years ago."

NOVEMBER 1853

THE MOSQUITO'S TRAIL—"There certainly is a greater proneness to disease during sleep than in the waking state. Those who pass the night in the Campagna di Roma inevitably become infected with its noxious air, while travelers who go through without stopping escape the miasmi."

WHAT IS HEAT?—"What do we know of heat as a substance? Has any man seen it with his eyes, handled it with his hands (like a stone) or weighed it in a balance? No. We have no positive proofs then that it exists as matter at all, and know nothing about it as such; but as a *quality* belonging to all matter, and developed under certain conditions, we know a great deal. Heat is a property with which the Great Creator has endowed all matter, the same as he has endowed all matter with the quality of *gravity*."

ELECTRICITY

Healing the Grid

SEVERAL NEAR-TERM SOLUTIONS CAN KEEP THE JUICE FLOWING **BY JR MINKEL**

DAWN'S EARLY LIGHT shines on a blacked-out New York City.

If the electric power grid is the nation's circulatory system, then it suffered a massive heart attack on the afternoon of August 14 when lights winked out from Ohio and Ontario to New York. Although no one knows precisely why a seemingly mundane local system failure cascaded so far, researchers have long seen tension in the grid and are pondering ways to minimize the chance of big blackouts.

The grid represents a delicate balancing act: the amount of electricity sucked from the lines (the load) at every moment has to match the electricity being generated. If generation slows too much, system controllers have to shed load, causing a blackout. Further complicating matters, electricity flows through the grid primarily as alternating current. So AC frequencies at each station must match but be offset in a precise manner to keep power flowing in the right direction.

Partial deregulation during the early 1990s allowed some states to separate their generation and transmission industries. Generation systems boomed, but transmission lagged behind because of the patchwork of interstate regulations and jurisdictions. Many policy and grid experts say that in the short term, the Federal Energy Regulatory Commission should enact nationwide policies covering transmission system operation, capacity and investment. The commission could force transmission owners to join Regional Transmission Organizations that would implement the policies.

Once the government decides how the grid should operate, "we have the technology to implement it almost on the shelf or coming down the pipe," says Paul Grant, science fellow at the Electric Power Research Institute (EPRI), an industry consortium in Palo Alto, Calif. Currently protective relays shut down power lines if high currents threaten to make them overheat and sag, but those lines could be kept functioning with more heat-resistant lines, which are already available. Generators, which are basically giant flywheels, switch off if the AC frequency or phase changes rapidly (because

KEEPING IT SIMPLE

In the long run, reduced grid complexity could be attractive. Direct current lines, which have no frequency associated with them, act as shock absorbers to disturbances in today's AC system. DC lines already separate the Texas power grid from the eastern and western grids. Adding more could help make the whole system more stable, although high-voltage DC is expensive, and replacing the right lines amid the tangle of interconnections would not be trivial.

CHIP EAST *Reuters*

news SCAN

the generators can damage themselves trying to respond); so-called breaking resistors, which exchange electricity for heat, could help generators make smoother transitions.

Better communication among power stations would also aid in stabilizing the grid. Protective relays rely on local information and can be fooled into disconnecting a line unnecessarily. Dedicated fiber optics would permit fast comparisons of conditions at adjacent stations, forestalling needless shutdowns. The Global Positioning System (GPS) could put a time stamp on each station reading, allowing operators to make better decisions by looking at successive snapshots of grid conditions. The Bonneville Power Administration, based in Portland, Ore., and Ameren Corporation, a St. Louis–based utility, use GPS time stamping.

Once operators get a picture of grid conditions, they could disseminate the information to faster, smarter switches. Flexible AC transmission system devices can tune power flow up or down, and superconducting valves called fault current limiters could enable circuit breakers to disconnect lines in a safer way. Installing more AC lines or more powerful superconducting lines alone would increase transmission capacity but could lead to bigger ripples in the grid if something went wrong. "You've got to be able to contain a major disturbance, and the most common way to do that" is to disconnect lines, explains electrical engineer Peter Sauer of the University of Illinois.

Ideally, Grant states, a master computer with a bird's-eye view would serve as air traffic control for the grid. Postmortem studies by the industry suggest that such a global view would have prevented about 95 percent of customers from losing power during the 1996 blackouts in the western U.S., he says. Although experts differ on the feasibility of constructing an über-computer, most agree that a slightly less ambitious scheme might work.

One such scheme involves an improved control method designed to automatically quarantine trouble spots and gerrymander the remaining grid into islands of balanced load and generation. EPRI commissioned computer-modeling studies of the technique, called adaptive islanding, which concluded that it could preserve more load than conventional responses. Massoud Amin, an electrical engineer at the University of Minnesota who headed the EPRI program that co-funded the research, says adaptive islanding could be implemented within five years.

Nobody familiar with the power grid expects blackouts to disappear entirely. If chaos or network theories are right, a chance of large cascading failures is inherent to stressed or highly interconnected systems. And with every incremental increase in grid reliability, the cost of the next increment goes up. So keeping a stash of fresh batteries will make sense for a long time.

JR Minkel retreated to the local park when his Brooklyn apartment lost power.

NEED TO KNOW: GRID TIMES

Outages affecting more than 500,000 customers:
1991 to 1995: 41
1996 to 2000: 58

Outages exceeding 100 megawatts:
1991 to 1995: 66
1996 to 2000: 76

Percent increase in total U.S. electricity demand:
1988 to 1998: 30
1999 to 2009 (projected): 20

Percent increase in transmission network capacity:
1988 to 1998: 15
1999 to 2009 (projected): 3.5

Industry R&D spending in the U.S. as a percentage of net sales, 1999:
Communications equipment: 12
Computer/electronics: 11
Electric utilities: less than 0.1

SOURCES: North American Electric Reliability Council; Energy Information Administration; National Science Foundation.

COMPUTERS

Malcode Melee

IN THE WAR OF THE WORMS, WAS ONE WEARING A WHITE HAT? **BY W. WAYT GIBBS**

Amid the several viral and wormy outbreaks that buffeted the Internet this past August, one had a peculiar modus operandi. Whereas the Sobig.F virus jammed up networks with virulent e-mail and the Blaster worm forced its host machines to reboot every few minutes, Welchia seemed to have honorable intentions. Some observers dubbed it a "white hat" worm.

After it enters a new PC, the Welchia worm forces the computer to contact Microsoft's Windows Update Web site and download a patch for the very hole that it and Blaster exploit. Welchia next attempts to remove the Blaster worm if the host machine is afflicted with it. Welchia then scans the local network for more vulnerable systems and attempts to procreate. But it contains an unusual subroutine: come New Year's Day 2004, the Welchia program deletes itself.

Through Welchia, maybe some well-meaning hacker attempted to clean up the mess caused by other bugs. The consequences of Welchia's rapid spread—it hobbled the U.S.

VIRAL ATTACK: Spam awaited tens of thousands of unwary victims of the Sobig.F e-mail virus.

Navy–Marine Corps intranet, shut down part of Lockheed Martin's network for 12 hours, and choked Air Canada's ticketing system, forcing the airline to cancel flights—were perhaps just unintended side effects.

But is that what really happened? Marty Lindner, who leads the incident-handling team at the CERT/CC Internet security organization in Pittsburgh, Pa., thinks an ecological explanation is much more likely.

Almost from birth, the Internet has been infested with viruses, Trojan horses, worms and other malicious software, or "malcode." To these synthetic pests, the Net is like the patchwork of cornfields that dot the Midwest, forming a vast reservoir of hosts for opportunistic bugs. As fields become larger, more connected, and more of a monoculture, the harm that any given parasite can inflict grows, too. But eventually the pests start competing with one another.

By closing the hole behind it, Lindner points out, Welchia guarantees that other worms can't follow it in. By deleting Blaster, he notes, Welchia stops the machine from rebooting and using its network connections for a competing task. These strategies make the compromised host a better platform for launching further Welchia attacks.

Lindner's theory is supported by the discovery that Welchia performs a fifth, distinctly hostile, job. "It installs a surreptitious file transfer server," Lindner reports, which gives the author of Welchia "a backdoor into the system." The suicide subroutine could simply be a scheme to remove evidence of the infection after that door has been propped open.

Whether Welchia was meant for good or ill, it does invite a question: Might it someday make sense to fight one worm with another? Farmers, after all, sometimes release one species of insect to thwart a burgeoning invasion by a second species.

"It is a very interesting notion," says Michael Liljenstam, who develops simulations of malcode epidemics at the University of Illinois. "A 'good' worm is not necessarily doomed to failure." But it would have to be released very quickly and remove itself in clever ways. "And it could be a difficult balancing act between spreading quickly enough to prevent infection and using up so much bandwidth that the cure is worse than the disease," Liljenstam concludes.

FAST FACTS: ONE WORMY WEEK

Pathogen	First Detected	Vulnerable Operating Systems
Blaster worm	August 11	Windows 2000, XP
Sobig.F virus	August 18	Windows (all versions)
Welchia worm	August 18	Windows 2000, XP

SOURCE: Symantec

OPTICS

All Screwed Up

AN OBSCURE PROPERTY OF LIGHT PUTS A SPIN ON ASTRONOMY BY GEORGE MUSSER

You'd think we'd have figured out light by now. Kids learn about prisms and lenses in elementary school, people wear Maxwell's equations on T-shirts, and the quantum version of those equations is the most precise theory in science. Yet knotted up within the theory is a phenomenon that physicists are still unraveling: an unexplored property of light.

In addition to color (which depends on the wavelength of the electromagnetic wave) and polarization (the orientation of the wave), light beams can also possess orbital angular momentum (the shape of the wave fronts). Optics researchers discovered this property a decade ago, but for some reason this realization has failed to propagate much beyond a small community of specialists [see "Hands of Light," Innovations, SCIENTIFIC AMERICAN, August]. It has barely been noticed even by those with the greatest need to exploit every conceivable aspect of light—namely, astronomers.

An astronomer has now taken it upon himself to spread the word. In the November 10 *Astrophysical Journal*, Cornell University emeritus professor Martin Harwit suggests that the orbital angular momentum of light could convey new information about celestial bodies—information unavailable by looking just at color and polarization. "The paper was mainly meant to be provocative," he says. "People are flabbergasted that this should even be possible."

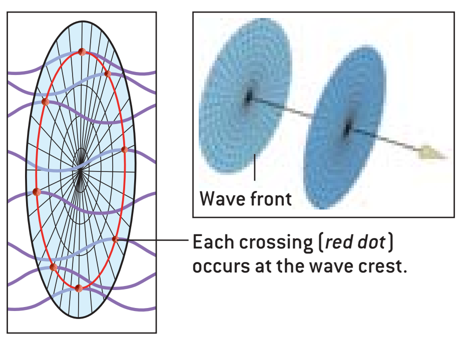
In an ideal beam of light, produced by a laser or a distant star, the wave fronts are flat. On each slice through the beam, the wave is at the same phase in its oscillation cycle: crests line up with crests, troughs with troughs. But in a slightly more complicated beam, the phase changes with the angle around the beam's axis. The 12 o'clock position on a slice might correspond to a crest, the 6 o'clock position to a trough [*see illustrations below, left*]. If you connect the wave crests, they form a helix. The next most complicated possibility is a double helix, in which the phase changes twice as rapidly (with troughs at 3 o'clock and 9 o'clock); beyond that is a fusilli-like triple helix (2 o'clock, 6 o'clock and 10 o'clock), and so on.

Like polarized light, twisted light carries angular momentum: in lab experiments, it has set small plastic beads spinning. If you think of light in terms of particles (photons) rather than waves and neglect some quantum-mechanical caveats, it is as though the photons were zipping along a corkscrew path.

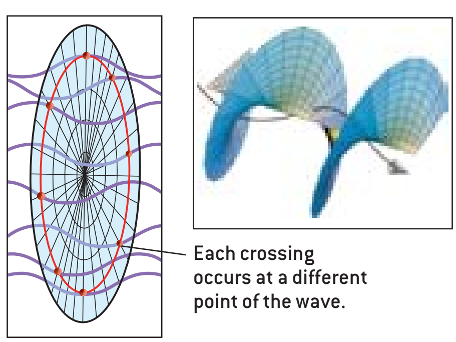
To create twisted light, physicists shine a laser through a helical lens or a special diffraction grating. Harwit argues that light could also be twisted by natural processes in the universe, such as lenslike density variations in interstellar gas or the warped spacetime around rotating black holes. Alien civilizations might transmit information by twisting light rather than using other encoding methods (as indeed physicists have proposed for terrestrial free-space communications). The most sensitive way to measure the twist would be a series of interferometers, as demonstrated last year by a team led by physicists Jonathan Leach and Miles Padgett of the University of Glasgow.

One peculiar aspect of twisted light could prove especially endearing to astronomers.

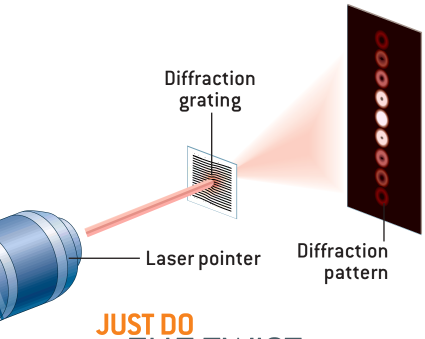
POLARIZED



TWISTED AND POLARIZED



TWIST SO FINE: In linearly polarized light (*top*), the electric component of the light wave oscillates up and down everywhere in sync, yielding wave fronts that are parallel slices. Add a twist to the light (*bottom*), and the wave gets out of sync in a particular pattern. In this case, the crests trace out a helix.



JUST DO THE TWIST

A simple laser pointer can demonstrate twisted light:

1. Download the diffraction grating pattern from departments.colgate.edu/physics/research/optics/oamgp/gp.htm. The fork at the center of the pattern is what twists the light.

2. Using a photocopier, reduce the pattern to about half a centimeter on a side and transfer it to an overhead transparency.

3. Shine the laser through the pattern, ensuring that the beam passes through the fork, and project it onto a wall a few meters away. The grating splits the laser beam into a row of circles. Each of the circles flanking the central circle should have a small hole in the middle. The holes are a sign that light is being twisted.

SAMUEL VELASCO (*left top and bottom and margin*); SOURCE: KIKO GALVEZ *Colgate University*; JOHANNES COURTIAL *University of Glasgow* (*right top and bottom*)

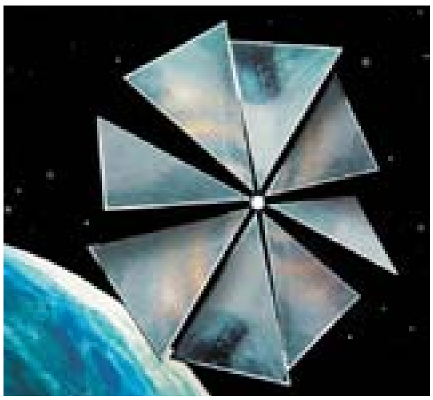
Just as Earth's North Pole sits in every time zone, the central axis of the beam contains waves of every phase. All those waves cancel one another out, leaving utter blackness. As a result, a lens focuses twisted light to a ring instead of a point. In 2001 physicist Grover Swartzlander of the University of Arizona proposed using this feature to look for extrasolar planets. Installed in a telescope, one of the special diffraction gratings would smear starlight into a ring, leaving a hole so dark that a nearby object millions or billions of times as faint could become visible. "It's a completely original idea," Padgett says. "When I first read the paper, I said, 'Gosh, that's a cute idea.'" Contemporaries of Newton probably thought it pretty cute that white light could be split into a rainbow of colors. Maybe one day twisted light will come to seem just as commonplace.

SPACEFLIGHT

Light Sails to Orbit

NASA WATCHES FROM THE SIDELINES AS COSMOS 1, THE FIRST SOLAR SAIL, GOES UP BY PHILIP YAM

Shiny and crinkly, the material looks more like something meant to wrap frozen foods than to provide a new way to travel through space. The aluminized Mylar reflects sunlight, thereby deriving a little kick from the recoiling photons. In principle, big sheets could act as solar sails that over time would reach speeds exceeding 100 kilometers a second—far faster than chemical rockets.



COSMOS 1 in flight—as an artist sees it.

COURTESY OF THE PLANETARY SOCIETY AND COSMOS STUDIOS; IMAGE PREPARED BY BABAKIN SPACE CENTER

The first solar sail, called Cosmos 1, will go for its test flight in early 2004. The demonstration of a revolutionary way to travel to the planets and maybe even to the stars would seem to be a natural activity for NASA, which spends several million dollars every year researching advanced propulsion systems. Yet in this case, the space agency has chosen to be a bystander.

The successful flight of Cosmos 1 would mark the culmination of three years of effort by the Planetary Society, a space-interest group, and the entertainment media firm Cosmos Studios [see "Sailing on Sunlight," News Scan, Scientific American, July 2001]. Both organizations, which can trace their roots to the late Carl Sagan, used their connections with Russian space officials and engineers. They enlisted the Babakin Space Center in Moscow as the prime contractor for Cosmos 1, which cost $4 million—cheap in the space-travel world. The craft consists of eight triangular Mylar panels 14 meters long stretched across inflatable spars. The goal is to have Cosmos 1 ride atop a modified ballistic missile launched from a Russian submarine. Once in orbit, the spacecraft would inflate the spars to unfurl the sails. The panels would spread out like flower petals and cover about 600 square meters. Then sunlight should push the sails, lifting Cosmos 1 into a higher orbit from its initial 800-kilometer altitude.

Russian involvement may be one reason NASA has shied away, suggests Louis D. Friedman, executive director of the Planetary Society. Informal discussions had NASA supplying the sail material, which is tougher and, at 2.5 microns thick, half the thickness (and therefore half the weight) of the Russian film being used. "We would have gotten it for free and tested it for them," Friedman says. But NASA management never gave the go-ahead. Bureaucracy might have been a problem, he surmises, with the "upper echelons fearing private companies working with the Russians on a submarine launch."

In any case, strict rules govern how closely NASA can work with other countries, remarks Hoppy Price, who was the lead solar-sail engineer for NASA at the Jet Propulsion Laboratory in Pasadena, Calif. "Possibly NASA is worried about the transfer of technology," he notes. Moreover, solar sails may provide some military advantage that the U.S. would rather not share. One proposed application, for instance, has solar sails hovering over the poles to provide valuable uplinks to anyone at the earth's communications-starved extremities.

Risk, though, is probably the main reason for NASA's noninvolvement. Battered by a bruising report about the *Columbia* disaster as well as by the loss of two Mars-bound spacecraft in 1999, the agency "can't spend taxpayer money with the level of risk" that the Cosmos 1 team is taking, notes Neil Murphy, who currently coordinates the solar-sail work at JPL. Plenty of pitfalls abound. "Concern lies with what happens to an ultrathin material over tens of meters," Friedman says, noting that engineers have no good way on the earth to test the behavior of the material in zero gravity. "You can imagine all sorts of problems—take Saran Wrap and wave it around," he offers. Ripping, fluttering and sagging would all undermine the sail's ability to reflect photons.

READY TO GO: Louis D. Friedman, Cosmos 1 project director, gives the craft a once-over. He had hoped for a test flight in October; scheduling conflicts with the Russian navy has pushed the date to early 2004.

COURTESY OF THE PLANETARY SOCIETY AND COSMOS STUDIOS

NASA would also want a solar-sail launch to have science-based goals to refine models and to plan the next mission, Murphy explains. Cosmos 1 is mostly a demonstration, and the components are not suitable for an extended voyage. The inflatable spars, for example, will not remain rigid for long because of the inevitable micrometeoroid impacts.

NASA is working on a more advanced solar-sail craft, probably to be configured as four square panels, but it won't be ready for at least another few years. That leaves the privately organized Cosmos 1 as the lone player—and NASA engineers in the cheering section.

GROUNDED THOUGHTS

Solar sails cannot fly, argues astronomer Thomas Gold of Cornell University. Gold is known for controversial ideas—for example, he has postulated that crude oil comes from geologic activity, not from dinosaurs and other past life. In the case of solar sails, he relies on thermodynamics: he notes that perfect mirrors do not create temperature differences, which are necessary to convert heat into kinetic energy.

Gold's analysis created a stir among solar-sail scientists, who think that 19th-century physics is the wrong reasoning to apply. Rather "it's the quantum-mechanical interaction between photons and sails" that must be examined, states Hoppy Price of the Jet Propulsion Laboratory. The flight—or nonflight—of Cosmos 1 should settle the matter.

NEUROLOGY

Brain Not Inflamed?

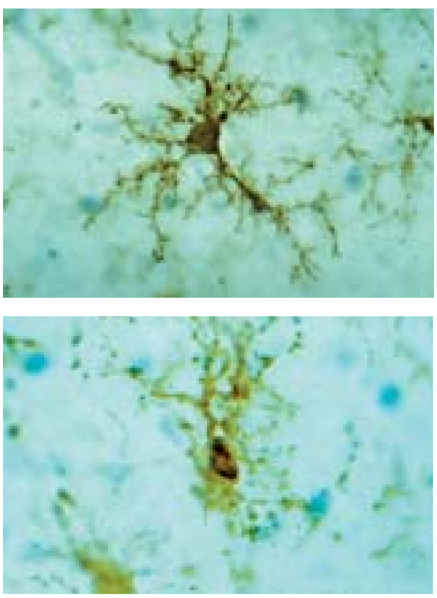
ALZHEIMER'S MAY NOT BE AN INFLAMMATION AFTER ALL BY DENNIS WATKINS

Researchers have for years observed that patients regularly taking ibuprofen, naproxen or other nonsteroidal anti-inflammatory drugs seem to have less risk of developing Alzheimer's disease. Some researchers hypothesize that the Alzheimer's-diseased brain is actually inflamed and that damage happens when the microglia, the brain's immune cells, become overactive and attack healthy neurons. New research, however, indicates that the opposite may be happening—that, as microglia age, they lose their ability to protect the brain.

Wolfgang J. Streit and his colleagues at the University of Florida compared autopsy tissue from two nondemented brains, one of a 38-year-old man and the other of a 68-year-old man. Many of the microglia in the older man's brain had lost their fine branches or were otherwise deformed. Streit found even more of these withered microglia in the brains of people who also had high levels of beta-amyloid protein—a hallmark of Alzheimer's. Streit hypothesizes that beta-amyloid may cause the deformities in microglia.

Moreover, Streit's lab examined in vitro cultures of rat microglia and determined that over time, their telomeres shorten (as they do for most other aging cells). Telomeres are end caps on chromosomes that help to maintain the integrity of the genes; as they shorten, the cells lose the ability to replicate and begin to die off. So "if we can keep our microglial cells healthy, then our neurons will be in good shape," Streit suggests. (Telomeres of neurons do not shorten.)

In further defense of his theory that aging microglia are associated with Alzheimer's, Streit points to a drug trial in the June 4 *Journal of the American Medical Association.* Contrary to previous preliminary findings, the study showed that Alzheimer's patients taking anti-inflammatory drugs fared no better than those taking a placebo. "I'm discouraged by this class of drugs on the disease," admits Paul S. Aisen, a neurologist at Georgetown University and lead author of the work. "Personally, I'm looking at other



WOLFGANG J. STREIT *University of Florida*

MICROGLIA from a young brain (*top*) appears healthier than that from an aged one (*bottom*). Deformed microglia are tied to Alzheimer's.

approaches to treatment that are not related to inflammation."

Aisen believes that microglia could potentially act both as protector and attacker. "There is evidence for both viewpoints, even though they are exclusive," Aisen says. "I just don't think we have evidence of what the net effect is of microglia during Alzheimer's." The key to this puzzle, he explains, lies in the interaction between microglia and beta-amyloid protein. In Alzheimer's patients, the proteins form tangled plaques in the brain. Microglia could be clearing away these harmful plaques.

Increasing the number of microglia, however, may have dangerous side effects. In January 2002 trials of a drug called AN1792, which was designed to immunize patients against the accumulation of beta-amyloid, were stopped because four subjects developed encephalitis. One woman was so debilitated after treatment stopped that doctors could not even give her a psychological examination, and she died less than two years after beginning therapy. "Anybody who stimulates inflammation is playing with physiological matches," warns Patrick McGeer, a neurologist at the University of British Columbia. McGeer adds that if Alzheimer's resulted from the aging of microglia, then giving a patient anti-inflammatory drugs to further suppress the immune response would exacerbate the disease. Streit, on the other hand, argues that the microglia were not functioning to begin with, so there was nothing to suppress.

John Breitner, an epidemiologist at the University of Washington, is studying whether anti-inflammatory drugs can prevent the disease from developing in the first place. Even if he discovers that the drugs are effective, Breitner says, that will still not explain exactly how they work, leaving the door open to a wide variety of theories on Alzheimer's and microglia. "We may all be barking up the wrong tree," he speculates. "It may be something that none of us has looked at."

Dennis Watkins is a science writer based in Woodbine, Md.

ECOLOGY

Restoring the Rio

EFFORTS TO KEEP THE RIO GRANDE FILLED WITH WATER **BY KRISTA WEST**

With its multiple dams, flood-control mechanisms and crop-irrigating structures, the Rio Grande has provided residents of Colorado, New Mexico and Texas with a reliable source of freshwater for nearly a century. And for almost as long, farmers, municipalities and conservationists have tussled over who has the right to use it. Now a new player has entered the disputes, one that could raise national awareness of the conflicts draining the Rio Grande—the federal government. Department of the Interior Secretary Gale A. Norton has proposed an $11-million congressional initiative to improve Southwest water management, measurement, storage and delivery and is leading a series of regional water conferences. If the initiative passes, it would become the first federal funding of its kind for the region and the river.

The fifth-longest river in North America, the Rio Grande begins in Colorado and winds 1,900 miles through New Mexico, Texas and Mexico—bisecting the northern half of the ecologically rich Chihuahuan Desert—before emptying in the Gulf of Mexico. Approximately 10 million people live along the Rio Grande's banks, and no single state or country has management authority or responsibility for the health of the river. Currently more than 80 percent of the Rio Grande's southern flows are diverted for agriculture, says agricultural engineer J. Phillip King of New Mexico State University. Historically, in fact, undiverted water was considered wasted.

But this spring an unknown, un-

news
SCAN

H_2 WOE: Water fights among multiple interest groups take their toll on the Rio Grande, shown here in Big Bend National Park in Texas. A new federal push to save a minnow species may help restore the river.

TIM FITZHARRIS *Minden Pictures*

threatening and relatively unimpressive endangered species called the Rio Grande silvery minnow forced river managers, for the first time in history, to leave previously allocated water in the river. This two-inch gray fish once swam abundantly throughout the river but today remains restricted to a 100-mile stretch in central New Mexico, representing about 5 percent of its historic range. This year scientists predicted that the stretch would run dry because of drought and overallocation of water for human activities, potentially sealing the fate of any remaining wild minnows.

In June a federal court ruled that the agency managing river flows, the Bureau of Reclamation, must, under federal law, provide the fish with water regardless of existing obligations to other water users. The bureau holds standing contracts to deliver water to the state's cities and farmers that, with this ruling, are unlikely to be met.

The decision, which was not welcome by state political leaders, coincided conveniently with efforts by the Department of the Interior. In addition to the initiative that Norton has proposed, the department has sponsored a series of meetings known as Water 2025 that began in June in Denver. The Interior has not been active in Southwest water issues since the early 1900s, when it helped to construct many of the Rio Grande's dams, levees and canals.

Many conservation groups are hoping the new federal interest will do something that they have been unable to accomplish despite years of effort—put the Rio Grande on the national radar screen as a place worth protecting. The river system (which includes the Rio Grande Basin and part of the overlapping Chihuahuan Desert) matches up well with a long-recognized national treasure—the Florida Everglades. The two regions are surprisingly similar: they are each home to roughly the same number of protected species, both consist of a river system that feeds a well-known national park (the Rio Grande flows through Big Bend National Park), and both are valuable agricultural regions. Yet the Rio Grande does not have the national status of the Everglades.

"The tremendous challenge for the Rio Grande," says Ron Tipton, vice president of programs for the National Parks Conservation Association, "is getting the country to notice the region." Tipton points out that efforts to protect the Everglades began as early as the 1960s, but it was federal attention and funding ($8 billion) obtained by Florida governor Bob Graham in 1984 that established the Everglades as a national asset.

Bob Irvin of the World Wildlife Fund agrees. "State cooperation was essential to the restoration of the Everglades, and it will be essential to the Rio Grande as well," he notes. "But federal leadership will be the key ingredient."

Krista West writes about conservation issues from Las Cruces, N.M.

DIVERSITY IN THE DESERT

The region represented by the Rio Grande Basin and the Chihuahuan Desert rivals the Florida Everglades in terms of ecological uniqueness. The approximate numbers of species are:

Everglades

Birds: 300
Mammals: 40
Reptiles/amphibians: 51
Fish: 150

Rio Grande/Chihuahuan Desert

Birds: 350
Mammals: 100
Reptiles/amphibians: 100
Fish: 29

Protected species

Everglades: 60
Rio Grande/Chihuahuan Desert: 64

SOURCES: World Wildlife Fund and Everglades National Park

BY THE NUMBERS

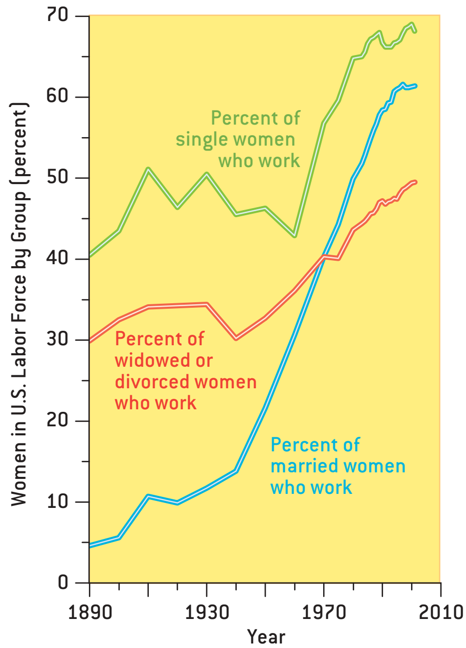
Why Women Work

THE EVOLUTION OF WOMEN IN THE U.S. JOB MARKET **BY RODGER DOYLE**

In the early Republic, most American women worked at home making soap, candles, clothes, shoes and other necessities for their families. But with the coming of the industrial revolution early in the 19th century, some worked for pay at home, using the machines and textiles supplied by merchants to produce clothes for the market. The first women to work outside the house in substantial numbers were single farm girls who took jobs in the new textile mills of New England beginning in the 1820s. Thereafter, women expanded into sales, domestic service, teaching and other occupations. Hardly any became doctors, lawyers or college professors, and most gave up their jobs after marriage.

Near the start of the 20th century, the emerging notions about women's roles, the greater availability of white-collar jobs and increasing pay lured married women into the labor market. Perhaps the most interesting explanation for the rise of married women in the workplace comes from the late Winifred D. Wandersee, a historian who taught at Hartwick College. Beginning early in the 20th century and with growing force in the 1920s, Americans had higher expectations of what constituted the good life. Everyone wanted the latest things—electric lighting, indoor bathrooms, telephones, refrigerators, washers, dryers and, above all, automobiles.

The old psychology of scarcity was giving ground to a psychology of abundance, and this trend accelerated in the 1920s thanks to several developments, including the consumer advertising that accompanied the advent of radio and the growth of consumer credit, when techniques such as installment buying were perfected. Expectations flowered in the even more prosperous 1950s and 1960s. Nearly every family, Wandersee contended, defined its standard of living in terms of an income that they hoped to achieve, rather than actual income, and thus the economy was propelled ever upward on a sea of consumer debt. Wandersee noted that, at least before World War II, women who worked were motivated primarily by economic need, not by career aspirations.



SOURCES: 1890 to 1960: U.S. Decennial Census data. 1970 forward: Bureau of Labor Statistics annual data. Data refer to women 16 and older.

Things changed substantially after the ascent of feminism in the 1960s and 1970s, when home economics was dropped as a requirement for high school girls. In 1970 women were awarded 43 percent of bachelor degrees and 9 percent of professional degrees; by 2001 these percentages had risen to 57 and 45 percent, respectively.

According to social critic Sally Helgesen, a change in the nature of corporate enterprise beginning in the 1970s made it easier for women to get better jobs. Corporate management was almost exclusively male, but as foreign competition and new technologies destabilized the economic environment, organizations had to change radically to survive, which meant drawing on the widest pool of talent. As a result, women increasingly occupied positions of authority in business, law, medicine, the military and politics. Still, fewer women than men work, and the glass ceiling remains in place, to judge by the number of Fortune 500 companies with women CEOs: in 2002, seven.

Rodger Doyle can be reached at rdoyle2@adelphia.net

RODGER DOYLE

MEN VERSUS WOMEN

Percent of population 16 years of age and older in labor force, 2001

	Women	Men
Total	59.6	74.1
Age		
16–19	47.3	47.5
20–24	72.1	80.7
25–34	75.1	92.4
35–44	76.4	92.1
45–54	76.0	88.5
55–64	55.2	69.2
65+	9.8	17.9

FURTHER READING

Women's Work and Family Values, 1920–1940. Winifred D. Wandersee. Harvard University Press, 1981.

Everyday Revolutionaries: Working Women and the Transformation of American Life. Sally Helgesen. Doubleday, 1998.

Working Women in America: Split Dreams. Sharlene Hesse-Biber and Gregg Lee Carter. Oxford University Press, 2000.

news SCAN

DATA POINTS: CUTTING CALORIES

A dose of an intestinal hormone called peptide YY_{3-36} (PYY) dramatically suppresses the urge to eat, without side effects, according to a study by Stephen R. Bloom and his colleagues at Imperial College London. Like the better-known hormone leptin, this peptide regulates the biochemical pathways in the hypothalamus that govern appetite. But unlike with leptin, obese subjects were not resistant to the effects of PYY. A natural deficiency of PYY may contribute to weight gain.

Percent reduction in food calories consumed at a buffet two hours after taking PYY:

Obese subjects: 29.9

Lean subjects: 31.1

Total daily food calories consumed:

Obese subjects given placebo: 2,456

Obese subjects given PYY: 1,810

Lean subjects given placebo: 2,312

Lean subjects given PYY: 1,533

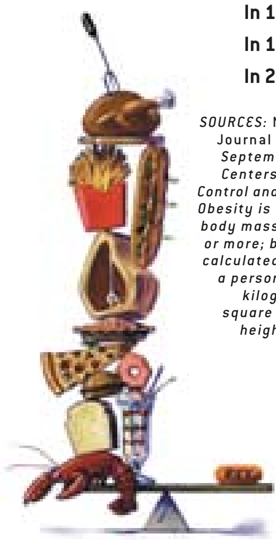
Percent of U.S. adults who are obese:

In 1991: 12.0

In 1995: 15.3

In 2001: 20.9

SOURCES: New England Journal of Medicine, September 4, 2003; Centers for Disease Control and Prevention. *Obesity is defined as a body mass index of 30 or more; body mass is calculated by dividing a person's weight in kilograms by the square of his or her height in meters.*



OBITUARY

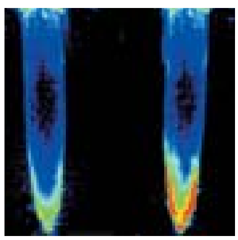
Edward Teller, 1908–2003

A YOUNG Edward Teller lectures.

When I interviewed Edward Teller in 1999, he was already suffering from myriad health problems, his memory impaired by a stroke, his vision clouded by ocular ulcerations [see "Infamy and Honor at the Atomic Café," Profile, SCIENTIFIC AMERICAN, October 1999]. I worried that he may have lapsed into a gerontological stupor. But after a few moments, the same voice that had made the case for thermonuclear weapons, Lawrence Livermore National Laboratory and the Star Wars missile defense emerged as strong and unmistakable as it had been to J. Robert Oppenheimer, Nelson Rockefeller and Ronald Reagan.

Teller is best known as the father of the hydrogen bomb. But his technological optimism—trying to teach the world, as Dr. Strangelove did, to love the bomb—combined with an unrelenting anti-Communism, occasioned by the experiences of his youth in Hungary to project him relentlessly into the eye of the maelstrom. Bad-mouthing Oppenheimer. Militating for bomb shelters to survive a fusion-induced holocaust. Hyping the x-ray laser. His style of hawkishness may have helped push the Soviet Union over the brink, but it also risked global thermonuclear annihilation. A whole generation could have done without duck-and-cover drills.

His death at age 95 after another stroke will give historians and journalists an opportunity to ponder Nobel physicist Isidor I. Rabi's famous comment that the world would have been a better place without Teller. Rabi's judgment was unquestionably harsh. And not to everyone's concurrence—certainly he had many admirers: George W. Bush awarded him the nation's highest civilian honor, the Presidential Medal of Freedom, earlier this year. —*Gary Stix*



VICTIMS of plant warfare soak up toxin (*red*) at their roots.

ECOLOGY

Killing the Competition

Invading species are commonly believed to succeed by outcompeting natives for vital resources. At least one notorious invader, however, wins out by poisoning the competition. The spotted knapweed is an intruder from eastern Europe that over the past century has displaced indigenous grasses and degraded pastures in North America. Investigators have found that the knapweed's roots exude a toxin that builds up in the soil. The toxin generates a wave of cell death and inhibits sprouting and growth. Plants in Europe seem to have limited the weed's spread after evolving resistance to the poison. These findings, in the September 5 *Science,* could help determine whether introduced plants, such as those created through genetic engineering, would overrun habitats. —*Charles Choi*

CHEMISTRY

Cleaner Living

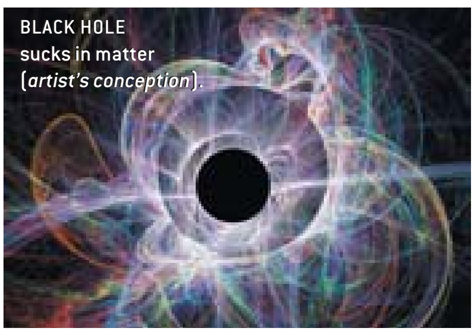
Best known for cleansing wounds and bleaching hair, hydrogen peroxide can be transformed into a supercleaner with a class of environmentally friendly catalysts called Fe-TAML activators. Each molecule of the catalyst consists of an iron atom surrounded by a ring molecule called a tetra-amido macrocyclic ligand. Fe-TAML binds to oxygen atoms in hydrogen peroxide, forming reactive intermediates that attack pollutants, converting them into harmless or less toxic substances. Investigators at Carnegie Mellon University, whose tinkering since 1980 led eventually to the Fe-TAML family, continue to refine catalyst lifetime, reactivity and selectivity with molecular attachments. They presented findings at the September American Chemical Society meeting that suggest that the catalysts can also scrub hard-to-remove sulfur compounds from fuel to prevent acid rain and improve its efficiency, as well as eliminate paper and textile dyes, which can cloud natural waterways. —*Charles Choi*

CORBIS (*top*); JORGE M. VIVANCO (*bottom*); ILLUSTRATION BY MATT COLLINS

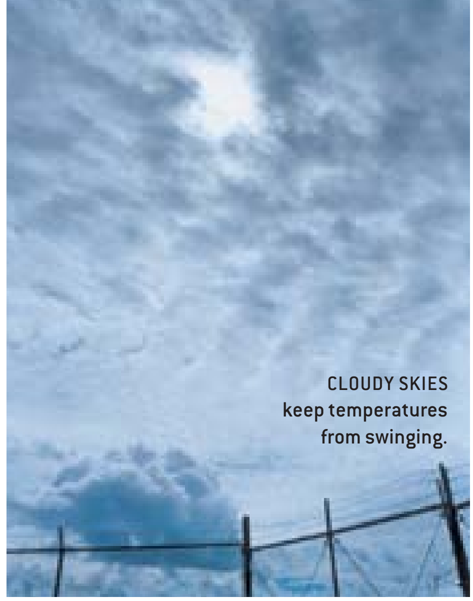
ASTROPHYSICS

Black Hole Life Preserver

Daredevils have risked trips in barrels over Niagara Falls since 1901, with 11 of 16 even surviving. Now scientists have figured out how to prolong the survival of anyone plummeting into a black hole. With a feet-first dive, your toes would experience a stronger pull than your head as your sides got crushed together. Such "spaghettification" would take just under 0.1 second, long enough for a pain signal to reach your brain. In a report submitted to *Physical Review D,* J. Richard Gott of Princeton University and Deborah L. Freedman of Harvard University suggest that the gravity exerted by a massive ring encircling your waist would counteract that of the black hole by pulling up on your feet and down on your head. This girdle would give you 0.09 second more life by cutting spaghettification time down by a factor of 26—"so fast you really wouldn't know what hit you," Gott explains. The life preserver's mass would have to be more than 12,800 trillion metric tons, roughly equal to an asteroid 100 miles wide. *—Charles Choi*



BLACK HOLE sucks in matter (*artist's conception*).



CLOUDY SKIES keep temperatures from swinging.

CLIMATE

Weekend Weather

Working for the weekend seems to affect daytime highs and nighttime lows. In the past few decades this diurnal temperature range has been narrowing. Now Piers M. de F. Forster and Susan Solomon of the National Ocean and Atmospheric Administration find that the daily swing follows a weekly pattern, based on 40 years of worldwide temperature data. Especially in urban settings in the U.S., Mexico, Japan and China, the diurnal range was a few tenths of a degree less from Wednesday through Friday than from Saturday through Monday. Because no natural phenomenon follows a seven-day cycle, the researchers suspect that human activity causes this "weekend effect." Specifically, soot and sulfate aerosols from motor vehicles and especially from coal-burning power plants may be affecting local cloud cover, which can dampen temperature swings. The study was published online September 18 by the *Proceedings of the National Academy of Sciences USA.* *—Philip Yam*

ENERGY

Bacterial Batteries

Trash and sewage are loaded with sugars that researchers have strived for decades to convert into fuels, such as ethanol, that can be burned to make electricity. Swades Chaudhuri and Derek R. Lovley of the University of Massachusetts at Amherst have cut out the middle step with an efficient way to turn those sugars directly into electricity. They used the sweet-loving microbe *Rhodoferax ferrireducens* dredged from marine mud. The bacterium strips the electrons off sugar molecules and transfers the negative charges to a graphite electrode, producing electricity for days with more than 80 percent energy conversion efficiency. Previously, microbial fuel cells showed at most 50 percent efficiency and required unstable components, rendering them unsuitable for long-term power generation. Improving the new battery's electrodes should increase power output, the researchers note in the October *Nature Biotechnology.* *—Charles Choi*

BRIEF BITS

- Conventionally chilled platelets die soon after transfusion, but a new refrigeration method could extend by more than a week the viability of transfused platelets.

  Science, September 12, 2003

- Disruptions of a gene called *DYX1C1* substantially raise the odds of a person becoming dyslexic. This gene, one of many thought to play a role in the learning disorder, could lead to more accurate diagnoses.

  Proceedings of the National Academy of Sciences USA online, September 3, 2003

- Mercury in fish may not be as harmful as thought, because the form of the metal in seafood (methylmercury cysteine) differs from that used in toxicology models (aqueous methylmercury chloride).

  Science, August 29, 2003

- In species where the female has multiple mates, the offspring tend not to get any fatherly care. Male savanna baboons, however, seem to look out for their progeny while intervening in squabbles between juveniles, favoring offspring of females with whom they frequently consorted and those who showed physical similarities.

  Nature, September 11, 2003

BRAD WILSON *Photonica* (*top*); MEHAU KULYK *Photo Researchers, Inc.* (*bottom*)

Baffling the Bots

Anti-spammers take on automatons posing as humans By LEE BRUNO

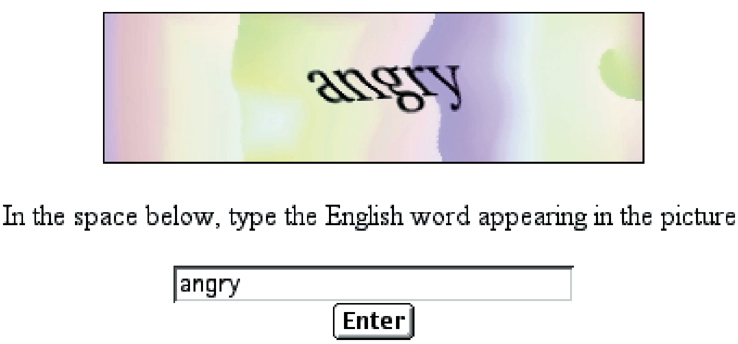
Three years ago rogue computer software programs called bots posed as teenagers in Yahoo's chat rooms on the Web. There they created mischief by collecting personal information about the teens who visited or by pointing chat participants to advertisements. The bots operated by waiting until a visitor typed a question mark. They would then automatically create a response about where a person could find an answer and provide a URL that would deliver the visitor to an advertising site.

Bots are well known for helping to generate millions of spam messages advertising printer cartridges, septic systems, Viagra and Nigerian money scams. They disseminate junk information by opening up new e-mail accounts and then automatically delivering a flood of messages. During 2001 estimates of the volume of spam reached more than six times that of a year earlier. And last year the volume was 21 times greater than in 2000, according to the Coalition against Unsolicited Bulk Email, an Australia-based organization.

E-mail filters are still rudimentary cures and pretty ineffective in curtailing the deluge of unwanted messages. After the bot incursion, Yahoo's technical staff realized that it needed to create a software gatekeeper that would allow human users in and keep automatons out. Udi Manber, Yahoo's chief scientist, went looking for help. He offered a challenge to Manuel Blum and his graduate students at the School of Computer Science at Carnegie Mellon University. Blum had an interest in investigating whether image-degradation models, which distort some part of a word or image, could be used to build a computer Turing test (named after the brilliant mathematician and a founding figure of computing Alan Turing). In 1950 Turing proposed a behavioral approach to determine whether a system could "think": a machine would pass the test if human interrogators could not tell whether replies to a series of typed questions they were asking were coming from a computer or a human.

In the course of his research, Blum came into contact with Henry Baird, a renowned figure in the computer-vision field. Baird had become familiar with the limits of computer vision from his years of work on building and analyzing systems at Lucent Technologies's Bell Labs, where he developed new software algorithms for document imaging. In 1998 he left the quiet Murray Hill, N.J., campus of Bell Labs to join another fabled institution: Xerox PARC in Palo Alto, Calif. There the armies of smart Internet bots roaming the Web to harvest information became an intellectual obsession for him.

During the fall of 2000 Baird conducted a trial at the University of California at Berkeley. The resulting paper dealt with a new image-degradation model named Pessimal Print. Concurrently, Yahoo and Blum and his team at Carnegie Mellon were working on a similar model, one version of which is called EZ-Gimpy. It is a kind of reverse Turing test, which has come to be known as a CAPTCHA, or "completely automated public Turing test to tell computers and humans apart."



READ THIS: A type of CAPTCHA, or image-degradation model, known as EZ-Gimpy tries to outwit computer bots with distorted letters and busy backgrounds. A human user easily recognizes the word and types it in the blank, allowing entry to a Web area.

CARNEGIE MELLON UNIVERSITY

These Turing tests for Internet bots are a cognitive puzzle that can be solved by humans but not by computers. “Humans are very good at reading very strange stuff,” says Baird, whose formal title is principal scientist and area manager of statistical pattern and image analysis at PARC (no longer Xerox PARC).

As an example, EZ-Gimpy selects a word from an 850-word dictionary and then disfigures the letters by warping the font or leaving gaps in the letters and placing them on a busy background. In doing so, the CAPTCHA presents a human verification test to the person trying to obtain a free e-mail account or entrance to a chat room. EZ-Gimpy quickly went to work at Yahoo. And other Internet mail services, such as Microsoft’s Hotmail, also use CAPTCHAs, based on EZ-Gimpy.

BAFFLETEXT: This latest generation of CAPTCHA, designed to fool particularly clever bots, employs nonsense words and type-obscuring tricks.

HENRY S. BAIRD *PARC*

EZ-Gimpy has worked well, but next-generation bots are getting wise to it. They are getting better at recognizing the distorted words contained in the dictionary. But Baird, along with Monica Chew of Berkeley, co-developed BaffleText, a new CAPTCHA scheme that goes beyond the 850-word dictionary of EZ-Gimpy. It randomly generates a few degraded words each time a person logs onto a Web site to establish an e-mail account or other service. The person has to recognize the word and type it into the blank space on the page in order to progress to the next stage.

Two principal ideas guided the researchers in their quest to create a stronger deterrent for bots. BaffleText incorporates nonsense words to overcome the problem of a small dictionary. Also, it leverages Gestalt psychology, or a human’s innate ability to infer the whole picture of an image from only partial information (something machines can’t do). For example, BaffleText uses non-English character strings like “inchem” and “scotter” to defend against dictionary-driven attacks. What’s more, its Gestalt-inspired images of words masked or degraded in appearance make it nearly impossible for a bot to decipher. Simply put, to crack BaffleText, bot programmers must solve perplexing computer-vision and pattern-recognition problems that have eluded them for decades.

To test the CAPTCHAs, other researchers from Berkeley and Carnegie Mellon are laboring to break them. And whereas the bulk of work done to date has taken place on text-based CAPTCHAs, research is under way on developing auditory and visual CAPTCHAs. All the while, the artificial-intelligence community views the challenge of trying to break CAPTCHAs as a kind of mind sport.

Baird continues to build, test and crack bots. “This is our arms race,” he says. “There’s no question that bots are going to become more and more sophisticated.” CAPTCHAs are expected to become important to businesses in protecting their networks from smart bot intruders. In effect, they have become new electronic guardians for Web services, helping to immunize and prevent attacks from increasingly smarter bots written by people intent on abusing the services for their own gain. Meanwhile programmers are expected to unleash fleets of bots bent on breaking CAPTCHAs, thus promulgating a game of one-upmanship. That is why, for the artificial-intelligence community, building ever more powerful CAPTCHAs has provoked the same excitement once elicited by the creation of ever more sophisticated chess programs. And this work should ultimately yield a more cogent answer to the question of whether it is a human or a machine knocking at the virtual door. SA

Lee Bruno is an editor at Red Herring, *an online magazine that covers business and technology.*

Shrink-Wrapping the World

A law that would crimp the rights of software buyers suffers a major defeat By GARY STIX

"You accept the terms of this agreement." In the eyes of a software vendor, the simple act of removing the plastic shrink-wrap from a software package is tantamount to signing a contract with the manufacturer that severely limits consumer rights. If buyers would read those licenses carefully, they might have second thoughts. That is, if they could. Most of the time the contract is buried in the box.

Until recently, though, things looked like they were about to get markedly worse for software buyers. The Uniform Computer Information Transactions Act (UCITA) was crafted as the first attempt to standardize nationally the commercial licensing of software and other information products. If all had gone according to plan, UCITA would ultimately have been adopted by every state legislature. But the proposed law contained provisions that critics perceived could have been imagined by George Orwell.

The original law would have let vendors turn off software remotely for breach of a license. Adversaries feared that the press would not have been able to review a software package without the publisher's approval and that reverse engineering to address bugs, security breaches and communications issues could have been prohibited. A later version of the law tried to deal with some of these concerns. Unchanged, however, was a stipulation that a vendor could alter the terms of a license at any time by sending an e-mail or by posting changes on a Web site. And, most important, foes argued that UCITA would let software providers run roughshod over current copyright law, which sets out certain rights for purchasers of a creative work.

The draconian nature of UCITA brought together a broad coalition of opponents, ranging from librarians and consumer groups to the insurance industry. The biggest blow to the law—perhaps a fatal one—came in early August. The National Conference of Commissioners on Uniform State Laws, UCITA's sponsoring organization, bowed to concerted opposition and decided at its annual meeting to drop its push to have state legislatures pass the law, an action that may undercut any further consideration by the states.

Critics assert that the UCITA battle has not reached closure, despite tremendous progress in the campaign. Since UCITA was released four years ago, only two states, Virginia and Maryland, have adopted the law. The National Conference of Commissioners on Uniform State Laws failed to gain adoption of the act in any state during the most recent legislative sessions. But UCITA's influence on information technology licensing may live on. Software companies could choose Virginia's or Maryland's as the state law that governs a particular software contract and attempt to make it binding throughout most of the U.S.—or vendors might simply use parts of UCITA as a model for how they draft licenses. "A lot of people say it's dead, but we'd rather say it's dormant," says Carol Ashworth, coordinator for the Americans for Fair Electronic Commerce Transactions (AFFECT), an umbrella group of opponents.

AFFECT will continue to push for "bomb shelter" laws, like those already enacted in Iowa, North Carolina, Vermont and West Virginia, that prevent software vendors from applying UCITA provisions in a given state. The defeat of the legislation marks part of a larger trend. Consumers and scholars have succeeded recently in expanding the dialogue on otherwise esoteric intellectual-property issues such as the patenting of basic biomedical research and fair use of digital content. Now at least the public has a chance to hear both sides of these critical debates. SA

JENNIFER KANE

Candle in the Dark

Instead of cursing the darkness of pseudoscience on television, light a candle with Cable Science Network   By MICHAEL SHERMER

Ever since Galileo began the tradition of communicating science in the vernacular so that all might share in its fruits, a tension has existed between those—call them "excluders"—who think science is for professionals only and regard its dissemination to wider audiences as infra dig and those—call them "includers"—who understand that all levels of science require clear composition and public understanding of process and product.

Throughout much of the 20th century the excluders have ruled the roost, punishing those in their flock who dared to write for those paying the bills. Cornell University astronomer Carl Sagan, for example, whose PBS television series *Cosmos* was viewed by more than half a billion people, was denied membership in the National Academy of Sciences primarily (his biographers have demonstrated through interviews with insiders) because he invested too much time in science popularization.

Over the past two decades, however, a literary genre has arisen in which professional scientists are presenting original research and theories in books written for both their colleagues and the public. Most of Stephen Jay Gould's works are in this mode, as are those of Edward O. Wilson, Ernst Mayr, Jared Diamond, Richard Dawkins, Steven Pinker and others. In fact, if you want to be considered a cultured person in today's society, it is not enough to be steeped in literature, art and music. You need to know something about science.

The problem is that most people do not get their science through books and PBS documentary series. Although science junkies can fill their trough with such outstanding series as PBS's *Nova* and *Scientific American Frontiers,* most folks pick up bits and pieces from short newspaper articles or evening news sound bites, which typically alternate between scary medical findings and stunning Hubble Space Telescope images, leaving out the subtleties of how science is really done and why contradictory findings do not mean that the process has failed. Worse still, most networks pander to the ratings game and air a mélange of pseudoscience about ESP, UFOs and moon landing hoaxes.

Like most scientists, I complain bitterly and often about such dismal programming. We write letters to network executives, but to no avail. One solution is to create our own network. Thus, Cable Science Network, or CSN, is in the offing. Roger Bingham of the Center for Brain and Cognition at the University of California at San Diego is spearheading a movement (of which I am a part, along with Sagan's widow, Ann Druyan, and Salk Institute neuroscientist Terry Sejnowski) to launch a nonprofit organization modeled on the ubiquitous C-SPAN (Cable Satellite Public Affairs Network), now available in more than 85 million homes. CSN would be science 24/7—all science, all the time—freeing us, in Bingham's words, from "the tyranny of the sound bite."

Cable Science Network would be science 24/7—all science, all the time—freeing us from "the tyranny of the sound bite."

Wouldn't it be great to watch congressional hearings on cloning, bioterrorism, global warming and aging? Wouldn't it be fabulous to attend—via cable—cutting-edge lectures given by scientists at various annual scientific conferences? Every year tens of thousands of neuroscientists, for example, converge to exchange data on how the brain works. Wouldn't you love to sit in on some of those presentations rather than waiting to hear about one of them in a 30-second encapsulation on network TV? Science luminaries who today may have an audience of a couple hundred people in a university lecture hall could instead reach a couple hundred *thousand*.

With CSN, all this will bring science to the people—and to scientists, legislators, teachers and students—as never before. Sagan called science "a candle in the dark." CSN is still in the developmental stage (see www.csntv.org), but if we can switch it on, it will be a candle whose light will illuminate a path toward the globalization of science. SA

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

BRAD HINES

TGGGATAGCGACGAGCCAGTCTGCTCTAGACAGACGTAGCATATGGGATAGCGACAGACAGACGTAGCATATGGGAG

FLECKS OF DARK BROWN in an iris may be a telltale sign of the hidden genome at work. Certain traits are transmitted not through ordinary genes but rather through chemical modifications to the chromosomes, changes that are regulated in part by bits of "junk" DNA. Unlike genetic mutations, these heritable traits are often reversible and appear in some cells but not others. (The white sphere on the iris is a reflection of the light shining on the eye.)

Just when scientists thought they had DNA almost figured out, they are discovering in chromosomes two vast, but largely hidden, layers of information that affect inheritance, development and disease

The Unseen Genome: Gems among the Junk

BY W. WAYT GIBBS

JAMIE KRIPKE

TGGGATAGCGACGAGCCAGTCTGCTCTAGACAGACGTAGCATATGGGATAGCGACGAGCCAGTCTGCTCTAGACAGT

About 20 years ago astronomers became convinced

that distant galaxies were moving in ways that made no sense, given the laws of gravity and the fabric of celestial objects visible in the sky. Gradually they were forced to conclude that the universe is not as empty as it appears, that in fact it must be dominated by some dark kind of matter. Although no one knew what the stuff is made of or how it works, scientists could see from its effects that it is out there. The quest to understand dark matter (and more recently, dark energy) meant revising or replacing theories, but it reenergized astrophysics and cosmology.

A similar revelation is now unfolding in molecular genetics. This year biologists celebrated the 50th anniversary of the discovery of the double helix, and the Human Genome Project announced its completion of a "final draft" of the DNA sequence for *Homo sapiens*. Scientists have clearly mastered DNA in the lab. Yet as they compare the DNA of distantly related species and look more closely at how chromosomes function in living cells, they are increasingly noticing effects that current theories cannot explain.

Journals and conferences have been buzzing with new evidence that contradicts conventional notions that genes, those sections of DNA that encode proteins, are the sole mainspring of heredity and the complete blueprint for all life. Much as dark matter influences the fate of galaxies, dark parts of the genome exert control over the development and the distinctive traits of all organisms, from bacteria to humans. The genome is home to many more actors than just the protein-coding genes.

Overview/*Hidden Genes*

- Geneticists have long focused on just the small part of DNA that contains blueprints for proteins. The remainder—in humans, 98 percent of the DNA—was often dismissed as junk. But the discovery of many hidden genes that work through RNA, rather than protein, has overturned that assumption.
- These RNA-only genes tend to be short and difficult to identify. But some of them play major roles in the health and development of plants and animals.
- Active forms of RNA also help to regulate a separate "epigenetic" layer of heritable information that resides in the chromosomes but outside the DNA sequence.

The extent of this unseen genome is not yet clear, but at least two layers of information exist outside the traditionally recognized genes. One layer is woven throughout the vast "noncoding" sequences of DNA that interrupt and separate genes. Though long ago written off as irrelevant because they yield no proteins, many of these sections have been preserved mostly intact through millions of years of evolution. That suggests they do something indispensable. And indeed a large number are transcribed into varieties of RNA that perform a much wider range of functions than biologists had imagined possible. Some scientists now suspect that much of what makes one person, and one species, different from the next are variations in the gems hidden within our "junk" DNA.

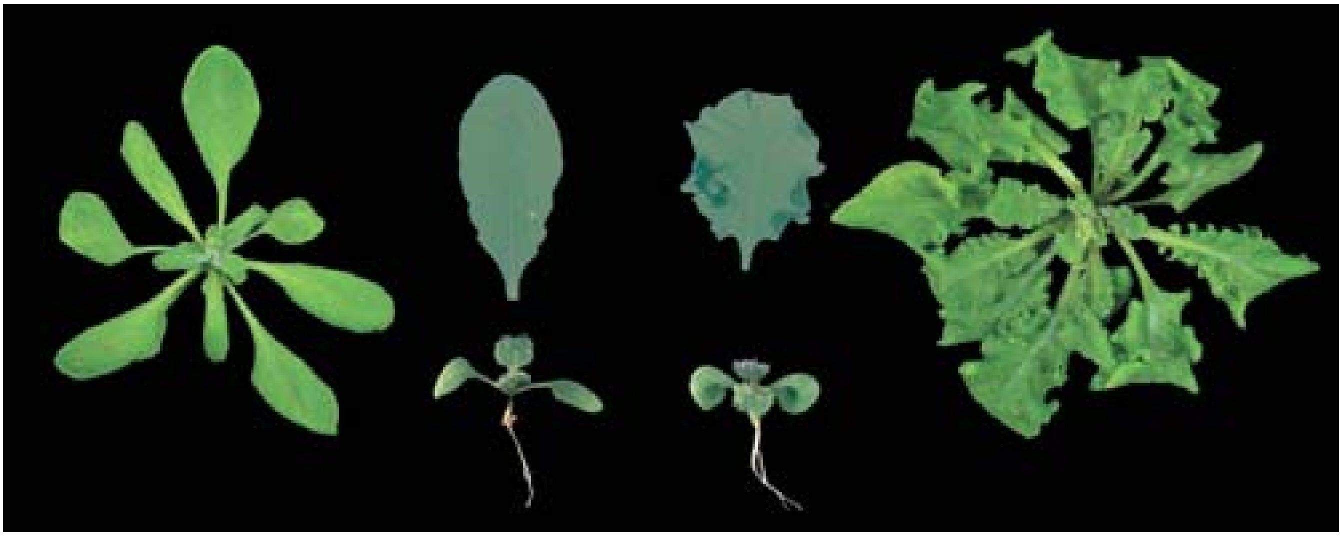
Above and beyond the DNA sequence there is another, much more malleable, layer of information in the chromosomes. "Epigenetic" marks, embedded in a mélange of proteins and chemicals that surround, support and stick to DNA, operate through cryptic codes and mysterious machinery. Unlike genes, epigenetic marks are routinely laid down, erased and rewritten on the fly. So whereas mutations last a lifetime, epigenetic mistakes—implicated in a growing list of birth defects, cancers and other diseases—may be reversible with drugs. In fact, doctors are already testing such experimental treatments on leukemia patients.

Researchers are also coming to realize that just about anything that can happen in the genome does happen, says Carmen Sapienza of Temple University, who started investigating epigenetic phenomena back when they were dismissed as minor anomalies. "There may even be fundamental mechanisms still to discover," Sapienza considers. "I think we are entering the most interesting time yet in genetics."

The Perils of Dogma

IT WILL TAKE YEARS, perhaps decades, to construct a detailed theory that explains how DNA, RNA and the epigenetic machinery all fit into an interlocking, self-regulating system. But there is no longer any doubt that a new theory is needed to replace the central dogma that has been the foundation of molecular genetics and biotechnology since the 1950s.

The central dogma, as usually stated, is quite simple: DNA makes RNA, RNA makes protein, and proteins do almost all the real work of biology. The idea is that information is stored in the twisted ladders of DNA, specifically in the chemical bases



BIG DIFFERENCES in the appearance and health of organisms can arise from small changes to tiny, unconventional genes. *Arabidopsis* plants, for example, normally have spoon-shaped leaves (*left*). But when scientists interfered with the action of a microRNA, produced by an RNA-only gene, the mutant *arabidopsis* plants developed gross defects (*right*). The microRNA appears to control the activity levels of numerous genes.

(commonly labeled A, T, G and C) that pair up to form the rungs of the ladders. A gene is just a particular sequence of bases on one side of the ladder that specifies a protein.

The dogma holds that genes express themselves as proteins, which are made in four steps: First an enzyme docks to the chromosome and slides along the gene, transcribing the sequence on one strand of DNA into a single strand of RNA. Next, any introns—noncoding parts of the initial RNA transcript—are snipped out, and the rest is spliced together to make a piece of messenger RNA. The RNA message then moves out of the nucleus to the main part of the cell, where molecular machines translate it into chains of amino acids. Finally, each chain twists and folds into an intricate three-dimensional shape.

It is their shapes that make proteins so remarkably versatile. Some form muscles and organs; others work as enzymes to catalyze, metabolize or signal; and still others regulate genes by docking to specific sections of DNA or RNA. No great wonder, then, that many biologists (and journalists) have taken the central dogma to imply that, with very few exceptions, a DNA sequence qualifies as a gene only if it can produce a protein.

"Typically when people say that the human genome contains 27,000 genes or so, they are referring to genes that code for proteins," points out Michel Georges, a geneticist at the University of Liège in Belgium. But even though that number is still tentative—estimates range from 20,000 to 40,000—it seems to confirm that there is no clear correspondence between the complexity of a species and the number of genes in its genome. "Fruit flies have fewer coding genes than roundworms, and rice plants have more than humans," notes John S. Mattick, director of the Institute for Molecular Bioscience at the University of Queensland in Brisbane, Australia. "The amount of noncoding DNA, however, does seem to scale with complexity."

In higher organisms (such as humans), genes "are fragmented into chunks of protein-coding sequences separated by often extensive tracts of nonprotein-coding sequences," Mattick explains. In fact, protein-coding chunks account for less than 2 percent of the DNA in human chromosomes. Three billion or so pairs of bases that we all carry in nearly every cell are there for some other reason. Yet the introns within genes and the long stretches of intergenic DNA between genes, Mattick says, "were immediately assumed to be evolutionary junk."

That assumption was too hasty. "Increasingly we are realizing that there is a large collection of 'genes' that are clearly functional even though they do not code for any protein" but produce only RNA, Georges remarks. The term "gene" has always been somewhat loosely defined; these RNA-only genes muddle its meaning further. To avoid confusion, says Claes Wahlestedt of the Karolinska Institute in Sweden, "we tend not to talk about 'genes' anymore; we just refer to any segment that is transcribed [to RNA] as a 'transcriptional unit.'"

Based on detailed scans of the mouse genome for all such elements, "we estimate that there will be 70,000 to 100,000," Wahlestedt announced at the International Congress of Genetics, held this past July in Melbourne. "Easily half of these could be noncoding." If that is right, then for every DNA sequence that generates a protein, another works solely through active forms of RNA—forms that are not simply intermediate blueprints for proteins but, rather, directly alter the behavior of cells.

What is true for mice is probably true for people and other animals as well. A team of scientists at the National Human Genome Research Institute (NHGRI) recently compared excerpts from the genomes of humans, cows, dogs, pigs, rats and seven other species. Their computer analysis turned up 1,194 segments that appear with only minor changes in several species, a strong indication that the sequences contribute to the species' evolutionary fitness. To the researchers' surprise, only 244 of the segments sit inside a protein-coding stretch of DNA. About two thirds of the conserved sequences lie in introns, and the rest are scattered among the intergenic "junk" DNA.

"I think this will come to be a classic story of orthodoxy derailing objective analysis of the facts, in this case for a quarter of a century," Mattick says. "The failure to recognize the full implications of this—particularly the possibility that the intervening noncoding sequences may be transmitting parallel information in the form of RNA molecules—may well go down as one of the biggest mistakes in the history of molecular biology."

More Than a Messenger

NOW THAT BIOLOGISTS have turned their attention back to RNA, they are finding it to be capable of impressive feats of cellular chemistry. Like proteins, some RNA transcripts can interact with other bits of RNA, with DNA, with proteins and even with small chemical compounds. Proteins are analog molecules, however; they bind to targets in much the way keys fit in locks. "The beauty of RNA is that it has a specific sequence, so it's digital, like a zip code," Mattick points out. A bit of RNA can float around until it bumps into a DNA (or another RNA) that has a complementary sequence; the two halves of the ladder then join rungs. (Two segments are complementary when all C bases mate with G's and all T or U bases join to A's.)

As an example of the unappreciated power of RNA, consider pseudogenes. Surveys of human DNA have found in it almost equal numbers of genes and pseudogenes—defective copies of functional genes. For decades, pseudogenes have been written off as molecular fossils, the remains of genes that were broken by mutation and abandoned by evolution. But this past May a group of Japanese geneticists led by Shinji Hirotsune of the Saitama Medical School reported their discovery of the first functional pseudogene.

Hirotsune was genetically engineering mice to carry a fruit fly gene called *sex-lethal*. Most mice did fine with this foreign gene, but in one strain *sex-lethal* lived up to its name; all the mice died in infancy. Looking closer, the scientists discovered that in those mice *sex-lethal* happened to get inserted right into the middle of a pseudogene, clobbering it. This pseudogene (named *makorin1-p1*) is a greatly shortened copy of *makorin1*, an ancient gene that mice share with fruit flies, worms and many other species. Although researchers don't know what *makorin1* does, they do know that mice have lots of *makorin1* pseudogenes and that none of them can make proteins. But if pseudogenes do nothing, why were these mice dying when they lost one?

For some reason, *makorin1*—and apparently only *makorin1*—all but shuts down when its pseudogene *p1* is knocked out. RNA made from the pseudogene, in other words, controls the expression of the "real" gene whose sequence it mimics, even though the two lie on different chromosomes. There is nothing pseudo about that.

It is too early to say whether many pseudogenes give rise to active RNA. But there are plenty of other sources scattered about the dark parts of the genome. Every normal protein-making gene, for instance, has a complementary DNA sequence that sits on the other side of the ladder and usually is not transcribed into RNA. Biologists like to think of this as a backup copy, because the cell can use it to repair damage to the gene.

In some cases, however, the backup has its own agenda. While the gene is producing a sensible RNA message, its alter ego can churn out an "antisense" RNA that has a complementary sequence. Whenever matched sense and antisense RNAs meet, they mesh to form their own double-stranded ladders—effectively interfering with the gene's ability to express its protein.

Biologists knew that bacteria and plants can produce antisense, but most thought that mammals rarely do. In April, Galit Rotman and her co-workers at CompuGen, a biotech firm in Tel Aviv, dashed that assumption. They screened human genome databases and concluded that at least 1,600 human genes (and probably many more) have a mate that yields antisense RNAs.

These competing RNAs may suppress a gene just by tying up the gene's messenger RNA. But Rotman speculates that they employ a built-in genome censor, known as the RNA interference machinery. Scientists are still enthralled by the discovery several years ago of this scheme for selectively silencing individual genes. When double-stranded RNA appears in a cell, enzymes dice it up, peel the two strands apart, and use one RNA fragment to seek out and destroy any other RNA messages that stick to its sequence. The system protects cells against viruses, which often deliver their payloads in the form of double-stranded RNA. But the censor also provides a handy way for scientists to shut off any gene at will [see "Censors of the Genome," by Nelson C. Lau and David P. Bartel; SCIENTIFIC AMERICAN, August].

Neither pseudogenes nor antisense RNAs, however, can explain the crinkled leaves that Detlef Weigel of the Max Planck Institute for Developmental Biology in Tübingen, Germany, and his collaborators saw in their *arabidopsis* plants this summer. These weeds of the mustard family normally have smooth, spoon-shaped leaves. The plants owe their gentle symmetrical curves, Weigel's group showed in *Nature* this past August, in part to a kind of active RNA called microRNA.

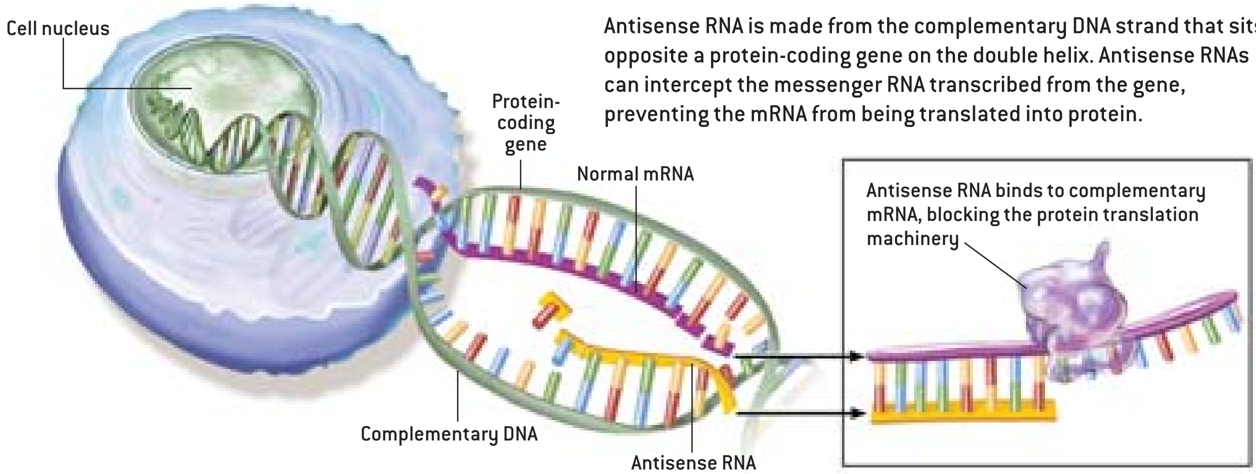
MicroRNAs, first observed a few years ago in roundworms, are short noncoding RNAs that fold back on themselves, like hairpins. In *arabidopsis*, the JAW microRNA doubles over and is then captured by the RNA interference machinery, just as if it had come out of a virus. But the JAW sequence matches a handful of different protein-making genes, members of a family that control the shape and size of the plant. The censor dutifully represses each of them by chopping up much (but not all) of the messenger RNA they produce. Thus, JAW, a tiny RNA-only gene, serves as the main lever by which *arabidopsis* cells adjust the volume of a suite of crucial protein genes. When Weigel's

The failure to recognize the importance of introns "may well go down as one of the biggest mistakes in the history of molecular biology."

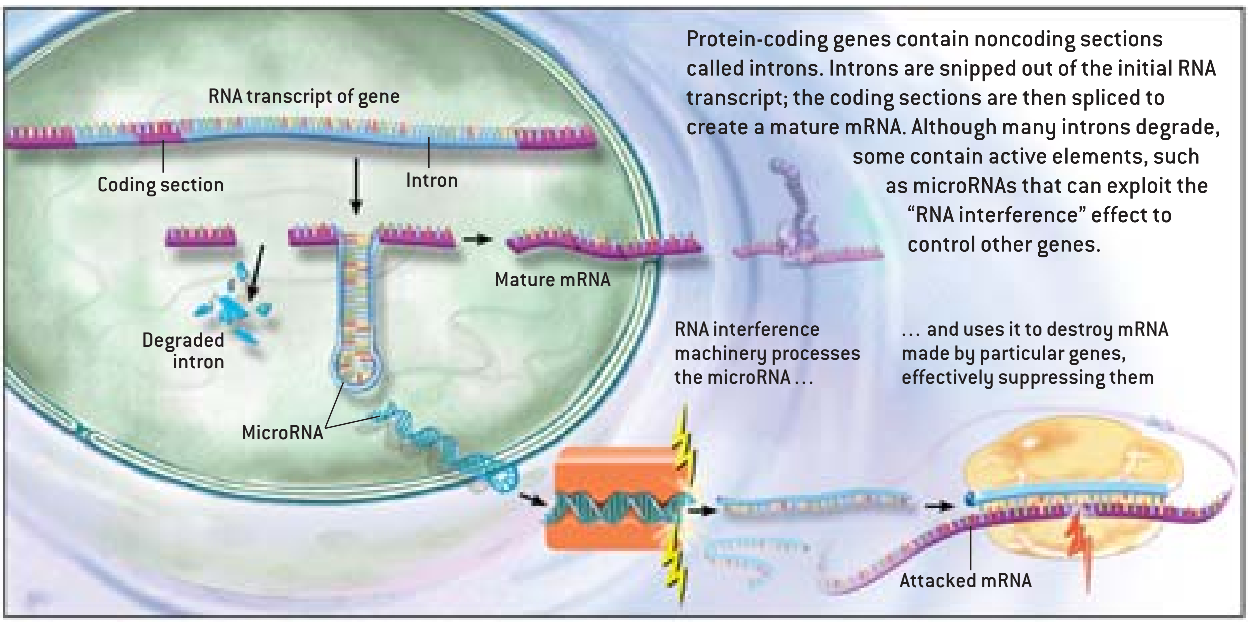
A BESTIARY OF UNCONVENTIONAL GENES

GENES, according to conventional wisdom, are those sections of the DNA that encode functional proteins. Such sequences make up only about 2 percent of the human genome, however. The rest of the human genome is filled with DNA that is "noncoding"—but not useless. Scientists are discovering many noncoding genes that give rise to surprisingly active RNAs, including varieties that can silence or regulate conventional genes.

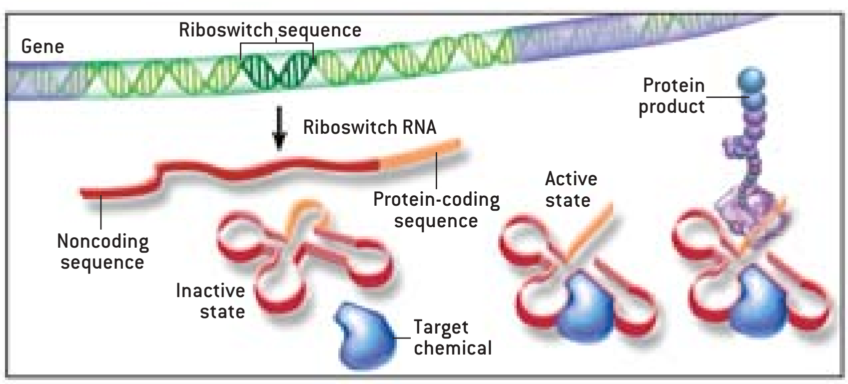
Antisense RNA is made from the complementary DNA strand that sits opposite a protein-coding gene on the double helix. Antisense RNAs can intercept the messenger RNA transcribed from the gene, preventing the mRNA from being translated into protein.



Protein-coding genes contain noncoding sections called introns. Introns are snipped out of the initial RNA transcript; the coding sections are then spliced to create a mature mRNA. Although many introns degrade, some contain active elements, such as microRNAs that can exploit the "RNA interference" effect to control other genes.



Riboswitches are a newly discovered form of RNA that act as precision genetic switches. Produced in many cases from noncoding DNA between known genes, a riboswitch folds into a complex shape. One part of the folded RNA can bind to a specific target protein or chemical. Another part contains the RNA code for a protein product. The riboswitch turns "on" and produces the protein it encodes only when in the presence of its target.



TERESE WINSLOW

group engineered plants in which the microRNA could not do its job, the plants became sick and deformed.

In just the three years since researchers started looking in earnest, they have found hundreds of microRNAs—more than 150 in humans alone. They seem to be a well-established means for organisms to wrangle genes; about half the microRNAs in humans also appear, in nearly identical form, in the DNA of pufferfish, even though the two species went their separate ways some 400 million years ago.

Just what those 150-plus microRNAs do in people is a mystery. Anna M. Krichevsky of Harvard Medical School suspects that, among other things, they play an important role in brain development. Her lab used a "gene chip" to screen mouse neurons for 44 different kinds of microRNA. Krichevsky reported in September that levels of at least nine distinct microRNAs are precisely regulated in the mice as their brains grow. The link is still indirect, but as Diya Banerjee of Yale University noted last year in a review of microRNA science, "it seems that we are on the verge of an explosion of knowledge in this area."

Digital *and* Analog

PROTEINS MAY BE the draft horses of the cell, but active RNA sometimes wields the whip. And several kinds of RNA have turned up doing mules' work as well: catalyzing, signaling and switching as competently as any protein. In fact, some inherited diseases have stumped researchers because, in their diligent search for a mutant protein, the investigators ignored the active RNA right under their noses.

> "What was damned as junk because it was not understood may, in fact, turn out to be the very basis of human complexity."

Doctors struggled for more than nine years, for example, to nail down the gene responsible for cartilage hair hypoplasia. This recessive disease was first identified in the Amish, one in 19 of whom carries a copy of the defective gene, which causes an unusual kind of dwarfism. People with CHH are not only small in stature but also at high risk for cancer and immune disorders. Geneticist Maaret Ridanpää of the University of Helsinki tracked the gene to chromosome nine, sequenced a large region and then proceeded to check all 10 protein-making genes in the area, one by one. None caused the disease.

Finally, in 2001, Ridanpää and his co-workers identified the culprit, an RNA-only gene called *RMRP*. The RNA transcribed from *RMRP* links up with proteins to form an enzyme that works inside a cell's energy generators, the mitochondria. A change to just a single base at a critical spot on this RNA can mean the difference between a full-size, healthy life and a short, abbreviated one (if the same mutation is inherited from both parents). Such "analog" RNAs, which fold up into complex shapes just as proteins do, have also been discovered recently to be essential to the function of enzymes that protect the chromosomes and that escort secreted protein signals out of cells' portholes.

Perhaps the most intriguing form of RNA yet discovered is the riboswitch, isolated last year by Ronald R. Breaker's lab at Yale. He and others have long wondered how, billions of years ago, the very earliest chemical precursors to life got along in the RNA world before DNA and proteins existed. They speculated that such proto-organisms would need to use RNA as sensors and switches to respond to changes in the environment and in their metabolism. To test the idea, they tried to create RNAs with such capabilities.

"Our laboratory successfully produced a number of synthetic RNA switches," Breaker recalls. Dubbed riboswitches, these long RNAs are both coding and noncoding at once. As the

Moving Genetics Forward

EVER SINCE THE INVENTION of recombinant DNA technology made genetic engineering feasible, most research in genetics has been run in "reverse." Reverse genetics begins with a particular gene of interest. The scientist fiddles with that gene in a cell culture or a living organism, watches what happens, and then tries to deduce the gene's function. It is a classic reductionist approach, and it can be very powerful.

But the gradual realization that the genome includes hidden genes—functional sequences that were misclassified as junk—highlights a major problem with reverse genetics: it can lead to tunnel vision. So recently a number of geneticists have been returning to the older practice of "forward" genetics as a way to identify the genes, both conventional and unconventional, that they don't know about.

Phenomix, a biotechnology company in La Jolla, Calif., founded last year by several prominent genetics teams, hopes to make a business out of the approach. The firm has set up a kind of production line for making mutant mice. In each group of mice, mutations to random points in the genome disable not just standard protein-coding genes but also hidden genes that make only active forms of RNA.

Phenomix starts with both healthy mice and mice that have diseases analogous to common human illnesses, such as diabetes, asthma, arthritis and Parkinson's disease. Some mutations induce or alleviate symptoms of these disorders in the mice. Researchers then do genetic screening to determine which mutations accounted for the effects. Whether the approach will inspire better drug designs remains to be seen. But forward genetics has already unearthed genetic phenomena, such as a functional pseudogene (*see main text*), that no one knew were possible. —*W.W.G.*

CLONES IN ALL BUT NAME, these littermates from a highly inbred strain of mice share practically identical DNA. Yet their coat colors run the spectrum from golden yellow to mahogany brown because of variations in the "epigenetic" chemical attachments each has to a particular segment of DNA that lies outside any known gene. The hair color of these mice cannot be predicted by current theories of genetics.

HUGH MORGAN

RNA folds up, the noncoding end becomes a sensitive receptor for a particular chemical target. A collision with the target flips the switch, causing the other end, which contains a standard blueprint for a protein, to change shape. The riboswitch thus gives rise to a protein, much like a normal gene does—but only when it senses its target.

Breaker's group started hunting for riboswitches in the wild and soon found them hiding in intergenic DNA. These precision genetic switches have been extracted now from species in all three kingdoms of life. "This implies that they were probably present in the last common ancestor," not long after the dawn of evolution, Breaker argues.

In August, Breaker and his co-workers reported that one family of riboswitches regulates the expression of no fewer than 26 genes in *Bacillus subtilis,* a common kitchen bacterium. These are not once-in-a-blue-moon genes, either, but genes that the bacterium relies on to metabolize such basic staples as sulfur and amino acids. Breaker estimates that *B. subtilis* has at least 68 genes, nearly 2 percent of its total, under the control of riboswitches. His lab has already begun engineering the hybrid digital-analog molecules to do useful things, such as selectively kill germs.

The Big Picture

AS BIOLOGISTS SIFT more and more novel kinds of active RNA genes out of the long-neglected introns and intergenic stretches of DNA, they are realizing that science is still far from having a complete parts list for humans or any other higher species. Unlike protein-making genes, which have standard "start" and "stop" codes, RNA-only genes vary so much that computer programs cannot reliably pick them out of DNA sequences. To spur the technology on, the NHGRI is launching this autumn an ambitious $36-million project to produce an "Encyclopedia of DNA Elements." The goal is to catalogue every kind of RNA and protein made from a select 1 percent of the human genome—in three years.

No one knows yet just what the big picture of genetics will look like once this hidden layer of information is made visible. "Indeed, what was damned as junk because it was not understood may, in fact, turn out to be the very basis of human complexity," Mattick suggests. Pseudogenes, riboswitches and all the rest aside, there is a good reason to suspect that is true. Active RNA, it is now coming out, helps to control the large-scale structure of the chromosomes and some crucial chemical modifications to them—an entirely different, epigenetic layer of information in the genome.

The exploration of that epigenetic layer is answering old conundrums: How do human beings survive with a genome horribly cluttered by seemingly useless, parasitic bits of DNA? Why is it so hard to clone an adult animal yet so easy to clone an embryo? Why do certain traits skip generations in an apparently unpredictable way? Next month the conclusion to this article will report on the latest discoveries about how the chromosomal layer of epigenetic phenomena works and on the initial attempts to exploit epigenetics in medicine and biotechnology. SA

W. Wayt Gibbs is senior writer.

MORE TO EXPLORE

Non-Coding RNA Genes and the Modern RNA World. Sean R. Eddy in *Nature Reviews Genetics,* Vol. 2, pages 919–929; December 2001.

An Expanding Universe of Noncoding RNAs. Gisela Storz in *Science,* Vol. 296, pages 1260–1263; May 17, 2002.

Widespread Occurrence of Antisense Transcription in the Human Genome. Rodrigo Yelin et al. in *Nature Biotechnology,* Vol. 21, pages 379–385; April 2003.

Challenging the Dogma: The Hidden Layer of Non-Protein-Coding RNAs in Complex Organisms. John S. Mattick in *BioEssays,* Vol. 25, pages 930–939; October 2003.

To prevent an asteroid from hitting Earth, a space tug equipped with plasma engines could give it a push

The Asteroid Tugboat

By Russell L. Schweickart, Edward T. Lu, Piet Hut and Clark R. Chapman

On an average night, more than 100 million pieces of interplanetary debris enter Earth's atmosphere. Luckily, most of these bits of asteroids and comets are no bigger than small pebbles; the total weight of the 100 million objects is only a few tons. And our planet's atmosphere is thick enough to vaporize the vast majority of these intruders. So the debris usually streaks harmlessly overhead, leaving the bright trails popularly known as shooting stars.

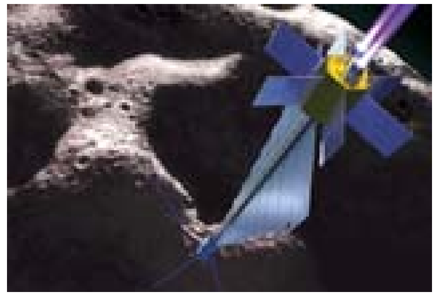
SPACE TUG is shown pushing an asteroid in this artist's highly speculative rendering of a deflection mission. The tug could use plasma engines to steadily thrust the asteroid in the desired direction. An array of radiator panels would dissipate the heat from the craft's nuclear reactor, located in the section closest to the asteroid's surface. The segmented arms on the surface attach the tug to the asteroid and stabilize the craft.

When bigger objects slam into the atmosphere, however, they explode rather than vaporize. In January 2000, for example, a rock about two to three meters wide exploded over Canada's Yukon Territory with a force equivalent to four or five kilotons of TNT. This kind of event occurs once a year, on average. Less frequently, larger rocks produce even more powerful explosions. In June 1908 a huge fireball was seen descending over the Tunguska region of Siberia. It was followed by an enormous blast that flattened more than 2,000 square kilometers of forest. The consensus among scientists today is that a rocky asteroid about 60 meters in diameter exploded some six kilometers above the ground with a force of about 10 megatons of TNT. The blast wave devastated an area approximately the size of metropolitan New York City.

Recent observations of near-Earth objects—asteroids and comets whose paths could intersect Earth's orbit—suggest that the chance of a similar event happening in this century is about 10 percent. Asteroids 100 meters across and larger pose an even more ominous threat because they will penetrate deeper into the atmosphere or hit the surface. Such an impact, which has a 2 percent chance of occurring before 2100, would cause an explosion equivalent to 100 megatons or more of TNT. If a large asteroid crashes into the ocean, which happens in about 70 percent of impacts, it could create a tsunami that might kill millions of people by inundating coastal cities. Events of this kind happen once every 40,000 years or so. And an asteroid with a diameter bigger than one kilometer would strike Earth with the energy equivalent of 100,000 megatons of TNT, far greater than the combined energy of all the nuclear weapons in existence. Impacts of this size and larger have the potential to wipe out human civilization, and there is a chance of perhaps one in 5,000 that such a strike will occur in this century.

> Rather than giving an asteroid a brief, powerful shove, the tug would deliver gentle pressure.

Can humanity prevent these catastrophes? Over the past decade scientists and engineers have proposed a variety of schemes to deflect an asteroid that is heading toward Earth [*see box on page 58*]. Several researchers have advocated detonating a nuclear weapon on or near the asteroid to either break it up or change its course, but the effects of a nuclear blast are difficult to predict, and that uncertainty has led many experts to view this option as a last resort at best. Recently interest has focused on more controlled options for shifting an asteroid's trajectory. For the past two years we have been studying the concept of an unmanned space tug that would rendezvous with an incoming asteroid, attach to its surface and slowly push the body so that it misses Earth. (Because of the unique characteristics of comets, we do not address them in this proposal. New studies indicate that comets constitute only about 1 percent of the overall impact threat to Earth.) To push the asteroid, the space tug would use nuclear-powered engines that expel jets of plasma, a high-temperature mix of ions and electrons. We believe that a mission to demonstrate the asteroid-tug concept could be accomplished by 2015.



Why develop such a spacecraft now, before astronomers have identified any asteroids on a collision course with Earth? Because the system should be tested before it is urgently needed. By attempting to deflect an asteroid that is not on, or even close to, a collision trajectory, researchers will acquire the experience necessary to build a reliable defense. Potentially hazardous asteroids have not yet been studied in any detail; because we do not know much about their interior makeup, surface characteristics or structural integrity, we cannot know what will happen when a space tug nudges one. The best way to learn about these crucial aspects is to land a spacecraft on an asteroid and then try to move it. As a bonus, the mission would add to our understanding of asteroids, pioneer the way to asteroid mining, and demonstrate critical technologies for future exploration of the solar system.

What is more, NASA is already working on the key technologies needed for the asteroid tug. As part of the Prometheus Project, the space agency is trying to design nuclear reactors that could power ion-propulsion systems for interplanetary

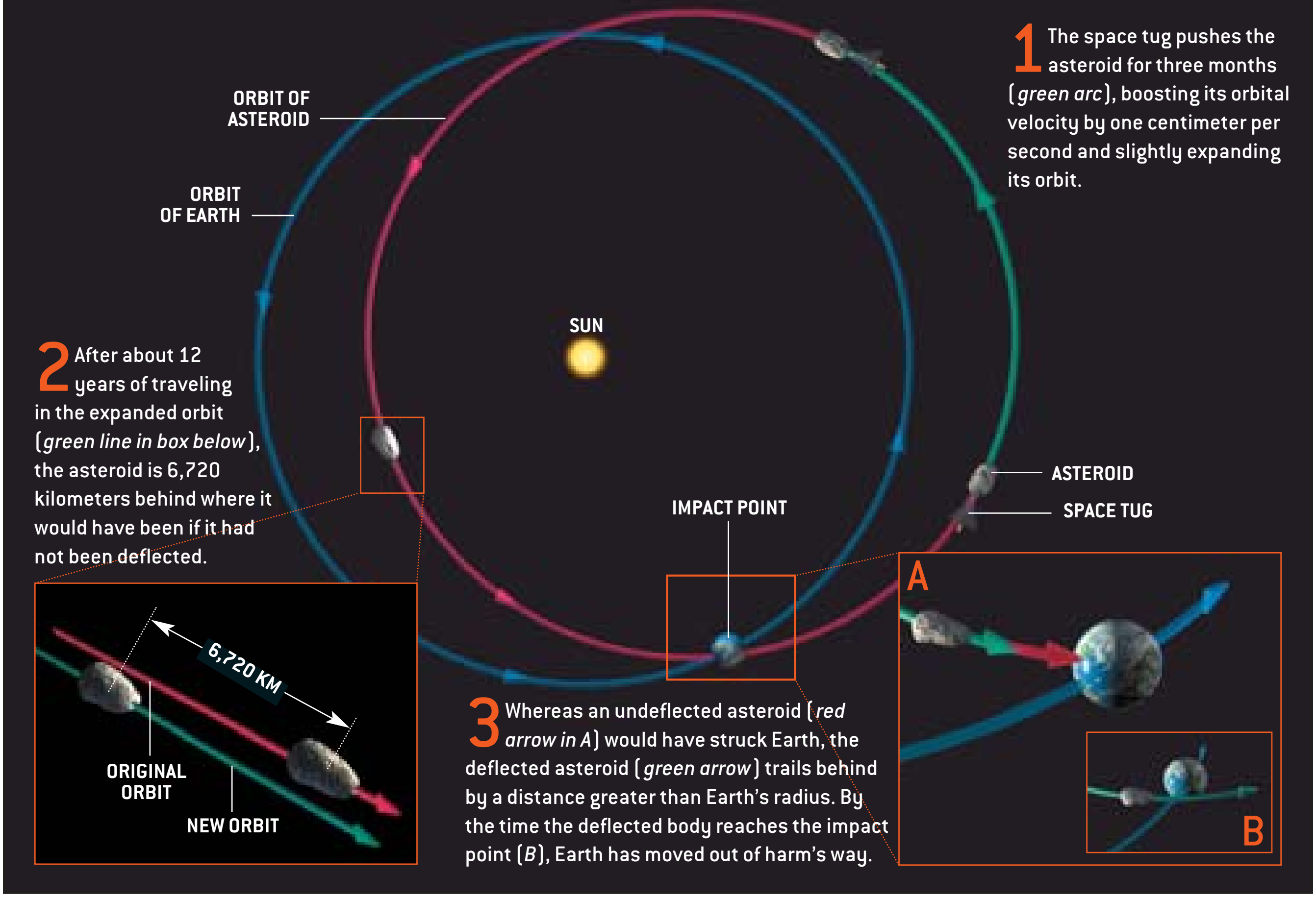
Overview/*Nudging an Asteroid*

- Near-Earth asteroids pose a threat to humanity. A direct hit by a 100-meter-wide asteroid would destroy a large city, and a one-kilometer-wide object could wipe out our civilization.
- Previous proposals for deflecting Earth-bound asteroids, such as nuclear explosions or kinetic-impact schemes, are unreliable. But a space tug equipped with plasma engines could provide a gentle push that would cause the asteroid to miss its rendezvous with Earth (assuming sufficient warning time).
- A mission to demonstrate the asteroid-tug concept could be accomplished by 2015. NASA is already developing nuclear reactors and propulsion systems that could be used by the space tug.

PAT RAWLINGS/SAIC (*preceding pages*)

AVERTING A COLLISION

A SPACE TUG can alter an asteroid's orbit by pushing in the direction of its orbital motion. This diagram assumes that the tug begins pushing 12 years before the projected impact and that the asteroid has an orbital period of 1.15 years.



spacecraft. NASA plans to integrate these systems into the Jupiter Icy Moons Orbiter (JIMO), a spacecraft that is expected to visit the Jovian moons of Ganymede, Callisto and Europa in the next decade. The same technologies could be applied to the greatest public safety project in history: warding off the doomsday rock that will sooner or later threaten humanity.

The B612 Mission

THE PROBLEM OF DEFLECTING an asteroid resolves into a timing issue. First, astronomers must detect the asteroid at least a decade before impact to provide time for the actions to take effect. Fortunately, with continued improvement in ongoing asteroid-detection programs, this is a reasonable expectation. To prevent the rock from hitting Earth, the most efficient plan is to either speed up the body by pushing it in the direction of its orbital motion or slow it down by pushing in the opposite direction. Changing the asteroid's velocity alters its orbital period—the time it takes to go around the sun. Because Earth moves along its orbit at an average speed of 29.8 kilometers per second and its diameter is 12,800 kilometers, our planet takes 215 seconds to move half its diameter. If an asteroid were headed for a bull's-eye collision with Earth, the challenge would be to change the asteroid's orbital period so that it arrives at the rendezvous site at least 215 seconds before or after Earth does, allowing the body to whiz safely by our planet [*see illustration above*].

Applying a soft but prolonged push on the asteroid about 10 years before it is expected to hit Earth, the tug would need to boost the asteroid's velocity by only about one centimeter per second. This change would slightly expand the asteroid's orbit and lengthen the time it takes to travel around the sun. For example, for an asteroid with an orbital period of two years, a one-centimeter-per-second velocity change would increase its period by 45 seconds and create a delay of 225 seconds over 10 years—enough for the asteroid to miss Earth by a small margin. Alternatively, the space tug could slow down the asteroid, shrinking its orbit and reducing the period by 45 seconds; after 10 years, the asteroid would arrive at the rendezvous site 225 seconds before Earth does. Of course, if the space tug reaches the asteroid when it is closer to striking Earth, it would need to give the body a bigger push. This fact underscores the importance of early and accurate detection of all near-Earth asteroids [*see box on page 60*].

To demonstrate this concept and the technologies involved, we have proposed

ALFRED T. KAMAJIAN

the development of a space tug that could deflect a 200-meter-wide asteroid, which would cause regional devastation if it hit Earth. We have dubbed this test project the B612 mission (B612 is the name of the asteroid in *The Little Prince*, the well-known children's book by Antoine de St. Exupéry). A rocky 200-meter asteroid has a mass of about 10 billion kilograms. Rather than giving the asteroid a brief, powerful shove—which might shatter the body instead of altering its course—the B612 tug would deliver gentle pressure. The force would be only about 2.5 newtons, approximately equivalent to the force required to hold up a glass of milk. But if this light nudge were applied for just over three months, it would be enough to change the asteroid's velocity by 0.2 centimeter per second. Should we be faced with an actual threat by a 200-meter asteroid, our small demonstration mission would either have to be scaled up by a factor of five or more to prevent the body from smashing into Earth, or else we would have to act at least 50 years before impact.

Because the force must be provided continuously for an extended period, the space tug's engines would require a significant amount of fuel. An additional large supply of propellant would be needed to get the tug to rendezvous with the asteroid. The average velocity change to get from our planet to a typical near-Earth asteroid is about 15 kilometers per second—one third more than the velocity change required to escape Earth's gravity. The standard chemical rocket engines, which mix fuel with oxidizer in a combustion chamber, would be hard-pressed to propel a substantial spacecraft (and all the fuel needed to push the asteroid) to these speeds. Such a vehicle would require so much propellant to per-

Asteroid Roundup

THE VARIOUS PLANS for deflecting an Earth-bound asteroid fall into two categories: those that rely on brief but intense applications of force and those that involve gently pushing or pulling the body over a long time. The most frequently mentioned concepts are described below.

NUCLEAR EXPLOSIONS have been proposed in two schemes. The obvious one is to destroy the asteroid by blasting it to smithereens. The less obvious approach would be to detonate a nuclear device off to one side of the asteroid, which would intensely heat the surface facing the explosion. The vaporization of surface rocks on that side of the asteroid would accelerate it slightly in the opposite direction. The advantage of these options is that the technology already exists and could be rapidly deployed. Theoretically, a powerful nuclear explosion could deflect a large asteroid that is just months from hitting Earth, a capability beyond that of any other technique. The problem, however, is that the results are neither predictable nor controllable. The explosion could split the asteroid into several large pieces, which might compound the problem rather than solve it.

NUCLEAR EXPLOSION might split an asteroid instead of changing its course.

KINETIC IMPACT plans also take advantage of existing technology. Simply launch the largest spacecraft available and smash it into the threatening asteroid at as high a velocity as can be mustered. Given the extremely high relative velocities necessary to deflect a substantial asteroid, a major challenge would be guiding the spacecraft so that all its impact energy goes into moving the asteroid off course and not spinning the body or knocking off a small chip. And as with nuclear explosions, splitting the asteroid is also a concern.

A MASS DRIVER is a device built on the surface of the asteroid that would repetitively hurl rocks into space, causing the asteroid to accelerate slowly in the opposite direction. Throwing enough rocks in the right direction would change the velocity of the asteroid enough to avoid a collision with our planet. The advantage of the mass driver is that it ejects materials from the asteroid itself, obviating the need to carry propellant from Earth. Throwing rocks, however, still requires a substantial energy source. The design of such a machine and its robotic installation on the asteroid's surface would be daunting tasks.

ABLATION is similar in concept to the standoff nuclear explosion but much slower. A small area on one side of the asteroid would be heated by a powerful laser flying near the asteroid or by sunlight reflected from a very large space mirror. Vaporized surface material would propel the asteroid in the desired direction. The attractive aspect of this option is that the asteroid's rotation is of no concern. But the laser or mirror must be able to maintain its position accurately to the side of the asteroid for a long period and therefore would require a substantial fuel supply. The optical elements of such concepts would also be vulnerable to coating by the ablating material from the asteroid.

SOLAR PRESSURE is another possible mechanism. A spacecraft would coat the asteroid's surface with highly reflective paint, which would change the radiation pressure caused by solar heating and very gradually alter the asteroid's course. But it is difficult to see this technique as a workable option given the massive amount of paint required and the difficulty of applying it to the surface.

LAND AND PUSH, the concept behind the asteroid tug, is very straightforward. The propulsion system required to get to the asteroid, which would also have to be developed for the other alternatives, is used to deflect the rock as well. The greatest advantage of this option is that it is fully controllable. The challenge lies in maneuvering the spacecraft and attaching it to the asteroid.

PAT RAWLINGS/SAIC

form the B612 mission that it could not be launched by a single rocket; dozens of heavy-lift rockets would be needed to boost all the components into low Earth orbit. Then the spacecraft would have to be assembled in orbit, which would drastically raise the mission's cost and delay the journey to the asteroid.

Our goal is to design a space tug that could be launched on a single heavy-lift rocket, such as a Proton, Ariane 5 or Titan 4. Because the tug must have a total mass less than about 20 tons, it needs extremely fuel-efficient engines. The primary measure of rocket efficiency is specific impulse, which is the thrust generated for each unit of fuel consumed per second. The most efficient chemical rockets have a specific impulse of up to 425 seconds when operating in the vacuum of space. (The units of specific impulse are seconds.) But the engines of our asteroid tug must have a specific impulse of 10,000 seconds.

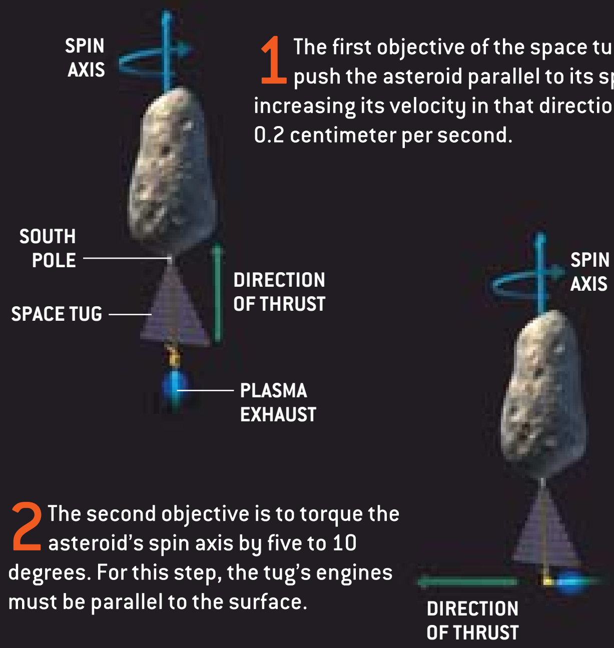
This performance is not feasible for standard chemical rockets but is comfortably within the range of electric engines, which use electrical or magnetic fields to accelerate ions out the exhaust nozzle of the rocket. In this way, the engines can achieve much higher exhaust velocities than chemical rockets, which simply burn fuel and allow the expanding hot gases to escape out the nozzle. Ion engines with a specific impulse of 3,000 seconds have successfully flown in space. A promising new engine known as the VASIMR (Variable Specific Impulse Magnetoplasma Rocket) uses radio waves to ionize a gas and accelerate the plasma to even higher exhaust velocities [see "The VASIMR Rocket," by Franklin R. Chang Díaz; Scientific American, November 2000]. Rather than using a conventional nozzle, the VASIMR employs magnetic fields to direct the expanding stream of ions out of the rocket at specific impulses between 3,000 and 30,000 seconds.

Of course, there is a price to be paid for such high performance. Although plasma and ion engines are more efficient than chemical rockets, their thrust is much lower (because the high-temperature exhaust is so tenuous). Several ion engines now under development could achieve specific impulses approaching the target of 10,000 seconds, but with the exception of the VASIMR, most electric engines generate less than 0.1 newton of force. Thus, many such engines would have to be ganged together to reach the desired thrust level of 2.5 newtons. Even when combined, the engines must push on the asteroid for a very long time to alter its orbit. Long-term operation has already been demonstrated, however: the ion engine on the Deep Space 1 spacecraft, launched in October 1998, accumulated 677 days of operating time.

To provide the required thrust, the plasma engines would need about 250 kilowatts of electrical power (assuming an engine efficiency of 50 percent). This amount of power is considerably beyond the capability of the solar arrays typically used for small spacecraft. Even the enormous solar arrays of the International Space Station, when completed, will produce less than half this amount (and they will weigh more than 65 tons). Clearly, such an array is infeasible for a spacecraft that must weigh less than 20 tons in total. The only current technology that can steadily supply this much power for several years in a package that weighs just a few tons is nuclear fission.

THE B612 MISSION

THE GOAL of the B612 mission is to significantly alter an asteroid's orbit in a controlled manner by 2015. The space tug would need to rendezvous with a target asteroid, attach itself to the surface and show its ability to maneuver the object.



1 The first objective of the space tug is to push the asteroid parallel to its spin axis, increasing its velocity in that direction by 0.2 centimeter per second.

2 The second objective is to torque the asteroid's spin axis by five to 10 degrees. For this step, the tug's engines must be parallel to the surface.

ALFRED T. KAMAJIAN

THE AUTHORS

In October 2002 *RUSSELL L. SCHWEICKART, EDWARD T. LU, PIET HUT* and *CLARK R. CHAPMAN* formed the B612 Foundation, a nonprofit group dedicated to developing and demonstrating the capability to deflect asteroids from Earth. Schweickart, chair of the foundation's board, is a former NASA astronaut who piloted Apollo 9's lunar module in 1969 and served as the backup commander for the first Skylab mission in 1973. Lu, the foundation's president, is a current astronaut who e-mailed his contributions to this article while onboard the International Space Station. Hut is a professor at the Institute for Advanced Study in Princeton, N.J., whose main research interests are computational astrophysics and the study of dense stellar systems. Chapman, a scientist at the Southwest Research Institute in Boulder, Colo., is a member of the science team for the upcoming MESSENGER mission to Mercury.

The asteroid tug needs a simple, small and safe nuclear reactor. Fortunately, NASA has already proposed some new designs for spacecraft reactors, and one has undergone preliminary testing. An important safety feature in these new designs is that the nuclear fuel is minimally radioactive until the reactor has produced power for a significant amount of time. Because the reactor would be launched cold—that is, inactive—even a catastrophic launch accident would pose little environmental danger. If the entire uranium core of the SAFE-1000, an advanced space reactor being developed at Los Alamos National Laboratory, were dispersed in a launch explosion, the radiation released into the environment would be only six to 10 curies—less than the total radiation contained in the walls of New York City's Grand Central Station. Ground controllers would send the command to activate the reactor only after it was safely in space.

The Problem of Spin

A MAJOR CHALLENGE for the B612 mission will be maneuvering around the target asteroid, landing on the body and attaching to its surface. In 2000 the NEAR Shoemaker spacecraft successfully maneuvered into orbit around Eros, the second largest of the known near-Earth asteroids, and even managed an impromptu landing on the 34-kilometer-long body. Japan's Hayabusa spacecraft (formerly Muses-C) is now on its way to near-Earth asteroid 1998 SF36 using ion propulsion. Once there it will lightly touch the asteroid's surface several times to pick up samples that will be returned to Earth. But the asteroid tug would be far larger than either of these spacecraft, and it would have to attach itself firmly to the asteroid because the gravitational attraction at the surface of such a body is only a hundred-thousandth of the gravity on Earth. Researchers are considering several concepts for a mechanism to hold the tug to the asteroid's surface, but the final design will most likely depend on the results of upcoming missions that will study the composition and structure of small asteroids.

To speed up or slow down the asteroid, the space tug must keep the direction of thrust parallel to the body's orbital motion. Small asteroids, though, often spin at rates of 10 rotations or more a day. One way to solve this problem would be to stop the rotation before pushing the asteroid. The tug would land on the asteroid's equator (the ring midway between the two poles of the axis of rotation), point its engines horizontally along the equator and fire them until the thrust brought the rotation to a halt.

This method could be risky, however, because most rocky asteroids appear to be porous, low-density "rubble piles," collections of many large and small boul-

Scouring the Sky

ON MARCH 18, 2002, newspapers and TV news shows around the world reported that Earth had just survived a near miss with a newly discovered asteroid named 2002 EM7. Astronomers observed the 70-meter-long rock four days after it passed within 461,000 kilometers of our planet, about 1.2 times the distance between Earth and the moon. Although it received quite a bit of attention, 2002 EM7 is just one of hundreds of thousands of asteroids that have come close to or crossed Earth's orbit. The international effort to detect and track these potentially threatening objects is called the Spaceguard Survey.

In 1998 NASA, at the urging of Congress, adopted the goal of detecting 90 percent of the 1,100 or so near-Earth objects (NEOs) larger than one kilometer in diameter. Halfway into the 10-year program, astronomers have found more than 660 NEOs of this size and more than 1,800 smaller bodies. Many of the asteroids currently being tracked were first seen as they were leaving Earth's vicinity, just as 2002 EM7 was. Fortunately, any asteroid destined to smash into Earth will most likely pass within a few lunar distances of our planet thousands of times before finally striking it. If researchers identify an object headed toward us, destined for an Earth impact, they will probably spot it decades or even centuries before it actually hits. The short-warning scenario, as dramatized in the 1998 movies *Armageddon* and *Deep Impact,* is exceedingly improbable.

SPACEGUARD SURVEY uses telescopes like this one at the White Sands Missile Range in New Mexico.

Every time Spaceguard detects a new NEO, scientists make projections based on its orbit to determine if it might strike Earth in the next 100 years or so. The vast majority of the objects discovered so far (more than 99 percent) do not seem to pose a threat. On rare occasions Spaceguard finds a NEO that is predicted to swing close by Earth in several decades. Because the procedure for determining future orbits, like all predictions, has only limited precision, one of these objects might actually be on a collision course. So Spaceguard monitors these few NEOs very carefully, gradually improving the accuracy of the predictions of their trajectories.

An asteroid with a diameter of 200 meters would not wreak the planetwide devastation that a one-kilometer-long rock could, but with an explosive force of 600 megatons or so it would still completely obliterate a city should it make even a nearby hit. Although Spaceguard has found many asteroids of this size, larger telescopes will be required to efficiently detect all the 100,000 smaller but still dangerous asteroids that cross Earth's orbit. Scientists have made a number of proposals to extend the asteroid search down to objects of about 200 meters, but no commitment yet exists. At best, such an augmented survey will not be complete until 15 to 20 years from now.

ders, interspersed with pebbles and smaller grains loosely held together by the body's weak gravity. Although this type of structure could withstand a force of several newtons distributed over two to five square meters of its surface, the same cannot be said for the internal stresses created by slowing down and stopping the body's rotation. It seems highly likely that altering the finely balanced gravitational and centripetal forces associated with asteroid rotation would cause significant and possibly destructive rearrangements—in other words, asteroid quakes.

For this reason, a better alternative might be to allow the asteroid to continue rotating but to torque the spin axis gradually until it is parallel with the body's orbital motion and keep it there. With the axis properly aligned, the tug would push the spinning asteroid along its orbit like a pinwheel. For the B612 demonstration mission, we plan to choose an asteroid spinning at about four rotations a day (typical of asteroids this size) and torque its spin axis by five to 10 degrees [*see illustration on page 59*]. Using 2.5 newtons of thrust applied at either the asteroid's north or south pole, the task would require a couple months of steady torquing. Although this result would clearly demonstrate the capability to maneuver an asteroid, an actual deflection would require many months, and perhaps even years, to properly orient the asteroid and accelerate it in the desired direction.

Another important challenge would be to deflect the asteroid in such a manner that it does not simply return again several years later on a new collision path. Bodies passing close to Earth are often gravitationally deflected into resonance orbits that have periods that are proportional to Earth's period; as a result, the bodies may periodically return to our planet's vicinity. We must therefore precisely deflect the asteroid onto a trajectory that ensures it will not end up in a resonance orbit. This requirement for precision is one of the best arguments for the asteroid-tug concept. The tug provides a carefully controlled maneuver, whereas most of the other deflection schemes yield an approximate, uncontrolled velocity change at best, thereby risking a boomerang scenario.

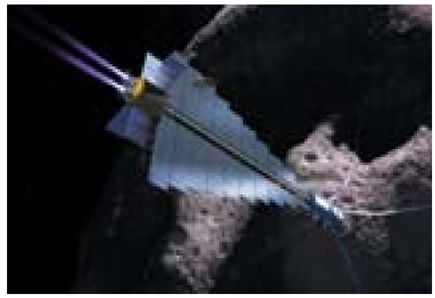
Although the use of an asteroid deflection system would be rare, its value would be beyond measure.

Protecting Our Planet

THE MISSION we are proposing would cost about $1 billion—a bit more than half of 1 percent of NASA's expected spending over the next 10 years—provided that off-the-shelf power and propulsion systems are used and a single existing launch vehicle can lift the spacecraft. Is this project worth the expense? Although the actual use of an asteroid deflection system would be rare—never in our lifetimes, we hope—its value would be beyond measure. An asteroid collision with Earth would be so potentially devastating that preventing it would be worth almost any cost. By practicing an asteroid deflection, the B612 mission would show whether the asteroid-tug concept is feasible and, if so, how it should be refined in the event of a real impact threat.

The scientific benefits of the demonstration mission would also be significant. Asteroids are remnants of the early solar system and have much to tell us about the formation of the planets and perhaps even the origins of life. Researchers have already learned a great deal by studying meteorites, the pieces of asteroid debris that survive the fiery plunge through Earth's atmosphere, but a much greater payoff would come from visiting the source of these fragments.

In addition, asteroids are believed to contain large amounts of metals, minerals and water ice. Experts on space exploration claim that taking advantage of these resources could dramatically reduce the cost of future interplanetary flights [see "Tapping the Waters of Space," by John S. Lewis; SCIENTIFIC AMERICAN PRESENTS, Spring 1999]. The B612 mission would vividly show that spacecraft could access these materials; using the same maneuvering and docking techniques developed for the asteroid tug, other vehicles could land on asteroids and begin mining operations. And these efforts may eventually pave the way for a manned mission to a near-Earth asteroid. Indeed, many experts contend that sending astronauts to an asteroid would be quicker, less costly and more worthwhile than a human mission to Mars.



Most important, the B612 demonstration would fulfill NASA's stated mission, "To protect our home planet ... as only NASA can." A better match could hardly be found. SA

MORE TO EXPLORE

Rain of Iron and Ice: The Very Real Threat of Comet and Asteroid Bombardment. John S. Lewis. Perseus, 1997.

Cosmic Pinball: The Science of Comets, Meteors, and Asteroids. Carolyn Sumners and Carlton Allen. McGraw-Hill Trade, 1999.

Report of the Workshop on Scientific Requirements for Mitigation of Hazardous Comets and Asteroids. Michael J. S. Belton. National Optical Astronomy Observatory, March 2003. Available online at **www.noao.edu/meetings/mitigation/report.html**

More information about the B612 mission can be found at **www.b612foundation.org**

New reports on near-Earth objects are available at **neo.jpl.nasa.gov/neo/report.html**, **impact.arc.nasa.gov/** and **neo.jpl.nasa.gov/neo/pha.html**

FOR ROBOT DESIGNERS THESE DAYS,

TOO SMALL TO DO MUCH as individuals, the authors' "millibots" must work as a team. The white snakelike contraption in the center is a chain of millibots linked together to climb stairs. Arrayed around it are other types of millibots, each customized for a specific task.

SMALL IS BEAUTIFUL

An Army of Small Robots

By Robert Grabowski, Luis E. Navarro-Serment and Pradeep K. Khosla

A group of terrorists has stormed into an office building and taken an unknown number of people hostage. They have blocked the entrances and covered the windows. No one outside can see how many they are, what weapons they carry or where they are holding their hostages. But suddenly a SWAT team bursts into the room and captures the assailants before they can even grab their weapons. How did the commandos get the information they needed to move so confidently and decisively?

The answer is a team of small, coordinated robots. They infiltrated the building through the ventilation system and methodically moved throughout the ducts. Some were equipped with microphones to monitor conversations, others with small video cameras, still others with sensors that sniffed the air for chemical or biological agents. Working together, they radioed this real-time information back to the authorities.

This is roughly the scenario that the Defense Advanced Research Projects Agency (DARPA) presented to robotics researchers in 1998. Their challenge was to develop tiny reconnaissance robots that soldiers could carry on their backs and scatter on the floor like popcorn. On the home front, firefighters and search-and-rescue workers could toss these robots through windows and let them scoot around to look for trapped victims or sniff out toxic materials. For now, these scenarios—let alone the lifelike robots depicted in science-fiction movies such as *Minority Report*—remain well beyond the state of the art. Yet the vision of mini robots has captured the attention of leading robot designers. Rather than concentrate on a few large platforms bristling with sensors (like Swiss Army knives on wheels), the focus these days is shifting toward building fleets of small, light and simple robots.

In principle, lilliputian robots have numerous advantages over their bulkier cousins. They can crawl through pipes, inspect collapsed buildings and hide in inconspicuous niches. A well-organized group of them can exchange sensor information to map objects that cannot be easily comprehended from a single vantage point. They can come to the aid of one another to scale obstacles or recover from a fall. Depending on the situation, the team leader can send in a bigger or smaller number of robots. If one robot fails, the entire mission is not lost; the rest can carry on.

But diminutive robots require a new design philosophy. They do not have the luxury of abundant power and space, as do their larger cousins, and they cannot house all the components necessary to execute a given mission. Even carrying something as compact as a video camera can nearly overwhelm a little robot. Consequently, their sensors, processing power and physical strength must be distributed among several robots, which must then work in unison. Such robots are like ants in a colony: weak and vulnerable on their own but highly effective when they join forces.

Whegs, Golf Balls and Tin Cans

RESEARCHERS HAVE TAKEN various approaches to the problems of building robots at this scale. Some have adopted a biological approach to mimic the attributes of insects and animals. For example, robot designers at Case Western Reserve University have developed a highly mobile platform modeled after a cockroach. It uses a hybrid of wheels and legs ("whegs") to scoot across uneven terrain. A team from the University of Michigan at Ann Arbor has come up with a two-legged robot with suction cups at the ends of its articulated limbs that allow it to climb walls, much like a caterpillar.

Biology has inspired not only the physical shape of the robots but also their control systems. Roboticists at the Massachusetts Institute of Technology have invented robots the size of golf balls that forage for food in the same fashion as ants. They use simple light sensors to express "emotions" to one another

MARK PORTLAND

and to make decisions collectively. This type of research takes its cue from the work of famous robot scientist Rodney A. Brooks. In the behavior-based control algorithms that he pioneered, each robot reacts to local stimuli. There is no central plan, no colonel commanding the troops. Instead the team's action emerges as a consequence of the combination of individuals interacting with one another. As innovative as this approach is, many problems remain before it can bear fruit. Deliberate missions require deliberate actions and deliberate plans—something that emergent behavior cannot reliably provide, at least not yet.

On the more deliberate side, researchers at the University of Minnesota have developed scouts, robots that can be launched like grenades through windows. Shaped like tin cans, these two-wheeled devices are equipped with video cameras that allow them to be teleoperated by a controlling user. Similarly, PARC (formerly known as Xerox PARC) in Palo Alto, Calif., has created a highly articulated snake robot that can be guided via remote video by a user. It literally crawls over obstacles and through pipes. Like the scouts, though, these robots currently lack sufficient local sensing and must rely on a human operator for decision making. This handicap currently makes them unwieldy for deployment in large numbers.

A few small robot platforms have become commercially available over the past few years. Khepera, a hockey-puck-size robot developed in Switzerland, has become popular among researchers interested in behavior-based control. Hobbyists, too, are experimenting with the technology. Living Machines in Lompoc, Calif., puts out a tiny programmable robot known as Pocket-Bot. Along the same lines, Lego Mindstorms, an extension to the popular Lego toy bricks, allows the general public to build and operate simple robots. Already they are being used in science projects and college contests. But the sensing and control for these commercial designs remain extremely rudimentary, and they lack the competence for complex missions.

Power Shortage

HERE AT CARNEGIE MELLON UNIVERSITY, the emphasis is on flexibility. We have built a team of about a dozen "millibots," each about five centimeters on a side. This is the scale at which we could still use off-the-shelf components for sensing and processing, although we had to custom-design the circuit boards and controllers. Each robot consists of three main modules: one for mobility, one for control and one for sensing. The mobility module sits on the bottom. Its two motors drive treads made from small O-rings. The present version can move across office floors and rugs at a maximum speed of about 20 centimeters a second, or about a sixth of normal human walking speed. As we develop new mobility platforms, we can snap them into place without having to redesign the rest of the robot.

The middle module provides processing and control. The current design contains an eight-bit microcontroller akin to the ones used in personal computers of the early 1980s. Though no match for modern desktop computers, these processors can still perform real-time control for the robot. The sensing module, which sits on top, includes sonar and near-infrared sensors for measuring the distance to nearby obstacles; a mid-infrared sensor (like those used in motion detectors) for detecting warm bodies; a video camera for surveillance; and a radio modem for communicating with other robots or the home base.

Perhaps the most severe limitation on these and other small robots is power. Batteries are bulky and heavy. They do not scale well: as its size is reduced, a battery reaches a threshold at which it cannot supply the power needed to move its own weight. The two rechargeable NiMH cellular-phone batteries on our millibots take up about a third of the available space. They provide enough power for only a limited array of sensors and a run time of between 30 and 90 minutes, depending on the complexity of the mission. Larger batteries would increase the run time but crowd out necessary components. Small-robot design is all about compromise. Speed, duration and functionality compete with weight, size and component availability.

To deal with these constraints, we have adopted two design methodologies for the millibots: specialization and collaboration. The former means that a robot is equipped with only enough sensing and processing for a specific task, allowing it to make optimal use of the available room and power. In a typical mission, some millibots are charged with making maps of the surroundings. Others provide live feedback for the human operator or carry sensors specific to that mission. To get the job done, the robots must collaborate.

Overview/*Millibots*

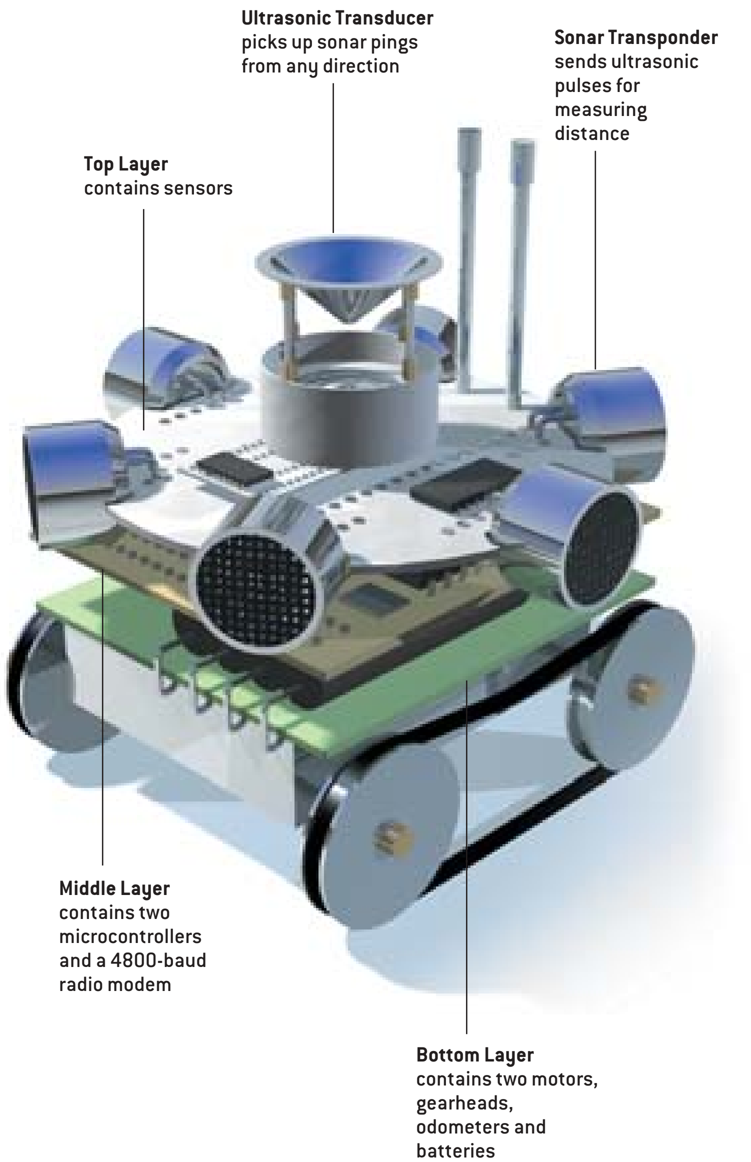
- Small robots will one day complement their larger, pricier cousins. Bots the size of Matchbox cars could scurry down pipes and crawl through the debris of collapsed buildings—very useful skills in espionage, surveillance, and search and rescue.
- Limited by size and battery power, small robots do not have the capabilities of a single larger robot. They must divvy up tasks and work together as a team, which is not as easy as it might sound. Engineers have had to develop new techniques for tasks such as ascertaining position and mapping territory.

HE'S GOT THE WHOLE ROBOT in his hands: One of the authors (Grabowski) holds a millibot. This particular design is about as small as designers could make it from off-the-shelf components.

MARK PORTLAND

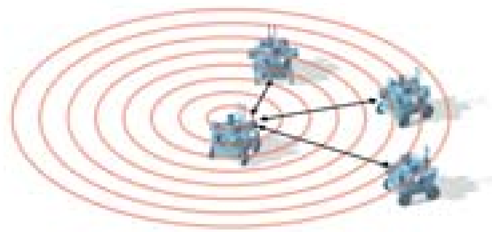
FINDING THEIR WAY IN THE WORLD

ANATOMY OF A MILLIBOT

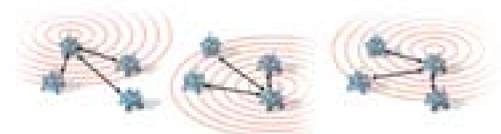


LOCALIZATION

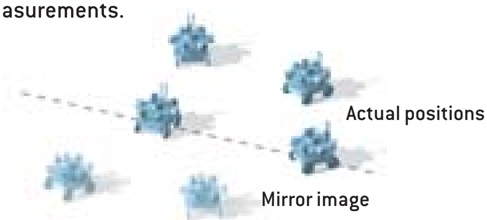
One robot simultaneously sends out an ultrasonic and radio pulse. The others receive the radio pulse instantaneously and the sound pulse shortly after. The time difference is a measure of the distance.



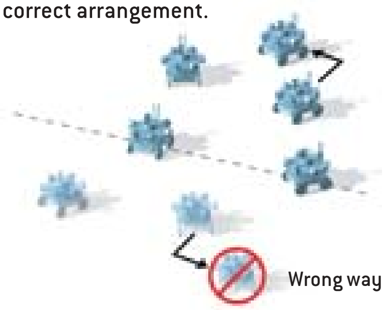
The robots take turns sending and receiving pulses.



A computer uses the distance measurements to deduce the position of each robot. One caveat is that mirror-image arrangements give the same set of measurements.

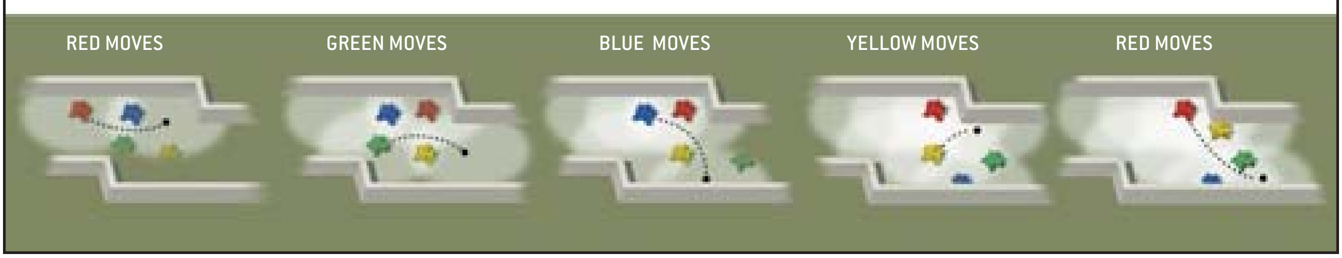


This ambiguity is resolved by having one of the robots take a left turn and measuring its new position, which will differ depending on which mirror image is the correct arrangement.



MAPPING STRATEGY

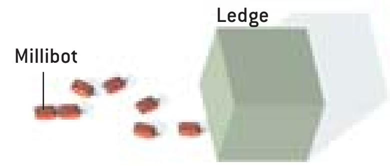
By using one another as reference points, millibots can find their way through an unknown space. In this example, three robots fix themselves in place and act as beacons. The fourth robot surveys the area using its sonar. When it is done, the robots switch roles. The lead robots become the new beacons, and the rearmost millibot begins moving around and taking data. The maps thus collected can be stitched together into a larger composite map of the entire area.

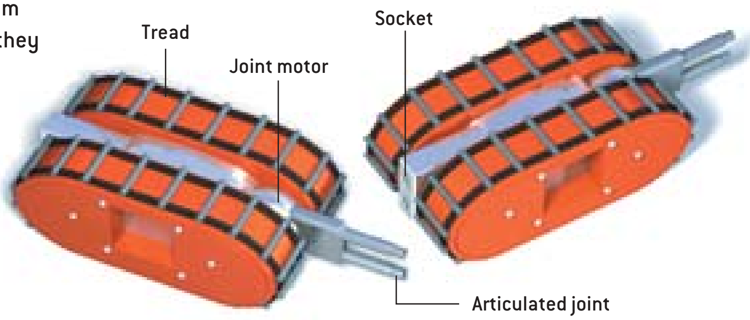


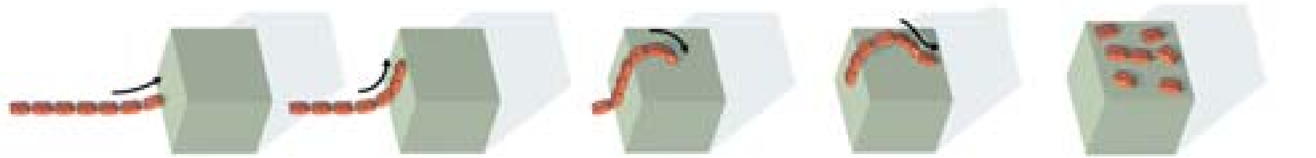
BRYAN CHRISTIE DESIGN

MILLIBOT CHAIN

DURING NORMAL OPERATION, individual millibots explore their space and share information to build maps. When the team reaches an impasse, such as a ledge or a flight of stairs, they come together to form an articulated train.







An assembled train pushes up against the edge of the obstacle.

Motorized joints lift the forwardmost robots against the side of the obstacle.

As the topmost robots reach over the lip of the object, they bend forward to gain traction.

Once the lead robots have gained enough traction, they pull up the bottom ones.

Having scaled the obstacle, the robots go their separate ways.

Where Are We?

ONE VITAL TASK that requires collaboration is localization: figuring out the team's position. Larger robots have the luxury of several techniques to ascertain their position, such as Global Positioning System (GPS) receivers, fixed beacons and visual landmark recognition. Moreover, they have the processing power to match current sensor information to existing maps.

None of these techniques works reliably for midget robots. They have a limited sensor range; the millibot sonar can measure distances out to about two meters. They are too small to carry GPS units. Dead reckoning—the technique of tracking position by measuring the wheel speed—is frustrated by their low weight. Something as seemingly inconsequential as the direction of the weave of a rug can dramatically influence their motion, making odometry readings inaccurate, just as a car's odometer would fail to give accurate distances if driven on an ice-covered lake.

So we have had to come up with a new technique. What we have developed is a miniaturized version of GPS. Rather than satellites, this technique utilizes sound waves to measure the distances between robots in the group. Each millibot is equipped with an ultrasonic transducer in addition to its radio modem. To determine distance, a millibot simultaneously emits a radio pulse and an ultrasonic signal, which radiate in all directions. The other robots listen for the two signals. The radio wave, traveling at the speed of light, arrives essentially instantaneously. The sound, moving at roughly 340 meters a second, arrives a few milliseconds later, depending on the distance between the robot sending the signal and the robot receiving it. A cone-shaped piece of metal on the sensing module reflects ultrasound down onto a transducer, allowing the robots to detect sound from any direction. The process is analogous to measuring the distance to an approaching storm by timing the interval between lightning and thunder.

By alternating their transmitting and listening roles, the robots figure out the distances between them. Each measurement takes about 30 milliseconds to complete. The team leader—either the home base or a larger robot, perhaps the mother bot that deployed the millibots—collects all the information and calculates robot positions using trilateration. Trilateration resembles the better-known technique of triangulation, except that it relies on distances rather than compass headings to get a fix on position. In two dimensions, each range estimate indicates that another robot lies somewhere on a circle around the transmitting robot. The intersection of two or more circles marks the po-

THE AUTHORS

ROBERT GRABOWSKI, LUIS E. NAVARRO-SERMENT and *PRADEEP K. KHOSLA* began working together on the millibot project in the summer of 1999. Khosla is chair of the electrical and computer engineering department at Carnegie Mellon University. He made his name in robotics by developing the first direct-drive manipulator arms, which are now used in most automated factories. Grabowski and Navarro-Serment are Ph.D. students. Grabowski served eight years in the U.S. Navy working with nuclear reactors. He has tinkered with electronics all his life and still enjoys playing with Legos and taking apart old VCRs. Navarro-Serment's background is in industrial automation and control systems; he used to head the electrical engineering department of the Guadalajara campus of the Monterrey Institute of Technology and Higher Education in Mexico. He is an avid amateur astronomer.

The authors thank the rest of the millibot team—Chris Paredis, Ben Brown, Curt Bererton and Mike Vande Weghe—for their invaluable contributions.

BRYAN CHRISTIE DESIGN

tential location of other robots [*see box on page 65*]. The algorithm finds the arrangement of robots that best satisfies all the circle intersections and range measurements.

One thing that complicates the procedure is that more than one arrangement of robots may match the data. Another is that range measurements are prone to error and uncertainty. Ultrasonic signals echo off floors and walls, creating ambiguity in the distance readings. In fact, depending on the geometry, wave interference can cause the signal to vanish altogether. For this reason, we developed an algorithm that combines the ultrasonic ranging with dead reckoning, which, despite its problems, provides enough additional information to resolve the ambiguities. The algorithm estimates the measurement error and computes the set of robot positions that minimizes the overall error.

The advantage of this localization method is that the millibots do not need fixed reference points to navigate. They can enter an unfamiliar space and survey it on their own. During mapping, a few selected millibots serve as beacons. These robots remain stationary while the others move around, mapping and avoiding objects while measuring their position relative to the beacons. When the team has fully explored the area around the beacons, the robots switch roles. The exploring robots position themselves as beacons, and the previous set begins to explore. This technique is similar to the children's game of leapfrog, and it can be executed without human intervention.

Chain of Command

OBSTACLES PRESENT small robots with another reason to collaborate. By virtue of its size, a little robot is susceptible to the random clutter that pervades our lives. It must deal with rocks, dirt and loose paper. The standard millibot has a clearance of about 15 millimeters, so a pencil or twig can stop it in its tracks. To get around these limitations, we have come up with a newer version of the millibots that can couple together like train cars. Each of these new millibots, about 11 centimeters long and six centimeters wide, looks like a miniature World War I–style tank. Typically they roam around independently and are versatile enough to get over small obstacles. But when they need to cross a ditch or scale a flight of stairs, they can link up to form a chain.

What gives the chain its versatility is the coupling joint between millibots. Unlike a train couple or a trailer hitch on a car, the millibot coupling joint contains a powerful motor that can rotate the joint up or down with enough torque to lift several millibots. To climb a stair, the chain first pushes up against the base of the stair. One of the millibots near the center of the chain then cantilevers up the front part of chain. Those millibots that reach the top can then pull up the lower ones [*see illustration on opposite page*]. Right now this process has to be remotely controlled by humans, but eventually the chain should be able to scale stairs automatically.

Already researchers' attention has begun to turn from hardware development toward the design of better control systems. The emphasis will shift from the control of a few individuals to the management of hundreds or thousands—a fundamentally different challenge that will require expertise from related fields such as economics, military logistics and even political science.

RICHARD VOYLES *University of Minnesota*

YOU COULD THROW THIS ROBOT through a window, and after it landed, it would start zipping around. Designed by a University of Minnesota team, TerminatorBot (a variant of the "scout" robots) is a bit smaller than a beer can. It has two arms that can pull it along, climb stairs and manipulate objects.

One of the ways we envision large-scale control is through hierarchy. Much like the military, robots will be divided into smaller teams controlled by a local leader. This leader will be responsible to a higher authority. Already millibots are being directed by larger, tanklike robots whose Pentium processors can handle the complex calculations of mapping and localization. These larger robots can tow a string of millibots behind them like ducklings and, when necessary, deploy them in an area of interest. They themselves report to larger all-terrain-vehicle robots in our group, which have multiple computers, video cameras, GPS units and a range of a few hundred kilometers. The idea is that the larger robots will deploy the smaller ones in areas that they cannot access themselves and then remain nearby to provide support and direction.

To be sure, small robots have a long way to go. Outside of a few laboratories, no small-robot teams are roaming the halls of buildings searching for danger. Although the potential of these robots remains vast, their current capabilities place them just above novelty—which is about where mobile phones and handheld computers were a decade ago. As the technology filters down from the military applications and others, we expect the competence of the small robot to improve significantly. Working as teams, they have a full repertoire of skills; their modular design allows them to be customized to particular missions; and, not least, they are fun to work with. SA

MORE TO EXPLORE

Behavior-Based Robotics (Intelligent Robotics and Autonomous Agents). Ronald C. Arkin. MIT Press, 1998.

Heterogeneous Teams of Modular Robots for Mapping, and Exploration. Robert Grabowski, Luis Navarro-Serment, Christopher J. J. Paredis and Pradeep K. Khosla in special issue on heterogeneous multirobot systems, *Autonomous Robots*, Vol. 8, No. 3, pages 293–308; June 2000.

Millibot Trains for Enhanced Mobility. H. Benjamin Brown, J. Michael Vande Weghe, Curt A. Bererton and Pradeep K. Khosla in *IEEE/ASME Transactions on Mechatronics*, Vol. 7, No. 4, pages 452–461; December 2002.

For more information and links to various projects: **www.andrew.cmu.edu/~rjg/army.html**

THE FUTURE of STRING THEORY

A Conversation with Brian Greene

String theory used to get everyone all tied up in knots. Even its practitioners fretted about how complicated it was, while other physicists mocked its lack of experimental predictions. The rest of the world was largely oblivious. Scientists could scarcely communicate just why string theory was so exciting—why it could fulfill Albert Einstein's dream of the ultimate unified theory, how it could give insight into such deep questions as why the universe exists at all. But in the mid-1990s the theory started to click together conceptually. It made some testable, if qualified, predictions. The outside world began to pay attention. Woody Allen satirized the theory in a *New Yorker* column this past July—probably the first time anyone has used Calabi-Yau spaces to make a point about interoffice romance.

Few people can take more credit for demystifying string theory than Brian Greene, a Columbia University physics professor and a major contributor to the theory. His 1999 book *The Elegant Universe* reached number four on the *New York Times* best-seller list and was a finalist for the Pulitzer Prize. Greene is now host of a three-part Nova series on PBS and has just completed a book on the nature of space and time. SCIENTIFIC AMERICAN staff editor George Musser recently spoke with him over a plate of stringy spaghetti. Here is an abridged, edited version of that conversation.

RANDY HARRIS

SCIENTIFIC AMERICAN: Sometimes when our readers hear the words "string theory" or "cosmology," they throw up their hands and say, "I'll never understand it."

BRIAN GREENE: I've definitely encountered a certain amount of intimidation at the outset when it comes to ideas like string theory or cosmology. But what I have found is that the basic interest is so widespread and so deep in most people that I've spoken with, that there is a willingness to go a little bit further than you might with other subjects that are more easily taken in.

SA: I noticed that at several points in *The Elegant Universe,* you first gave a rough idea of the physics concepts and then the detailed version.

BG: I found that to be a useful way of going about it, especially in the harder parts. It gives the reader permission: If the rough idea is the level at which you want to take it in, that's great; feel free to skip this next stuff. If not, go for it. I like to say things more than one way. I just think that when it comes to abstract ideas, you need many roads into them. From the scientific point of view, if you stick with one road, I think you really compromise your ability to make breakthroughs. I think that's really what breakthroughs are about. Everybody's looking at a problem one way, and you come at it from the back. That different way of getting there somehow reveals things that the other approach didn't.

The difference between making a breakthrough and not can often be just a small element of perception.

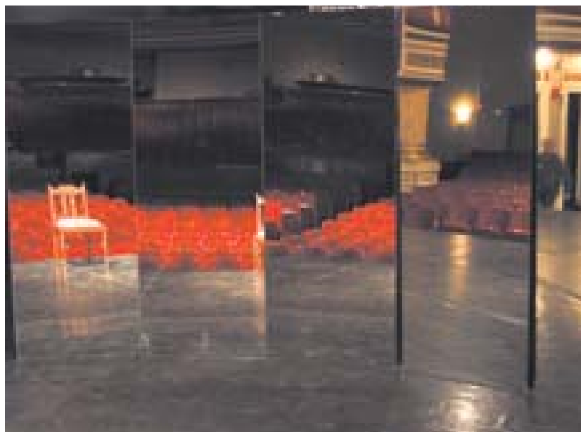
SA: What are some examples of that back-door approach?

BG: Well, probably the biggest ones are Ed Witten's breakthroughs. Ed [of the Institute for Advanced Study in Princeton, N.J.] just walked up the mountain and looked down and saw the connections that nobody else saw and in that way united the five string theories that previously were thought to be completely distinct. It was all out there; he just took a different perspective, and bang, it all came together. And that's genius.

To me that suggests what a fundamental discovery is. The universe in a sense guides us toward truths, because those truths are the things that govern what we see. If we're all being governed by what we see, we're all being steered in the same direction. Therefore, the difference between making a breakthrough and not often can be just a small element of perception, either true perception or mathematical perception, that puts things together in a different way.

SA: Do you think that these discoveries would have been made without the intervention of genius?

BG: Well, it's tough to say. In the case of string theory, I think so, because the pieces of the puzzle were really becoming clearer and clearer. It may have been five or 10 years later, but I suspect it would have happened. But with general relativity, I don't know. General relativity is such a leap, such a monumental rethinking of space, time and gravity, that it's not obvious to me how and when that would have happened without Einstein.



NOVA

SA: Are there examples in string theory that you think are analogous to that huge leap?

BG: I think we're still waiting for a leap of that magnitude. String theory has been built up out of a lot of smaller ideas that a lot of people have contributed and been slowly stitching together into an ever more impressive theoretical edifice. But what idea sits at the top of that edifice, we still don't really know. When we do have that idea, I believe that it will be like a beacon shining down; it will illuminate the edifice, and it will also, I believe, give answers to critical questions that remain unresolved.

SA: In the case of relativity, you had the equivalence principle and general covariance in that beacon role. In the Standard Model, it's gauge invariance. In *The Elegant Universe* you suggested the holographic principle could be that principle for string theory [see also "Information in the Holographic Universe," by Jacob D. Bekenstein; SCIENTIFIC AMERICAN, August]. What's your thinking on that now?

BG: Well, the past few years have only seen the holographic principle rise to a yet greater prominence and believability. Back in the mid-'90s, shortly after the holographic ideas were suggested, the supporting ideas were rather abstract and vague, all based upon features of black holes: Black hole entropy resides on the surface; therefore, maybe the degrees of freedom reside on the surface; therefore, maybe that's true of all regions that have a horizon; maybe it's true of cosmological horizons; maybe we're living within a cosmological region which has its true degrees of freedom far away. Wonderfully strange ideas, but the supporting evidence was meager.

But that changed with the work of Juan Maldacena [of the Institute for Advanced Study in Princeton, N.J.], in which he found an explicit example within string theory, where physics in the bulk—that is, in the arena that we consider to be real—would be exactly mirrored by physics taking place on a bounding surface. There'd be no difference in terms of the ability of either description to truly describe what's going on, yet in detail the descriptions would be vastly different. One would be in five dimensions, the other in four. So even the number of dimensions seems not to be something which you can count on, because there can be alternative descriptions that would accurately reflect the physics you're observing.

So to my mind, that makes the abstract ideas now concrete; it makes you believe the abstract ideas. And even if the details of string theory change, I think, as many others do—not everyone, though—that the holographic idea will persist and will guide us. Whether it truly is *the* idea, I don't know. I don't think so. But I think that it could well be one of the key stepping-stones towards finding the essential ideas of the theory. It steps outside the details of the theory and just says, Here's a very general feature of a world that has quantum mechanics and gravity.

SA: Let's talk a bit about loop quantum gravity and some of the other approaches. You've always described string theory as the only game in town when it comes to quantum gravity. Do you still feel that way?

BG: Well, I think it's the most fun game in town! But to be fair, the loop-quantum-gravity community has made tremendous progress. There are still many very basic questions that I don't feel have been answered, not to my satisfaction. But it's a viable approach, and it's great there are such large numbers of extremely talented people working on it. My hope—and it has been one that Lee Smolin [of the Perimeter Institute in Waterloo, Canada] has championed—is that ultimately we're developing the same theory from different angles. It's far from impossible that we're going down our route to quantum gravity, they're going down their route to quantum gravity, and we're going to meet someplace. Because it turns out that many of their strengths are our weaknesses. Many of our strengths are their weaknesses.

Relativity is a monumental rethinking of space and time. We're still waiting for another leap of that magnitude.

One weakness of string theory is that it's so-called background-dependent. We need to assume an existing spacetime within which the strings move. You'd hope, though, that a true quantum theory of gravity would have spacetime emerge from its fundamental equations. They [the loop-quantum gravity researchers], however, do have a background-independent formulation in their approach, where spacetime does emerge more fundamentally from the theory itself. On the other hand, we are able to make very direct contact with Einstein's general relativity on large scales. We see it in our equations. They have some difficulty making contact with ordinary gravity. So naturally, you'd think maybe one could put together the strengths of each.

SA: Has that effort been made?

BG: Slowly. There are very few people who are really well versed in both theories. These are both two huge subjects, and you can spend your whole life, every moment of your working day, just in your own subject, and you still won't know everything that's going on. But many people are heading down that path and starting to think along those lines, and there have been some joint meetings.

SA: If you have this background dependence, what hope is there to really understand, in a deep sense, what space and time are?

BG: Well, you can chip away at the problem. For instance, even with background dependence, we've learned things like mirror symmetry—there can be two spacetimes, one physics. We've learned topology change—that space can evolve in ways that we wouldn't have thought possible before. We've learned that the microworld might be governed by noncommutative geometry, where the coordinates, unlike real numbers, depend upon the order in which you multiply them. So you can get hints. You can get isolated glimpses of what's truly going on down there. But I think without the background-independent formalism, it's going to be hard to put the pieces together on their own.

SA: The mirror symmetry is incredibly profound, because it divorces spacetime geometry from physics. The connection between the two was always the Einsteinian program.

BG: That's right. Now, it doesn't divorce them completely. It simply says that you're missing half of the story. Geometry is tightly tied to physics, but it's a two-to-one map. It's not physics and geometry. It's physics and geometry-geometry, and which geometry you want to pick is up to you. Sometimes using one geometry gives you more insight than the other. Again, different ways of looking at one and the same physical system: two different geometries and one physics. And people have found there are mathematical questions about certain physical and geometrical systems that people couldn't answer using the one geometry. Bring in the mirror geometry that had previously gone unrealized, and, all of a sudden, profoundly difficult questions, when translated, were mind-bogglingly simple.

SA: Can you describe noncommutative geometry?

BG: Since the time of Descartes, we've found it very powerful to label points by their coordinates, either on Earth by their latitude and longitude or in three-space

The theory seems to be able to give rise to many different universes, of which ours seems to be only one.

by the three Cartesian coordinates, *x, y* and *z,* that you learn in high school. And we've always imagined that those numbers are like ordinary numbers, which have the property that, when you multiply them together—which is often an operation you need to do in physics—the answer doesn't depend on the order of operation: 3 times 5 is 5 times 3. What we seem to be finding is that when you coordinatize space on very small scales, the numbers involved are not like 3's and 5's, which don't depend upon the order in which they're multiplied. There's a new class of numbers that *do* depend on the order of multiplication.

They're actually not that new, because for a long time we have known of an entity called the matrix. Sure as shooting, matrix multiplication depends upon the order of multiplication. *A* times *B* does not equal *B* times *A* if *A* and *B* are matrices. String theory seems to indicate that points described by single numbers are replaced by geometrical objects described by matrices. On big scales, it turns out that these matrices become more and more diagonal, and diagonal matrices do have the property that they commute when you multiply. It doesn't matter how you multiply *A* times *B* if they're diagonal matrices. But then if you venture into the microworld, the off-diagonal entries in the matrices get bigger and bigger and bigger until way down in the depths, they are playing a significant part.

Noncommutative geometry is a whole new field of geometry that some people have been developing for years without necessarily an application of physics in mind. The French mathematician Alain Connes has this big thick book called *Noncommutative Geometry*. Euclid and Gauss and Riemann and all those wonderful geometers were working in the context of commutative geometry, and now Connes and others are taking off

NOVA

and developing the newer structure of noncommutative geometry.

SA: It is baffling to me—maybe it *should* be baffling—that you would have to label points with a matrix or some nonpure number. What does that mean?

BG: The way to think about it is: There is no notion of a point. A point is an approximation. If there is a point, you should label it by a number. But the claim is that, on sufficiently small scales, that language of points becomes such a poor approximation that it just isn't relevant. When we talk about points in geometry, we really talk about how something can move through points. It's the motion of objects that ultimately is what's relevant. Their motion, it turns out, can be more complicated than just sliding back and forth. All those motions are captured by a matrix. So rather than labeling an object by what point it's passing through, you need to label its motion by this matrix of degrees of freedom.

SA: What is your current thinking on anthropic and multiverse-type ideas? You talked about it in *The Elegant Universe* in the context of whether there is some limit to the explanatory power of string theory.

IF YOU WERE A STRING, spacetime might look something like this: six extra dimensions curled into a so-called Calabi-Yau shape.

BG: I and many others have never been too happy with any of these anthropic ideas, largely because it seems to me that at any point in the history of science, you can say, "Okay, we're done, we can't go any further, and the final answer to every currently unsolved question is: 'Things are the way they are because had they not been this way, we wouldn't have been here to ask the question.' " So it sort of feels like a cop-out. Maybe that's the wrong word. Not necessarily like a cop-out; it feels a little dangerous to me, because maybe you just needed five more years of hard work and you would have answered those unresolved questions, rather than just chalking them up to, "That's just how it is." So that's my concern: that one doesn't stop looking by virtue of having this fallback position.

But you know, it's definitely the case that the anthropic ideas have become more developed. They're now real proposals whereby you would have many universes, and those many universes could all have different properties, and it very well could be that we're simply in this one because the properties are right for us to be here, and we're not in those others because we couldn't survive there. It's less of just a mental exercise.

SA: String theory, and modern physics generally, seem to be approaching a single logical structure that *had* to be the way it is; the theory is the way it is because there's no other way it could be. On the one hand, that would argue against an anthropic direction. But on the other hand, there's a flexibility in the theory that leads you to an anthropic direction.

BG: The flexibility may or may not truly be there. That really could be an artifact of our lack of full understanding. But were I to go by what we understand today, the theory seems to be able to give rise to many different worlds, of which ours seems to be potentially one, but not even necessarily a very special one. So yes, there is a tension with the goal of absolute, rigid inflexibility.

SA: If you had other grad students waiting in the wings, what would you steer them to?

BG: Well, the big questions are, I think, the ones that we've discussed. Can we understand where space and time come from? Can we figure out the fundamental ideas of string theory or M-theory? Can we show that this fundamental idea yields a unique theory with the unique solution, which happens to be the world as we know it? Is it possible to test these ideas through astronomical observations or through accelerator-based experiment?

Can we even take a step further back and understand why quantum mechanics had to be part and parcel of the world as we know it? How many of the things that we rely on at a very deep level in any physical theory that has a chance of being right—such as space, time, quantum mechanics—are truly essential, and how many of them can be relaxed and potentially still yield the world that appears close to ours?

Could physics have taken a different path that would have been experimentally as successful but completely different? I don't know. But I think it's a real interesting question to ask. How much of what we believe is truly fundamentally driven in a unique way by data and mathematical consistency, and how much of it could have gone one way or another, and we just happened to go down one path because that's what we happened to discover? Could beings on another planet have completely different sets of laws that somehow work just as well as ours? SA

On the Web . . .

The full transcript of this conversation, with comments on everything from television to the arrow of time, is available at www.sciam.com

NOVA

Stranger in a New Land

By Kate Wong

Stunning finds in the Republic of Georgia upend long-standing ideas about the first hominids to journey out of Africa

We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.
—*T. S. Eliot,* Four Quartets: "Little Gidding"

In an age of spacecraft and deep-sea submersibles, we take it for granted that humans are intrepid explorers. Yet from an evolutionary perspective, the propensity to colonize is one of the distinguishing characteristics of our kind: no other primate has ever ranged so far and wide. Humans have not always been such cosmopolitan creatures, however. For most of the seven million years or so over which hominids have been evolving, they remained within the confines of their birthplace, Africa. But at some point, our ancestors began pushing out of the motherland, marking the start of a new chapter in our family history.

JOHN GURCHE

PORTRAIT OF A PIONEER: With a brain half the size of a modern one and a brow reminiscent of *Homo habilis,* this hominid is one of the most primitive members of our genus on record. Paleoartist John Gurche reconstructed this 1.75-million-year-old explorer from a nearly complete teenage *H. erectus* skull and associated mandible found in Dmanisi in the Republic of Georgia. The background figures derive from two partial crania recovered at the site.

SKULL SURPRISES

Early *Homo* from Dmanisi

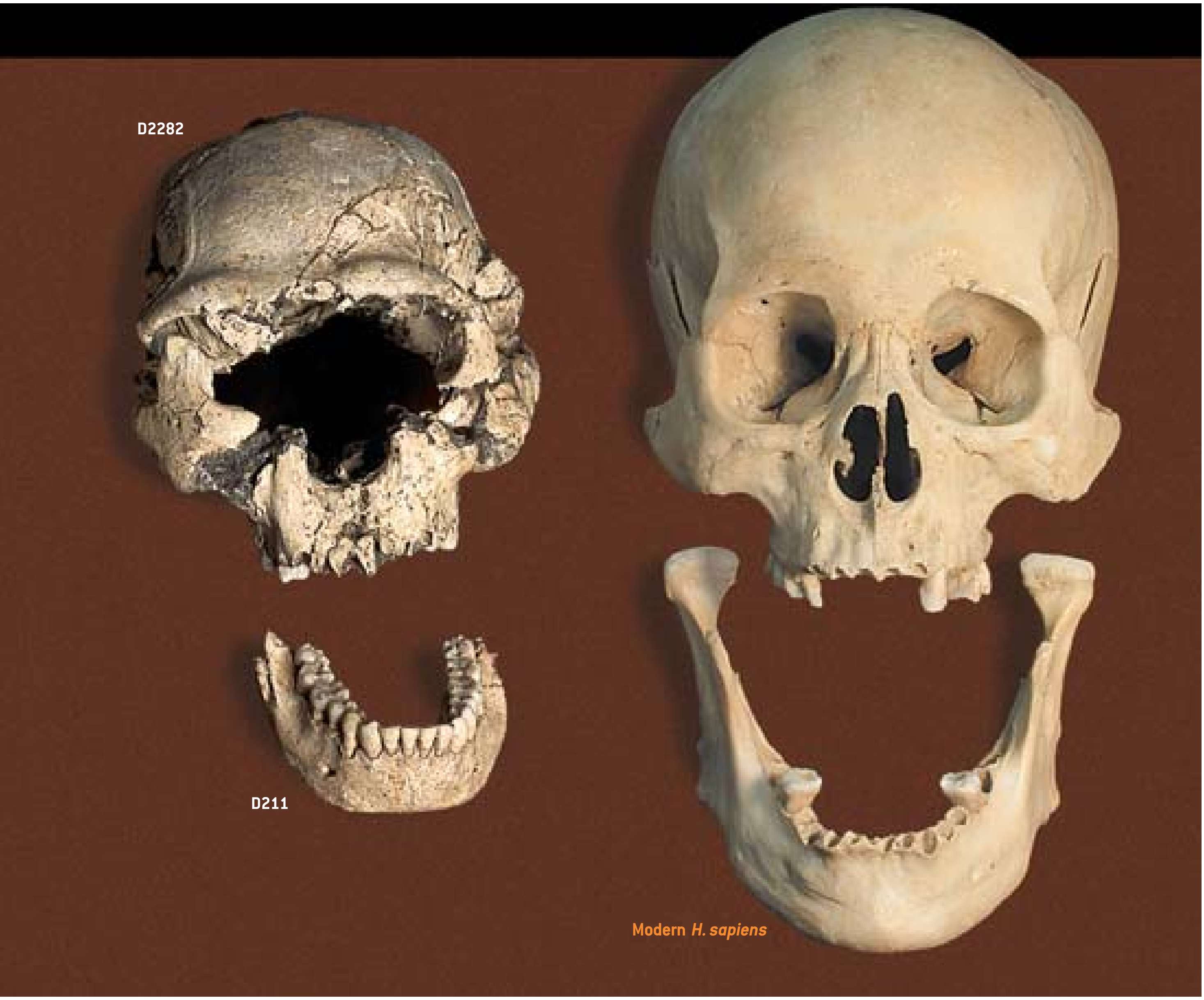
FOSSIL TROIKA hints at a variable *H. erectus.* These specimens from Dmanisi exhibit characteristic *H. erectus* features, such as a heaping up of bone along the midline of the skull known as a sagittal keel and marked constriction of the skull behind the eyes. But they stop short of the classic morphology of that hominid in several ways—their small brain size, for example, which was about half that of a modern human (*right*). Specimen D2700 (*left*), from a teenager, is especially primitive, resembling *H. habilis* not only in size but in the thinness of its brow, the projection of its face and the rounded contour of the rear of the skull. Some researchers propose that these fossils might represent a new species of *Homo.* Others suggest that the remains belong to more than one species, pointing to the enormous lower jaw known as D2600 that was unearthed in 2000. Indeed, this mandible is far too large to fit comfortably with any of the crania yet discovered (only D2700 turned up with an associated mandible, D2735; the other fossils were isolated finds). For now, the Dmanisi team considers all the fossils as members of the same, mutable species, *H. erectus.*

It was, until recently, a chapter the fossil record had kept rather hidden from view. Based on the available evidence—a handful of human fossils from sites in China and Java—most paleoanthropologists concluded that the first intercontinental traveling was undertaken by an early member of our genus known as *Homo erectus* starting little more than a million years ago. Long of limb and large of brain, *H. erectus* had just the sort of stride and smarts befitting a trailblazer. Earlier hominids, *H. habilis* and the australopithecines among them, were mostly small-bodied, small-brained creatures, not much bigger than a modern chimpanzee. The *H. erectus* build, in contrast, presaged modern human body proportions.

Overview/*The First Colonizers*

- Conventional paleoanthropological wisdom holds that the first humans to leave Africa were tall, large-brained people equipped with sophisticated stone tools who began migrating northward around a million years ago.
- New fossil discoveries in the Republic of Georgia are forcing scholars to rethink that scenario in its entirety. The remains are nearly half a million years older than hominid remains previously recognized as the most ancient outside of Africa. They are also smaller and accompanied by more primitive implements than expected.
- These finds raise the question of what prompted our ancestors to leave their natal land. They are also providing scientists with a rare opportunity to study not just a single representative of early *Homo* but a population.

GOURAM TSIBAKHASHVILI (*fossils*); CHRISTIAN SIDOR *New York College of Osteopathic Medicine* (*modern skull*)



Curiously, though, the first representatives of *H. erectus* in Africa, a group sometimes referred to as *H. ergaster,* had emerged as early as 1.9 million years ago. Why the lengthy departure delay? In explanation, researchers proposed that it was not until the advent of hand axes and other symmetrically shaped, standardized stone tools (a sophisticated technological culture known as the Acheulean) that *H. erectus* could penetrate the northern latitudes. Exactly what, if anything, these implements could accomplish that the simple Oldowan flakes, choppers and scrapers that preceded them could not is unknown, although perhaps they conferred a better means of butchering. In any event, the oldest accepted traces of humans outside Africa were Acheulean stone tools from a site called 'Ubeidiya in Israel.

Brawny, brainy, armed with cutting-edge technology—this was the hominid hero Hollywood would have cast in the role, a picture-perfect pioneer. Too perfect, it turns out. Over the past few years, researchers working at a site called Dmanisi in the Republic of Georgia have unearthed a trove of spectacularly well preserved human fossils, stone tools and animal remains dated to around 1.75 million years ago—nearly half a million years older than the 'Ubeidiya remains. It is by paleoanthropological standards an embarrassment of riches. No other early *Homo* site in the world has yielded such a bounty of bones, presenting scientists with an unprecedented opportunity to peer into the life and times of our hominid forebears. The discoveries have already proved revolutionary: the Georgian hominids are far more primitive in both anatomy and technology than expected, leaving experts wondering not only why early humans first ventured out of Africa but how.

A Dubious Debut

AS THE CROW FLIES, the sleepy modern-day village of Dmanisi lies some 85 kilometers southwest of the Georgian capital of Tbilisi and 20 kilometers north of the country's bor-

STONE TOOL TRICK

UNTIL RECENTLY, experts believed that humans could not leave Africa until they had developed an advanced technology known as the Acheulean, in which tools were symmetrically shaped and standardized (*see hand ax at right*). The tools found at Dmanisi, however, are simple flakes and choppers (*left* and *center*) manufactured according to much the same primitive Oldowan tradition that hominids in Africa were practicing nearly a million years earlier.

der with Armenia, nestled in the lower Caucasus Mountains. During the Middle Ages, Dmanisi was one of the most prominent cities of the day and an important stop along the old Silk Road. The region has thus long intrigued archaeologists, who have been excavating the crumbling ruins of a medieval citadel there since the 1930s. The first hint that the site might also have a deeper significance came in 1983, when paleontologist Abesalom Vekua of the Georgian Academy of Sciences discovered in one of the grain storage pits the remains of a long-extinct rhinoceros. The holes dug by the citadel's inhabitants had apparently opened a window on prehistory.

The next year, during paleontological excavations, primitive stone tools came to light, bringing with them the tantalizing possibility that fossilized human remains might eventually follow. Finally, in 1991, on the last day of the field season, the crew found what they were looking for: a hominid bone, discovered underneath the skeleton of a saber-toothed cat.

Based on the estimated ages of the associated animal remains, the researchers judged the human fossil—a mandible, or lower jaw, that they attributed to *H. erectus*—to be around 1.6 million years old, which would have made it the oldest known hominid outside of Africa. But when David Lordkipanidze and the late Leo Gabunia, also at the Georgian Academy of Sciences, showed the specimen to some of the biggest names in paleoanthropology at a meeting in Germany later that year, their claims met with skepticism. Humans were not supposed to have made it out of Africa until a million years ago, and the beautifully preserved mandible—every tooth in place—looked too pristine to be as old as the Georgians said it was. Many concluded that the fossil was not *H. erectus* but a later species. Thus, rather than receiving the imprimatur of paleoanthropology's elite, the jaw from Dmanisi came away with question marks.

Undaunted, team members continued work at the site, refining their understanding of its geology and searching for more hominid remains. Their perseverance eventually paid off: in 1999 workers found two skulls just a few feet away from where the mandible had turned up eight years prior. A paper describing the fossils appeared in *Science* the following spring. "That year the fanfare began," recollects Lordkipanidze, who now directs the excavation. The finds established a close relationship between the Dmanisi hominids and African *H. erectus*. Unlike the earliest humans on record from eastern Asia and western Europe, which exhibited regionally distinctive traits, the Dmanisi skulls bore explicit resemblances—in the form of the browridge, for example—to the early African material.

By this time, geologists had nailed down the age of the fossils, which come from deposits that sit directly atop a thick layer of volcanic rock radiometrically dated to 1.85 million years ago. The fresh, unweathered contours of the basalt indicate that little time passed before the fossil-bearing sediments blanketed it, explains C. Reid Ferring of the University of North Texas. And paleomagnetic analyses of the sediments signal that they

GOURAM TSIBAKHASHVILI

were laid down close to 1.77 million years ago, when Earth's magnetic polarity reversed, the so-called Matuyama boundary. Furthermore, remains of animals of known antiquity accompany the hominid fossils—a rodent called *Mimomys,* for instance, which lived only between 1.6 and 2.0 million years ago—and a second, 1.76-million-year-old layer of basalt at a nearby site caps the same stratigraphy.

Together the new fossils and dating results clinched the case for Dmanisi being the oldest unequivocal hominid site outside of Africa, pushing the colonization of Eurasia back hundreds of thousands of years. They also toppled the theory that humans could not leave Africa until they had invented Acheulean technology. The Dmanisi tool kit contained only Oldowan-grade implements fashioned from local raw materials.

Pint-Size Pioneer

THE GREAT AGE of the Georgian hominids and the simplicity of their tools came as a shock to many paleoanthropologists. But Dmanisi had even more surprises in store. Last July, Lordkipanidze's team announced that it had recovered a third, virtually complete skull—including an associated mandible—that was one of the most primitive *Homo* specimens on record. Whereas the first two skulls had housed 770 cubic centimeters and 650 cubic centimeters of gray matter, the third had a cranial capacity of just 600 cubic centimeters—less than half the size of a modern brain and considerably smaller than expected for *H. erectus*. Neither was the form of the third skull entirely *erectus*-like. Rather the delicacy of the brow, the projection of the face and the curvature of the rear of the skull evoke *H. habilis,* the presumed forebear of *H. erectus*.

The discovery of the third skull has led to the startling revelation that contrary to the notion that big brains were part and parcel of the first transcontinental migration, some of these early wayfarers were hardly more cerebral than primitive *H. habilis*. Likewise, the Georgian hominids do not appear to have been much larger-bodied than *H. habilis*. Only isolated elements from below the neck have turned up thus far—namely, ribs, clavicles, vertebrae, as well as upper arm, hand and foot bones—and they have yet to be formally described. But it is already clear that "these people were small," asserts team member G. Philip Rightmire of the University of Binghamton.

"This is the first time we have an intermediate between *erectus* and *habilis,*" Lordkipanidze observes. Although the fossils have been provisionally categorized by the team as *H. erectus* based on the presence of certain defining characteristics, he thinks the population represented by the Dmanisi hominids may have been more specifically the rootstock of the species, a missing link between *erectus* and *habilis*.

Other scholars have proposed a more elaborate taxonomic scheme. Noting the anatomical variation evident in the skulls and mandibles recovered so far (including a behemoth jaw unearthed in 2000), Jeffrey Schwartz of the University of Pittsburgh suggested that the Dmanisi fossils might represent two or more early human species. "If that's the case, I'll eat one of them," retorts Milford H. Wolpoff of the University of Michigan at Ann Arbor. A more likely explanation, he offers, is that the rogue mandible comes from a male and the rest of the bones belong to females.

For his part, Lordkipanidze acknowledges that the massive mandible "is a bit of a headache," but given that the fossils all come from the same stratigraphic layer, he reasons, they are probably members of the same population of *H. erectus*. Indeed, one of the most important things about Dmanisi, he says, is that it "gives us an opportunity to think about what variation is." Perhaps some researchers have underestimated how variable *H. erectus* was—a notion that recent discoveries from a site called Bouri in Ethiopia's Middle Awash region and another locality known as Ileret in Kenya support. Lordkipanidze suspects that as the Georgian picture becomes clearer, the sex and species of more than a few African fossils will need reassessing, as will the question of who the founding members of our lineage were. "Maybe *habilis* is not *Homo,*" he muses. In

> Brawny, brainy, armed with cutting-edge technology—this was the hominid hero Hollywood would have cast.

fact, a number of experts wonder whether this hominid may have been a species of *Australopithecus* rather than a member of our own genus.

"It is not cladistically compelling to place *habilis* in *Homo,*" comments Bernard Wood of George Washington University. Considering its brain and body proportions, characteristics of its jaws and teeth and features related to locomotion, "*habilis* is more australopithlike than it has been made out to be." If so, the emergence of *H. erectus* may well have marked the birth of our genus. What is unclear thus far, Wood says, is whether the Dmanisi hominids fall on the *Homo* side of the divide or the *Australopithecus* one.

Taxonomic particulars aside, the apparently small stature of the Dmanisi people could pose further difficulty for paleoanthropologists. Another popular theory of why humans left Africa, put forth in the 1980s by Alan Walker and Pat Shipman of Pennsylvania State University and elaborated on more recently by William R. Leonard of Northwestern University and his colleagues, proposes that *H. erectus*'s large body size necessitated a higher-quality diet—one that included meat—than that of its smaller predecessors to meet its increased energy needs. Adopting such a regimen would have forced this species to broaden its horizon to find sufficient food—an expansion that might have led it into Eurasia. The exact proportions of these primitive Georgians are pending, but the discovery of individuals considerably smaller than classic *H. erectus* outside of Africa could force experts to rethink that scenario.

Georgia on Their Minds

HOWEVER EARLY HOMINIDS got out of Africa, it is not hard to see why they settled down in southern Georgia. For one, the presence of the Black Sea to the west and the Caspian Sea to the east would have ensured a relatively mild, perhaps even Mediterranean-like, climate. For another, the region appears to have been incredibly diverse ecologically: remains of woodland creatures, such as deer, and grassland animals, such as horses, have all turned up at the site, suggesting a mosaic of forest and savanna habitats. Thus, in practical terms, if the going got tough in one spot, the hominids would not have had to move far to get to a better situation. "The heterogeneity of the

DIGGING DMANISI

DMANISI, REPUBLIC OF GEORGIA, JULY—From the Republic of Georgia's capital, Tbilisi, the village of Dmanisi is just a two-hour drive, yet it seems a world apart from the bustle of the diesel- and dust-choked city. Here in the foothills of the Caucasus Mountains, donkey-drawn carts outnumber cars and the air is fragrant with hay. The locals farm the rich soil and raise sheep, pigs and goats; children spend summer afternoons racing down a stretch of paved road on homemade scooters. Even the roosters appear to lose track of time, crowing not only at daybreak but in the afternoon and evening as well.

The leisurely pace of modern life belies the region's storied past, however. Centuries ago Dmanisi was a seat of great power, situated at a crossroads of Byzantine and Persian trading routes. Today the region is littered with reminders of that bygone era. Haystacklike mounds resolve into ancient Muslim tombs on closer inspection; medieval burials erode out of a hillside after heavy rains; and looming above it all are the imposing ruins of a citadel built on a promontory that once overlooked the Silk Road.

That much about Dmanisi's past has been known for decades. Only recently have scholars learned that long before the rise and fall of the city, this was the dominion of a primitive human ancestor, the first known to march out of Africa and begin colonizing the rest of the Old World some 1.75 million years ago—far earlier than previously thought. It is a realization that still gives David Lordkipanidze pause. Just a dozen years ago he helped to unearth the first hominid bone at Dmanisi. Four skulls, 2,000 stone tools and thousands of ancient animal fossils later, the 40-year-old is deputy director of the Georgian State Museum and head of an excavation many paleoanthropologists regard as the most spectacular in recent memory. "It is big luck to have these beautiful fossils," he reflects. But it is also "a big responsibility." Indeed, equal parts paleontologist and politician, Lordkipanidze

GOURAM TSIBAKHASHVILI

REMAINS OF THE DAY: Excavations of Dmanisi's medieval city led to the discovery of the much older fossils. So far paleoanthropologists have thoroughly probed only 100 square meters of the site, which is estimated to span 11,000 square meters.

environment may have promoted occupation," Ferring says. The Dmanisi site in particular, located on a promontory formed by the confluence of two rivers, may have attracted hominids with its proximity to water, which would have not only quenched their thirst but lured potential prey as well.

"Biologically this was a happening place," remarks Martha Tappen of the University of Minnesota. Of the thousands of mammal fossils that workers have unearthed along with the hominid remains, many come from large carnivores such as saber-toothed cats, panthers, bears, hyenas and wolves. Tappen, whose work centers on figuring out what led to the accumulation of bones at the site, suspects that the carnivores may

seems to work around the clock, talking on his cell phone late into the night with colleagues and prospective sponsors.

Largely as a result of those efforts, what started as a 10-person team of Georgians and Germans has mushroomed into a 30-strong collaboration of scientists and students from around the world, a number of whom have gathered here for the annual field season. For eight weeks every summer, the Dmanisi field crew surveys, digs and analyzes new finds. It is a shoestring operation. Team members live in a no-frills house a couple miles from the site, typically sleeping four to a tiny room. Electricity is ephemeral at best, hot running water nonexistent.

Every morning at around 8:30, after a breakfast of bread and tea at the picnic tables on the porch, the groggy workers pile into a Russian army-issue lorry left over from the days of Soviet occupation and drive up to the site. In the main excavation area—the 20-meter-by-20-meter square that in 2001 yielded an extraordinarily complete skull and associated lower jaw—each person tends a square-meter plot, meticulously recording the three-dimensional position of each recognizable bone and artifact uncovered during removal of the sediments. These items are then labeled and bagged for later study. Even nondescript pebbles and sediments are saved for further scrutiny: rinsing and sieving them may expose shells, minuscule mammal bones and other important environmental clues.

On this particular day the fossil hunters are in especially good spirits. A rare bout of soggy weather left them housebound yesterday (waterlogged bones are too fragile to extract), and this morning's skies threatened to do the same. But the mist draping the mountains has finally burned off, eliciting a chorus of Johnny Nash's "I Can See Clearly Now," sung over the taps and scrapes of trowels, hammers and spackle knives against the chalky sediments. They progress slowly. The excavators are now working in the dense upper layer, which does not yield its bones and stones easily. They must take care not to scratch the remains with their implements, lest the fresh marks be mistaken for ancient ones in later analyses. When noon arrives, the diggers break eagerly for lunch—tomatoes, cucumbers, bread, hard-boiled eggs and pungent, brine-soaked cheese (an acquired taste)—and a catnap on the grass before returning to their squares.

Meanwhile, in a makeshift lab back at camp, other crew members sort through remains brought back earlier by the excavators. Seated at metal-topped wooden tables and sharing an outmoded microscope, they identify the species to which each bone belongs and inspect it for telltale breaks, cut marks and tooth

MARC MEYER *University of Pennsylvania*

SCRAPING AND BRUSHING away the chalky sediments, crew members expose stone tools and animal remains—the work of hungry hominids.

marks. Such data should eventually disclose how the bones accumulated. Preliminary findings from the main excavation suggest that denning saber-toothed cats may have collected them. In contrast, early data from another dig spot about 100 meters away, known as M6, hint that humans worked there—the abundance of smashed bone in this locale is more characteristic of hominid activity than carnivore activity. If so, M6 could provide critical insight into how the primitive Dmanisi hominids eked out an existence in this new land.

When the fossil hunters return with the day's haul at around 4:00, camp is once again the center of activity. An early dinner leaves time for a shower, a game of chess or a trip down the road to visit the enterprising village woman who vends candy, soda, cigarettes and other luxury goods from a small whitewashed building affectionately dubbed the Mall, before a final hour of lab work and the evening tea.

For Lordkipanidze, the work has come full circle. Here at the site where he cut his teeth on paleoanthropology, he hopes to establish a preeminent field school to train aspiring young archaeologists and anthropologists. In the meantime he and his colleagues have plans to test promising spots elsewhere in the region for hominid fossils. Perhaps Georgia's biggest surprises are yet to come. *—K.W.*

have been using the water-lined promontory as a trap. "The question," she says, "is whether hominids were, too."

So far Tappen has identified a few cut marks on the animal bones, indicating that, at least on occasion, the Dmanisi settlers ate meat. But whether they scavenged animals brought down by the local carnivores or hunted the beasts themselves is not known. The matter warrants investigation. One of the few remaining hypotheses for what allowed humans to expand their range into northern lands holds that making the transition from the mostly vegetarian diet of the australopithecines to a hunter-gatherer subsistence strategy enabled them to survive the colder winter months, during which plant resources were scarce, if not altogether unavailable. Only further analyses of the mammal bones at the site can elucidate how the Dmanisi humans acquired meat. But Tappen surmises that they were hunting. "When you're a scavenger, the distribution of animals is so unpredictable," she remarks. "I don't think it was their main strategy."

> Contrary to expectations, some of these early wayfarers were hardly more cerebral than primitive *Homo habilis.*

That does not mean that humans were the top carnivores, however. "They could have been both the hunters and the hunted," Tappen observes. Telltale puncture wounds on one of the skulls and gnaw marks on the large mandible reveal that some of the hominids at Dmanisi ended up as cat food.

Outward Bound

THE GEORGIAN REMAINS prove that humans left Africa shortly after *H. erectus* evolved around 1.9 million years ago. But where they went after that is a mystery. The next oldest undisputed fossils in Asia are still just a bit more than a million years old (although controversial sites in Java date to 1.8 million years ago), and those in Europe are only around 800,000 years of age. Anatomically, the Dmanisi people make reasonable ancestors for later *H. erectus* from Asia, but they could instead have been a dead-end group, the leading edge of a wave that washed only partway across Eurasia. There were, scientists concur, multiple migrations out of Africa as well as movements back in. "Dmanisi is just one moment," Lordkipanidze says. "We need to figure out what happened before and after."

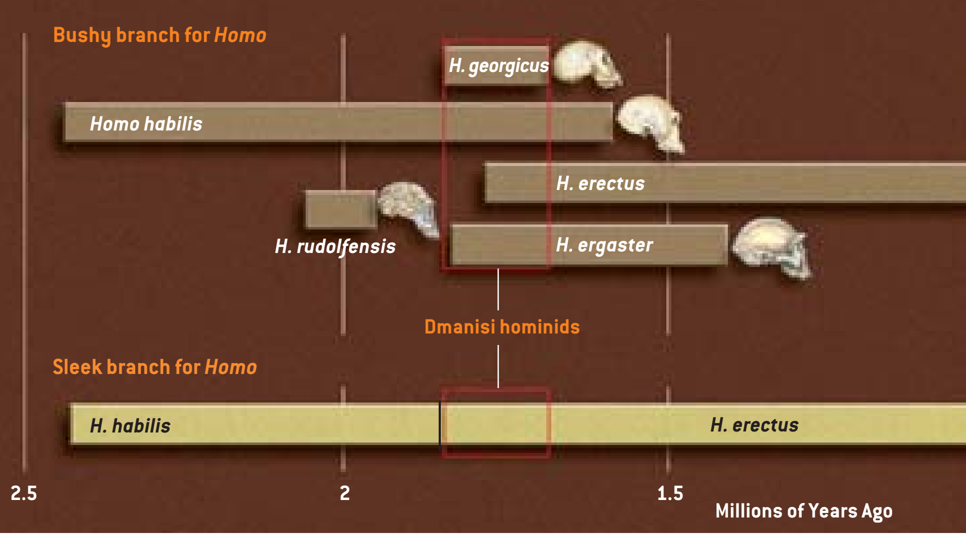
Echoing what has become a common refrain in paleoanthropology, the Dmanisi discoveries in some ways raise more questions than they answer. "It's nice that everything's been shaken up," Rightmire reflects, "but frustrating that some of the ideas that seemed so promising eight to 10 years ago don't hold up anymore." A shift toward meat eating might yet explain how humans managed to survive outside of Africa, but what prompted them to push into new territories remains unknown. Perhaps they were following herd animals north. Or maybe it was as simple and familiar as a need to know what lay beyond that hill, or river, or tall savanna grass—a case of prehistoric wanderlust.

The good news is that scientists have only begun plumbing Dmanisi's depths. The fossils recovered thus far come from just a fraction of the site's estimated extent, and new material is emerging from the ground faster than the researchers can formally describe it—a fourth skull unearthed in 2002 is still undergoing preparation and analysis and a new jaw, tibia and ankle bone were unearthed this summer. Topping the fossil hunters' wish list are femurs and pelvises, which will reveal how

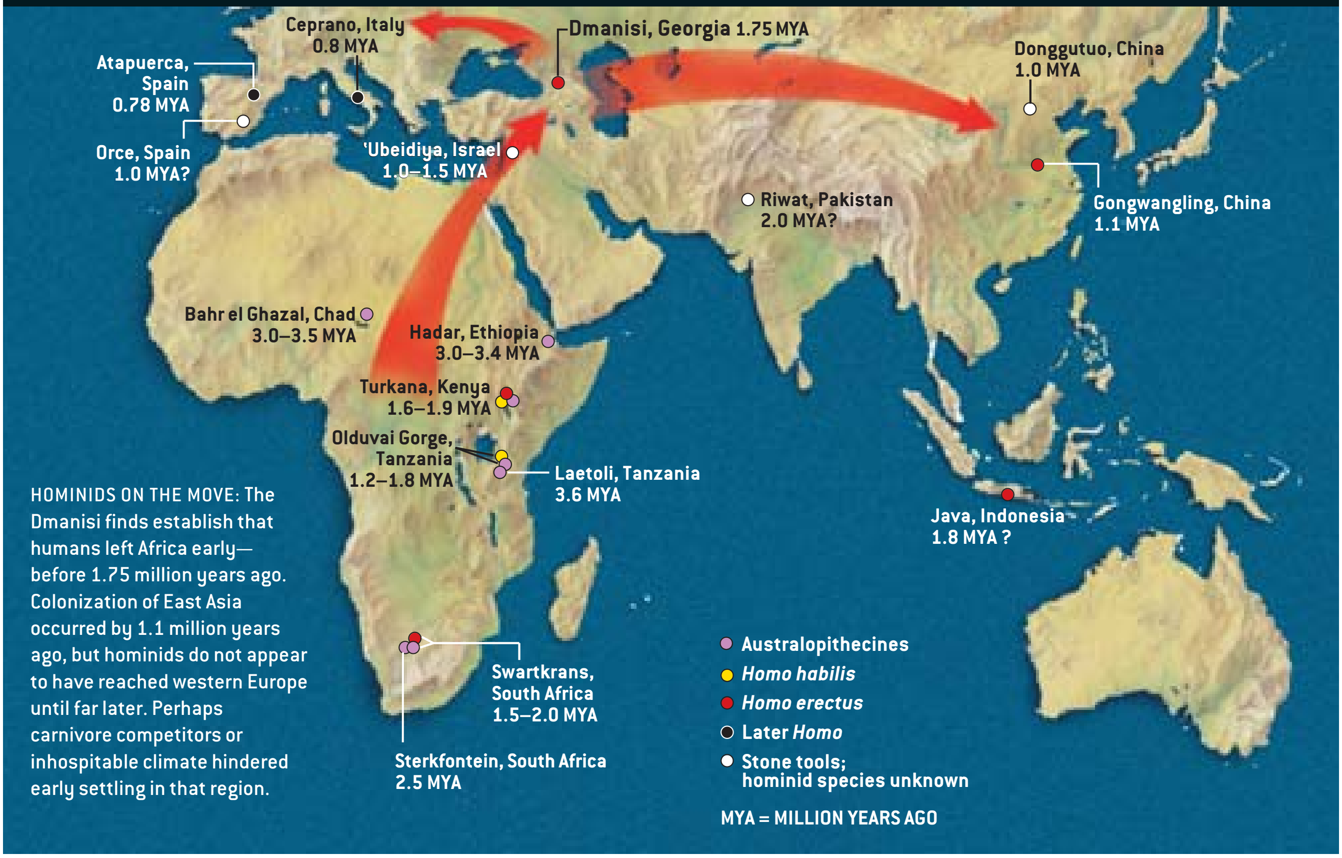
CORNELIA BLIK, PATRICIA J. WYNNE AND EDWARD BELL

TRIMMING THE FAMILY TREE

SPURIOUS SPECIES? Experts vigorously debate just how many species our genus, *Homo,* comprises. The bushiest representations of the *Homo* branch of the family tree contain up to eight species, a number of which were evolutionary dead ends (*top*). Other renditions appear as a streamlined succession of just a few forms (*bottom*). The fossils from Dmanisi—categorized variously as *H. habilis, H. erectus, H. ergaster* and a new species, *H. georgicus*—could be compatible with scenarios of substantial hominid diversity. Alternatively, the anatomical range evident in the Georgian remains could just underscore how variable a species can be. Viewed that way, some pruning may be in order.



AFRICAN EXODUS

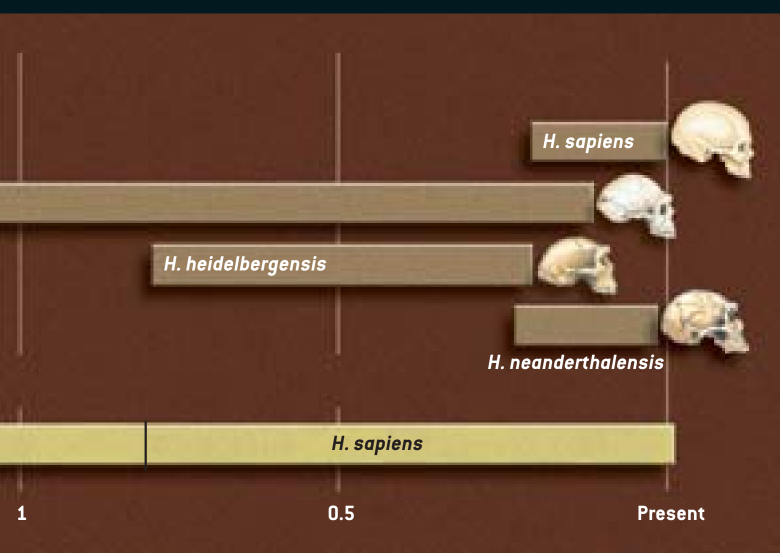


HOMINIDS ON THE MOVE: The Dmanisi finds establish that humans left Africa early—before 1.75 million years ago. Colonization of East Asia occurred by 1.1 million years ago, but hominids do not appear to have reached western Europe until far later. Perhaps carnivore competitors or inhospitable climate hindered early settling in that region.

EDWARD BELL

these early colonizers were proportioned and how efficiently they covered long distances. There is every reason to expect that they will find them. "They've got the potential to have truckloads of fossils," Wolpoff says enthusiastically. "There is work for generations here," Lordkipanidze agrees, noting that he can envision his grandchildren working at the site decades from now. Who knows what new frontiers humans will have explored by then? SA

Kate Wong is editorial director of ScientificAmerican.com



MORE TO EXPLORE

The Human Career: Human Biological and Cultural Origins. Second edition. Richard G. Klein. University of Chicago Press, 1999.

Earliest Pleistocene Hominid Cranial Remains from Dmanisi, Republic of Georgia: Taxonomy, Geological Setting, and Age. Leo Gabunia, Abesalom Vekua, David Lordkipanidze et al. in *Science,* Vol. 288, pages 1019–1025; May 12, 2000.

The Environmental Context of Early Human Occupation in Georgia (Transcaucasia). Leo Gabunia, Abesalom Vekua and David Lordkipanidze in *Journal of Human Evolution,* Vol. 38, No. 6, pages 785–802; June 2000.

A New Skull of Early Homo from Dmanisi, Georgia. Abesalom Vekua, David Lordkipanidze, G. Philip Rightmire et al. in *Science,* Vol. 297, pages 85–89; July 5, 2002.

Food for Thought. William R. Leonard in *Scientific American,* Vol. 287, No. 6, pages 106–115; December 2002.

A segment based on this article will air October 23 on *National Geographic Today,* a program on the National Geographic Channel. Please check your local listings.

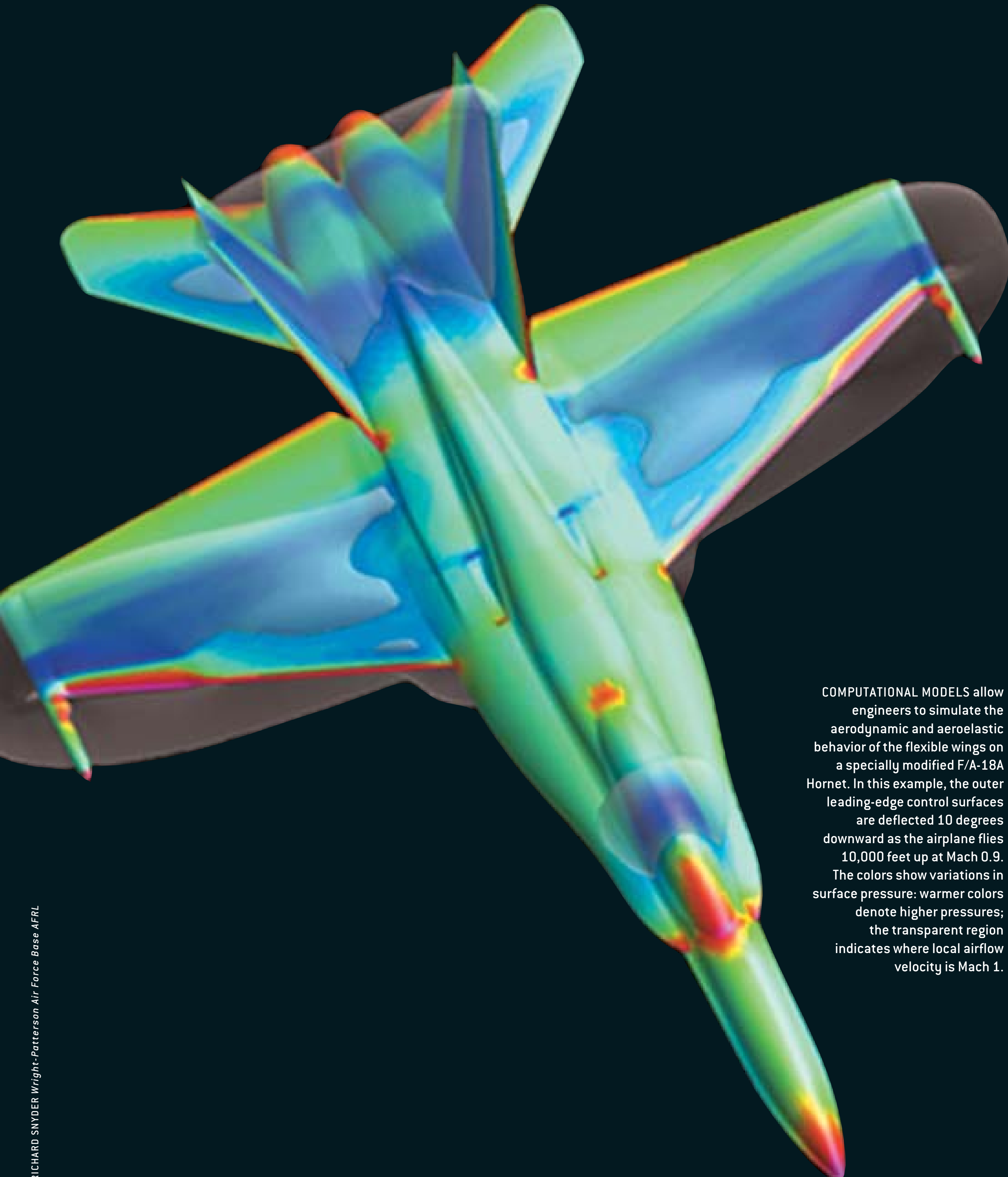
FUTURE AIRCRAFT MAY FLY more like birds, adapting the geometries

Flying on FLEXIBLE WINGS

By STEVEN ASHLEY

Airplanes typically look the same whether they are in the air or on the ground. Most wings extend from the fuselage at a fixed angle, and they are sufficiently rigid that they do not move or twist much in flight—certainly a reassuring feature for pilots and passengers alike.

In years to come, however, radical wing designs for advanced aircraft may change that. So-called morphing wings will be sophisticated structures that automatically reconfigure their shapes and surface textures to adapt to monitored changes in flying conditions. Such capabilities will in some of their wings **TO BEST SUIT CHANGING FLIGHT** conditions



COMPUTATIONAL MODELS allow engineers to simulate the aerodynamic and aeroelastic behavior of the flexible wings on a specially modified F/A-18A Hornet. In this example, the outer leading-edge control surfaces are deflected 10 degrees downward as the airplane flies 10,000 feet up at Mach 0.9. The colors show variations in surface pressure: warmer colors denote higher pressures; the transparent region indicates where local airflow velocity is Mach 1.

RICHARD SNYDER *Wright-Patterson Air Force Base AFRL* ways mimic the subtle, nearly instantaneous adjustments that birds make instinctively to their wings, tails and feathers when aloft.

Morphing wings will conform themselves to accommodate the mission and maneuvers at hand. A next-generation combat drone, for example, might loiter for hours above a potential target on elongated, low-drag wings. When the time comes to drop its weapons, the wings could sweep back and telescope down to a shorter length, preparing the craft for a fast attack run. Passenger airliners might similarly transform their wings in flight to save fuel or reach their destinations more rapidly.

Aircraft with that kind of flexibility are still far off and will involve materials and mechanisms that are barely off the drawing board now. A forerunner of those vehicles, though, is already being developed and tested. It takes advantage of aerodynamic forces generated by relatively conventional flaps and flight-control surfaces to cause less robust, lightweight wings to twist in flight to achieve optimal performance. These flexible wings mark a step forward for aviation—but also, in a sense, a step back.

TWISTING WINGS act to create more lift on one side of this 1911 Wright brothers' glider, permitting the pilot to roll the aircraft. This pioneering wing-warping technique continued in use until elevated flight speeds necessitated stronger, more rigid wing structures.

BETTMANN/CORBIS

Back to the Future

EVERY BEGINNING cyclist must learn to balance the vehicle to keep it upright and stable, no matter the maneuver. The same was true for the first airplane pilots.

Long before Wilbur and Orville Wright started building glider models, those determined bicycle mechanics realized that achieving controlled flight depended largely on finding means to regulate lateral (side-to-side) stability in the air. In particular, the brothers needed some way to govern rolling motions—rotations about a plane's longitudinal axis. Only through control of roll moments could a pilot maintain a level trajectory or dip a wing into a smooth banking turn.

By 1900 Wilbur had reported observing that "when partly overturned by a gust of wind," buzzards "regain their lateral balance . . . by a torsion of the tips of the wings." Twisting a wing's tip, he saw, alters its angle of attack into the oncoming air, which raises or lowers the wing's lift. This change redistributes the balance of lifting forces along the entire wingspan. Just as a rider's weight shifts can help sustain the equilibrium of a wobbly bicycle, actively balancing the lift forces on each wingtip is key to aircraft roll control. In time, Wilbur conceived an elegant cable-and-pulley mechanism that twisted a flexible fabric wing across its span so that one wingtip could generate greater lift and the other less. The force differential would tilt the craft toward the side with the reduced lift. Thus, the concept of wing warping was born.

Although aircraft designers and engineers continued to employ wing-warping

Overview/*Transforming Wings*

- Noting how birds change the shape of their wings in flight, the Wright brothers developed flexible aircraft wings, using cables and pulleys to modify wing lift to control rolling motions. As speeds increased during the succeeding years, builders adopted rigid wings that could better withstand the resulting aerodynamic stresses. These wings employed aileron flaps to create the differential lift required to roll aircraft.
- Today aeronautical engineers are investigating how lightweight, flexible wings could be twisted (using forces generated by leading-edge wing flaps) to boost the performance of modern aircraft.
- In the future, airplanes may radically alter wing shape as needed to better handle changing flying conditions, thus mimicking avian flight.

techniques for years afterward (especially on early monoplanes such as the Bleriot XI), ever faster flight speeds eventually led them to beef up wing structures to resist the resulting stresses. Stiff struts and ribs meant that wings were too rigid to flex much torsionally, so builders installed ailerons—flaps located on the outer, trailing-edge portions of wings. These hinged flight surfaces could deflect airflow to modify the lift, thereby supplying roll control. Most makers soon cast flexible wings aside. Only in slow-moving, lightweight flying machines, including model planes and pedal-powered aircraft (such as the English Channel–crossing Gossamer Albatross of 1979), has the use of wing warping continued.

After the early decades of flight, aircraft designers shifted from exploiting wing flexibility to actively avoiding it—and with good reason: the large aerodynamic pressures produced by high speeds can make wings with ailerons twist the wrong way. At lower velocities, the airflow-induced torquing forces lower a pilot's ability to induce a roll because wing twist lessens the amount of airflow the ailerons can deflect, explains James Guffey, a project manager at Boeing who used to oversee the modern effort to develop flexible wings for high-performance aircraft. At higher speeds, increased wingtip bending can bring about the disastrous phenomenon known as control reversal—when a pilot's stick command to roll in one direction results in the exact contrary maneuver. "Flexible wings can produce a rolling moment opposite to and greater than that applied by the control surfaces," Guffey says.

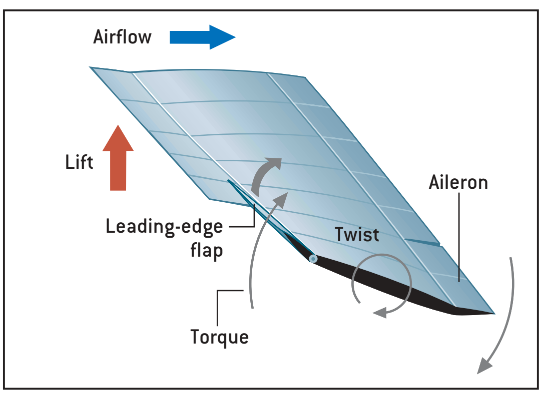
For these reasons, the wings on recent aircraft have generally been more robust and rigid than those of the early 20th century. As Guffey notes, "A long, slender wing, by its nature, would be lightweight and would provide very good lift and drag characteristics. The reason we don't have long, slender wings is people are concerned the inner structure of the wing will be overstressed. In the past, the response to twisting has been to remove it."

Elastic Wings

BUT A GREATER understanding of aerodynamics, along with space-age materials and electronics, has prompted researchers

A FLEXIBLE APPROACH TO FLIGHT

LEADING EDGE of wing on this experimental plane can be seen deflecting downward.



THE ACTIVE AEROELASTIC WING on this reengineered F/A-18A Hornet began its initial evaluation flights in November 2002 at the NASA Dryden Flight Research Center. The experimental aircraft—a joint project of NASA, the U.S. Air Force Research Laboratory and Boeing's Phantom Works—is designed to use deflection of modified leading-edge outer wing flaps to twist the entire wing for enhanced roll control. The demonstration vehicle is a reworked preproduction McDonnell Douglas F/A-18A fighter from the early 1980s. The prototype Hornet was known for poor roll performance at higher velocities, a phenomenon caused by its overly flexible wing. The active aeroelastic wing (AAW) program returned the prototype aircraft's rigid wing structure to its original, more supple state so the wing-warping controls would be effective. Engineers believe the new technology should boost maneuvering performance.

The diagram at the left describes how an airplane maneuvers using basic AAW technology. To roll the plane right, the left wing is twisted so the forward wing edge rises by deflecting the outer leading-edge flap upward. Greater roll force is produced by flexing the right wing leading-edge flap downward. Note that the aircraft's conventional ailerons can be rotated to enhance the twisting effect.

CARLA THOMAS NASA (photograph); SAMUEL VELASCO (illustration)

CRUISING ON CONVERTIBLE WINGS

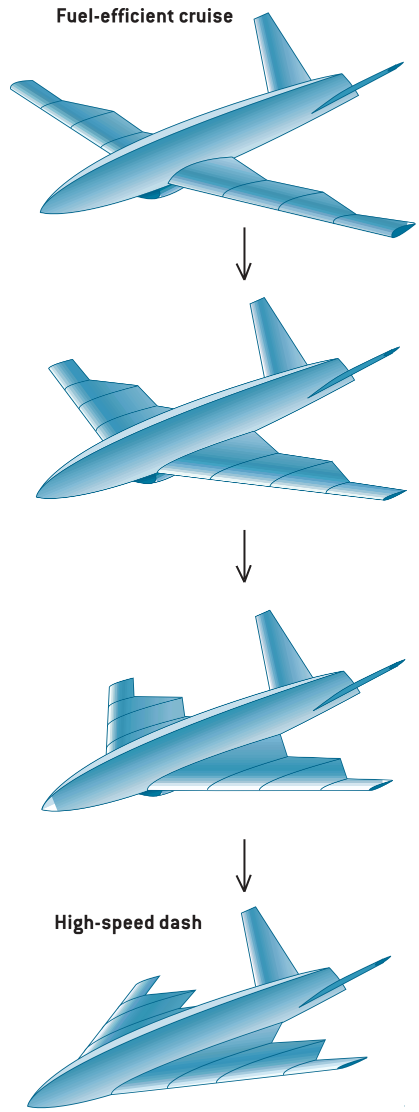
A PEREGRINE FALCON folds its wings tightly to its body as it dives—at remarkable speed—on its prey. When cruising, the raptor extends its wings fully outward, a conformation that saves energy. Engineers developing future aircraft designs would like to mimic these kinds of transformations to achieve better flight performance. They envision a revolutionary airplane that can morph its wing shape to convert the structure to a form that best exploits flight conditions.

NASA researchers are working on morphing-aircraft concepts [*below*] they say could become reality by 2030. They anticipate a machine that will be able to respond to constantly varying conditions using its sensors (somewhat as a bird uses its sensory nerves) to monitor pressure variations over the entire surface of the wing. The feedback response to these measurements will direct internal motion-generating actuators (which will function like muscles in a bird's wing) to alter wing shape to an optimal position.

ARTIST'S conception of morphing-wing plane

For high-speed flight, for instance, the wings would sweep back and change shape to reduce drag and quiet sonic booms. The engine inlets and nozzles could morph as well. Small jets of air and featherlike control surfaces could provide additional control forces for extreme maneuvers and added safety. To convert to a low-speed configuration, the wings would unsweep, as well as grow in thickness and span to improve efficiency. Instead of a vertical tail, the vehicle could use directed jet thrusts. As the vehicle prepares for a landing, the wingtips would split to better control energy-sapping tip vortices while the wings lengthened to permit a shorter landing on the runway. A tail could deploy to provide increased lift and additional control.

Nearer-term morphing-wing technology is the aim of a $25-million R&D program being sponsored by the Defense Advanced Research Projects Agency. One main contractor, NextGen Aeronautics, is working on an aircraft concept that is designed to transform from a low-speed cruise mode to a fast "dash" configuration and back again (*at right*).



to reconsider the advantages of flexibility for high-performance aircraft. A team from the U.S. government and the aviation industry has modified the wings of an F/A-18A Hornet fighter-bomber to make them less stiff, transforming the Hornet into the first conventional airplane in modern times with deliberately twisty wings. Late last year the team started putting the specially engineered jet through its paces high above the dry lake beds of California's Mojave Desert. In further flight tests over the next few years, the researchers plan to establish the benefits offered by the flexible—aeroelastic—wings. One of the first goals, for example, is to demonstrate the ability of the F/A-18's altered leading-edge wing flaps to generate torque forces. Engineers expect that precisely controlled front-edge flap deflections will twist the entire wing structure enough to roll the Hornet more effectively at high speeds than a standard F/A-18 wing can.

The active aeroelastic wing (AAW) project is an unusual joint effort involving the U.S. Navy, the U.S. Air Force Research Laboratory (AFRL) at Wright-Patterson Air Force Base, the NASA Dryden Flight Research Center and Boeing's Phantom Works. The $45 million in funding for the experimental craft comes from AFRL's Air Vehicles Directorate and NASA's Office of Aerospace Technology, with Boeing and subcontractor personnel performing the modifications to the navy-donated F/A-18A under contract to the air force.

Staying Flexible

A QUARTER OF a century ago, when the F/A-18 was first being tested, the prototype had problems rolling at high speeds because its wings twisted excessively. Test pilots suffered loss of control

NASA (*image*); SAMUEL VELASCO (*illustration*) effectiveness and finally aileron control reversal. The trouble stemmed from engineers' inexperience with the new lightweight composite materials used to substitute for metals in the wing structure, recalls David Riley, who succeeded Guffey as the AAW program manager at Boeing. "The F/A-18 was the aerospace industry's first big foray into the structural use of polymer composites, and the result came up somewhat short in terms of torsional stiffness," Riley says.

Managers of the Hornet program sent the flawed preproduction wings into storage and replaced them with a design that incorporated stronger ribs and thicker, stiffer wing-skin panels. Unfortunately, the extra structural weight degraded the plane's performance. "In essence, the airplane had a frustrated wing for the past 25 years," Guffey comments. The engineers also added to the prototype what is called a rolling tail, or stabilator (a combination stabilizer and elevator), a device employed by most other top-line fighter jets to provide further roll authority at high dynamic pressures.

What was unacceptable in a front-line fighter jet in the 1980s, however, is ideal for the experimental aerodynamics platform today. The AAW study team saw that they might take advantage of the past liabilities of the first F/A-18's floppy wings. The navy was willing to donate the airplane; the too-flexible preproduction wings were still available, which saved the cost of designing new ones, and the plane's structure was already flight-tested. Moreover, the researchers had access to a wealth of flight-control data on the F/A-18 from previous research programs. Finally, the F/A-18A features wing-fold mechanisms that save deck space during aircraft carrier operations, which means it has separate leading-edge flaps rather than one large one. Although the leading-edge flaps on the inner and outer segments of the standard Hornet wing act in unison, these could be readily reworked to function independently—a key to flying with active aeroelastic wings.

The approach being used to create an AAW is based on an idea conceived in 1983 by Jan Tulinius of the North American Aircraft Operations division of Rockwell International, a company that eventually merged with Boeing. "He came up with the notion that the wing's going to twist anyway, so why don't we see if we can make it work to our benefit?" says Peter Flick, air force AAW program manager. "Rather than counteracting the twisting effect, let's use it." From 1984 to 1988 a team of researchers succeeded in proving out the concept in the wind tunnel.

The reasoning behind the AAW is straightforward: First, a rigid wing is heavier than a flexible one because extra reinforcement must be added to stiffen it. Second, a fully flexible wing can provide much more surface area to divert airflow, and so create roll force, than can comparatively small conventional ailerons. By employing the energy available in the airflow field to simultaneously control the twist and the camber (the slight arch) of the wing, team members think they can attain better performance. Because such wings would require fewer moving parts for controlling flight, they could be made thinner, lighter and more aerodynamically efficient than today's counterparts, thus allowing for greater range, payloads and fuel efficiency. Surprisingly, an AAW may actually twist less overall than a conventional wing during maneuvering, the researchers claim.

Estimates of projected weight savings are encouraging. One design exercise suggested that aeroelastic wings might cut the weight of a future transonic fighter (one that flies near the speed of sound) by 7 to 10 percent or that of an advanced supersonic fighter by as much as 18 percent. The researchers also believe that AAW concepts may allow drag-inducing and radar-reflecting aircraft tail surfaces to be eliminated altogether because their functions can be accomplished by the wings.

"The WING'S GOING TO TWIST anyway, so why don't we see if we can make it work to OUR BENEFIT?"

Let's Do the Twist

THE FIRST FLIGHTS of the revamped Hornet followed a three-year period of reengineering and ground tests at the NASA Dryden facility. Technicians at Boeing's Phantom Works modified the wings of the F/A-18A test aircraft with additional actuators, a split leading-edge flap and thinner wing skins, which allow the outer wing panels to twist as much as five degrees (much more than normal).

The key to successful modern wing warping, Guffey says, is to ensure that the mechanical stresses are watched closely and negated quickly. "We have a couple hundred sensors distributed on both the wings and fuselage, so we can monitor all the structural loading in flight. By using the control surfaces—the flaps, ailerons and so forth—we can redistribute any loading the wing feels from twisting, so we don't overload it." Without load alleviation, material fatigue would damage the structure over time.

The preliminary data-gathering flights completed earlier this year are allowing engineers to create control laws that exploit the wings' aeroelasticity, explains Larry Myers, NASA's AAW project manager. "One by one, we evaluated the effect of activating each surface singly," he says. Team members measured the performance of the wing and its structural response to levels of deflection. "On the leading-edge flaps, for instance, we went from plus three degrees [up] to minus three degrees [down]." Then, skilled pilots put the experimental plane through a gut-wrenching series of dogfighting maneuvers intended to test out the technology.

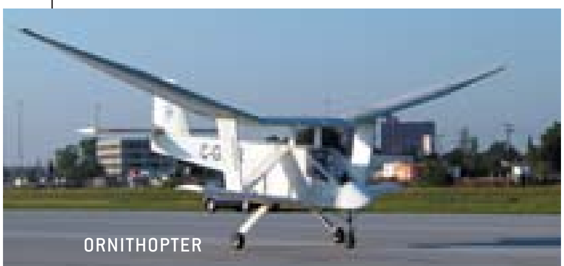
Through careful measurement of the aeroelastic load and maneuverability effects observed during the test flights,

FLAPPING LIKE A BIRD

FLAPPING-WING FLIGHT ENTHUSIASTS based at the University of Toronto Institute for Aerospace Studies are using a specialized flexible-wing structure to achieve humanity's oldest aeronautical dream—to (really) fly like a bird. Led by aeronautics professor James DeLaurier, the underfunded, student-driven team has been working since 1995 to get its ornithopter off the ground (*photograph*). After rebuilding some structures damaged in a previous attempt last year, team members hope to make another try this year—or the next.

Key to their flying machine is a "shearflexing" wing, an innovative concept that was conceived by DeLaurier's long-term collaborator in ornithopter research, Jeremy M. Harris, whom he met in 1973 when they both worked at Battelle Memorial Institute in Columbus, Ohio. By 1976, Harris writes, the pair had begun "to consider the interrelated problems of airfoil section shape and pitching freedom"—the ability of an airfoil to change its angle of attack into the airflow.

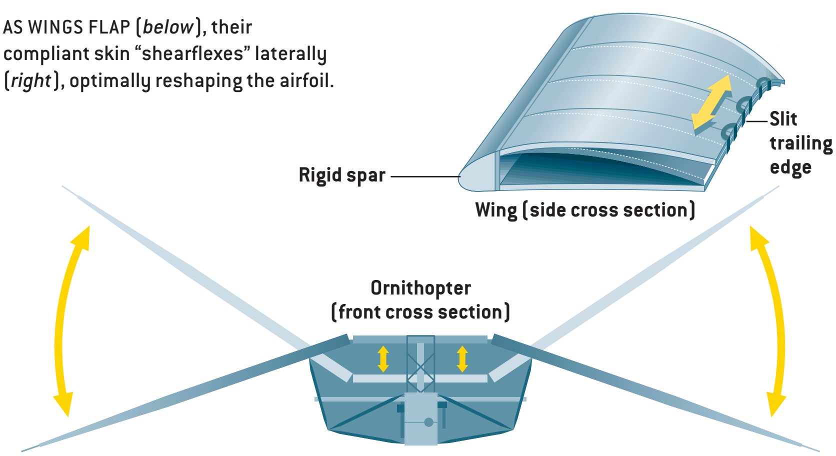
In an ornithopter, Harris states, the ideal mode for pitching along the outer wing sections takes the form of a linear twist that makes each part of the wing just avoid stalling (losing lift because its angle of attack drops below a critical point) at every instant during the flapping (up-down) cycle. "The action has to occur dynamically—that is, the wing has to go from positive to negative twist and back each time the wings flap," he explains. Just how one could achieve that complex motion with something approaching a conventional wing structure was the question.


ORNITHOPTER

In 1979, according to Harris, "we pondered the well-known fact that a circular tube loses a lot of its torsional rigidity if it is slit lengthwise. When a carefully slit tube is slipped over a supporting shaft and torqued, the mating edges of the slit slide smoothly along each other with a shearing action, giving torsional deflections unobtainable with an unslit tube." His solution? Turn the slit cylindrical tube into the entire wing so that the slit occurs at the trailing edge of a "hollow" airfoil with a compliant skin (*diagram*).

Subsequent testing showed that the concept was valid—the angle of attack of the airfoil varies as desired as the wing swings up and down because shearflexing allows the structure to twist around the rigid wing spar to maintain lift.

AS WINGS FLAP (*below*), their compliant skin "shearflexes" laterally (*right*), optimally reshaping the airfoil.



AAW engineers can now digitally model the dynamic phenomena associated with flexible wings, says David Vorcek, NASA AAW chief engineer. Armed with these data, they can develop appropriate control laws and fast-acting, computerized flight controls that can avoid adverse effects and produce beneficial ones, he explains.

In the next round of flight tests, slated for the spring and summer of 2004, the AAW team plans to demonstrate aeroelastic effects with the novel wings, achieving, it is hoped, higher roll rates than the standard F/A-18 can without engaging its stabilator tail. If those tests are successful, team members expect to develop benchmark design criteria for next-generation aircraft designs. Says Flick: "It just so happens that when speeds increase to the point that conventional trailing-edge surfaces [ailerons] are becoming ineffective, the leading-edge surfaces become effective in twisting the wing."

Another possibility, according to Riley, is to employ active feedback to suppress wing flutter—a potentially dangerous cyclic reaction to aerodynamic forces that can worsen quickly. He thinks that some kind of digital learning system (such as a neural network) might be used to observe the onset of the problem and then counteract it.

"AAW is applicable to a wide variety of future air vehicle concepts that are under study—not just to supersonic flight," Flick notes. The researchers can envision using AAW on high-flying, long-distance, high-endurance surveillance and attack aircraft, both manned and unmanned. Aeroelastic wing technologies may also find use on commercial airliners.

The Shapes of Wings to Come

IN SOME WAYS, experts see the AAW program as a first step on the lengthy path to a true morphing wing that can sense its environment and adapt its shape to perform optimally in a wide range of flight conditions. Clearly, the complex landing flap and lift-augmenting front slot systems of current airplanes, as well as the angle-changing swing wings of the F-14 Tomcat and the F-111 Aardvark, provide a degree of in-flight reconfiguration capability. But these technologies and even the

JAMES DeLAURIER *Project Ornithopter*

AAW fall short of what is contemplated for the future. The benefits of swing-wing designs, for example, were hobbled by the large weight penalty associated with the necessary moving parts. A multimission aircraft with a shape-changing wing would require the development of "smart" materials for sensing and motion generation and, perhaps, entirely new varieties of flight-control mechanisms.

That kind of technology is the goal of a $25-million research and development program being sponsored by the Defense Advanced Research Projects Agency (DARPA). (Work of this nature is also ongoing at NASA and the German Aerospace Center.) Various types of materials are being eyed to warp wings from the inside rather than relying on the outside air. Among them are shape-memory alloys, which respond to temperature changes, and piezoelectrics, electroactive and magnetostrictive materials, which contract or expand with the application of magnetic fields or electric current [see "Artificial Muscles," by Steven Ashley; SCIENTIFIC AMERICAN, October]. Researchers caution that these technologies are not sufficiently mature to depend on just yet.

The scale of geometric adjustments that DARPA managers envision includes a 200 percent change in aspect ratio (the square of the wingspan divided by the wing area), a 50 percent change in wing surface area, a five-degree alteration in wing twist, and a 20-degree change in wing sweep (angle to the fuselage). Furthermore, the final weight of the wing should not exceed that of a conventional wing structure. DARPA expects that the subsystems and components for the adaptive wing mechanisms will be integrated into prototype aircraft by late 2004, ready for subsequent wind-tunnel testing.

One of the attractions of using shape changes rather than traditional flight-control surfaces in an aircraft is that flaps and the like are large radar reflectors. The radar-evading B-2 Spirit, for instance, uses split trailing-edge flaps for control, which compromise the bomber's stealth characteristics when deployed.

Stretching One's Wings

THREE MAIN CONTRACTORS—Lockheed Martin, Raytheon Missile Systems and NextGen Aeronautics in Torrance, Calif.—as well as several university-based research programs are studying morphing-wing technology. Lockheed Martin is developing a small unmanned aerial vehicle (UAV) with folding wings for an air force mission; the Raytheon project focuses on a navy Tomahawk cruise missile with telescoping wings.

The NextGen research effort has more general aims and uses a Northrop Grumman Firebee drone as its base design platform. "We are working on techniques that can change the wing area and morph a low-drag, high-aspect-ratio wing into something that can sustain high maneuver rates," says Jayanth Kudva, NextGen's president. "We plan to tuck in the wings, then sweep them back and at the same time reduce the wing area" [*see illustration on page 88*].

Although NextGen researchers are considering multiple design concepts to accomplish these in-flight alterations, Kudva says that his company's "somewhat organic approach to stretching and expanding the wing has several building blocks." These include developing a flexible skin for the wing that could smoothly accommodate desired changes in the surface area; a nontraditional substructure that could carry the load yet still expand or contract as needed; a distributed system of actuators (movement generators); and an appropriate control system.

"If we succeed in doing this cleverly enough, it may be possible to get rid of conventional control surfaces and use the differential area change to control the airplane, which would compensate for any additional weight penalty caused by the new mechanisms," he adds. For rolling maneuvers, one could expand the area of one wing slightly to get more lift on that side and thus roll the plane.

"Although I'm somewhat biased," Kudva concludes, "I would expect that a UAV with a basic morphing wing would be in production within a decade or so." Passenger-carrying aircraft with morphing wings could follow soon thereafter. If that happens, birdlike manned flight will have come full circle in a bit more than a century. SA

> "We plan to tuck in the wings, then sweep them back, and at the same time REDUCE THE WING AREA."

Steven Ashley is a staff writer and editor.

MORE TO EXPLORE

Application of AFW Technology to the Agile Falcon. E. Pendleton, M. Lee and L. Wasserman in *Journal of Aircraft,* Vol. 29, No. 3, pages 444–457; May-June 1992.

An Active Flexible Wing Multi-Disciplinary Design Optimization Method. G. D. Miller. AIAA Paper No. 94-4412-CP, 1994.

Summary of an Active Flexible Wing Program. B. Perry III, S. R. Cole and G. D. Miller in *Journal of Aircraft,* Vol. 32, No. 1, pages 10–31; January-February 1995.

Letter from Wilbur Wright to aviation pioneer Octave Chanute: **http://invention.psychology.msstate.edu/i/Wrights/library/Chanute_Wright_correspond/May13-1900.html**

U.S. Air Force report on AAW program: **www.afrlhorizons.com/Briefs/Mar03/HQ0210.html**

Boeing report on AAW program: **www.boeing.com/news/frontiers/archive/2002/may/i_pw.html**

NASA report on morphing-wing research: **www.dfrc.nasa.gov/Newsroom/X-Press/stories/043001/new_morph.html**

NASA report on AAW and morphing-wing research: **www.nasa.gov/missions/research/twist_wing.html**

University of Toronto ornithopter project: **www.ornithopter.net**

WHY WE SLEEP

The reasons that we sleep are gradually becoming less enigmatic

By Jerome M. Siegel

Birds do it, bees do it, and, in a departure from the Cole Porter song lyrics, even fruit flies appear to do it. Humans certainly do it. The subject is not love, but sleep. Shakespeare's Macbeth said it "knits up the raveled sleave of care" and was the "balm of hurt minds, great nature's second course, chief nourisher in life's feast." Cervantes's Sancho Panza sang its praises as "the food that cures all hunger, the water that quenches all thirst, the fire that warms the cold, the cold that cools the heart … the balancing weight that levels the shepherd with the king, and the simple with the wise."

The simple and the wise have long contemplated two related questions: What is sleep, and why do we need it? An obvious answer to the latter is that adequate sleep is necessary to stay alert and awake. That response, however, dodges the issue and is the equivalent of saying that you eat to keep from being hungry or breathe to ward off feelings of suffocation. The real function of eating is to supply nutrients, and the function of breathing is to take in oxygen and expel carbon dioxide. But we have no comparably straightforward explanation for sleep. That said, sleep research—less than a century old as a focused field of scientific inquiry—has generated enough insights for investigators to at least make reasonable proposals about the function of the somnolent state that consumes one third of our lives.

What Is Sleep?

U.S. SUPREME COURT JUSTICE Potter Stewart's famous quote about obscenity—"I know it when I see it"—is a useful, if incomplete, guideline about sleep. Despite the difficulty in strictly defining sleep, an observer can usually tell when a subject is sleeping: the sleeper ordinarily exhibits relative inattention to the environment and is usually immobile. (Dolphins and other marine mammals swim while sleeping, however, and some birds may sleep through long migrations.)

In 1953 sleep research pioneer Nathaniel Kleitman and his student Eugene Aserinsky of the University of Chicago decisively overthrew the commonly held belief that sleep was simply a cessation of most brain activity. They discovered that sleep was marked by periods of rapid eye movement, commonly known now as REM sleep. And its existence implied that something active occurred during sleep. All terrestrial mammals that have been examined exhibit REM sleep, which alternates with non-REM sleep, also called quiet sleep, in a regular cycle.

More recently, the field has made its greatest progress in characterizing the nature of sleep at the level of nerve cells (neurons) in the brain. In the past 20 years, scientists have mastered techniques for guiding fine microwires (only 32 microns wide, comparable to the thinnest of human hair) into various brain regions. Such wires produce no pain once implanted and have been used in humans as well as in a wide range of laboratory animals while they went about their normal activities, including sleep. These studies showed, as might be expected, that most brain neurons are at or near their maximum levels of activity while the subject is awake. But neuronal doings during sleep are surprisingly variable. Despite the similar posture and inattention to the environment that a sleeper shows during both

MINDY JONES

REM and non-REM sleep, the brain behaves completely differently in the two states.

During non-REM sleep, cells in different brain regions do very different things. Most neurons in the brain stem, immediately above the spinal cord, reduce or stop firing, whereas most neurons in the cerebral cortex and adjacent forebrain regions reduce their activity by only a small amount. What changes most dramatically is their overall pattern of activity. During the awake state, a neuron more or less goes about its own individual business. During non-REM sleep, in contrast, adjacent cortical neurons fire synchronously, with a relatively low frequency rhythm. (Seemingly paradoxically, this synchronous electrical activity generates higher-voltage brain waves than waking does. Yet just as in an idling automobile, less energy is consumed when the brain "idles" in this way.) Breathing and heart rate tend to be quite regular during non-REM sleep, and reports of vivid dreams during this state are rare.

A very small group of brain cells (perhaps totaling just 100,000 in humans) at the base of the forebrain is maximally active only during non-REM sleep. These cells have been called sleep-on neurons and appear to be responsible for inducing sleep. The precise signals that activate the sleep-on neurons are not yet completely understood, but increased body heat while an individual is awake clearly activates some of these cells, which may explain the drowsiness that so often accompanies a hot bath or a summer day at the beach.

On the other hand, brain activity during REM sleep resembles that during waking. Brain waves remain at low voltage because neurons are behaving individually. And most brain cells in both the forebrain and brain stem regions are quite active, signaling other nerve cells at rates as high as—or higher than—rates seen in the waking state. The brain's overall consumption of energy during REM sleep is also as high as while awake. The greatest neuronal activity accompanies the familiar twitches and eye motion that give REM sleep its name. Specialized cells located in the brain stem, called REM sleep-on cells, become especially active during REM sleep and, in fact, appear to be responsible for generating this state.

Our most vivid dreams occur during REM sleep, and dreaming is accompanied by frequent activation of the brain's motor systems, which otherwise operate only during waking movement. Fortunately, most movement during REM sleep is inhibited by two complementary biochemical actions involving neurotransmitters, the chemicals that physically carry signals from one neuron to another at the synapse (the contact point between two neurons). The brain stops releasing neurotransmitters that would otherwise activate motoneurons (the brain cells that control muscles), and it dispatches other neurotransmitters that actively shut down those motoneurons. These mechanisms, however, do not affect the motoneurons that control the muscles that move the eyes, allowing the rapid eye movements that give the REM sleep stage its name.

REM sleep also profoundly affects brain systems that control the body's internal organs. For example, heart rate and breathing become irregular during REM sleep, just as they are during active waking. Also, body temperature becomes less finely regulated and drifts, like that of a reptile, toward the environmental temperature. In addition, males often get erections and females experience clitoral enlargement, although most dream content is not sexual.

This brief description of sleep at the gross and neuronal levels is both accurate and as unsatisfying as being awakened before the completion of a good night's slumber. The tantalizing question persists: What is sleep for?

The Function of Sleep

AT A RECENT SLEEP conference, an attendee commented that the function of sleep remains a mystery. The chair of the session argued vehemently against that position—she did not, however, provide a concrete description of exactly *why* sleep's function was no longer mysterious. Clearly, no general agreement yet exists. But based on the currently available evidence, I can put forth what many of us feel are some reasonable hypotheses.

One approach to investigating the function of sleep is to see what physiological and behavioral changes result from a lack of it. More than a decade ago it was found that total sleep deprivation in rats leads to death. These animals show weight loss despite greatly increased food consumption, suggesting excessive heat loss. The animals die, for reasons yet to be explained, within 10 to 20 days, faster than if they were totally deprived of food but slept normally.

In humans, a very rare degenerative brain disease called fatal familial insomnia leads to death after several months. Whether the sleep loss itself is fatal or other aspects of the brain damage are to blame is not clear. Sleep deprivation studies in humans have found that sleepiness increases with even small reductions in nightly sleep times. Being sleepy while driving or during other activities that require continuous vigilance is as dangerous as consuming alcohol prior to those tasks. But existing evidence indicates that "helping" people to increase sleep time with long-term use of sleeping pills produces no clear-cut health benefit and may actually shorten life span. (About seven reported hours of sleep a night correlates with longer life spans in humans.) So inexorable is the drive to sleep that achieving total sleep deprivation requires repeated and intense stimulation. Researchers employing sleep deprivation to study sleep function are therefore quickly confronted with the difficulty of distinguishing the effects of stress from those of sleep loss.

Researchers also study the natural sleep habits of a variety of organisms. An important clue about the function of sleep is the huge variation in the amount that different species need. For example, the opossum sleeps for 18 hours a day, whereas the elephant gets by with only three or four. Closely related species that have genetic, physiological and behavioral similarities might also be expected to have similar sleep habits. Yet studies of laboratory, zoo and wild animals have revealed that sleep times are unrelated to the animals' taxonomic classification: the range of sleep times of different primates extensively overlaps that of rodents, which overlaps that of carnivores, and so on across many orders of mammals. If evolutionary relatedness does not determine sleep time, then what does?

The extraordinary answer is that size is the major determinant: bigger animals simply need less sleep. Elephants, giraffes and large primates (such as humans) require relatively little sleep; rats, cats, voles and other small animals spend most of their time sleeping. The reason is apparently related to the fact that small animals have higher metabolic rates and higher brain and body temperatures than large animals do. And metabolism is a messy business that generates free radicals—extremely reactive chemicals that damage and even kill cells. High metabolic rates thus lead to increased injury to cells and the nucleic acids, proteins and fats within them.

Free-radical damage in many body tissues can be dealt with by replacing compromised cells with new ones, produced by cell division; however, most brain regions do not produce significant numbers of new brain cells after birth. (The hippocampus, involved in learning and memory, is an important exception.) The lower metabolic rate and brain temperature occurring during non-REM sleep seem to provide an opportunity to deal with the damage done during waking. For example, enzymes may more efficiently repair cells during periods of inactivity. Or old enzymes, themselves altered by free radicals, may be replaced by newly synthesized ones that are structurally sound.

Last year my group at the University of California at Los Angeles observed what we believe to be the first evidence for

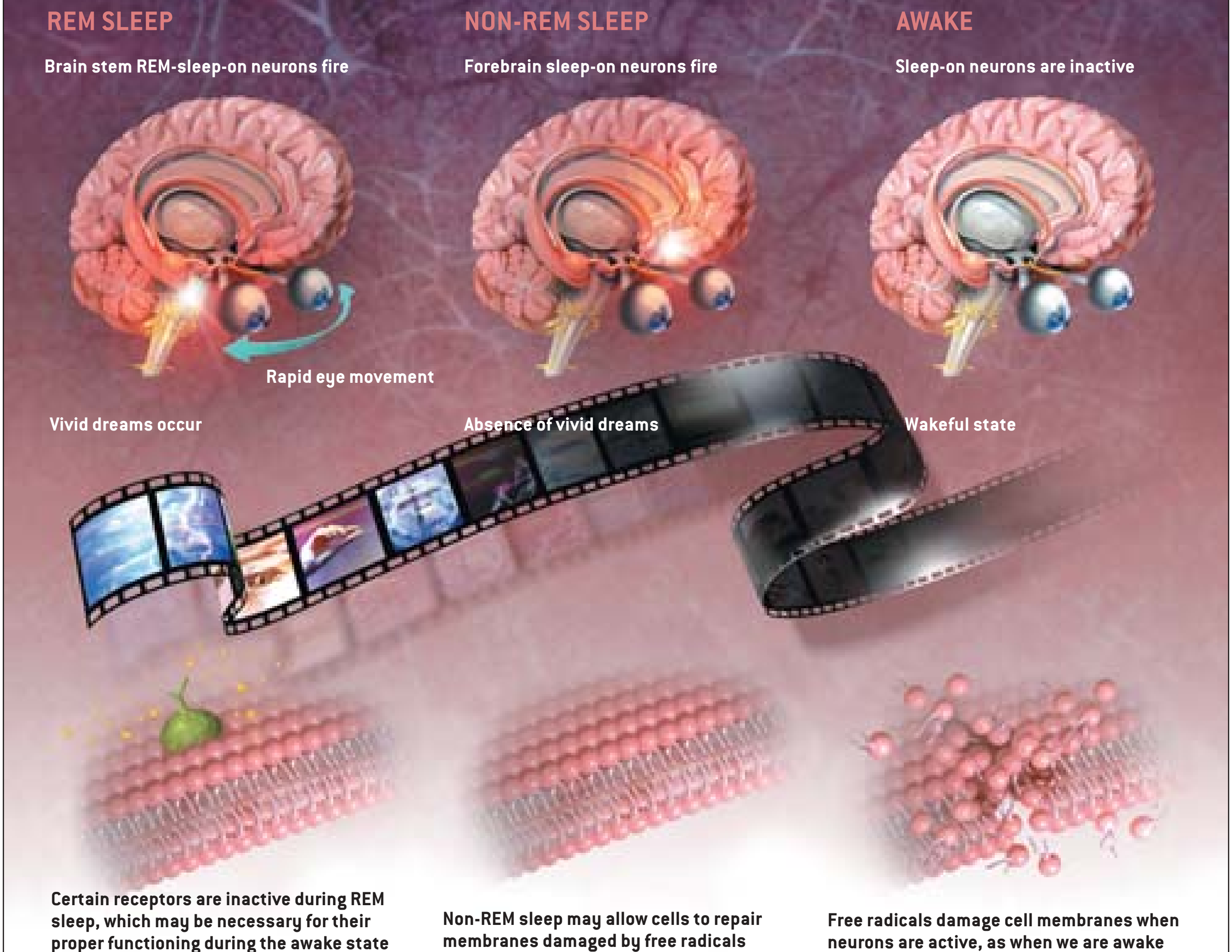
REM sleep is the proverbial riddle wrapped in a MYSTERY inside an ENIGMA.

Overview/*Uncovering Sleep*

- Researchers are still debating the function of REM and non-REM sleep and why we need both, but new findings suggest several reasonable hypotheses.
- One is that reduced activity during non-REM sleep may give many brain cells a chance to repair themselves.
- Another is that interrupted release of neurotransmitters called monoamines during REM sleep may allow the brain's receptors for those chemicals to recover and regain full sensitivity, which helps with regulation of mood and learning.
- The intense neuronal activity of REM sleep in early life may allow the brain to develop properly.

Sleeping, Dreaming, Waking

REM AND NON-REM SLEEP differ in several ways, some of which are illustrated below, along with one of the proposed functions of each type of sleep.



KEITH KASNOT

brain cell damage, in rats, occurring as a direct result of sleep deprivation. This finding supports the idea that non-REM sleep wards off metabolic harm.

REM sleep, however, is the proverbial riddle wrapped in a mystery inside an enigma. The cell-repair hypothesis could explain non-REM sleep, but it fails to account for REM sleep. After all, downtime repair cannot be taking place in most brain cells during REM sleep, when these cells are at least as active as during waking. But a specific group of brain cells that goes against this trend is of special interest in the search for a purpose of REM sleep.

Recall that the release of some neurotransmitters ceases during REM sleep, thereby disabling body movement and reducing awareness of the environment. The key neurotransmitters affected—norepinephrine, serotonin and histamine—are termed monoamines, because they each contain a chemical entity called an amine group. Brain cells that make these monoamines are maximally and continuously active in waking. But Dennis McGinty and Ronald Harper of U.C.L.A. discovered in 1973 that these cells stop discharging completely during REM sleep.

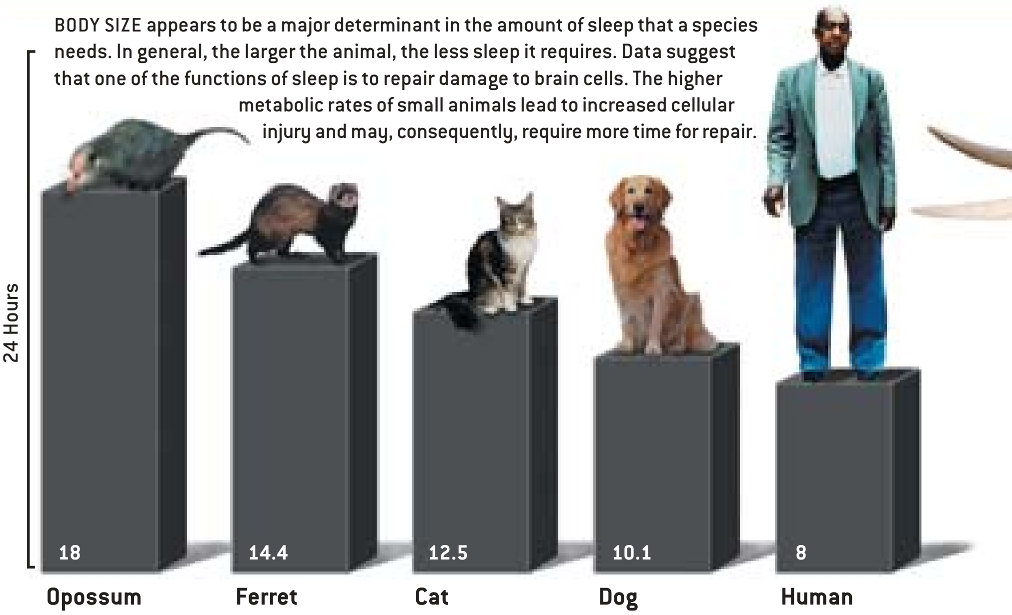
In 1988 Michael Rogawski of the National Institutes of Health and I hypothesized that the cessation of neurotransmitter release is vital for the proper function of these neurons and of their receptors (the molecules on recipient cells that relay neurotransmitters' signals into that cell). Various studies indicate that a constant release of monoamines can desensitize the neurotransmitters' receptors. The interruption of monoamine release during REM sleep thus may allow the receptor systems to "rest" and regain full sensitivity. And this restored sensitivity may be crucial during waking for mood regulation, which depends on the efficient collaboration of neurotransmitters and their receptors. (The familiar antidepressants Prozac, Paxil, Zoloft and other so-called selective serotonin reuptake inhibitors—SSRIs—work by causing a net increase in the amount of serotonin available to recipient cells.)

The monoamines also play a role in rewiring the brain in response to new experiences. Turning them off during REM sleep then may be a way to prevent changes in brain connections that might otherwise be inadvertently created as a result of other brain cells' intense activity during REM.

Interestingly, in 2000 Paul J. Shaw and his colleagues at the Neurosciences Institute in La Jolla, Calif., noted a connection in fruit flies between monoamine levels and sleeplike periods, during which the insects are relatively inactive. They found that disrupting the flies' downtime led to increased levels of monoamines, as is the case in humans. This discovery suggests that restoration of neurotransmitter function, eventually to become an attribute of what we now know as sleep, came into being well before mammals even evolved on the earth.

Counting Sleep

BODY SIZE appears to be a major determinant in the amount of sleep that a species needs. In general, the larger the animal, the less sleep it requires. Data suggest that one of the functions of sleep is to repair damage to brain cells. The higher metabolic rates of small animals lead to increased cellular injury and may, consequently, require more time for repair.



Other Possibilities

WHAT ELSE MIGHT REM sleep do? Researchers such as Frederick Snyder and Thomas Wehr of the National Institutes of Health and Robert Vertes of Florida Atlantic University have proposed that the elevated activity during REM sleep of brain cells that are not involved in monoamine production enables mammals to be more prepared than reptiles to cope with dangerous surroundings. When waking in a cold environment, reptiles are sluggish and require an external heat source to become active and responsive. But even though mammals do not thermoregulate during REM sleep, the intense neuronal activity during this phase can raise brain metabolic rate, helping mammals to monitor and react more quickly to a given situation on waking. The observation that humans are much more alert when awakened during REM sleep than during non-REM periods supports this idea.

Sleep deprivation studies indicate, however, that REM sleep must do more than prime the brain for waking experience. These studies show that animals made to go without REM sleep will undergo more than the usual amount when they are finally given the opportunity. They apparently seek to make up the "debt"—yet another clue that REM sleep is important. Of course, if brain arousal were the only function of REM sleep, being awake should also pay back the debt, because the waking brain is also warm and active. But wakefulness clearly does not accomplish this task. Perhaps REM sleep debt results from the need to rest monoamine systems or other systems that are "off" in REM sleep.

Old ideas that REM sleep deprivation led to insanity have been convincingly disproved (although studies show that depriving someone of sleep, for example by prodding him or her awake repeatedly, can definitely cause irritability). In fact, REM sleep deprivation can actually alleviate clinical depression. The mechanism for this phenomenon is unclear, but one suggestion is that the deprivation mimics the effects of SSRI antidepressants: because the normal decrease in monoamines during REM does not occur, the synaptic concentration of neurotransmitters that are depleted in depressed individuals increases.

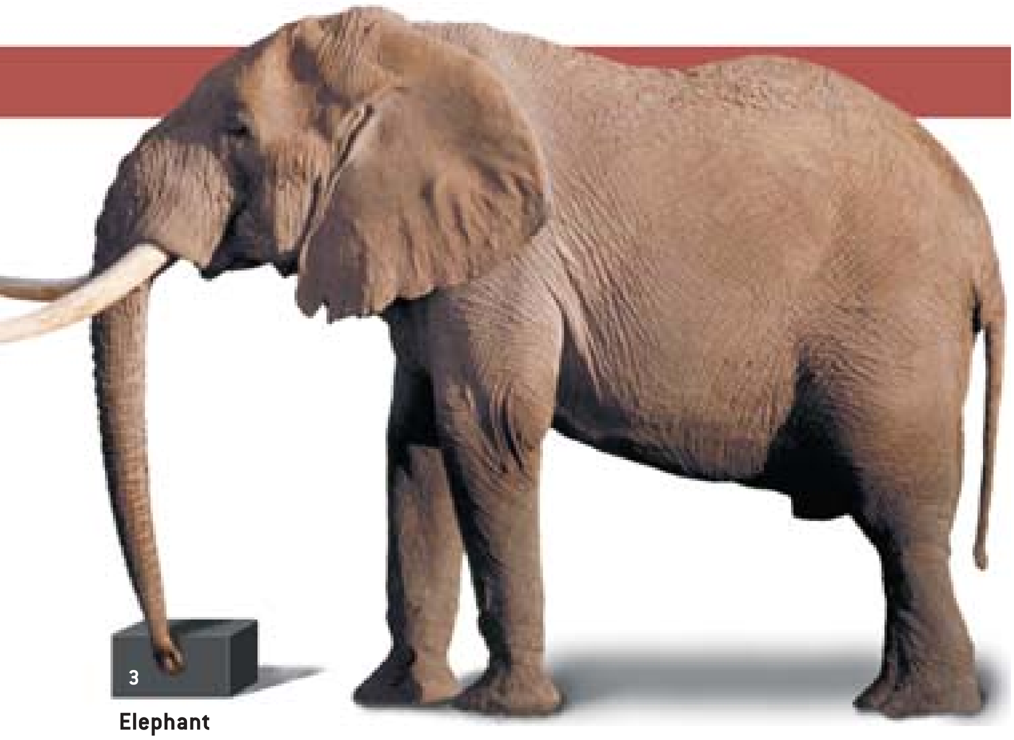
Some researchers are pursuing the idea that REM sleep might have a role in memory consolidation, but as I examined in detail in a 2001 article in *Science* [see "More to Explore" on opposite page], the evidence for that function is weak and contradictory. The findings that argue against memory consolidation include the demonstration that people who have brain damage that prevents REM sleep, or who have a drug-induced blockade of REM sleep, have normal—or even improved—memory. And although sleep deprivation before a task disturbs concentration and performance—sleepy students do not learn or think well—REM deprivation after a period of alert learning does not appear to interfere with retaining the new information. In addition, dolphins experience little or no REM sleep yet exhibit impressive reasoning and learning ability.

In fact, learning ability across species does not appear to be related to total REM sleep duration. Humans do not

THE AUTHOR

JEROME M. SIEGEL, professor of psychiatry and a member of the Brain Research Institute at the University of California at Los Angeles Medical Center, is chief of neurobiology research at Sepulveda Veterans Affairs Medical Center. Siegel is a former president of the Sleep Research Society and chair of the Associated Professional Sleep Societies. His recent nightly sleep time has been limited to about six hours so that he can take his daughter to a 7 A.M. class.

NINA FINKEL (*chart*); W. PERRY CONWAY *Corbis* (*opossum*); RENEE LYNN *Photo Researchers, Inc.* (*elephant*)



Elephant

have particularly long REM sleep times—90 to 120 minutes each night—compared with other mammals. (And humans with higher IQs or school performance do not have more, or less, REM sleep than those with lower IQs.) The amount of time spent in REM does change over an individual's life, however. In all animals studied, the portion of each day devoted to REM sleep is highest early in the subject's life and falls gradually to a steady, lower level in adulthood. An additional, fascinating fact emerges from comparing numerous species: the best predictor of the amount of REM sleep time for an adult in a given species is how immature the offspring of that species are at birth.

In 1999 Jack Pettigrew and Paul Manger of the University of Queensland in Australia and I were able to study an unusual research subject, the platypus. This evolutionarily earliest of extant mammals surprised us by revealing itself to be the champion REM sleeper: about *eight hours* a day. The platypus is born completely defenseless and blind, cannot thermoregulate or find food on its own, and stays attached to its mother for weeks after birth. At the other extreme, the newborn dolphin can and must thermoregulate, swim, follow its mother and avoid predators. And adult dolphins, as previously noted, do almost no REM sleeping.

Time spent in REM sleep is highest early in LIFE and falls GRADUALLY.

Michel Jouvet, the pioneering sleep researcher who discovered four decades ago that the brain stem generates REM sleep, has a provocative suggestion for the large amounts of REM in immature animals. REM sleep's intense neuronal activity and energy expenditure, Jouvet believes, have a role early in life in establishing the genetically programmed neuronal connections that make so-called instinctive behavior possible. Before birth, or in animals that have delayed sensory development, REM sleep may act as a substitute for the external stimulation that prompts neuronal development in creatures that are mature at birth. Work by Howard Roffwarg, director of the Sleep Disorders Center at the University of Mississippi Medical Center, and his colleagues support this idea. Roffwarg found that preventing REM sleep in cats during this early period can lead to abnormalities in the development of the visual system.

Animals that engage in a lot of REM sleep shortly after birth continue to experience relatively large amounts when mature. What is it about immaturity at birth that causes REM sleep duration to be high later in life? In simple evolutionary terms, animals that have low REM time should need less fuel and leave more descendants than animals that experience long periods of high energy consumption. From that perspective, it is most likely that animals that still have high REM times must have evolved a use for REM sleep that is not found in precocial animals. But that function remains to be identified. Sleep researchers are confident that progress in identifying the brain regions that control REM and non-REM sleep will soon lead to a more comprehensive and satisfying understanding of sleep and its functions. As we further study the mechanisms and evolution of sleep, we will probably gain insights into exactly what is repaired and rested, why these processes are best done in sleep, and why knitting up Shakespeare's raveled sleave of care ultimately helps us to stay awake. SA

MORE TO EXPLORE

Encyclopedia of Sleep and Dreaming. Edited by Mary A. Carskadon. Macmillan, 1993.

Narcolepsy. Jerome M. Siegel in *Scientific American,* Vol. 282, No. 1, pages 76–81; January 2000.

Principles and Practice of Sleep Medicine. Edited by Meir H. Kryger, Thomas Roth and William C. Dement. W. B. Saunders, 2000.

Sleep and Dreaming. Allan Rechtschaffen and Jerome M. Siegel in *Principles of Neural Science.* Fourth edition. Edited by Eric R. Kandel, James H. Schwartz and Thomas M. Jessell. McGraw-Hill/ Appleton & Lange, 2000.

The REM Sleep-Memory Consolidation Hypothesis. Jerome M. Siegel in *Science,* Vol. 294, pages 1058–1063; November 2, 2001.

Center for Sleep Research at U.C.L.A.: **www.npi.ucla.edu/sleepresearch**

NAILS AND STAPLES

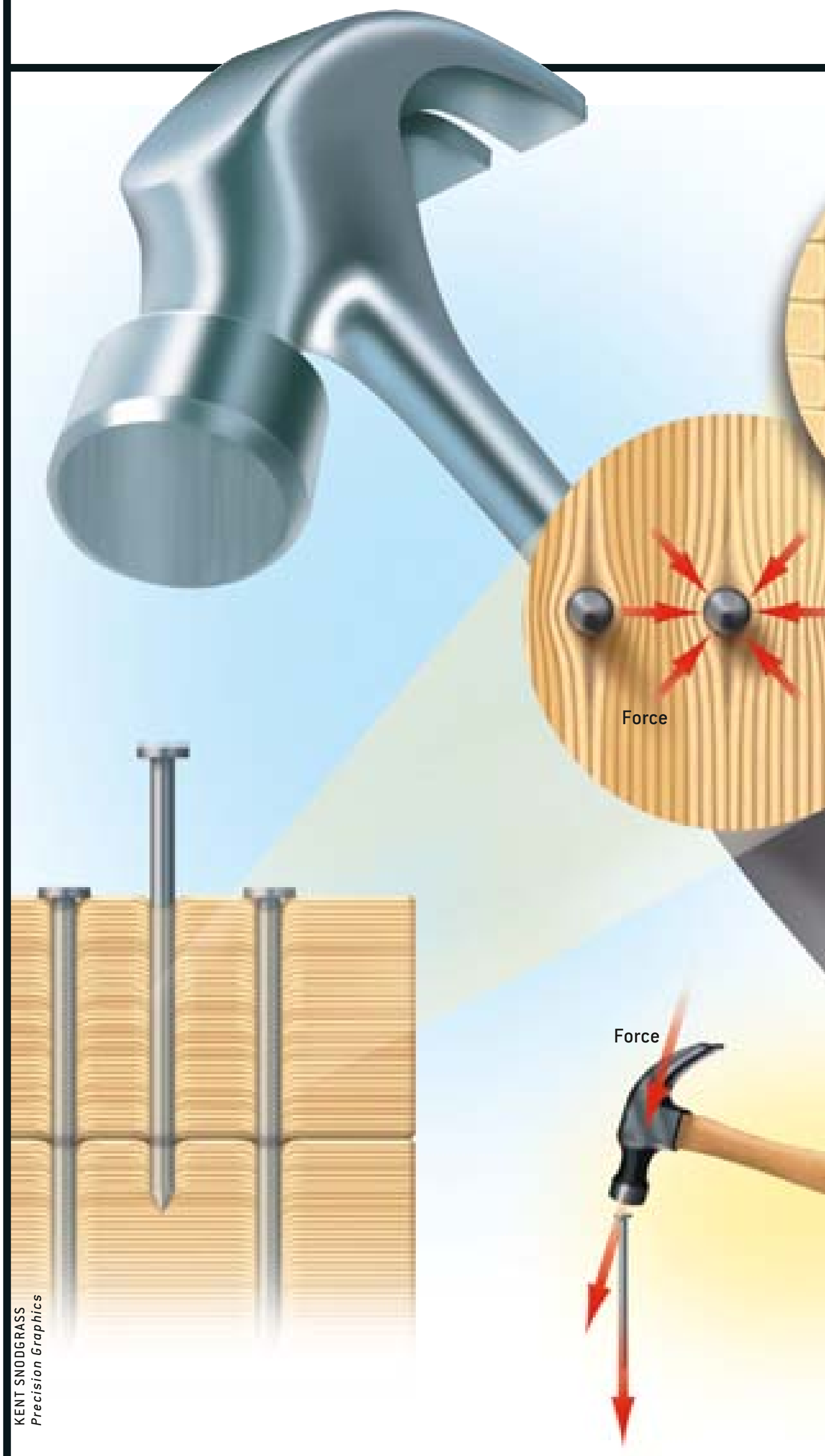
Staying Power

People have pondered the power of nails since the Roman days. But only in the past 100 years have nails come into widespread use in homes and furniture. For centuries, blacksmiths made nails one at a time, at considerable expense, by drawing a short rod of red-hot iron, hammering one end to a point and pounding the other end to a head. By the late 1700s nailsmiths had devised hand-operated machines that could cut nails from flat iron sheets. By the 1880s steam-powered machines sped up the process and "cut nails" became less expensive; however, their strength was still variable.

That changed in the early 1900s, when steel became both flexible and very strong. Machines cut nails and formed the tip and head in one step from a long spool of steel wire. That process allows manufacturers to craft many types of nail points and shanks that improve performance.

Simple friction against a nail shank [*see illustrations at right*] holds two pieces of wood together and prevents the nail from loosening as vibrations and changes in temperature and humidity expand and contract the wood's fibers. The same manufacturing techniques and holding traits apply to staples, which are essentially two nails joined by a crossbar. To help a nail stay put, manufacturers may etch micropits into what appears to be a smooth shank or add rings or barbs, all to better grip the fibers. Certain coatings, such as resin, may increase friction, too, although others may not. "Different manufacturers make different claims, but little scientific research has been done on coatings," says Ron Wolfe, a research engineer at the U.S. Forest Products Laboratory in Madison, Wis., which tests the properties of nails.

Most people take nails for granted, but their simple yet powerful physics make them vital tools. English poet George Herbert reminded his contemporaries of their worth in his early 1600s work *Jacula Prudentum* (first three lines that follow), extended by other, unknown enthusiasts (final two lines): "For want of a nail the shoe is lost/For want of a shoe the horse is lost/For want of a horse the rider is lost/For want of a rider the battle is lost/For want of a battle the kingdom is lost." —*Mark Fischetti*



KENT SNODGRASS *Precision Graphics*

DID YOU KNOW . . .

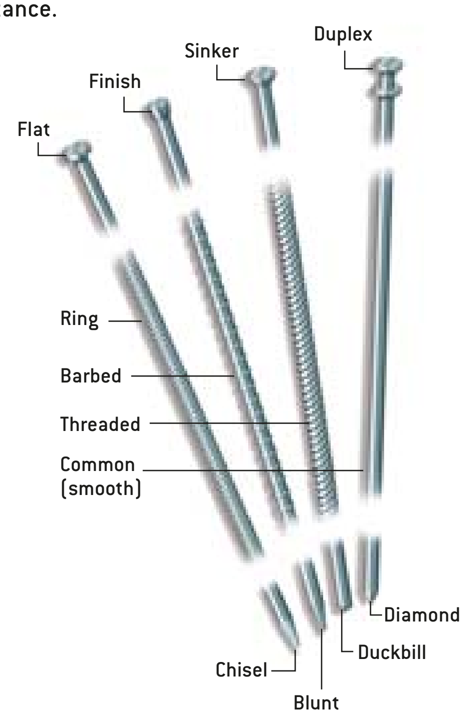
➤ 4D, 8D, 12D: Centuries ago nails were expensive, handmade by blacksmiths. Carpenters could buy 100 one-inch nails for two pennies, or pence, abbreviated as "d." One hundred three-inch nails cost 10d. In time, the cost per 100 came to represent a common nail's size, from 2d to 60d (six inches). The standard persists in nonmetric countries today. Nails shorter than 2d are called brads, longer than 60d are spikes, and both extremes are measured only in inches.

➤ NOT CREATED EQUAL: Could Brand X's 10d nail be better than Brand Y's? Yes. The steel in nails is highly variable, and the industry has shown little interest in standards related to stiffness of the steel. Michael O'Connor, president of Pacific Steel & Supply in San Leandro, Calif., explains that wholesalers order nails from mills according to the degree of carbon content in the steel—more carbon means greater tensile strength but higher cost. Virtually no hand-driven nails (as opposed to those for power nailers) are made in the U.S. any longer; most come from Asian mills.

➤ CORROSION: Rainwater will rust steel nail heads, which can stain a house's siding and shingles, so stainless-steel, aluminum or galvanized nails should be used. The so-called pressure-treated wood popular for decks resists rot because it is treated with ammoniacal copper arsenate or chromated copper arsenate. These chemicals corrode steel and zinc; copper, bronze and stainless-steel nails hold up much better.

A NAIL TIP BENDS wood cells in the direction of the incoming shank as it penetrates. Removing the nail requires enough force to break cells against the bend. A longer or thicker nail provides more surface area and therefore withdrawal resistance.

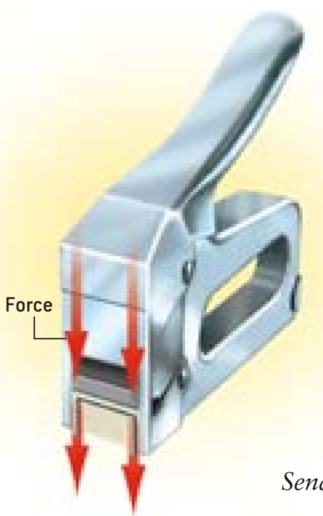
A NAIL SPREADS wood's dense fibers, causing them to compress and therefore push back, holding the nail in place with considerable friction. A longer or thicker nail creates more compression and thus holding power.



FLAT HEADS add some clamping power and ease hammering. Finish and sinker heads can be recessed and covered to improve aesthetics; scaffolding can be held on duplex heads.

RINGED, BARBED OR THREADED SHANKS are less likely than common (smooth) shanks to gradually back out, because the deformations mechanically hook into wood fibers.

SHARP POINTS SPREAD wood fibers cleanly, so all can push back to maximize holding power, but they are more likely to split wood. Blunt tips crush fibers, which reduces splitting but lessens the number of fibers that push back. A diamond point offers a compromise.



NAILS WOULD RARELY BEND if humans could hammer a nail in parallel with its shank every time, but even a few degrees off line creates sideways force that can buckle the steel. Nail guns have guide rails that align the plunger and shank to prevent buckling.

A STAPLE BENDS with the gentle pushing of a finger. But when struck hard with the plunger of a stapler it will not, because rails keep the force in line with the legs, and channels prevent the legs from buckling.

Nicholas William suggested this month's topic. Send your ideas to workingknowledge@sciam.com

Waiting for Liftoff

A ROCKET LAUNCH IS A RIVETING SIGHT. JUST DON'T COUNT ON THE COUNTDOWN **BY W. WAYT GIBBS**

CAPE CANAVERAL—It has taken half an hour to wiggle my way through the throng on Jetty Park pier to just within sight of the spacecraft. Someone's elbow is in my back, and a space junkie keeps blocking my view as he bobs up to check the tripod on his huge telephoto camera. But this is as close as I'm getting to launchpad 17-B. Lesson one for rocket watching: arrive early.

Stake out your spot well ahead of time, even if, as it is tonight, the event is scheduled for a few strokes before midnight and thunderstorms are lighting the late June sky just offshore. Even if T minus zero has been postponed so many times (four so far) that only lucky or very persistent tourists could add this spectacle to their vacation at nearby Kennedy Space Center or Disney World. And even if the payload includes no astronauts, just a large robotic rover.

None of that matters much to the thousands, many of them locals, who jam this pier and adjacent jetty just over two miles from Cape Canaveral Air Force Station. Under the moonless, cloud-covered sky, the stars shine from the crowd—reflections in dilated eyes of the Delta 2 rocket that glows white under the spotlights. It is 30 minutes to ignition, and people are excited. They know this will be a powerful blastoff, because this rover is going all the way to Mars, where it will join an identical twin launched earlier in June.

At Kennedy this afternoon, I learned about the Mars Exploration Rovers mission from an exhibit and an impressive computer animation of the voyage. If all goes well, on January 25 the rover, now snuggled atop the Delta, will bounce in its airbag to a soft landing on the Red Planet. Unfolding its thick metal shell to right itself, it will roll out over its deflated airbag onto the Martian outback. With its panoramic camera and microscope, its rock drill, spectrographs and magnetic dust collector, the robot will search for clues about the watery past of that desert world.

But first it has to get off the launchpad. Just minutes before liftoff, a boat drifts into the restricted area offshore, its pilot asleep at the wheel. The controllers abort and try for a second shot, at 12:37 A.M. By then, however, the winds have picked up, and they scrub the launch.

The next day engineers discover that cork bands on the rocket have come unglued. Departure is set back another week—a week I don't happen to have free. Lesson two for the space chaser: buy refundable plane tickets and keep your itinerary flexible. There is always another rocket, after all. I resolve to catch the late August launch of a new orbiting observatory, the Space Infrared Telescope Facility (SIRTF). (The rover finally did head into space on July 7.)

Meanwhile I salvage this trip by exploring Kennedy Space Center, which for space buffs is itself worth the airfare. One afternoon is just long enough to see some of the older permanent exhibits, such as the "rocket garden" of historic

ROCKET RISES to cheers, not a roar of fire and steam, for the first few seconds until the sound arrives at the viewing area, more than two miles away.

COURTESY OF BOEING

space vehicles and the walk-through replica of the space shuttle. But set aside a full second day to see the newer—and more entertaining—attractions.

An entire morning, for example, could go to a bus tour of some of the working parts of this spaceport. The route winds by the enormous vehicle assembly building, its doors so tall that the Statue of Liberty could be wheeled through upright. The bus then passes the motorized crawlers that haul the shuttle out to the launchpads at one mile an hour.

The highlight of the tour for me is the new Saturn 5 museum. Seeing the Apollo launches on television is no preparation for the awesome size of this retired rocket. Fully fueled, the vehicle contained the explosive power of an atomic bomb. It looks that dangerous.

Two short multimedia shows in the building's theater re-create the history of the moon shots in a visceral way that video alone just can't match. One replays the tense minutes before the first lunar landing, as the astronauts and controllers struggled with computer crashes, radio failures and a major navigational snafu. A second, even more evocative show takes place in a mock control room filled with real equipment salvaged from the Apollo 8 mission. Status lights, video screens and audio track are all synchronized to the countdown clock. As footage of an Apollo launch plays back in real time and from all angles, the faux windows at the back of the theater glow from the fire, their louvers rattling. It's the next best thing to witnessing an actual blastoff.

It only whetted my appetite for the real experience. In August I was back at Jetty Park, this time four hours rather than 30 minutes early. Before long I was joined at the end of the pier by half a dozen others—all scientists, it turned out, who had helped to design the telescope now perched inside the nose cone of a Delta 2 Heavy. SIRTF will complete NASA's 20-year Great Observatories program, joining the Hubble optical tele-

NEW EXHIBITS at Kennedy Space Center cover current and future missions to Mars (*above*) as well as the historic Apollo missions to the moon (*right*).

COURTESY OF KENNEDY SPACE CENTER

scope and the Chandra X-Ray Observatory. (The Compton Gamma-Ray sensor deorbited in June 2000 after completing its mission.)

If its predecessors are any guide, the infrared pictures that SIRTF delivers will be full of scientific surprises. "The atmosphere is more than 99 percent effective at screening out infrared light," explains David Cole, one of the SIRTF scientists sitting in beach chairs next to me on the pier. SIRTF will rise far above this fog, trailing Earth in its orbit about the sun.

The space junkie is back, this time with a telescope. His name is Bill Hughes. "I am an avid space chaser," he says. A parking attendant in Daytona Beach, he drives down for every shot he can; this will be his 28th. Hughes lets me peer through his scope at the condensation pouring off the rocket's body as it receives its liquid oxygen fuel. At T minus 30 minutes, the skies are clear and the winds are calm. The suspense is killing.

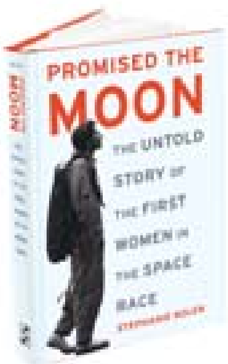
At last we get to the final minute, and then to "10, 9, 8 . . . the board is green . . . 3, 2, 1"—by now the throng is chanting in unison—and then the flash, a fireball, a huge burst of steam boiling off the pad like a magically invoked cloud. And it is rising, surreally, because the only roar is that of the whooping spectators, myself among them.

Then the sound arrives, so loud and deep that it is more felt than heard. The rocket passes through a thin cloud, seeming to set it aflame, and accelerates so quickly it is hard to believe that it weighs 283 metric tons. As it angles into the stars, we track it through our binoculars and see the nine solid boosters flame out and separate, just as they should. When the the main engine cuts off and the second stage takes over, beer bottles clank in a happy toast of relief, exhilaration and an expectation of great things to come.

Space shuttle launches have been postponed until at least next spring, but expendable rockets are set off every few months. For a schedule, browse to www-pao.ksc.nasa.gov/kscpao/schedule/mixfleet.htm. Information and tickets are available from www.ksctickets.com, or call 321-449-4444. No advance tickets are needed to enter Jetty Park; admission is $3 per car. SA

Limited Visibility

WHY DID IT TAKE NASA SO LONG TO LAUNCH A WOMAN INTO SPACE? **BY PHIL SCOTT**



PROMISED THE MOON: THE UNTOLD STORY OF THE FIRST WOMEN IN THE SPACE RACE
by Stephanie Nolen
Four Walls Eight Windows, New York, 2003 ($22.95)

Shortly after the Mercury 7 astronauts appeared on the September 14, 1959, cover of *Life,* the magazine featured a cover story about the women of the space program. Their photo cut the exact same pose. They were the astronauts' wives, whom *Life* always depicted as demure ladies fully behind their husbands while the men trained to become the first Americans in space.

Male astronauts. That's the way most Americans thought it should be. In those days a woman couldn't take out a bank loan unless her husband co-signed. In those days a woman had trouble getting a job—unless she wanted to be a secretary.

One story resurfaces every few years: the tale of the Mercury 13—13 women whom NASA recruited for the space program, then ditched with nary a launch or an explanation. Or so the story goes. According to *Promised the Moon,* the fine new book by Stephanie Nolen, a foreign correspondent for Canada's *Globe and Mail,* the whole truth went deeper.

Back in the 1950s America designed rockets to lift nuclear weapons. Because we had the technology to miniaturize components, our nukes weighed less and were more compact than Soviet nukes. Using their heavy-payload missiles, the U.S.S.R. could shoot Sputnik into orbit and were preparing to launch a "cosmonaut." Converting our pencil-thin missiles to carry a man, however, meant engineering a capsule with a circumference only inches wider than the length of a sitting pilot. All the essentials he needed—oxygen, maneuvering rockets and their fuel, food and so on—pushed the capsule payload to its weight limit.

For its astronauts the U.S. space agency recruited only military test pilots—male test pilots, of course. They flew higher and faster than anyone. But Randolph Lovelace II, chairman of NASA's Life Sciences Committee, theorized that women might make better astronauts. An average woman would be shorter and weigh less than the average man. She would breathe less oxygen and allow less fuel to be carried for those maneuvering rockets. She'd even eat less. Every ounce counted.

In 1959 Lovelace was introduced to Jerrie Cobb on a Florida beach. The 28-year-old record-setting pilot had 7,500 hours in the cockpit, "more … than any of the male astronauts," according to *Time* magazine. Intrigued by all her flying experience, Lovelace invited Cobb to undergo the same sadistic tests that scientists inflicted on the Mercury 7. She achieved scores on a par with those men or even higher. By early 1961 Lovelace submitted the names of Cobb and 12 other exceptional female pilots—including a senator's wife—to NASA as theoretical candidates for a woman-in-space program. Nothing official, mind you.

The newspapers sniffed out the story, but NASA honchos would say only that they had no real program. Cobb, though, took up a strenuous self-promotion campaign, pointing out that the Soviets were preparing to send a woman into orbit soon and that America could beat them to it. Her cage rattling annoyed nearly everyone, from NASA's old-boys' club, the military and the U.S. government to an earlier woman pilot: Jackie Cochran, a wealthy, strong-willed, well-connected woman too old to go into space but willing to go to any lengths to be put in charge of the prospective "astronettes."

On May 25, 1961 (not long after Alan Shepard's suborbital flight), President John F. Kennedy set the goal to send Americans to the moon, and the nation

JERRIE COBB, the first of the 13 "Fellow Lady Astronaut Trainees," in 1959.

COURTESY OF FOUR WALLS EIGHT WINDOWS

finally obtained a mission for the space program. That also gave NASA the perfect reason to ditch the ladies. We were in a race to the moon, not in a race to be the more egalitarian superpower. Still, Cobb refused to give up. At a congressional hearing, she and Janey Hart, the wife of Senator Philip A. Hart of Michigan, outlined the benefits of launching a woman into space. Then the influential Cochran—friend of both NASA officials and U.S. presidents—hijacked the hearing and enraged the two women by testifying that no female should launch into space, not at that time. The following day the men romped in, including astronaut hero John Glenn. "The fact that women are not in this field is a fact of our social order," Glenn said.

At that time his testimony was hard reality. NASA chose only jet test pilots as astronauts, but only the military had jet test pilot schools, and the military had no women pilots. So of course no female jet test pilots applied to NASA's astronaut training program. You couldn't beat 'em, and you couldn't join 'em.

Nolen points out further disadvantages that the women faced: military test pilots "already spoke the language of engineering.... They knew the designers and engineers.... They had shown they could work in secret and under discipline." Plus they "had also provided the largest body of data then extant on how the human body functioned in the outer reaches of the atmosphere." And, she writes, "in the larger picture, the women were grounded for one simple reason: they stepped outside the boundaries of the accepted roles for women in their time."

The book makes a quick, intriguing read, although it could have been written more tightly—especially in the long section recounting the history of female aviators. There is one error of fact: the author launches John Glenn on a Saturn rocket, when actually he rode an Atlas. But Nolen successfully returns the reader to an earlier America where sexism reigned, when the only jobs open to women in the aerospace field were to demonstrate airplanes "so easy to fly even a girl can do it."

It took NASA two more decades to launch a woman: Sally Ride became America's first in space, though as a nonpilot mission specialist. In 1995 NASA launched Eileen Collins, the first American female to pilot a shuttle. That was more than 30 years after the Mercury 13 disbanded. As the old cigarette ad said, You've come a long way, baby. SA

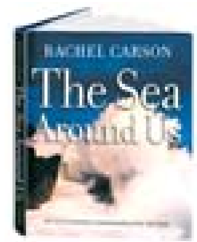
Phil Scott is the author of four books on aviation history. The most recent, The Wrong Stuff *(Hylas Publishing), was released in October.*

THE EDITORS **RECOMMEND**

THE SEA AROUND US: AN ILLUSTRATED COMMEMORATIVE EDITION

by Rachel Carson. Oxford University Press, New York, 2003 ($45)

When Carson wrote this classic book (it was first published in 1951, 11 years before *Silent Spring*), the concepts of seafloor spreading, subduction zones and plate tectonics were unknown. Yet her book remains fresh, in part because of her ability to convey scientific insight in vivid poetic language—but, perhaps more important, because what she has to say is still so relevant today. To bring the findings about the ocean basins up-to-date, this new edition offers an afterword by Brian Skinner, professor of geology at Yale University, which provides an extremely literate and succinct tutorial in the geologic developments of the past half a century; his explanation of plate tectonics is one of the most lucid anywhere. He also lays out before us the mineral wealth of the seafloor, ending with the question that would have been foremost in Carson's mind: Do we have the will and the vision to exploit these resources with a minimum of environmental disruption? The book has in addition some 130 color illustrations, an introduction by Robert D. Ballard, and a brief, moving foreword by Carl Safina, who concludes: "Carson's lasting power is that we still seek orientation by her moral compass."



THE DISCOVERY OF GLOBAL WARMING

by Spencer R. Weart. Harvard University Press, Cambridge, Mass., 2003 ($24.95)

This book tells the story of the long struggle to understand how humanity could be changing the weather. It is a complicated story because climate itself is irreducibly complicated. "We will never grasp it completely, in the way that one might grasp a law of physics," in the words of Weart, who is director of the Center for History of Physics at the American Institute of Physics. One of his goals is to help readers understand the present predicament by explaining how we got here: following the way scientists traced the uncertainties of climate change better prepares us to judge why they speak as they do today. In the end, the book is cautiously hopeful: "A few people, through ingenuity, stubborn persistence and a bit of luck, came to understand a grave problem even before any effects became manifest.... [M]any other people, defying the old human habit of procrastinating until a situation becomes unbearable, began working out solutions. For there are indeed ways to keep global warming within tolerable bounds with a reasonable effort."

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

Liquid Switchboard

BY DENNIS E. SHASHA

Imagine five vertical pipes arranged in a circle [*see illustration below*]. The pipes are labeled A, B, C, D and E, with each letter standing for the color of water that is poured into the top of the pipe: amber, blue, crimson, diamond and emerald. (Because the pipes are in a circle, pipe A is between pipes B and E.) Between any two adjacent pipes are three switches that can be used to divert the flows of colored water. For example, if the top switch between pipes A and B is open, the amber water from pipe A will be rerouted to pipe B and the blue water from pipe B will flow into pipe A. Further exchanges can take place at the middle and bottom switches; at each level, the water in a pipe can flow either to the left, to the right or straight down (if neither of the pipe's switches is open). But water cannot flow in two directions at once. For instance, the top switch between pipes A and B cannot be open if the top switch between pipes A and E is also open.

Here's a warm-up problem: Can you arrange the switches so that the colors of the liquids in pipes A, B, C, D and E become C, D, E, A and B at the very bottom? (That is, crimson at the bottom of pipe A, diamond at the bottom of pipe B, and so on.) As the illustration shows, you must first open the top switches between pipes A and B and pipes C and D, which changes the sequence of colors to B, A, D, C and E. Then open two of the middle switches to change the pattern to B, D, A, E and C. Last, open two bottom switches to create the desired arrangement.

This permutation of colors is so far from the original that one might think that any rearrangement is possible. Is that true? If not, which permutations are not possible, and how many more levels of switches would you need to make every rearrangement possible? SA

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

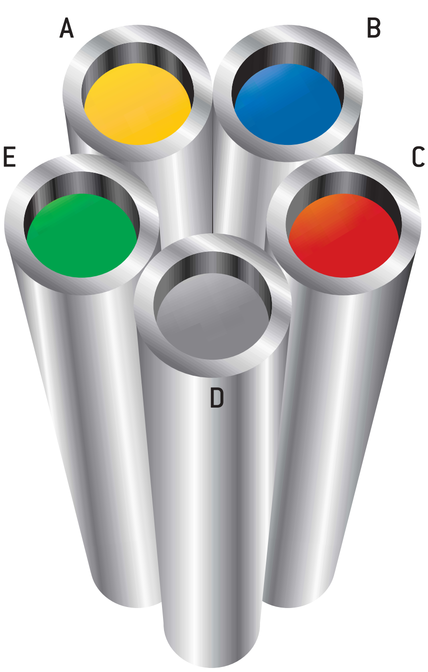
Answer to Last Month's Puzzle

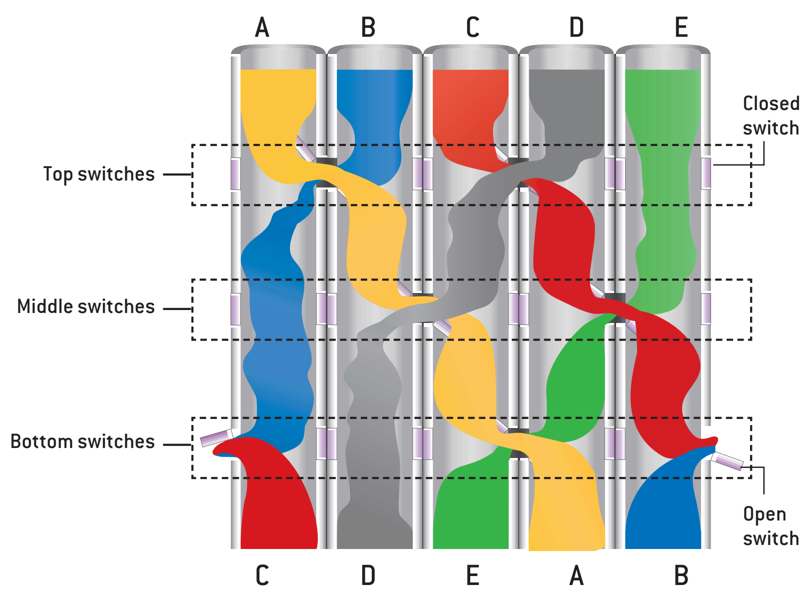
If the highest power value is 21, the largest set of distinct values forming a stable configuration is 1, 2, 3, 4, 5, 6 and 21. The largest set of nondistinct values has 21 1's and one 21.

A bee-sting capability (indicated by an asterisk) is not necessarily stabilizing. A configuration of 5, 4 and 3 is stable. But the set consisting of 5, 4* and 3* is unstable, because a coalition of 3* and 4* can attack 5 without consequences.

Web Solution

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





CORNELIA BLIK

A Bridge Too Far

A MAN, A PLAN, A RIVER, CAMBRIDGE **BY STEVE MIRSKY**

How can you tell that you're in Cambridge, Mass., over Labor Day weekend with the start of classes at the Massachusetts Institute of Technology only days away? For one thing, the streets teem with furniture-filled rental vans. Despite also carrying future physicists, these vehicles attempt to violate physical law by occupying the same space at the same time. On the other hand, a few drivers actually stop for red lights, which proves that they are not from the Boston area and are merely passing through to drop off freshmen.

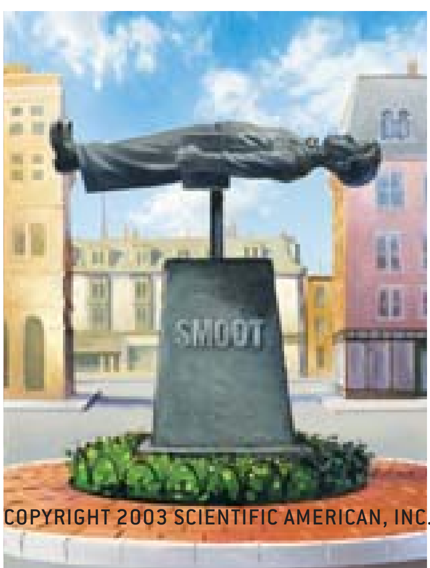
How do you know this is erudite Cambridge, also home to Lesley University, Cambridge College, the main headquarters of the American Academy of Arts and Sciences, Harvard University and Richdale A-Z Auto Service? Some checkout aisles in local supermarkets feature grammatically correct "10 Items or Fewer" signs instead of the commonly seen "10 Items or Less" notice. (Local folklore has it that anyone in the 10-items lane carrying 20 items is either a Harvard student who can't count or an M.I.T. student who can't read.)

How can you be reasonably sure that this is Cambridge? Eavesdrop on the two guys behind you in a restaurant. They discuss quantum foam during the appetizer, contemplate human evolution with the main course and accompany their dessert with an analysis of the fine points of an episode of *Star Trek*.

The clincher, however, is the strange and thrilling discovery I make upon consulting a map entitled "M.I.T. and Its Environs." While checking the legend to see if a destination is close enough to walk, I notice that the distance scales in familiar feet, meters and miles have been joined by an interloper: the smoot. A little Internet research turns up the glorious history of the smoot, a unit so specific to, and well known at, M.I.T. that a map of Harvard made by the same company fails to include it. (The Harvard map ignores those newfangled meters, too.)

First you need to know about the Harvard Bridge, "so named because it leads directly into the heart of MIT, which is near Harvard," notes Ken Nesmith, writing in the M.I.T. publication *The Tech* in 2001. The bridge crosses the Charles River and connects Boston with Cambridge. Now meet George Smoot, an M.I.T. graduate, famous physicist and author of the popular cosmology book *Wrinkles in Time*. George Smoot has nothing to do with the smoot.

Well, almost nothing—in an attempt to literally clear his name, George Smoot has made the history of the smoot available on his Web site. It seems that in 1958 an M.I.T. freshman named Oliver R. Smoot, Jr., hoped to join the Lambda Chi Alpha fraternity. Of course, inclusion in such august societies often involves ordeals that test the prospective member's fitness. Oliver Smoot's future frat brothers therefore commenced to roll him "head over heels the entire length of the bridge," according to the account published on George's site. Oliver's own version, published in *Nightwork: A History of Hacks and Pranks at MIT*, has him lying flat and being dragged. Either way, his five feet, seven inches is therefore one smoot. "Every ten smoots they calibrated the bridge, painting marks," the Web version continues. "The bridge was found to be exactly 364.4 smoots plus an ear," which was, miraculously, still attached.



Future frat pledges followed in Oliver's footsteps, handprints, etc., by repainting the 10-smoot lines. In 1987 the state announced plans to renovate the bridge, which would have smote the smoot, relegating the stripes to the stuff of nonmap legend. Local reporters tracked down Oliver, who removed himself from demarcation calibration replication consideration. Nevertheless, when I went to the bridge, I found relatively fresh 10-smoot markings, indicating that undergraduates had been active in the area in the recent past. Oliver's legacy will apparently endure. And the point, of course, is smoot. SA

For the next eight months, Steve Mirsky will file the Anti Gravity column from Cambridge, Mass., where he is, amazingly enough, a Knight science journalism fellow at M.I.T.

ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

ASK THE EXPERTS

What makes Kansas, Texas and Oklahoma so prone to tornadoes?

—T. Irwin, Kissimmee, Fla.

Harold Brooks, head of the Mesoscale Applications Group at the National Oceanic and Atmospheric Administration's National Severe Storms Laboratory in Norman, Okla., explains:

The central part of the U.S. gets many tornadoes, particularly strong and violent ones, because of the unique geography of North America. The combination of the Gulf of Mexico to the south and the Rocky Mountains to the west provides ideal conditions for tornadoes to develop more often than any other place on earth. The central U.S. experienced a record-breaking week from May 4 through May 10 this year, when close to 300 tornadoes occurred in 19 states, causing 42 deaths, according to NOAA's National Weather Service.

Storms that produce tornadoes start with warm, moist air near the ground. Dry air is aloft (between altitudes of about three to 10 kilometers). Some mechanism, such as a boundary between the two air masses, acts to lift the warm, moist air upward. The boundary can be a front, dryline or outflow from another storm—essentially any kind of difference in the physical properties of two air masses. "Kinks" in the boundary are locations where rotation could occur. An updraft (air going up) traveling over the kink will "stretch" and intensify the rotation, just like an ice skater pulling in her arms.

Strong tornadoes are also most likely to happen when the horizontal winds in the environment increase in speed and change direction with rising altitude. In the most common directional change of this kind, winds at the surface blow from the equator, and winds a few kilometers above the ground blow from the west. When this wind pattern occurs in the central part of the U.S., the surface winds flow from the direction of the Gulf of Mexico, bringing in warm, moist air. The winds aloft, in contrast, come from over the Rocky Mountains and are relatively dry. As a result, when the winds over the central part of the U.S are optimal for making thunderstorms, they often combine the right distribution of atmospheric temperature and moisture to produce tornadoes as well.

MATT COLLINS

Are humans the only primates that cry?

—C. Henderson, Winter Park, Colo.

Kim A. Bard, a researcher in comparative developmental psychology at the University of Portsmouth in England, offers this perspective:

The answer to this question depends on how you define "crying." If it is defined as tears coming from the eyes, then the answer is yes: tears appear to be unique to humans among the primates. If you define crying as a vocalization that occurs under conditions of distress, or what humans might describe as sadness, then you can find it in almost all primates.

Others argue that all mammals have feelings, because emotions are the product of deep-brain functioning with a long evolutionary history. Some researchers reserve such emotional terms for humans alone and will not use such words for other primates. Some scientists take a conservative stance and say that it is too difficult to tell whether or not nonhuman primates have feelings. Rather than broadly describing particular primate vocalizations as crying, scientists prefer specific names for certain conditions. For example, a young primate that is not in contact with its mother produces a separation call. Researchers also describe what the vocalization sounds like, as with the "smooth early high" coos of Japanese macaques. Or scientists note what the animal is trying to communicate, such as when infants try to satisfy their basic needs for food, social contact or relief from pain. SA

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


OUT-OF-MEMORY LANE
I'M MARJORIE, YOUR NEXT-DOOR NEIGHBOR, YOU IDIOT.
I'M STU. WE'VE MET SEVERAL TIMES.
I'M ONE OF YOUR SON'S FRIENDS, DUH.
YOU'VE BEEN TO MY HOUSE FOR DINNER- TWICE.
JUDY, THE DENTIST'S RECEPTIONIST
HELLO! MY NAME IS GREG... SOUND FAMILIAR?
I'M- OH, NEVER MIND.