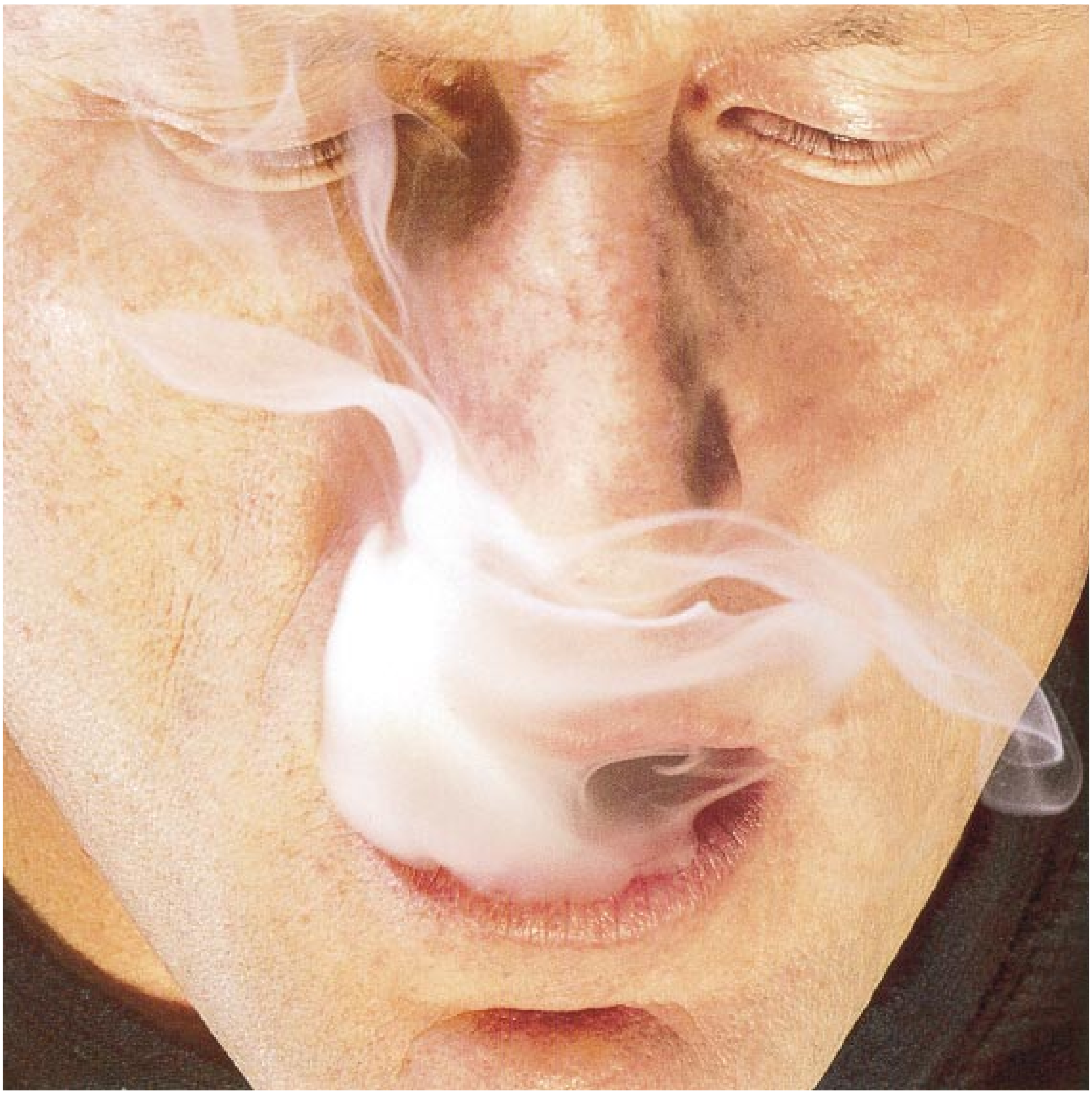
SCIENTIFIC AMERICAN

MAY 1995
$3.95

What found the top quark.
Archaeology in peril.
The Niels Bohr mysteries.



***Clouds** of tobacco smoke continue their spread, despite warnings.*

SCIENTIFIC AMERICAN

May 1995 Volume 272 Number 5

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 1995 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher. Second-class postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian GST No. R 127387652. Subscription rates: one year $36 (outside U.S. and possessions add $11 per year for postage). Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. Reprints available: write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; fax: (212) 355-0408 or send E-mail to SCAinquiry@aol.com. **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.**

DEPARTMENTS

THE COVER photograph depicts the very familiar habit of one in four American adults. Although cigarette use had been declining since the 1960s, the number of smokers in the U.S. has remained static during the 1990s—currently about 46 million. Globally, smoking is on the rise, outpacing the rate of the world's population growth. Aggressive marketing, low taxes and weak regulations are the main reasons (see "The Global Tobacco Epidemic," by C. E. Bartecchi, T. D. MacKenzie and R. W. Schrier, page 44). Photograph by Christopher Burke, Quesada/Burke.

  PRINTED IN U.S.A.

LETTERS TO THE EDITORS

Closing In on Zeno

In "Resolving Zeno's Paradoxes" [Scientific American, November 1994], William I. McLaughlin overlooks an important point—Zeno's paradoxes question the validity of our descriptions of physical reality. They are not simply mathematical puzzles and should not be considered solved unless there is reason to believe that space-time is accurately described by the mathematics used to formulate the solutions. Can one formulate all the known laws of physics using internal set theory? Can any experiments be performed to determine whether infinitesimal nonstandard points exist? Until these questions are addressed, McLaughlin's solutions must be understood as speculative.

Stephen G. Dillingham
Johns Hopkins University

McLaughlin replies:
I agree that my analysis does not constitute a physical theory. I also agree that Zeno did not raise his objections merely to create puzzles; he was addressing the way he thought the world was built. Surely, however, Dillingham asks too much when he requires us to map Zeno's objections to a modern empirical setting. I prefer to cage Zeno in a cosmos intelligible to a Greek geometer and test concepts within that context. This less ambitious program could still yield meaningful results. Mensuration limitations on the system of real numbers might prove relevant to the development of physical theory in dynamics or in a quite unrelated discipline.

Whither the Infobahn?

Despite the fears voiced in "The Speed of Write," by Gary Stix [Scientific American, December 1994], there will not be a decline in standards for refereed electronic journals. It is precisely because the number of E-journals on the Usenet will expand that the top E-journals will become more strict. In the competition for prestige in a drive-through marketplace of ideas, E-journals will raise their standards as high as possible while still having articles left to publish. There will be more "trash" on the Usenet as a whole, but the fear of becoming considered trashy themselves will keep the standards of serious journals high and push them higher.

Jason Fossen
University of Texas at Austin

In the news story "Pricing Internet" [Scientific American, November 1994], W. Wayt Gibbs raises the nettlesome question of how to deal fairly with the economics of a rather loose federation of computer networks. Balance sheets and payroll checks do not come close to providing a complete evaluation of the work done by sysops and assistants. The Internet exists because technical people approach it as a labor of love. If the likes of Pacific Bell, Sprint and Ameritech fail to account for these aspects of their "economic" ventures into the Internet, they may ultimately have very little to offer.

Robert I. Price
University of Nebraska at Kearney

Wringing the Bell Curve

In his review of *The Bell Curve* [Scientific American, February] Leon J. Kamin describes the Pioneer Fund as "nativist, eugenically oriented." In fact, Pioneer limits its activities to grant making. It does not suggest research projects, and it does not make grants to individual scientists, only to institutions. It does not oversee research, it does not comment on results, it does not have any publications and it does not take positions on political issues of any kind. The fund stays strictly hands-off. Twin and adoption studies funded by Pioneer have become famous and are reflected today in standard textbooks.

Harry F. Weyher
President, The Pioneer Fund, Inc.
New York, N.Y.

Kamin devotes the first part of his review to criticism of my work on the average IQ of black Africans. I assembled 11 studies of black African IQ, set out the results and proposed to rely primarily on what I considered the best study, one of black 16-year-olds by Ken Owen. I calculated their mean IQ as 69. Richard Herrnstein and Charles Murray preferred to adopt the median of the 11 studies, which gives a figure of 75.

Kamin points out that Owen reported black-white difference expressed in standard deviation units. This can be converted to an IQ difference on the basis of one standard deviation unit equaling 15 IQ points. Contrary to Kamin's assertion, it is an entirely valid procedure. Kamin criticizes me for omitting certain other studies of black African IQ. I ruled out those in which the sampling was clearly not representative. Whatever precise figure is adopted as the best estimate of the black African IQ, the evidence is solid that it is lower than that of American blacks. The most probable explanation is that most American blacks carry a number of Caucasian genes that raise their intelligence above that of Africans.

Richard Lynn
University of Ulster
Coleraine, Northern Ireland

Kamin replies:
The Pioneer Fund's white-supremacist history is well documented. It supports such scholars as Roger Pearson, who wrote that "if a nation with a more advanced, more specialized or in any way superior set of genes mingles with, instead of exterminating, an inferior tribe, then it commits racial suicide."

The rules by which Lynn eliminates "not representative" studies are murky. An example: based on the claim that testosterone causes prostate cancer, Lynn accounts for "the high rate of sexual activity in Negroids" by citing evidence "that Negroids have higher rates of cancer of the prostate than Caucasoids" and so must have higher testosterone levels. He presents data from a paper by D. G. Zaridze et al. to show that blacks have a higher incidence of prostate cancer than do whites in six American cities. But Lynn ignores other data in that paper showing the incidence for African blacks is far below that among American blacks (and American whites). Lynn seems to lose interest in comparing black Americans and black Africans when the evidence does not support his racial theories.

Letters selected for publication may be edited for length and for clarity. Unsolicited manuscripts and correspondence will not be returned or acknowledged unless accompanied by a stamped, self-addressed envelope.

50 AND 100 YEARS AGO

MAY 1945

A recent development in plastics and electronics is a wafer-thin Vinylite plastics record, only seven inches in diameter. Each side of the disk will record approximately 15 minutes of dictation. These records can be bent, rolled, dropped, and written on with a pencil without harming the sound track. The thin plastic can be stored indefinitely, without warpage, breakage, or distortion, in an ordinary filing cabinet—100 disks to the inch—and played back at least 100 times."

"A new type of Diesel engine will enable the operator to use either gas or oil as fuel without any electrical sparking device and will cut fuel consumption of gas engines by as much as 25 percent. The unit operates on a wide variety of fuels, including fuel oil, natural gas, manufactured and coke oven gases, sewage gas, and refinery by-products. Furthermore, the engine will have the same fuel economy regardless of the type of fuel used."

"When the problem of washing bearings came up at The Electric Auto-Lite Company, engineers whipped it by reverting to a regulation orange squeezer. The bearings are simply put in where the orange used to go, then a spray of oil is sent over them as they whirl around in the container. The bearings are taken out by tweezers, never handled by human hands. The cleaning fluid drains from the spout."

"In the new technique of electronically controlled vulcanization of rubber, high-frequency oscillation shakes the molecules of rubber and sulfur millions of times a second, creating uniform heat throughout the product being vulcanized in a fraction of the time required when steam is used. Sponge rubber mattresses and pads have been cured by this electronic method. Tires, molded rubber goods, brake bands, and many other products can also be cured much more rapidly by electronics."

MAY 1895

Spring colds usually occupy about a week of time, with the aid of various remedies. It is possible in the early stage of a cold, especially when such is of the nasal variety, to abort an attack by irrigating the nose twice a day with warm water in which a little borax has been placed. No syringe is necessary; but by simply immersing the nose in a basin of water, and making forcible inspiratory and expiratory movements, holding the breath at the epiglottis, the nasal passages may be thoroughly irrigated. Of course there are advantages in the syringe, which may be preferable from the standpoint of neatness."

"Prof. James E. Keeler has made the interesting discovery that the ring of Saturn is made up of many small bodies, and that the satellites of the inner edge of the ring move more rapidly than those of the outer edge."

"There is one aspect of the immigration question that appeals purely to business men. The social and moral influences on the American people of the unrestrained horde of Europeans pouring upon our shores are, of course, the most important, but the heavy tax in money thus levied upon the American people is not to be disregarded."

"The cocaine habit is a comparatively new addition to the evils by which humanity is beset, and it promises to excel even morphinism in the insidiousness of its growth, in its blasting destructiveness and in the number of its victims. Several distinct causes result in the acquirement of this habit. Prominent among these is the pernicious practice of a certain class of druggists (fortunately small in number) who offer cocaine when asked for something that will relieve toothache, neuralgia and countless other aches and pains."

"The Layman pneumatic boat is acquiring wide popularity among sportsmen and those fond of aquatic sports, as well as with ladies and children for use on the seashore. The bottom of the boat, which is made entirely of India rubber cloth, has a strong sheet of the same cloth from whose forward portion two boots or leg cases descend. The bottom of the boots consists of collapsing paddles, which open on the back stroke and close on the forward stroke, as does a duck's foot. This cut illustrates a passage through Hell Gate, East River, New York, which was made without difficulty in such boats, by a party including a lady. The experience is described as delightful, the waves of the steamers adding to the excitement."

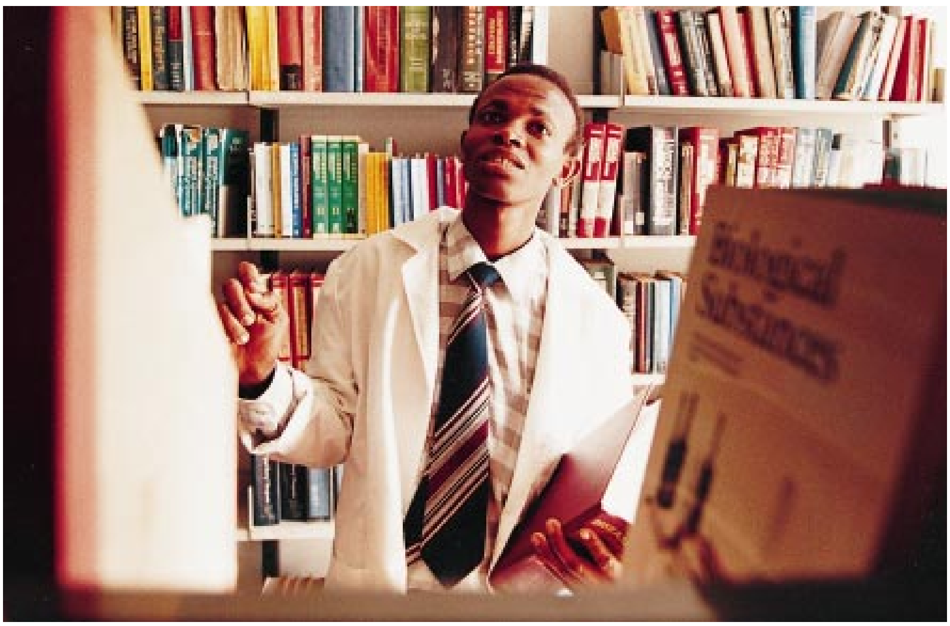
Party crossing Hell Gate in Layman boats

SCIENCE AND THE CITIZEN

Information Have-Nots

A vicious circle isolates many Third World scientists

Researchers at Addis Ababa University face a disheartening sight when they visit the library to catch up on advances in their fields. Shelves that just six years ago were filled with the latest issues of more than 1,200 academic journals lie barren. The elimination of its foreign currency budget in 1989 forced the library to cancel about 90 percent of its subscriptions, severing the conduit that conveyed news of discovery to scientists in the Ethiopian capital.



RICARDO O. MAZALAN *Associated Press*

LACK OF INFORMATION hinders scientists at the University of Nairobi, whose medical library received just 18 journal titles in 1992.

Throughout Africa and many other parts of the developing world, the flow of scientific information from the rich countries of the North has dried up over the past decade. The squeeze tightens a vicious circle that dooms many poor nations to waste precious investments in science and technology on duplicative research of dubious quality. *Scientific American*'s interviews with more than 40 scientists in 18 countries reveal that many believe poverty, cultural differences and a subtle prejudice against so-called Third World researchers combine to largely shut them out of major journals, important international conferences and critical databases.

An investigation of a handful of the most influential journals shows that nearly all the articles they published in 1994 include at least one author working in the U.S., Canada, Europe, Japan, Australia or Israel, even though those regions harbor only about 76 percent of the world's scientists and engineers. Even considering the lopsided funding of science—industrial nations footed 95 percent of the world's research bill in 1990—reports from the rest of the globe account for a surprisingly tiny proportion of articles: just 0.3 percent for *Science,* 0.7 percent for *Nature* and 2.7 percent for *The Lancet. Cell* and *Scientific American,* among many others, ran no such articles at all.

A stockade of barriers seems to prevent scientists in less developed countries from publishing in these journals. Foremost is the want of money: they receive smaller pieces of smaller pies than do their U.S. and European colleagues. As a result, says Mounir Laroussi, a Tunisian researcher at the University of Tennessee and assistant editor of *Physics Essays,* "few can afford to pay the fees of up to $150 per page that many mainstream journals charge authors to publish their papers." Laroussi was able to recruit only two Tunisian authors for his journal in the past year, and he had to loan both of them American dollars to meet the fees.

Small and unstable budgets force many investigators in sub-Saharan Africa and the poorer parts of Asia to communicate without the luxuries of fax machines and electronic mail. The explosive growth of networks and CD-ROM drives that promises to open up science publishing in the U.S. and Europe to a larger audience thus threatens to strangle the South's access. In a recent study of India's situation, Subbiah Arunachalam of the Central Electrochemical Research Institute observed that publishers tend to "adopt a pricing policy which makes the print-on-paper form more expensive than the [electronic] forms. Thus, the poor end up paying more for the same information than the rich!"

Increasing subscription rates and plummeting currency values have already priced academic libraries in many countries out of the market for journals. "We recently did a survey of 31 libraries in 13 African countries," reports Amy A. Gimbel, director of the sub-Saharan African program at the American Association for the Advancement of Science. "Not one has a viable serials collection." Eight of the libraries are completely dependent on donations for foreign subscriptions.

Elsewhere, Latin American scientists say their research libraries generally carry at least the top journals. But "India, which used to receive about 20,000 journals in 1983, now gets less than 11,000, and fewer copies of each," states Thiagarajan Viswanathan, director of the Indian National Scientific Documentation Center.

This lack puts authors at a serious disadvantage when they submit their work for publication. "If you don't have access to references and the current citations to related work in the North, you won't pass muster," Gimbel says. Autar S. Paintal, former director general of the Indian Council of Medical Research, notes that "an Indian is often unaware of the latest trends in science publishing [because] hardly 10 percent of our libraries get the top journals."

Institutional prejudice may play a role, too, according to a significant minority of researchers who believe that some editors give papers from poor countries second-class treatment. "Many of them feel discriminated and think their papers are rejected on the grounds that they are from developing countries," observes Abdus Salam, a Nobel Prize–winning physicist from Pakistan who founded and until recently chaired the Third World Academy of Sciences. Gursaran P. Talwar, former director of India's National Institute of Immunology, says that when a scientist whose paper has been rejected "goes abroad for postdoctoral study, the change of address makes all the difference." By all accounts, theoreticians fare better than experimentalists, who often lack sophisticated equipment. But Ana María Cetto, a physicist at the National Autonomous University of Mexico, reports that even in her field, "numerous colleagues have mentioned that their articles co-authored with collaborators in the U.S. are much more easily and promptly published than those of similar quality and content that they write alone."

All but excluded from the best-known international publications, many researchers in nonindustrial regions submit their work to local periodicals, few of which are included in the databases that Northern scientists rely on to keep abreast of their field. Of the 3,300 journals catalogued in 1993 by the Science Citation Index, the most popular such database, just 50 are published in less developed nations. The net result, says Ramsay Saunders, who recently stepped down as president of the Caribbean Academy of Sciences, is that in the West Indies and many other poor regions, "valuable advances in science and technology sometimes go unnoticed by researchers in the U.S. and Europe." He cites progress in scoliosis and timber research as examples.

"A lot of locally published literature is just lost," laments Bryan L. Duncan, who directs the International Center for Aquaculture at Auburn University in Alabama and has worked in 35 countries, including an eight-year stint in Southeast Asia. "The vast majority is not the quality we would want, but who is to say that it's not important?" As Northern scientists study increasingly global systems, they may find that Southern research deserves more attention. To scientific workers in poor regions struggling to solve fundamental health and development problems, the knowledge gained from foreign colleagues could make the difference between repetition and progress. *—W. Wayt Gibbs*

Additional research was supplied by Subhadra Menon in New Delhi.

The Sound of One Tree Breathing

As part of the Southern Oxidants Study, Environmental Protection Agency researchers and their colleagues at Duke University are conducting experiments to determine the amount of volatile organic compounds (VOCs) given off by some native tree species. Such natural hydrocarbons are of particular concern because they can react with oxides of nitrogen to form low-level ozone, a serious atmospheric pollutant.

In order for the EPA to formulate strategies to control levels of hydrocarbons and nitrogen oxides resulting from human activity, researchers must establish the rates at which trees release VOCs. Some studies have suggested that in the U.S., naturally occurring volatile organics might exceed those introduced by cars or manufacturing. But these estimates are highly uncertain, and more direct measurements of biogenic sources are sorely needed. So a few trees must suffer in temporary confinement while their effusions are collected and carefully measured (*right*). At least no one is trying to make gasoline this way. *—David Schneider*

ANN STATES *SABA*

The Cold War's Dirty Secrets

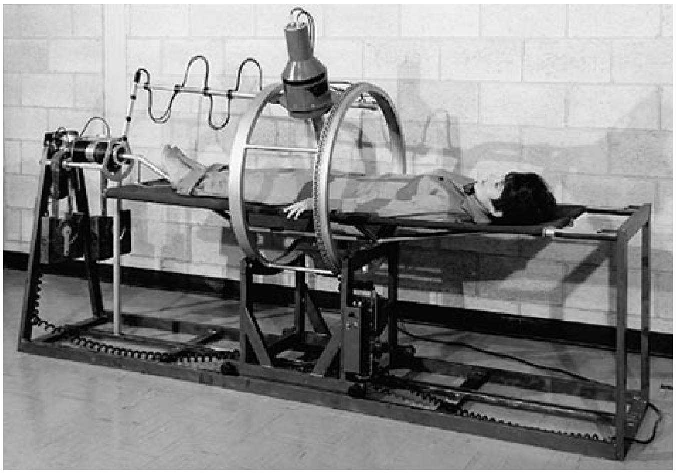
Radiation experiments ignored ethics guidelines

Over the past year a federal advisory committee has doggedly dragged into public view thousands of government-funded studies in which people were deliberately exposed to radiation. The details, to be released in a report next month, are chilling. Some of the tests—conducted between 1944 and 1974—exposed humans to levels of radioactivity now known to be dangerous, and the number of subjects appears to be far greater than previously realized. It is also coming to light that many patients were not well informed about possible dangers or were deceived outright. Perhaps most distressing of all, the Advisory Committee on Human Radiation Experiments has determined that informed consent was required—but ignored.

Some of these horror stories have been known for years. At the top of the list are studies conducted at the University of Rochester and elsewhere in which 18 people were injected with plutonium, 17 of them unknowingly. The tests were designed to determine the risks the substance posed to laboratory workers. Although some of the doses were considered lethal at that time, Wright Langham, then at Los Alamos Scientific Laboratories, justified the work by saying the subjects were hopelessly ill. Nevertheless, four of these "doomed" participants survived another 20 years.

Just as controversial is work that was undertaken by Eugene Saenger between 1961 and 1972 at the University of Cincinnati. Saenger exposed some 88 cancer patients to high levels of whole-body radiation; 62 were African-Americans, a high proportion for a clinical study at the time. According to David S. Egilman, a physician in Braintree, Mass., who is studying the topic, many of the subjects had cancers known to be resistant to whole-body radiation. They were deceived about the likely side effects, and radiation was given in intensities known to be too high for optimal therapeutic effect. The true intent, Egilman contends, was to gather data useful for the Defense Department's nuclear war–planning file. The University of Cincinnati, which is facing lawsuits from the families of victims, refuses to comment. The American College of Radiology defends the work, saying the patients had no alternative therapies available to them.

Other disturbing tales became public only after December 1993, when Energy Secretary Hazel R. O'Leary asked her department to release as much relevant data as possible. For instance, scientists from the Massachusetts Institute of Technology fed mentally retarded boys radioactive iron—yet the parental consent forms made no mention of radioactivity. The list goes on and on: the number of tests logged by the committee is close to 4,000, and, all told, it seems likely that more than 20,000 subjects nationwide were exposed.



U.S. DEPARTMENT OF ENERGY

WHOLE-BODY SCANNER was used to detect the amount and distribution of radiation in experimental subjects.

The recent findings have made untenable the defense that experimenters were simply following contemporary ethical codes. The general manager of the Atomic Energy Commission, the agency that preceded the Department of Energy, is on record as insisting that informed consent be obtained from subjects as early as 1947, when the Nuremberg Code was drafted for the trials of Nazi concentration camp doctors. The code advances informed consent as a requirement for medical research. The Defense Department had a similar directive in place by 1953.

Low-ranking officials seem to have ignored such orders. One possible explanation is that the codes were classified, so some administrators might not have been aware of them. But memorandums now being released suggest another reason. Although the American Medical Association endorsed informed consent in 1946, physicians said the requirement limited their authority. As a result, consent was watered down: two doctors were allowed to certify that a subject understood the setup and would cooperate.

The actual risks in most of the experiments were probably not excessive, notes Ruth R. Faden of Johns Hopkins University, the advisory committee chair. And the data led to procedures that are currently widely used. Faden also points out that some cancer victims may have been willing subjects. Others may have volunteered to help counter the Soviet threat. Nevertheless, no exemptions excusing military-related studies from informed consent have been discovered.

Subjects of medical experiments were not the only victims. Millions of people were exposed to radiation from intentional releases of radioisotopes into the atmosphere during bomb tests. The Department of Energy recently disclosed that there have been more than 250 such releases; soldiers in the 1940s were routinely exposed to fallout. Thousands have joined in class-action lawsuits against the government.

Can perpetrators be judged at 20 to 50 years' remove? Faden says the panel will focus on institutional failings rather than on blaming individuals. But the lessons, she says, carry force even now. The committee is taking a hard look at whether participants in medical research today always know what they are getting into. —*Tim Beardsley*

Tribal Struggle

Stone Age guardians of the Andaman Islands fight to survive

Over the 18 square miles of North Sentinel Island in the Bay of Bengal roams possibly the most isolated tribe on the earth. For centuries these 100-odd hunter-gatherers have enforced their seclusion by greeting approaching ships with arrows. Nearby, on other islands of the Andaman chain, related Negrito groups evince different hazards of battling civilization. Some, having lost, are dying of disease and mysterious sterility. Others pursue guerrilla warfare, vanishing into forests after moonlit raids on immigrant villages.

"Negrito tribes everywhere are declining," observes Ranjit K. Bhattacharya of the Anthropological Survey of India. Soon these remnants of a people who once ranged across Southeast Asia may be gone as well. But not without a fight.

Seafarers have long feared these Stone Age islanders. Wrecked ships (the crews of which they almost invariably killed) supplied them with iron for arrowheads. A practice of throwing the vivisected bodies of their enemies onto a fire—which they cannot make but preserve—appears to have earned them a reputation for cannibalism. (Marco Polo, in addition, declared that their heads resembled those of dogs.) In 1858, after one aborted attempt, British colonizers established a penal settlement on South Andaman Island.

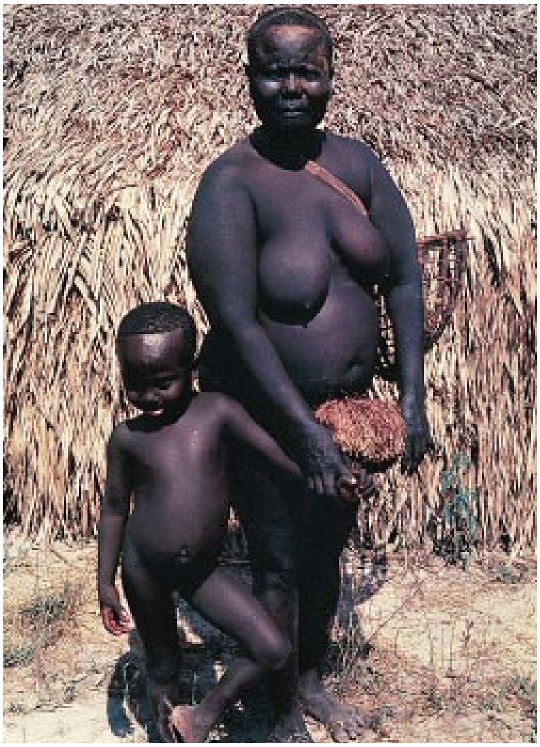
Ten tribes, known as the Great Andamanese, resisted the invasion and suffered high casualties. But peace proved deadlier than war. Alcohol—reward for returning an escaped prisoner—along with syphilis and measles, slashed the initial population of 3,500 to the current mixed-race group of 37. Their chief, Jirake, now wheedles rum from visitors.

Farther south, on Little Andaman, the 700-strong Onge tribe had made peace with the British after a few skirmishes. In 1947 the islands passed to independent India, and in the 1960s thousands of refugees from mainland conflicts were brought to Little Andaman. Luxuriant forests gave way to poor agricultural land, and the Onge way of life became unviable. The remaining 99, gathered in two settlements, depend on government dole.

Unused to clothes, which they wear even when wet, or to starchy foods (their original diet consisted mostly of wild pig, fish and mussels), the Onge suffer from tuberculosis and other ailments. The tribe is doomed by high sterility and infant mortality. Kanarss K. Jindal, the newly appointed director of tribal welfare, frets that the children "have sad eyes" and hopes to introduce them to soccer and volleyball.

Not unlike the fate of the Onge is that of the Shompen, an Indo-Mongoloid tribe on neighboring Great Nicobar Island. Their numbers diminished in the 1980s as a result of dysentery; the 161 survivors hide in dense forests, their health dependent on isolation and medicine men. The Shompen conduct unequal barter with another Mongoloid people, the Nicobarese. This group of 20,000 horticulturists endured Japanese labor camps (during an occupation from 1942 to 1945), converted to Christianity and now watches TV and votes as its leaders direct. Members continue to enjoy tribal privileges such as the right to hunt endangered species.

Unlike these tribes, the Jarawa, who now occupy the western half of Middle and South Andaman Islands, shun peace. Decades of relentless friendliness have induced one group to accept coconuts, iron rods and red ribbons from an occasional shipload of officials. (Such contacts have inherent risks for the exuberantly healthy Jarawa, who are free of even the common cold.) But on all other fronts, the tribe is at war. Its roadblocks and raids failed to stop the Indian government from building a Great Andaman Trunk Road through the Jarawa "reserve." Travelers sometimes fall to well-



DINODIA PICTURE AGENCY

ONGE WOMAN and child are among the last of the Negrito peoples who ruled the Andaman Islands.

Sponging off Shrimp

Sponges are not picky eaters: they dine on nearby particles or microorganisms. But the discovery of flesh-eating sponges in a Mediterranean cave suggests that the phylum Porifera may be more diverse—and perhaps more discerning—than scientists thought. The sponges, from the family Cladorhizidae, were found by Jean Vacelet and Nicole Boury-Esnault of the University of Aix-Marseilles II. They resemble sponges known to exist only in ocean depths.

Finding these creatures in shallower waters enabled the researchers to document their feeding process. Prey are held by filaments covered in small, hook-shaped spicules, which act like Velcro (*left*). Epithelial cells on the outer surface gradually migrate toward the captured food, in this case a shrimp, and envelop it (*micrograph at right*). Once absorbed, the meal is digested over the course of a few days, and new filaments grow in the place of old ones. —*Steven Vames*

BENOIT DECOUT *REA SABA*

aimed arrows, and settlers who venture into the forest for honey or game risk death. In February the tribe attacked a forest outpost, impaling a woman and slaying a calf.

The Jarawa also keep at bay timber merchants and building contractors (who eye the sand on their beaches), and they kill dogs and elephants, which they associate with settlers. In the process they have protected the pristine forests of their territory, along with its unique wildlife. Roughly 40 percent of the species and subspecies of fauna on the Andaman and Nicobar Islands are found nowhere else. Many of the creatures have been threatened by the unceasing development.

But some of the newcomers have guns, as do the bush police, who are charged with keeping the Jarawa and settlers apart. The casualties among the 200-odd Jarawa are not known. Some anxiety about food is, however, evident: villagers say that in their raids, in addition to iron implements, the Jarawa now carry off cooked rice, which the gift-dropping team taught them to eat. Moreover, they display inordinate pleasure on receiving food, often breaking into dance and song. (Given to giggles, they seem to derive much merriment from the ample girth of some officials.)

As yet, the North Sentinel islanders do not exhibit such paroxysms of glee but brandish weapons even as they retrieve their gifts, which are floated ashore. The closest contact with these people occurred in 1991, when a few men clambered onto a government boat and carried off bagfuls of coconuts. The offerings, Bhattacharya explains, are designed to open up channels of communication: in the event of shipwrecks or oil spills, mutual trust could help save the tribe. But in private, academics and administrators alike wonder if the Sentinelese do not know best what their survival entails: distance from all other humans. —*Madhusree Mukerjee*

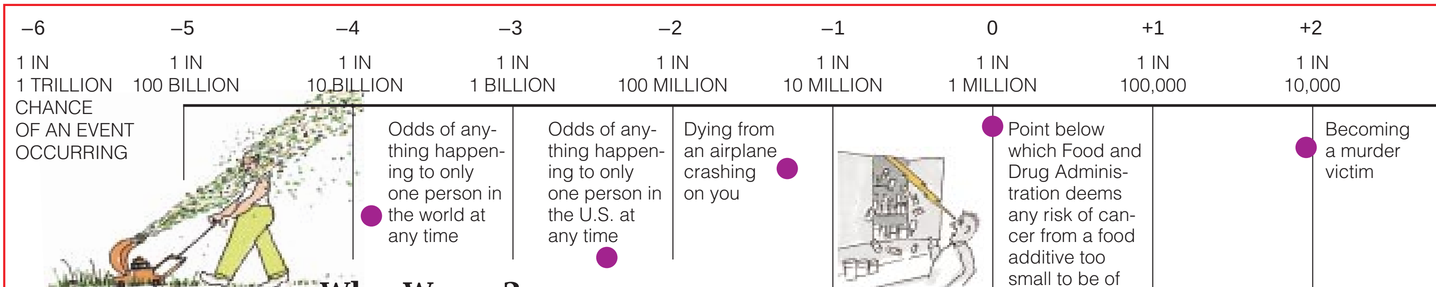
Top Price for the Top Quark

A critic decries the cost of particle physics

After years of rumored sightings, researchers at Fermilab in Batavia, Ill., finally, officially, found the fat but fleeting top quark—one of a class that combines to form neutrons and protons—this past March. Although most physicists considered the result a foregone conclusion, the *New York Times* saw fit to announce it on page one; in the story, Energy Secretary Hazel R. O'Leary called the finding a "major contribution to human understanding of the fundamentals of the universe."

O'Leary is hardly a neutral observer, since the Energy Department is the biggest supporter of U.S. particle physics. Rustum Roy, a materials scientist at Pennsylvania State University and a critic (to put it mildly) of particle physics, has a different perspective. The short version of his response to the news was: "Who gives a damn?" Roy charges that such findings do not justify their cost. Particle physics will receive $642 million this year from the Energy Department and $57.6 million from the National Science Foundation; Fermilab consumed more than $1 billion in the seven years it spent tracking down the top quark.

In assessing the importance of any scientific research, Roy applies the Weinberg criterion: How relevant is the work to anything else? (The criterion is

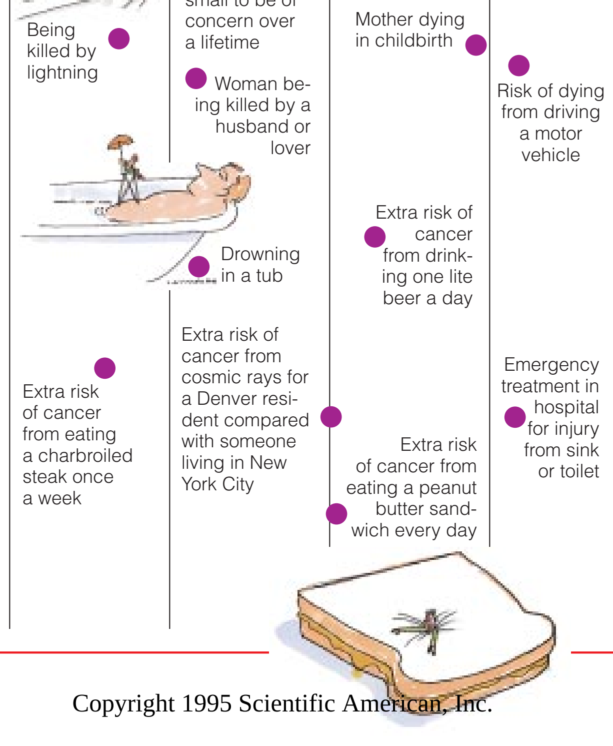


Why Worry?

We are all going to die. The likelihood of how and when becomes quickly muddled by the latest statistics on traffic deaths or on the risks of getting cancer from consuming a peanut butter sandwich every day. Because this barrage of information creates such confusion, John Paling, a former biology professor at the University of Oxford, came up with what he describes as a Richter scale to gauge the dangers of daily living. He got the idea after observing a woman smoking a cigarette while inquiring about the benefits of buying a water-purification kit.

Paling describes his scale in *Up to Your Armpits in Alligators? How to Sort Out What Risks Are Worth Worrying About*. Risks are identified with negative and positive numbers. The midpoint, 0, represents a one-in-a-million hazard, the point below which the cancer risk from a food additive is too small to be of concern to the Food and Drug Administration. Between –2 and –4 are one-of-a-kind risks, the chance of something happening once a year in the entire U.S., what Paling calls the "Bobbitt zone." Going up the scale are still rare threats such as drowning in a bathtub. Above +2, anxiety starts to rise; +6 represents a million-in-a-million risk—in other words, our days are numbered.

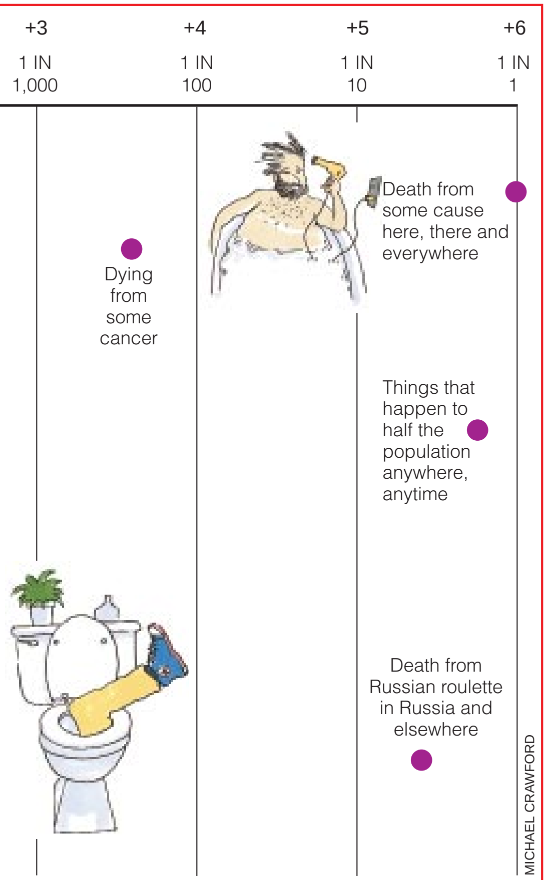
The measures can be used by hypochondriacs to prioritize preoccupations. Or perhaps Republicans in Congress might use the data to block new environmental regulations: a person stands more chance of being struck by lightning than of getting cancer from an organo-whatchamacallit. —*Gary Stix*



named, needless to say, not after Steven, the particle physicist, but Alvin, the nuclear-power engineer and administrator.) Particle physics, Roy argues, fares poorly on the test: the field has little significance for the rest of physics, let alone for biology and the social sciences—it is relevant only to itself. He thinks particle physics will lead not to a theory of everything, as some proponents have claimed, but a theory of nothing.

Roy is also upset that the new, supposedly tight-fisted Congress has not turned its knives on the field. “Why are Republicans taking money away from school lunch programs and keeping it for particle physics?” he cries. “Why aren’t we moving to privatize this?” Roy maintains that particle physicists, if cut off from the public dole, could tap into the riches of such high-tech entrepreneurs as Bill Gates or David Packard.

Roy offered his views to Robert Walker, a Republican who recently became chair of the powerful House Committee on Science—so far to no avail. But the researcher insists it is only a matter of time before Congress imposes “really draconian cuts” on particle physics. “I give them two more years, or maybe four at most,” he says. Seekers of a final theory had better hurry. *—John Horgan*



SOURCE: Up to Your Armpits in Alligators? *by John and Sean Paling; all figures are annual risks for the U.S. except where specified*

CHRIS HUSS *The Wildlife Collection*

COHO SALMON returning from the Pacific to spawn in North American rivers have been getting smaller since 1975, losing an average of 0.012 to 0.059 kilograms a year.

So Many Salmon, But So Little

Ocean warming may be shrinking the size of Pacific salmon

The annual return of salmon to the streams of their birth is one of nature’s great pageants and a dramatic prologue to the spectacle of seasonal change near the rugged edges of the earth’s temperate zones. In the Northern Hemisphere, however, evidence of fundamental changes in this ancient ritual has begun to accumulate.

For more than 20 years, various studies on Pacific Rim rivers have noted that the size of this fish, prized by anglers and epicures alike, has declined. In a study presented last October, biologists Brian Bigler and John H. Helle made the first thorough assessment of the problem: reduced sizes are being found throughout the North Pacific, in a vast area stretching from Japan to California. “It is astonishing and frightening,” says Bigler of Wards Cove Packing Company, a commercial fishing concern.

Previous problems with salmon, particularly reduced populations on specific rivers, have convincingly been tied to human activity—to hydroelectric dams and overfishing as well as to logging and pollution. In the latest findings, though, some more pervasive factor seems to be at work.

Helle, who is at the National Marine Fisheries Service, and Bigler reviewed data from government records, published reports and other sources. The two concluded that of 47 populations on specific rivers (“runs”) of the five salmon species in the North Pacific, 45 experienced decreases in average individual weight between 1975 and 1993. The losses were more than 25 percent for nine runs and less than 10 percent for 10 of the others.

The discovery is worrisome because studies of North Pacific salmon have linked smaller body size to reduced reproductive success. Besides being ill equipped to meet the demands of upstream migration, small fish build inferior nests. They produce smaller eggs that hatch diminutive, less hardy fry.

Unsettling trends have also been noticed among salmon in the North Atlantic. But, in general, the problem there is a decline in numbers, says Kevin D. Friedland of the National Marine Fisheries Service. Waterwheels in the 19th century, then hydroelectricity and pollution, ended runs on many rivers in New England and parts of Europe. Although restoration efforts had reestablished some runs by the 1970s, populations have continued to dwindle.

In the Pacific, size reductions coincide with increased numbers. Throughout the region, hatcheries serve to reestablish and sustain runs on rivers where no wild stocks remain or to enhance wild populations. Virtually all salmon stocks on Japanese rivers are entirely bred in hatcheries, whereas on North American and Russian rivers such fish tend to be a minority. Total hatchery contribution to the North Pacific is about 5.5 billion young salmon a year; the corresponding number of wild young is believed to be about 20 billion.

In recent years hatchery production may have reached such a level that it more than compensates for the reductions in annual returns caused by human activity. This fact, combined with relatively high survival rates of wild fish and record harvests, has led some fisheries experts to suggest that the total number of salmon in the Pacific is higher now than it has ever been.

Some biologists argue that hatcheries genetically weaken stocks by allowing unsuitable fish to survive. Their weaknesses then enter wild populations through interbreeding. But that notion is not rigorously supported by experimental data, and it is generally downplayed as an explanation for size reductions. There is also little evidence that another oft-cited culprit, commercial gill netting, is responsible either.

Instead the explanation that seems best to fit the facts concerns the amount of plankton, krill, young fish and other edibles the marine environment serves up. This so-called oceanic carrying capacity, some experts suggest, can no longer sustain the salmonid hordes. "You're getting older, smaller fish pervading the ocean," Bigler says. "It's a textbook example of population response to overgrazing of limited food resources." Supporting this thesis are recent findings of a precipitous drop in Pacific zooplankton populations over the past 44 years.

Carrying capacity is quite complex, however, and teasing apart its influence on salmon size is proving challenging. Whether fish find food depends on currents, temperature, light, chemical conditions and the mix of organisms in the food web. All these factors are, in turn, entangled with climate. "We're dealing with a very new idea in fisheries science: that climate and the marine environment can cause rather abrupt changes in ocean survival trends," states Dick Beamish of Canada's Department of Fisheries and Oceans.

Since the mid-1970s water flows on certain key rivers, such as the Fraser in British Columbia, have been abating, and water has become warmer. Such havoc, some researchers reason, could be caused only by climate changes—specifically ones traceable to the recurring El Niño Southern Oscillation in the Pacific and the North Atlantic Oscillation, because of their vast movements of warm ocean water.

Indeed, recent studies have correlated salmon population size to climate phenomena. In the Atlantic, a significant factor underlying sparse populations is fewer salmon that spend more than one winter at sea before returning to spawn. Such fish are important to the well-being of Atlantic salmon stocks because of their robustness and superior spawning. Friedland recently found that their populations rise and fall in proportion to the size of the area of the ocean that is between four and eight degrees Celsius, and his latest work suggests that the mechanism may be closely tied to variability in their annual miration pattern, as influenced by climate.

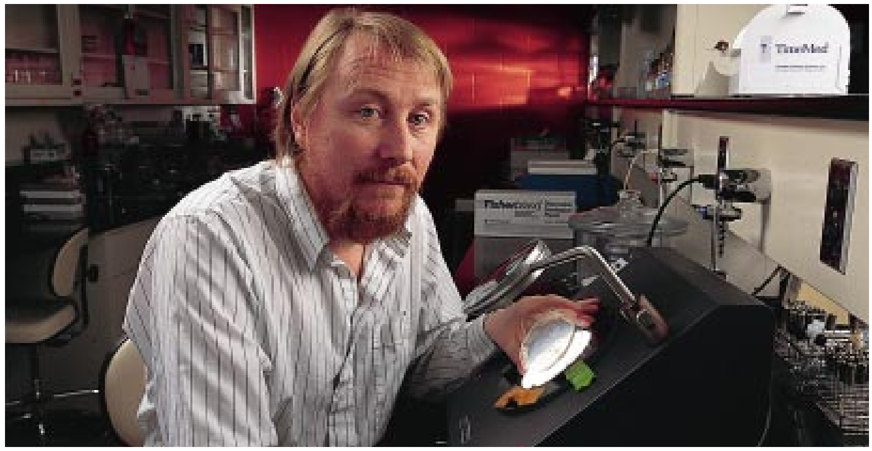
Similar correlations have been established between Pacific salmon and climate. In the late 1980s researchers found that the abundance of pink, chum and sockeye rose and fell with the expansions and contractions of the Aleutian low-pressure index, an enormous winter-weather system.

In the end, far from being another straightforward example of the consequences of human meddling, the case of the mysterious shrinking salmon may turn out to be much more complicated. "Nature's pretty tricky," says Ray Hilborn of the University of Washington. "A lot of changes going on out there we can't control." —*Glenn Zorpette*

Life's a Draw

Chance and survival of the fittest duke it out in bacteria

Arguments over the role of chance events in evolution have long divided biologists. One camp emphasizes the awesome power of natural selection to shape biological forms. Another group, whose most prominent member is Stephen Jay Gould of Harvard University, points out that random happenings—a drought here, an earthquake there—also play a key part.



PETER YATES SABA

***COUNTING COLONIES** of bacteria has led biologist Richard E. Lenski and his colleagues to evolutionary conclusions: fate and natural selection seem evenly matched.*

In principle, the role of chance could be determined by rerunning evolution. If it took much the same course the second time around, that would support the selectionist camp. If replaying life's tape generated an entirely different biota, it would indicate the importance of random events.

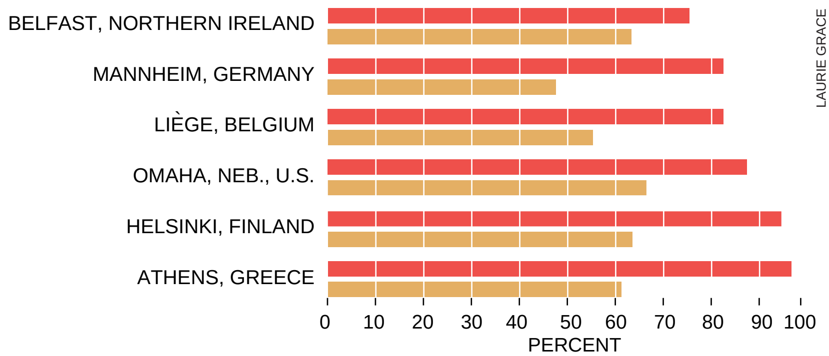
Gould has written, reasonably, that the experiment cannot be done. But Michael Travisano and Richard E. Lenski of Michigan State University and their colleagues have tried to simulate it. First, they propagated multiple colonies of the common bacterium *Escherichia coli.* They measured how quickly each colony could grow and the size of the cells produced. Next, the researchers divided each colony to make subcolonies and switched the food medium. Then they examined how fecundity and cell size in the subcolonies changed over time. Their findings were published in *Science* earlier this year.

When the type of food was first altered, the progeny of different colonies varied markedly in their rate of reproduction. Over time, however, the slowest caught up with the fastest, indicating that selection was in the driver's seat. Subcolonies derived from any one colony all increased their fecundity in lockstep, with little random wandering that could be ascribed to chance. Score one for the selectionists.

On the other hand, the size of individual bacterial cells depended more on blind chance than on selection, even after 1,000 generations in the different food medium. Size did not change overall during that period, and subcolonies varied at random. The shift in food apparently had not caused selection for a

The Naughtiest Teens in the World

Surprise: it is not America's youth. The first study using nearly identical survey methods to measure adolescent delinquency rates in five European nations and nine Western cities [*see excerpts below*] found that Athenian juveniles rank highest. Americans should not gloat, however: young Nebraskans led the world in violent attacks. —*W. Wayt Gibbs*



EVER DELINQUENT   EVER COMMITTED VIOLENT OFFENSE

SOURCE: Delinquent Behavior among Young People in the Western World, *Kugler Publications, 1994*

new optimum size, which a strict selectionist might have expected. One point for the random events school. Travisano and Lenski and their colleagues then held a rematch in which they adjusted the temperature regime rather than the food. The results were broadly the same.

Lenski points out that in a more lifelike setting, over longer periods, the experiment might have come up with different answers—although what they would be nobody knows. For the time being, biologists still have plenty to argue about. —*Tim Beardsley*

As They Lay Dying

Near the end, artificial neural networks become creative

Not too many personal computers are known to hallucinate. But the one belonging to Steven Thaler has been doing so, off and on, for the past couple of years. The physicist, at McDonnell Douglas in St. Louis, has been exploring what happens as an artificial neural network breaks down. But rather than allowing the network to peter out into oblivion, Thaler has a second network observe the last gasps of its dying sibling. Some of those near-death experiences, it turns out, are novel solutions to the problem the net was designed to solve. Thaler says he has found a kind of creativity machine that can function more quickly and efficiently than traditional computer programs can.

An artificial neural network is software written to mimic the function and organization of biological neurons. The system consists of units (representing neurons) connected by links (standing in for dendrites and axons). Like the brain, an artificial network can learn: the programmer presents it with training patterns, which it learns by adjusting the strengths, or weights, of the links. Many researchers use these networks to model brain function and, by destroying part of the net, to mimic disorders such as dyslexia.

Thaler took that concept one step further: he killed his networks. As the links between units were randomly severed over time, the net produced not only gibberish but also some of its original training patterns. For instance, a neural net taught to act as an "or" logic gate would often begin spitting out its trained patterns of 0 and 1 (yes or no) in addition to nonsense (that is, other numbers).

Nothing mystical is going on. Thaler's explanation is that in a fully functioning network, all the weighted links to a given unit are about the same in magnitude but opposite in sign. The sum of several weights to the unit, therefore, is often zero. Without any input, the unit might not notice the loss of those links, because it might not have been receiving any signals from them anyway. A few surviving units are often enough to generate coherent output. Indeed, Thaler used his earlier work to model human near-death visions, suggesting that the reported imagery may have some mathematical basis rather than being purely biochemical.

Thaler soon began experimenting with more sophisticated nets and found that the output contained some unusual juxtapositions of learned patterns and balderdash. To see if those combinations would be useful or esthetically pleasing, he drafted a second neural net to sort through the output and record the most interesting products.

By keeping the dying network partially alive, Thaler has been able to generate many kinds of novelties. For instance, after feeding 30 years' worth of top-10 musical tunes to the networks and letting them run for a few days, Thaler created 11,000 songs—which he has copyrighted. "This diabolical plot will make me the most prolific songwriter of all time," he jokes. From photographs of Thaler's own body movements, another net generated dances. More serious applications included searches for ultrahard materials and for plausible automobile designs.

But what can Thaler's net offer that more traditional programs cannot? "That's the big question," notes Andy Clark, who studies philosophy and neural science at Washington University. The network would have to be compared with classic creativity programs such as EURISKO, Clark observes, which established a benchmark. That algorithm, developed along more traditional programming lines in the 1980s by Douglas B. Lenat and his colleagues at Stanford University, defeated all other programs in various games by coming up with unorthodox solutions. In a military competition, for example, it sank its own disabled ships to improve the overall maneuverability of its fleet.

Nevertheless, EURISKO requires a human to update its heuristics, whereas Thaler's system functions automatically, so dying neural nets may have an advantage in some applications. Thaler also believes his software has philosophical implications. "I am claiming this is a model of consciousness," he asserts. "The images are triggered by internal noise—the network manufactures experiences from stored experiences."

But whether the net emulates the creative mind is debatable. "Creativity isn't a thing in itself," notes mathematical biologist Stephen Grossberg of Boston University. If the network were truly a model of consciousness, it would have to explain something about a particular function of the brain—such as its ability to tune in to only one conversation at a cocktail party. "It may be telling us something about hallucination," Clark echoes, "but creativity seems to be a long way away." —*Philip Yam*

THE ANALYTICAL ECONOMIST

Miracles for Export

In 1960 South Korea's gross domestic product per capita was lower than that of many sub-Saharan countries. During the next 30 years, South Koreans saw this measure of national output jump by an average of nearly 7 percent annually as they rocketed past once far wealthier Brazilians and Argentines. Other East Asian countries also tallied extraordinary growth statistics. Korea, Hong Kong, Taiwan and Singapore became collectively known as the Four Tigers, the Four Dragons or, with an occasional touch of derision or envy, the Gang of Four. Other members of this fast-track club include Japan, Malaysia, Indonesia and Thailand.

Economists, sociologists and political scientists have made careers out of studying the ingredients that shaped the region's economic accomplishments. Books, papers and doctoral theses have weighed in on the lessons that could be gleaned for a Paraguay or a Chad, countries that have yet to achieve an economic takeoff. But no final consensus has been reached on the secrets of success.

The continuing debate has largely focused on the role of government intervention in the marketplace. Most of these East Asian countries manipulated their domestic markets in ways that Washington-based international lending and development institutions considered anathema. During the 1980s, the World Bank and the International Monetary Fund were associated—in policy pronouncements and loan making—with the so-called neoclassical school. Adherents of this view believe that government should limit its exertions to building efficient health care and school systems as well as keeping budget deficits low and inflation in check.

Although none of the East Asian *wunderstadts* ignored the basics, they each did more than just construct classrooms and fret about interest rates. After World War II, Japan set protective tariffs and decided which industries and firms should receive financial credit from the government. Korea promoted steel and heavy industries. The governments of Indonesia, Malaysia and Thailand, among others, obligated banks to channel a portion of their loans to small and medium-size businesses.

Until the early 1990s, the World Bank ignored the economic significance of these events or dismissed them as irrelevant. At the same time, however, the historical record did not go unnoticed by the bank's second largest shareholder. As the world's leading supplier of foreign aid, Japan had become the de facto leader of the view that state intervention is needed in underdeveloped countries because markets cannot always be relied on to guide investment to the areas with the highest growth potential.

To get its point across, Japan's Ministry of Finance decided to give the World Bank a learn-by-doing exercise. It reasoned that the bank might best confront its own prejudices by analyzing the economic factors behind the East Asian boom, including the role of industrial policy and other government interventions. The ministry ponied up a reported $1.2 million for the bank to take a look at the region's experiences.

The 1993 report, *The East Asian Miracle: Economic Growth and Public Policy,* showed that the bank had moved off its neoclassical pedestal. The study acknowledged that state meddling in the market—for instance, directing credit to favored industries—had indeed brought some benefit. "We could no longer be exposed to the criticism that we were ostriches who had ignored the evidence," says John Page, a World Bank economist and the report's chief author.

Even while making this concession, the report did hasten to add that except for export policy, government engineering of the economy may hold few lessons for other developing countries. A critical factor in East Asia—absent from many other parts of the Third World—was a cadre of technocrats who could manage the economy undisturbed by and insulated from lobbying by special political interests.

The *Miracle* report has kept busy a small army of experts who continue to write rebuttals and clarifications to the arguments put forth by the World Bank. Critics contend that the report wrongly concludes that industrial policies and other government-led measures cannot serve as a strategy for the developing world. Where economies are weak, the argument goes, the government may need to promote specific industries or to intercede in financial markets.

The publication, others say, also glosses over the seeming link between "miracle" economies and authoritarian regimes. Leadership in those countries ranged from dominant political parties to outright dictators. But Stephan Haggard, a political scientist at the University of California at San Diego, denies that the enlightened dictatorships that have reigned in some East Asian countries were a prerequisite for an economic liftoff. "The problem can be seen by analyzing the strategies available to a dictator seeking to maximize personal and political power," Haggard wrote in an article for Overseas Development Corporation, a Washington-based policy organization. "He might achieve this objective through growth-enhancing policies, but he might also increase taxes and engage in extortion."

RICARDO AZOURY SABA

KOREAN TIGER, with its own domestic automobile plants, has witnessed phenomenal economic growth since 1960.

Miracles are also associated with luck, and the Asian variety may be no exception. An analysis of different measures—from per capita income growth to secondary school enrollment for some 100 countries—did not necessarily single out the Four Tigers as good candidates for "most likely to succeed," remarks William Easterly, a World Bank economist. A few extraordinary performers are not unusual in any sample.

Easterly emphasizes that policy measures are still important. The East Asian high-growth club members were unlikely to have become economic luminaries if they could not keep inflation in check and maintain good schools. But even if this approach was taken, one developing nation may become a tiger, another a mediocrity. There may be no substitute for the serendipity of being in the right place at the right time and, more disturbingly, a little to the right of center. —*Gary Stix and Paul Wallich*

Lithography Becomes Political Pork

Get it while it's hot

The Superconducting Super Collider is dead, but legislators with a taste for high-tech pork can still pig out on lithography. Like particle physics, lithography—the technique for making circuit patterns on microchips—requires focused beams of energy and large infusions of cash.

More than half of the nearly $60 million in the Department of Defense's main lithography program for the 1995 federal budget was targeted by Congress for pet projects—including the use of x-rays or short wavelengths of ultraviolet light to create a circuit pattern on a chip. Legislators either specified an amount or asked the department's Advanced Research Projects Agency (ARPA) to decide how much the designated research should receive. Either way, lawmakers who would have difficulty distinguishing a memory chip from a microprocessor have usurped at least some of the job responsibilities of ARPA's engineering wizards.

To be sure, some of the earmarked projects might have received ARPA's endorsement anyway. But certainly not all of them. For example, the House Armed Services Committee granted Brookhaven National Laboratory about $2 million for research on coronary angiography, a method of examining clogged arteries in the heart. Its only relationship to lithography is that x-rays are used for this imaging technique. "It was just a comfortable place to put it," says one congressional staffer, explaining why the funds ended up in the lithography budget. The money was earmarked by Congressman George J. Hochbrueckner, who was a member of the House Armed Services Committee before his defeat in the November election. Hochbrueckner hails from Long Island, where Brookhaven is located. (The Department of Defense was making an attempt to remove this item from its budget.)

Other ARPA money, more than $8 million so far, has gone to a new type of lithography that uses hydrogen or helium ions to create circuit patterns on chips. Ion-beam lithography, which has drawn a heatedly negative response from U.S. semiconductor manufacturers, has a true cold war legacy. One version of the technology got its start in Austria in the 1970s at a Vienna company, Sacher Technik Wien, that was working under contract to the East German government. The company went out of business in 1983. The secretive U.S. National Security Agency became interested in ion-beam lithography in the early 1990s, more than five years after two ex-Sacher employees set up in Vienna their own company, called Ion Microfabrication Systems (IMS).

CHIP SIMONS

CORONARY ANGIOGRAPHY EQUIPMENT was funded from the Advanced Research Projects Agency's lithography budget.

The National Security Agency says its curiosity about this type of lithography stems not from any cloak-and-dagger machinations but from a desire to find a technology for making small batches of chips with ultratiny circuit features. It makes its own specialized chips for secure electronic communications. Its officials helped to set up the Advanced Lithography Group (ALG), a Maryland consortium that has received ARPA funding to collaborate in development of the IMS technology. The Austrian firm, a member of the consortium, receives ARPA money through ALG.

ALG also found a friend in a politically conservative congresswoman, Helen Delich Bentley. The former Maryland representative is perhaps best remembered for smashing a Toshiba radio with a sledgehammer on the steps of the Capitol to protest that company's sale of machine tools to the Soviet Union that could make propellers that would have let submarines run more quietly. Bentley helped in funneling ARPA funds to ALG before she retired from Congress in December.

U.S. semiconductor manufacturers perceive ion-beam technology as technically the least promising alternative for making chips with very small circuit components. Instead the industry continues to pursue research on x-rays, electron beams and advanced forms of optical lithography using short wavelengths of ultraviolet light. A leading panel of industry and university lithography experts voted at a meeting last fall to take ion-beam lithography off a list of suggested technologies into which funding should be channeled for commercial development.

Ion-beam advocates point out the biases of their opponents. The current budget gave $15 million to IBM to develop an x-ray lithography component. That project received backing in the budget from Vermont senator Patrick J. Leahy. Vermont is where the IBM development facility is located.

Behind all the finagling lies a comedy of the absurd. Even if one technology prevails over the other, not much of a U.S. lithography industry remains to take advantage of the research. The once dominant U.S. manufacturers of lithography machines, called steppers, today account for less than 10 percent of the global market. American chip manufacturers, meanwhile, have flourished, using Japanese and European lithography equipment. An investment in ALG or IBM may turn out to be nothing more than money spent on Canon and Nikon, the leading Japanese lithography manufacturers who may choose to reverse-engineer the technologies from the U.S. Says G. Dan Hutcheson of VLSI Research: "We run the risk of the U.S.'s being a funding source for Japanese technology on the cutting edge." —*Gary Stix*

A Rogue's Routing

Hackers may ignore individual PCs and undermine the Net

The technique that hacker Kevin Mitnick used to break into a computer-security expert's machine (and onto the front page) was published almost 10 years ago by Cornell University graduate student Robert Morris—the father of the worm that shut down the Internet briefly in 1988—during his summer stint at AT&T Bell Laboratories. No one had used it before, says Bell Labs computer scientist Steven M. Bellovin, because there were so many easier ways of cracking most systems.

Bellovin and others have worked out a modification to Internet protocols that would prevent Mitnick's technique from working again. But malicious hackers have had a decade's worth of technical literature to draw on since then.

Bellovin has a strong idea of what form of sabotage could come next. Indeed, he grows quite animated as he predicts the kinds of debacles most likely to strike this year. Breaking into individual computers is passé, he explains; the new target is the Net itself: the thousands of connections that route data packets from source to destination.

By feeding false update information to routers, hackers can effectively redraw the map of the Internet. It would be as if rogue road builders could invisibly detour every car heading for Dallas so that it ended up in San Francisco. At least one company has already disrupted parts of the Internet by accidentally causing its routers to claim that they could deliver packets to destinations they had no connection to. Network protocols are designed so that routers in one domain must ask their counterparts in other domains how to send packets destined for distant locations—so a single incorrect source of information could cause widespread damage.

Such attacks completely bypass many of the methods computer-security experts use. A route hacker can simply wait until a "secure" connection has been established before detouring packets and taking over the connection. Even more dangerous, falsified routing could let an attacker act as an unwanted intermediary in exchanges of cryptographic keys, passing subtly altered information to each party, explains William R. Cheswick, also at Bell Labs.

So is this really "the death of the Net—film at 11," as doomsayers have been predicting for various reasons since the early 1980s? "I'm waiting for the first big lawsuit," Cheswick says. He foresees one of the pioneers now attempting to transact business over the Internet being shot full of arrows before the rest figure out how to arm themselves.

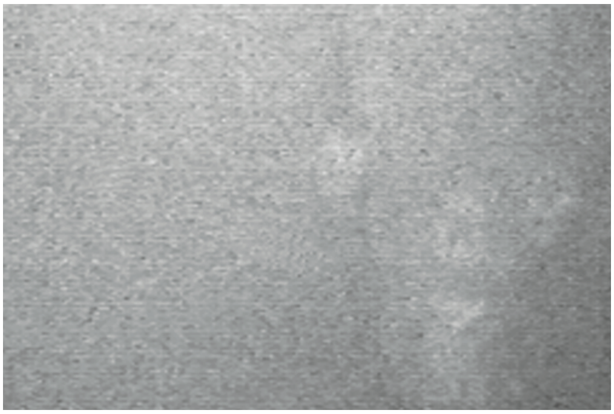
Bellovin believes the most likely debacle would be a class-action suit against a large software company whose bugs—or unintended features—place users at risk. He recounts his discovery that a colleague, who had just connected his PC to the Internet, was running an ftp (file-transfer protocol) server that would have allowed anyone in cyberspace to pull all the files off his hard drive. His associate had no idea that the server was turned on; the Internet software started the program automatically and by default left it open to all.

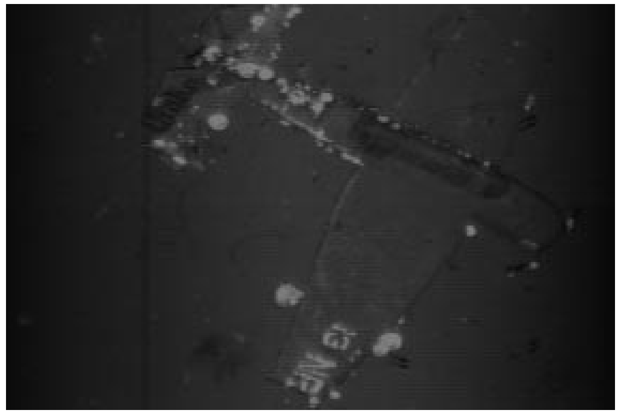
Nevertheless, Bellovin is sanguine. "The business will reach a stable state" once companies understand the risks that they are exposed to, he claims. For many network transfers, information that gets mangled, stolen or lost can be retransmitted. People who need to transact business securely, Bellovin suggests, will use sophisticated cryptographic techniques or some other communications medium. *—Paul Wallich*

Feature from the Dark Lagoon

Peering at shipwrecks in murky depths has been, until recently, a dim affair. But a new development in underwater sensing, the laser line scanner, is clearing things up. Normally, underwater imagery is hampered by the abundance of suspended particles that scatter light like dense fog—as in this video frame of a submerged World War II torpedo bomber (*left*). This limitation restricts subsurface photography to close-up views and makes it difficult to capture large objects.

The new system circumvents that problem, yielding sharp images of, for instance, the same bomber (*right*). The optical instrument uses a single blue-green laser to scan the subject, one line at a time, much like the electron beam of a television picture tube. Blue-green light penetrates seawater more effectively than do other colors, and because the illumination is focused in a single narrow swath, it does not scatter back from all directions as would light from conventional sources. —*David Schneider*





WESTINGHOUSE OCEANIC DIVISION

Putting the Mass Back in Transit

Technology for reviving the collective commute

The empty seats of the automobiles that U.S. commuters drive every day could hold nearly all 250 million Americans. This calculation is testament to the growth of the suburbs and the failure of public transportation to provide access to the vast tracts of housing that extend almost 100 miles away from urban centers.

A few pioneers are now trying to use computer and communications technologies to broaden the definition of mass transit to encompass everything except a car with only a driver. The work of these innovators is hidden away as a small piece of the hundreds of millions of dollars in annual federal and state research and development spending that goes by the name Intelligent Transportation Systems (ITS). The ITS has been investigating how drivers could use radar to detect hard-to-see objects on the road or even relinquish control of a car to a remote computer.

Some ITS projects go beyond making a car into a spaceship and seek to overcome the inherent disadvantage of living carless in the suburbs. (Bus routes often leave passengers miles from home—a reason why only about 2 percent of suburban trips employ buses or trains.) Some plans entail computerized ride-sharing systems that make commuting faster. The transit agency for the Houston metropolitan area expects to test a system this year that within 10 minutes can match riders and drivers who commute along one of its busiest highways, Interstate 10.

Certain other ITS projects that are still on paper sketch a broader framework for suburban transportation. Simple communications with telephones and pagers would give around-the-clock access not just to a job but also to the post office or a nearby shopping mall.

Robert W. Behnke, an Oregon-based transportation consultant, has nurtured for more than 15 years the notion of scheduling car pools, vans and buses with the same sophisticated computer algorithms that airlines employ in their flight reservation systems. Behnke foresees a suburbanite's being able to dial a computer using a touch-tone telephone (or perhaps a pager or hand-held computer) and then keying in a "trip"

MARTIN H. SIMON SABA

PICKING UP well-dressed hitchhikers going to and from work in Washington, D.C., lets Virginia drivers use special lanes that are reserved for cars with multiple occupants.

code that identifies the person and destination. A driver—who has indicated that he is going in the same direction—will retrieve that information by telephone or with a communications device.

As an incentive to participate, drivers would receive a portion of a $1 to $2 fare, which would be credited to their account by the computer system. Car poolers would also be registered in the database as a security check. To be able to guarantee a ride, Behnke envisions extending his suburban transit system beyond the private car. If the computer is unable to match a rider with a car, a "smart" jitney, or roving van, would be dispatched, and it could be tracked with inexpensive satellite-aided navigation systems.

These ideas lack the high-tech allure of remotely controlled vehicles detailed in other ITS projects. But they try to minimize capital expenditures for financially drained local governments.

Despite work on a number of planning studies, Behnke has yet to see his vision realized. He may get a chance to see at least some of his ideas put to the test in a $2-million project called Athena. This transit project—to take place in the city of Ontario, some 45 miles east of Los Angeles—will receive federal and state funds.

Even with such an experiment, transit may never work in the suburbs. There are liability concerns about strangers riding in the same car. And, in general, getting Americans onto buses or trains, or even into car pools, has been a losing proposition. The number of public-transit trips per person dropped from 114 in 1950 to 31 in 1990. Commuters have little inclination to make transit a communal experience: the percentage of U.S. trips to work by car pool fell from about 20 percent in 1980 to roughly 13 percent in 1990. More fundamental approaches to the problem, such as higher gas taxes, are politically unpopular.

Despite the antitransit collective unconscious, there are a few recent success stories. An informal ride-sharing system in suburban Virginia is working smoothly: Washington-area employees hitch rides with drivers who then use a high-occupancy vehicle lane. Van services nationwide take travelers from airports to their suburban doorsteps.

Changes in transportation patterns could have a dramatic impact. Removing just one of every 10 cars on the road during the morning rush hour could cut congestion delays by nearly half while easing suburbanites' dependence on the automobile. It would also have the effect of filling those empty seats with something other than the hot air of radio talk-show hosts. —*Gary Stix*

Electric Genes

Current flow in DNA could lead to faster genetic testing

As more and more of the human genetic blueprint is unraveled, the pressure to know what it means for people grows. Does the baby have any serious genetic problems? Does that teenager carry genes predisposing her to breast cancer? Does a particular adult have the DNA associated with diabetes or with Alzheimer's disease?

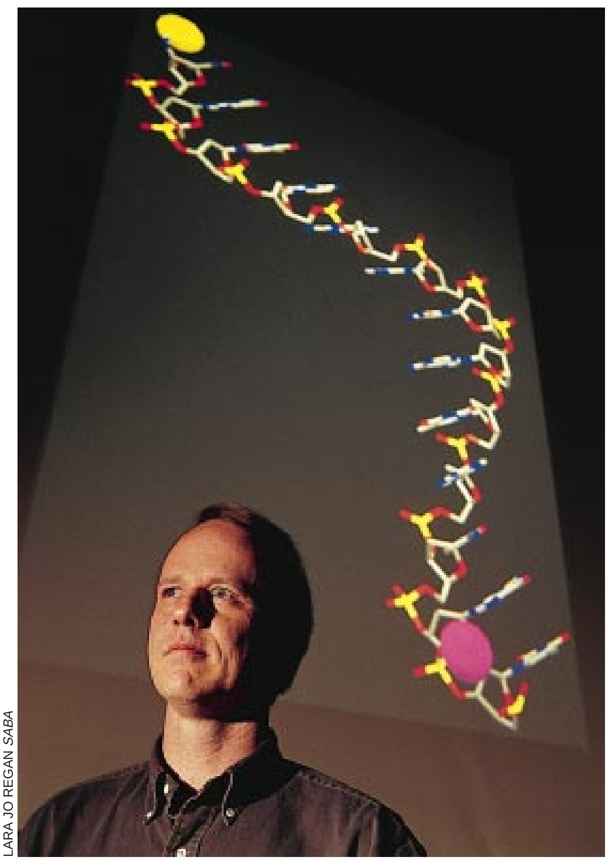
During the past few years, it has become possible to provide answers to more of these questions—to find, for example, the *Apo E4* gene that indicates a greater risk of Alzheimer's or the *BRCA1* gene associated with certain cases of breast cancer. But at present such testing is limited to patients in research projects or those who have a family history of the disease. Widespread speculative genetic screening of populations is too costly to consider—even were it ethically acceptable. This situation may be about to change, at least from a technical standpoint.

Imagine having a machine that could screen almost instantaneously for hundreds, maybe even thousands, of genes. A similar device could also detect the presence of viruses in a person's blood or toxic bacteria in food. These prospects have become realistic as a result of discoveries made in the past few months at the California Institute of Technology.

Chemist Thomas J. Meade and molecular biologist Jon F. Kayyem have been exploring how electrons move in large molecules. Such processes underlie many important biological phenomena; for instance, the conversion of sunlight into plant food by the magnesium chlorophyll molecule depends on stimulation of electron movement through the chlorophyll by the incoming photons. Meade and Kayyem's molecule of study was DNA. They devised a way of binding atoms of ruthenium, a heavy metal, to ribose, one of the backbone components of the helical chains of DNA. Ruthenium atoms act like electrical connectors into and out of the molecule; they have the added virtue of neither disrupting nor distorting its overall shape. Although there has been a long history of using such metals to understand DNA, the ruthenium-ribose combination revealed something extraordinary.

The researchers examined the electrical properties of short lengths of double-helix DNA in which there was a ruthenium atom at each end of one of the strands. Meade and Kayyem estimated from earlier studies that a short single strand of DNA ought to conduct up to 100 electrons a second. Imagine their astonishment when they measured the rate of flow along the ruthenium-doped double helix: the current was up by a factor of more than 10,000 times—over a million electrons a second. It was as if the double helix was behaving like a piece of molecular wire.

For some time, chemists have suspected that the double helix might create a highly conductive path along the axis of the molecule, a route that does



LARA JO REGAN *SABA*

CHEMIST THOMAS J. MEADE is one of a team that has electrified DNA. The technique could hasten cheaper, rapid genetic tests for certain diseases.

not exist in the single strand. Here was confirmation of this idea.

What Meade and Kayyem wanted to know next was whether this newly discovered property could be used to discriminate between DNA strands that were identical to the original and those that differed by one base pair out of the 15 (in other words, a match of 14 of the 15 base pairs). Practically, the test was to see if the perfect match carried significantly more current than the 14-out-of-15 match. To the scientists' delight, there was a large difference, although commercial implications inhibit candor when they are posed the question, "How big is the difference?" (Further, the work has not yet been peer-reviewed or published, so the team remains quite cautious about the details.)

Essentially what Meade and Kayyem have found is an electronic way to distinguish between different sequences of DNA. To convert this finding into a practical device will require concerted development, but even as is, it hints at useful technology. Workers can already build synthetic single DNA strands that can duplicate any known sequence. An amino acid sequence of the gp120 protein of HIV, for example, corresponds to a specific DNA sequence of bases.

Using Meade and Kayyem's invention, one could assemble this gp120 sequence base by base with a ruthenium-doped backbone on an electric current detector, such as a silicon chip. The ruthenium DNA strand could then be used to search for an HIV nucleic acid in a biological sample. If the matching complementary strand from the virus were present, it would bind tightly to the synthetic sequence, and a high flow of electrons would be possible along the molecule's axis. If there were no HIV sequence, there would be no perfect binding with the synthetic DNA, and no current would flow. The answer could be instantaneous—no waiting for gels, no electrophoresis—just a matter of waiting for an indicator to light up.

Meade suspects that the device would need between 15 and 20 bases of single-chain DNA deposited on a chip. Such a stretch of code would allow more than a billion different gene fragments to be specified. And a sophisticated indicator might allow the simultaneous detection of maybe even hundreds of genes. Kayyem is already installed in the Pasadena-based company Clinical Micro Sensors to exploit the discovery. Meade posits that the technique could be useful for any situation in which a rapid, accurate test for the presence or absence of a particular genetic sequence is important. No doctor's office, no farm, no kitchen may be without one. —*David Paterson*

Europeans On-Line

National boundaries still matter, even in cyberspace

Wiring Europe is providing the first real test of one of the more optimistic assumptions of cyberspace: there are no limits in the electronic realm, because national borders can be vaulted with a flick of the mouse. True, technology can render such boundaries meaningless, but do people want that to happen? The answer will determine the fortunes of the companies rushing to hook up Europe.

At first glance, Europe is ripe for the wiring. Although it has a slightly larger population than does the U.S. (where seven million Americans cavort in cyberspace), only a few hundred thousand Europeans have found an on-ramp to the information superhighway. Sales of personal computers and modems are rising; in Britain some 4,000 households sign on to the Internet every month. Nevertheless, corporations vying for Europe's attention may be in for a rude shock. Europeans long ago learned to cope with differing national preferences in everything from cheese to washing machines. But on-line services are new territory, dominated by Americans, who have not (yet) had to worry about internationalization.

So far attempts to win over the market have taken three approaches: global, local and in between. CompuServe, America's biggest on-line service with about 2.5 million customers, takes the first approach. It makes the same databases and discussions available on both sides of the Atlantic. CompuServe remains the largest European service—with about 200,000 customers.

At the other extreme exists a series of small, local bulletin boards. Few make any attempt to serve customers beyond their own country or dialing code. Black Dog in Britain offers ravers a chance to talk about tech-music. In Italy a Bologna bulletin board called Cybersex offers a lively advocacy of transsexuality.

Most vendors are trying to steer between such extremes by providing local appeal to a mass market. To this end, America Online entered a joint venture with Germany's Bertelsmann. The partners will spend $100 million or so of Bertelsmann's money to set up a Europeanized version of the service, to be launched later this year. Microsoft, meanwhile, is talking with virtually every major newspaper, television producer and database in Europe in hopes of tempting customers to sign up for the Microsoft Network. And a collection of European publishers recently created Europe On-Line (not to be confused with the new venture from America Online).

If Microsoft, Bertelsmann and other regionalizers are to succeed, they must overcome the contradiction between mass markets and local appeal—no mean feat. The European market is small: about 17 percent of households have personal computers (versus 37 percent in the U.S.); just 1.6 percent now use on-line services (versus 14 percent in the U.S.). Even local telephone calls cost money. Thus, the average residential telephone in Britain is used for only about five minutes a day (versus more than an hour in the U.S.).

Worse, already small markets are made smaller by a fragmentation of taste. Inteco, a research firm, surveyed more than 10,000 Europeans to determine what services they want from their information autoroutes, autobahns and motorways. Because those services do not yet exist, Inteco researchers looked at video rentals and other things that Europeans will do on networks.

In France, 91 percent of PC owners use the machines to play games; 38 percent admit to working on them. In Germany, in contrast, 48 percent play games, and 62 percent work. (National stereotypes are reinforced by tax laws allowing Germans to deduct home computers used for gainful employment.) In Britain the top 10 television shows are mostly dramas or comedies; in Italy the top 10 are almost entirely football (soccer) broadcasts. In Italy the television is often in the kitchen; in Germany it is in the family room. In Britain more than half of video rentals are accounted for by the 10 most popular films; in Italy, however, the top 10 account for about 15 percent.

Such diversity has economic consequences. It challenges the "department store" model of on-line services concocted by CompuServe and America Online, which attempts to supply all the information potential customers might want. The more diverse the demand, the harder it is to cater to it all.

In contrast, Microsoft and the Internet take a "shopping mall" approach to on-line services. They are establishing networks to open doors to information sources—not to the stuff itself. These networks simply require providers to connect their system to a central network. Microsoft reckons the network should be privately owned, like a mall, and that vendors should pay rent for the safe, well-maintained surroundings.

The Internet harkens back to the traditions of European market towns; it leaves responsibility for the safety and upkeep of the town square to merchants and inhabitants. Given that Microsoft wishes to charge rent and reserves the right to compete directly with any successful provider (it is already devising its own news service), sensible Europeans should try to move on to the Internet. Of course, in cyberspace "economically sensible" and "European" could prove contradictory. *—John Browning*

DAVID REED *Material World/Impact*

GUGLIELMO D' MICHELI *Material World*

NATIONAL DIFFERENCES are apparent even in Europeans' approach to television watching: the British generally view TV in the living room **(left),** ***whereas the Italians often dine in the kitchen at the same time*** **(right).**

PROFILE: BRIAN D. JOSEPHSON

Josephson's Inner Junction

Brian D. Josephson, Nobel laureate, stands at an incandescent intersection in Tucson, Ariz., squinting through thick black spectacles, lost. His floppy white hat has been pulled down so far that—intentionally?—it almost conceals his dark-browed, furtive face. He wears a black T-shirt bearing the digitized likeness of Alan S. Turing, another British prodigy whose relations with the scientific establishment were troubled.

"So, let's see," Josephson mutters, as traffic roars and squeals around him. Someone at the meeting Josephson is attending here has recommended a "very good" restaurant within a few blocks of the conference center, but he's not sure exactly where it is. We cross the street, wander some more, and finally Josephson exults, "Ah, that's it."

Following his finger, I see a squat brick building capped with a gigantic, yellow Mexican hat: Taco Bell. I point out that Taco Bells are more renowned for being fast than for being good, but Josephson, for all his surface diffidence, is stubborn at the core: he cannot be dissuaded. Inside, the restaurant is jammed with Tucsonites, each one seemingly young, blond and tanned, in stark contrast to Josephson.

He gawks at the billboard listing Taco Bell's fare as if it concealed the secret of existence. He confesses he has never eaten Mexican food. Could I explain the meaning of the terms? I expound on the difference between a taco and a burrito. Josephson expresses interest in the nachos. I inform him that nachos, although they do indeed look enticing as pictured on the menu, are more often consumed as a snack or appetizer than as a meal. After more cogitation, he orders a taco and a burrito.

I squelch an impulse to turn to the woman in the turquoise spandex shorts or the man in the yellow muscle shirt and tell them about this awkward little man so improbably in their midst. In 1962, when he was just a 22-year-old graduate student at the University of Cambridge, Josephson discovered that certain superconducting circuits, now known as Josephson junctions, exhibit a seemingly magical quantum property called the Josephson effect.

Josephson junctions have been fashioned into high-speed switches and computers; IBM alone spent more than $100 million investigating the potential of Josephson-junction computers before abandoning its effort a decade ago. The most successful application has been superconducting quantum interference devices, or SQUIDs. These ultrasensitive instruments measure phenomena ranging from the whispers of neurons in human brains to the seismic mumbles of the earth.

DAVID LEVENSON *Black Star*

NOBELIST Brian D. Josephson renounced conventional physics for the study of psychic phenomena.

To no one's surprise Josephson received a tenured position at Cambridge's legendary Cavendish Laboratory in 1972 and won a Nobel Prize a year later. But then he renounced conventional physics and dedicated himself to the study of psychic and mystical phenomena and other forbidden matters. Now he writes articles with titles such as "Physics and Spirituality: The Next Grand Unification?" His contributions to mainstream journals consist, for the most part, of letters denouncing science's narrow-minded attitude toward extrasensory perception and religion.

For years, I have heard physicists trade rumors about Josephson's metamorphosis. What happened? How could someone with so much scientific talent defect to the dark side? I have an opportunity to find out when I visit Tucson to attend a meeting on consciousness, that scientific swamp into which many venture and few return. The symposium has attracted a number of investigators pursuing "alternative" approaches to the mind. Josephson is scheduled to promote his view that music can serve as a key to the secrets of the psyche.

The physicist has apparently accepted my invitation to lunch so that he can rehearse his speech, but I hope to persuade him to talk a bit about his past, too. Josephson speaks haltingly, between nibbles, shunning all but the most fleeting eye contact. His face is framed by wads of charcoal hair and huge sideburns. He was born in Cardiff, Wales, in 1940. As a youth, he was a strict scientific materialist. "I was pretty well turned off religion by the rituals," he says. "I was exposed to the idea that you could explain everything on the basis of science."

Josephson's own genius for scientific explanation first seized the attention of the physics world when he was still an undergraduate. In 1960, his third year at Cambridge, he presented his startled professors with an improved method for calculating the relativistic influence of gravity on Doppler shifts. His paper on the Josephson effect appeared two years later. Just as cinematic ghosts pass through walls in seeming violation of the laws of physics, Josephson proposed, so might electrons "tunnel" through a barrier of insulating material placed in the middle of a superconducting circuit.

Josephson also surmised, based on his reading of quantum mechanics, that the current in such a circuit might actually flow in both directions at once. The interference of the counterflowing currents would create a kind of standing wave extremely sensitive to magnetic or electrical influences. The wave's amplitude would not change smoothly but, like electrons and other quantum entities, would leap between certain values.

Researchers at Bell Laboratories soon confirmed Josephson's predictions, and accolades showered down on him. Subsequent papers on phase transitions and other topics contributed to his reputation as a powerful, original thinker.

Unfortunately, the painfully shy young physicist was ill equipped to handle his fame, according to former colleagues. One remembers Josephson's bolting across the street to avoid encountering him and his wife. Josephson tells me nothing of such incidents, but he does recall feeling no great joy when he learned that he had received the Nobel Prize. "Mainly it was a nuisance, the amount of attention I got," he murmurs between sips of Dr Pepper.

By that time, moreover, Josephson had already begun taking less of an interest in physical matters and more in mental ones. His conversion stemmed at least in part from "the climate of the time," he recalls. Like many other physicists in the 1960s, he became entranced by apparent analogies between quantum mechanics—with its oddly subjective aspects—and Eastern mysticism. George Owen, a Canadian mathematician who was then working at Cambridge, aroused Josephson's interest in telekinesis, poltergeists and other paranormal phenomena.

After some hemming and hawing, Josephson reveals that his transformation also sprang from changes "within." I ask him to elaborate: Did he have psychic or mystical experiences himself? "Well, in some ways, but not—" He pauses. "I've had some strange experiences—" He prods his burrito with a plastic fork. Eventually he tells me that he began having "hallucinatory states" as a result of working too hard on a physics problem. "My experiences were basically a result of a long period of having very little sleep," he says. He took "major tranquilizers" to cope with his mental distress for several years.

In the early 1970s Josephson managed to quell his turmoil without the use of tranquilizers—through transcendental meditation. He still meditates for half an hour or so a day; the practice has given him "something like inner peace." His marriage in 1976 has been another anchor. He and his wife now have a teenage daughter. Josephson feels "her talents are really in a creative way, particularly writing." Discussing his daughter, Josephson permits himself a rare smile.

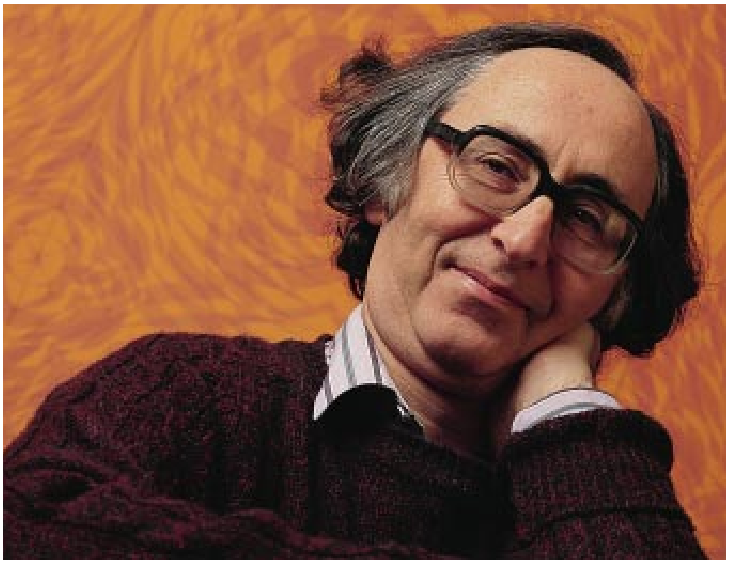
In his articles and published letters Josephson exudes self-assurance, even when making assertions that seem speculative at best. In 1993 the former scientific materialist argued in a letter to *Nature* that religion can help societies "function more harmoniously and more efficiently." He also proposed that religious practices stem from "genes linked to the potential for goodness." (Other letter writers promptly retorted that religions propagate intolerance and brutality at least as often as goodness.)

Josephson also excoriates the scientific community for refusing to accept the evidence for psychic phenomena, or "psi." Here in the clamorous Taco Bell, he seems less confident than in his writings. He calls the evidence for psi "fairly convincing" but admits that "there may always be some problem that may turn up" with the data.

With similar tentativeness, Josephson suggests that quantum mechanics may allow nonlocal "synchronicities" that "produce the appearance of psychic phenomena." But the current theory "doesn't allow the language of process or intention and so on. So I think we're going to have to extend quantum theory so we take that into account as well." Josephson says he feels some kinship with the late David J. Bohm, a physicist who advocated a more holistic approach to physics. (Bohm, in an interview shortly before his death in 1992, said he did not share Josephson's belief or interest in psi.)

Does he have any regrets about having abandoned conventional physics? "No," Josephson replies, firmly this time, "because I consider what I'm doing now to be more important." He has become accustomed to dealing with disapproval from other physicists. "It's not as bad as it used to be, so I guess it doesn't bother me so much at the present time. Occasionally I've arranged lectures on psychic phenomena at the Cavendish, and people on the whole have been quite impressed," he says. He adds, rather wistfully, that he wishes funding agencies were enlightened enough to help him form a psi study group at Cambridge.

Josephson would also like to explore the possibility that scientists can enhance their abilities through meditation. During ordinary consciousness, he informs me, the ego "dominates everything" and suppresses the intuitions available to a "pre-egoic" child. Through meditation "you gain the benefits of the processes that you were influenced by before the ego became dominant, while retaining some of the organizing ability of the ego." (Josephson also believes meditators can learn to levitate during their trances, although he has never mastered this skill.)



DAVID LEVENSON *Black Star*

JOSEPHSON says transcendental meditation, which he practices daily, has helped him to achieve "something like inner peace."

That brings us, finally, to Josephson's theory of music. His meditation has led him to propose that music stems less from superficial cultural influences than from timeless, universal "structures" of the mind. By probing the human response to music, researchers may discern these structures. "So my intuition is, that may have great significance for our understanding of mind."

Josephson's own tastes in music include classical and even a bit of rock and roll. "Some of that has considerable merit," he muses. "Something that may appear quite noisy, sometimes you get the feeling there is something quite deep to it." Any personal favorites? He purses his lips for a moment and then reveals that he likes Simon and Garfunkel's "Bridge over Troubled Water." "I don't know if that's particularly deep, but—"

In the background Whitney Houston is shrieking, "I will always love youuuuuu!" The Taco Bell lunch throng has come and gone. Josephson has consumed his burrito and taco, which were "quite good." He glances at his watch; he is keen to get back to the conference to hear a lecture on "information physics, neuromolecular computing and consciousness." We dump our leftovers, stack our trays and head back out into the blinding day. *—John Horgan*

SCIENTIFIC AMERICAN

The Global Tobacco Epidemic

Cigarette smoking has stopped declining in the U.S. and is rising in other parts of the world. Aggressive marketing and permissive regulations are largely to blame

by Carl E. Bartecchi, Thomas D. MacKenzie and Robert W. Schrier

Since the early 1960s, medical research, public information campaigns and government assessments have exposed the dangers of tobacco smoke. The result has been a substantial drop in the number of smokers in the U.S.—from a peak of 41 percent to its current level of about 25 percent. Yet despite considerable scientific evidence and continuing exhortations from the medical community, the trend has now mostly ceased: the number of adult smokers has remained static since 1990. Similarly, the proportion of adolescents who smoke has changed little in the past 10 years. Perhaps even more disconcerting is that in the global picture, cigarette production during the past two decades has increased an average of 2.2 percent each year, outpacing the annual world population growth of 1.7 percent. Because of growing cigarette consumption in developing nations, worldwide cigarette production is projected to escalate by 2.9 percent a year in the 1990s, with China leading the way with jumps near 11 percent a year.

To understand the driving forces behind modern directions in tobacco consumption and to formulate strategies to combat its pervasiveness, the medical community has had to extend observations beyond the individual smoker and the addictive power of nicotine. The focus of some recent work has been on the tobacco industry itself. In this context, changes in smoking behavior depend in large part on cigarette pricing, advertising, promotion and exportation. Researchers in preventive medicine and public health agree that education campaigns must be supplemented. The new strategies should aim to regulate the marketing of cigarettes, to raise taxes on tobacco and to rethink current trade practices.

CARL E. BARTECCHI, THOMAS D. MACKENZIE and ROBERT W. SCHRIER collaborate at the University of Colorado School of Medicine. Bartecchi, who helped to found the Southern Colorado Clinic in Pueblo, is a clinical professor in the department of medicine at the school. MacKenzie is a general internist with the Denver Department of Health and Hospitals and an assistant professor of medicine at the University of Colorado Health Sciences Center. Schrier is professor and chairman of the department of medicine at the University of Colorado School of Medicine.

A 1,000-Year-Old Habit

Although humans probably began sampling tobacco during the first millennium, based on Mayan stone carvings dated at about A.D. 600 to 900, physicians did not begin to suspect in earnest that the plant could produce ill effects until around the 19th century. The renowned colonial physician Benjamin Rush condemned tobacco in his writings as early as 1798. By the mid- to late 1800s, many prominent physicians were expressing concern about the development of certain medical problems connected with tobacco. They suggested a relation between smoking and coronary artery disease, even recognizing the potential association between passive smoking (inhaling smoke from the air) and heart problems. They also noted a correlation with lip and nasal cancer.

Although tobacco use was relatively common in that century, it did not produce the widespread illnesses it does today. Individuals of the time consumed only small amounts, mostly in the form of pipe tobacco, cigars, chewing tobacco or snuff. Cigarette smoking was rare. Then, in 1881, came the invention of the cigarette-rolling machine, followed by the development of safety matches. Both significantly encouraged smoking, and by 1945 cigarettes had largely replaced other forms of tobacco consumption. Smokers increased their average of 40 cigarettes a year in 1880 to an average of 12,854 cigarettes in 1977, the peak of American consumption per individual smoker.

The rise in tobacco use made the adverse effects of smoking more apparent. Medical reports in the 1920s strengthened the suspected links between tobacco and cancers. The connection to life span was first noted in 1938, when an article in the journal *Science* suggested that heavy smokers had a shorter life expectancy than did nonsmokers.

In 1964 U.S. Surgeon General Luther Terry released a truly landmark public health document. The work of an independent body of scientists, it was the country's first widely publicized official recognition that smoking causes cancer and other diseases. In many subsequent reports by the surgeon general's office, cigarette smoking has been identified as the leading source of preventable morbidity and premature mortality in the U.S. These statements enumerate many experimental studies in which animals have been exposed to tars, gases and other constituents in tobacco and tobacco smoke.

JAMES LUKOSKI

EXPORTATION OF CIGARETTES, such as the Gold Coast brand marketed by R. J. Reynolds in San'aa, Yemen, is one strategy tobacco companies are adopting in order to offset lowered consumption in the U.S.

A review of mortality statistics underscores the tobacco epidemic. Of the more than two million U.S. deaths in 1990, smoking-related illnesses accounted for about 400,000 of them and for more than one quarter of all deaths among those 35 to 64 years of age. When deaths from passive smoking are included, estimates near 500,000. A recent British study suggests that one half of all regular smokers will die from their habit. Statistically, each cigarette robs a regular smoker of 5.5 minutes of life.

Tobacco also drains society economically. The University of California and the Centers for Disease Control and Prevention (CDC) have calculated that the total health care cost to society of smoking-related diseases in 1993 was at least $50 billion, or $2.06 per pack of cigarettes—about the actual price of a pack in the U.S. That price figure greatly exceeds the average total tax on a pack of cigarettes in the U.S., now currently about 56 cents. Although a 1989 study suggested that smokers "pay their own way" at the current level of excise taxes (because they live long enough to contribute to their pensions and to Social Security but die before they enjoy the benefits), more recent estimates show otherwise. These newer calculations, which incorporate the effects of passive smoking, indicate that smokers take from society much more than they pay in tobacco taxes.

Moreover, because tobacco kills so many people between the ages of 35 and 64, the cost of lost productivity must be accounted for in the analysis. With this factor in mind, the average annual expense to an employer for a worker who smokes has been pegged at $960 a year. The total toll of tobacco consumption for the country may exceed $100 billion annually.

Staying Addicted

Expanding public awareness of tobacco's dangers is probably the reason for the decline of smoking in the U.S. Based on a 1993 count, an estimated 46 million adults (25 percent) in the U.S. smoke—24 million men and 22 million women. Smoking prevalence is highest among some minority groups—in particular, black males, Native Americans and Alaskan natives—and among those with the least education and those living below the poverty level. Perhaps most disheartening, an estimated six million teenagers and another 100,000 children younger than 13 years smoke.

Of greatest concern, however, are the most recent data from the CDC. They suggest that overall smoking prevalence among adults, at approximately 25 percent, was unchanged from 1990 to 1993. Moreover, smoking prevalence among adolescents has remained static since 1985.

On a global scale, the patterns are even more alarming. Although the smoking habit in most developed countries is being kicked, the rate of decline has been slower than it has been in the U.S. In developing countries, data suggest that cigarette smoking is up by 3 percent a year. Richard Peto of the University of Oxford has estimated that the total number of deaths attributable to smoking worldwide will increase from 2.5 million today to 12 million by the year 2050.

There are several reasons for the current pattern of cigarette consumption. In the U.S. the flattened decline since 1990 may have resulted from recent price wars between premium and discount brands. For years, tobacco companies have maintained a high profit margin despite dwindling consumption because smokers are willing to pay a stiff price to satisfy their craving. The addiction of their customers has allowed tobacco companies to boost the price of cigarettes with minimal fear of losing sales. Throughout the 1980s, for instance, the price of cigarettes outpaced inflation.

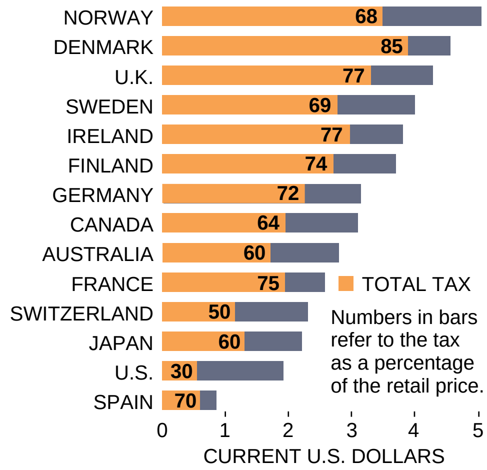
But the rapidly rising popularity of discount brands has made cigarettes cheaper and more accessible. The market share of these brands rose from 10 percent in 1987 to 36 percent in 1993. They earn about five cents per pack in profit, compared with 55 cents for a brand-name pack. This trend forced a series of price cuts by the major brands in 1993. If the cuts are sustained, smoking prevalence in the U.S., especially among young and poor populations (for whom price is often important), may actually increase.

Despite the recent price deductions, cigarette companies are likely to remain financially and politically potent entities. The two biggest corporations—Philip Morris and R. J. Reynolds—expanded their presence appreciably in the consumer market during the 1980s by acquiring many big, nontobacco-related firms. For instance, Philip Morris bought Kraft and General Foods, among others, and now sells more than 3,000 different products. In 1992 it ranked as the seventh largest industrial corporation in the U.S., with $50 billion in sales, and made more money that year than any other U.S. business. Almost half of its $4.9 billion in profits came from cigarette sales. The major tobacco companies will undoubtedly be able to afford a price war with discount competitors as well as establish their own discount brands. And unlike the discounters, the larger companies can market their products aggressively, both at home and abroad.

There has been little government restriction on the marketing of cigarettes in recent decades. The bulk of today's regulations stems from actions taken shortly after the 1964 surgeon general's report. In 1966 the Federal Trade Commission required that all cigarette packages carry warning labels and that tobacco advertising not be directed at

Socioeconomics of Smoking

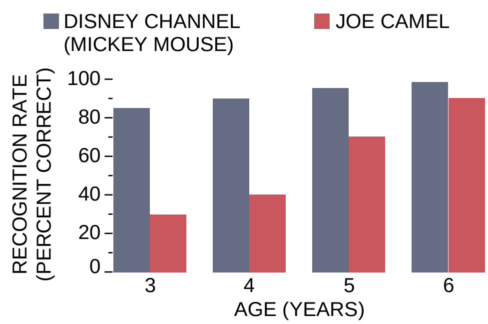
Average Price of 20 Cigarettes



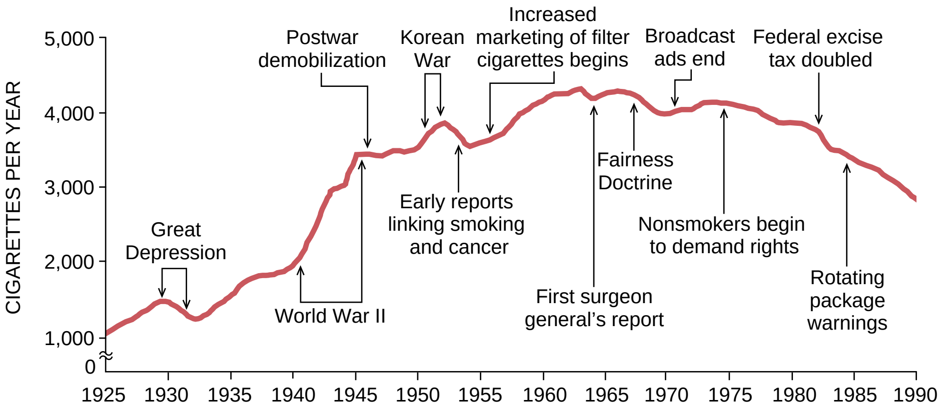
Deaths from Preventable Causes in the U.S. in 1990

CAUSE	ESTIMATED NUMBER OF DEATHS	PERCENT OF TOTAL DEATHS
Tobacco	400,000	19
Diet/activity patterns	300,000	14
Alcohol	100,000	5
Microbial agents	90,000	4
Toxic agents	60,000	3
Firearms	35,000	2
Sexual behavior	30,000	1
Motor vehicles	25,000	1
Illicit use of drugs	20,000	<1
TOTAL	1,060,000	50

Recognition of Logos by Children

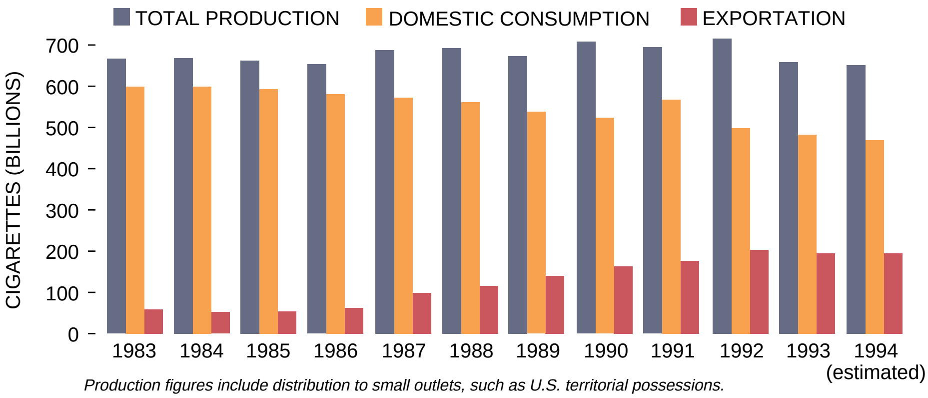


Cigarette Consumption per U.S. Adult



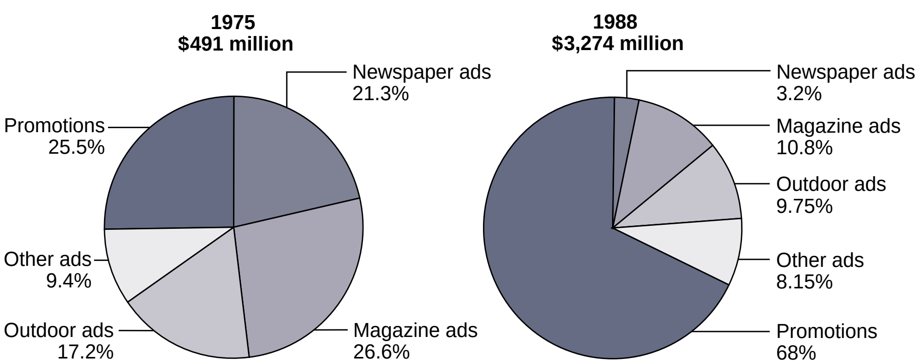
SOURCE: U.S. Department of Health and Human Services

U.S. Cigarette Distribution



Production figures include distribution to small outlets, such as U.S. territorial possessions.

Cigarette Marketing Expenditures



people younger than 25 years. In 1967 the Federal Communications Commission mandated that local television and radio stations that ran cigarette advertisements had to compensate by airing public service announcements about the product's bad effects. Cigarette advertising shifted completely from television and radio in 1971, when Congress banned all such advertising on electronic media. As a result, magazines, newspapers and billboards took over.

Magazines benefited substantially from the shift. For some, revenues from the tobacco industry increased by $5.5 million per magazine a year (figured in 1983 dollars). Moreover, coverage of smoking-related health issues decreased by 65 percent in magazines that carried cigarette advertisements, as compared with a 29 percent drop in similar stories in periodicals that did not carry them. During the three years following the electronic media ban, per capita cigarette consumption actually rose slightly before resuming its drop. Many analysts attribute the brief surge to the cessation of public service announcements that coincided with the electronic ban.

Several major health organizations, including the American Medical Association, have recommended barring tobacco advertising completely. In other developed countries, such antitobacco legislation is common. By mid-1986, 55 countries had enacted legislation to control advertising: 20 with total bans, 15 with strong partial bans and 20 with moderate ones.

In comparison, the U.S. has been lax. Since 1971 it has passed nothing to restrict cigarette advertising, despite many attempts to do so by several members of Congress. Instead tobacco has become the most displayed product on billboards and the second most marketed product in magazines. In 1989 Philip Morris spent $2 billion on advertisements—more than any other U.S. company. The industry as a whole increased expenditures on advertising from $500 million in 1975 to more than $5 billion in 1992, which represents a fourfold increase in constant 1975 dollars.

The tobacco industry has also concentrated on promotion. Sponsorship of sporting events, the distribution of free cigarettes and other strategies have increased from one quarter of the marketing budget in 1975 to two thirds in 1988. Of particular note are widely televised competitions such as the Camel motocross and the Virginia Slims tennis tournament (Philip Morris, however, voluntarily pulled out of sponsorship last year). Despite the advertising ban on electronic media, sponsorship of such tournaments has granted substantial airtime. For example, during the 93-minute broadcast of the 1989 Marlboro Grand Prix, the Marlboro name flashed on the screen or was mentioned by the announcers 5,933 times, for a total of 46 minutes. For 18 of those minutes, the Marlboro name was clear and in focus, which represents an estimated $1 million of commercial airtime.

Appealing to the Young

These marketing efforts have begun to focus on minorities, women and children, an approach that the medical community has strongly criticized. Recent work has found a link between the start of smoking and targeted advertising [*see box on pages 50 and 51*]. Children are probably the most vulnerable segment. The average age that habitual smoking begins has been dropping for decades and is currently 14.5 years. Approximately 90 percent of regular smokers start before the age of 21.

Data suggest that the tobacco industry recognizes these figures and develops advertisements to appeal to children and teenagers. For example, in 1988 R. J. Reynolds fashioned "Old Joe Camel," a cartoon character who shoots pool, rides motorcycles and associates with attractive women as he smokes cigarettes. Three years after the campaign began, several studies clearly demonstrated that children and teenagers easily recognized Joe Camel. One study showed that six-year-olds knew the character as often as they picked out Mickey Mouse. Teenagers were likewise influenced. Surveys done in 1988 and 1990 show that the proportion of teenage smokers who bought the Camel brand increased from 0.5 to 32 percent. In this same period, it is estimated that Camel cigarette sales to minors soared from $6 million to $476 million.

How can minors purchase cigarettes so easily? Although 46 states have laws prohibiting the sale of cigarettes to minors, compliance has been consistently poor in many communities. Furthermore, only nine states have stopped the sale of cigarettes in vending machines, and just 22 states prohibit the free distribution of cigarettes to underage individuals. Many legislators and health officials have suggested that the sale of cigarettes should require licensing similar to that for the sale of alcohol.

The tobacco industry may be relying on a more insidious strategy to gain new customers—that is, through smokeless tobacco. It is estimated that 7.5 million people in the U.S. use tobacco in this way, with snuff (shredded tobacco that is sucked but not chewed) being the most popular. A 1994 *Wall Street Journal* article reported that tobacco companies doctor their snuff products to increase the nicotine that the mouth can absorb—an alarming assertion, given that the average age of first-time snuff users is nine years. The article argued that these companies try to appeal to young people with pleasant-tasting, milder forms that are lower in free nicotine (that is, in a form immediately available for absorption) and then to graduate these consumers to very potent, very addictive brands high in free nicotine. Although the tobacco companies admit they can control the amount of nicotine in the product, they deny that they do so to addict individuals.

Despite the toxic effects of tobacco, the agencies primarily responsible for protecting the consumer—the Food and Drug Administration and the Consumer Product Safety Commission—have never subjected tobacco products to health and safety regulations commonly used for hazardous compounds. Their permissiveness very likely stems from the lobbying efforts by the tobacco industry, which is considered one of the most powerful at all levels of government today. In 1992 the tobacco industry donated more than $4.7 million to the leading political parties, representing three times the amount given in 1988. Few government representatives refuse these contributions. In 1989 it was reported that over a two-year period, 420 of 535 congressional representatives and 87 of 100 senators accepted tobacco campaign contributions, making the tobacco lobby one of the most influential forces in government.

Tobacco companies have also formed industry organizations to channel contributions and to give them a central voice. One such group, the Tobacco Institute, has consistently created a public smoke screen by questioning the association between smoking and human disease. As late as 1986, a Tobacco Institute publication stated that "eminent scientists believe that questions relating to smoking and health are unresolved."

The regulation of tobacco products may change because of allegations that the industry has knowingly manipulated the nicotine content of cigarettes to maximize addiction and has suppressed evidence pointing out the hazards. To many, the congressional testimony of tobacco executives last year—who stated their belief that nicotine was not addictive and that cigarettes were not proved to cause cancer—was designed to avoid any potential liability. The FDA is now considering regulating cigarettes as drug-delivery systems for nicotine (which can act as a stimulant or as a tranquilizer, depending on the amount used). Although the new Congress is much less enthusiastic about such regulations, the health care community regards the FDA's case to be strong enough to force passage of some kind of legislation. How new laws will alter the control of tobacco is unclear. But given current standards of consumer product safety, the introduction and sale of a similar product today would assuredly be denied.

ADDICTION TO NICOTINE can turn youngsters such as this Albanian boy in Kosovo into regular smokers (*top*). After having been sworn in before a congressional hearing in April 1994 (*bottom*), executives from major U.S. tobacco firms later testified that they believed nicotine was not addictive.

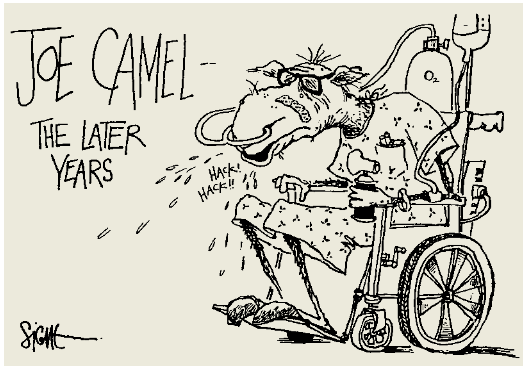
Taxing Tobacco

The political might of the tobacco industry has prevented significant rises in cigarette excise taxes, thus keeping the cost of the habit affordable. The federal tax has risen from eight cents a pack of 20 cigarettes in 1951 to only 24 cents today, a climb far less than inflation. (Adjusted for inflation, the 1951 tax would be approximately 40 cents today, meaning that the tax has actually declined.) With the addition of state and local taxes, the average total tax on a pack of cigarettes in the U.S. is 56 cents, or 30 percent of the average retail price. This amount is substantially lower than those in many other industrial nations.

The tobacco companies have employed a strategy of

The Medical Effects of Tobacco Consumption

Discovered in the early 1800s and named nicotianine, the oily essence now called nicotine is the main active ingredient of tobacco. Indeed, researchers recognized in 1942 that smoking dried tobacco leaves was basically a means of administering nicotine, just as smoking opium was a means of obtaining morphine. Nicotine, however, is but a small component of cigarette smoke, which contains more than 4,700 chemical compounds, including 43 cancer-causing substances. Condensates of tobacco smoke suspended in acetone and applied to the skin of mice for long periods cause papillomas or carcinomas at the site. Toxins in cigarette smoke cause breaks in the DNA of cultured human lung cells. In some cases, these carcinogens greatly accelerate the mutation rate in dividing cells, which in turn can lead to tumor formation.



SIGNE WILKINSON *Cartoonists & Writers Syndicate*

Unfortunately for the smoker, no threshold level of exposure to the toxins has been found. What is clear is that years of cigarette smoking vastly increase the risk of developing several fatal conditions. In addition to being responsible for more than 85 percent of lung cancers, smoking is associated with cancers of the mouth, pharynx, larynx, esophagus, stomach, pancreas, uterine cervix, kidney, ureter, bladder and colon. Cigarette smoking is thought to cause about 14 percent of all leukemias and 30 percent of new cases of cervical cancer in women. All told, cigarette smoking is responsible for 30 percent of all deaths from cancer and clearly represents the most important preventable cause of cancer in the U.S. today.

Smoking also increases the risk of cardiovascular disease, including stroke, sudden death, heart attack, peripheral vascular disease and aortic aneurysm. Cigarettes caused almost 180,000 deaths from cardiovascular disease in the U.S. in 1990. Components of cigarette smoke damage the inner lining of blood vessels, which can lead to the development of atherosclerosis. The toxins can also stimulate occlusive elements in coronary arteries, thus promoting clots to form and triggering spasms that close off the vessels. In this regard, the smoking of a single cigarette can profoundly disturb blood flow to the heart in patients with existing coronary artery disease.

Furthermore, cigarette smoking is the leading cause of pulmonary illness and death in the U.S. In 1990 smoking caused more than 84,000 deaths from pulmonary disease, mainly resulting from such problems as pneumonia, emphysema, bronchitis and influenza.

Passive smoking—the breathing of sidestream smoke (emitted from the burning tobacco between puffs) or of smoke exhaled by the smoker—poses a similar health risk. A 1992 Environmental Protection Agency report emphasized the dangers, especially of sidestream smoke. This type of smoke contains more particles of smaller diameter and is therefore more likely to be deposited deep in the lungs. On the basis of this report, the EPA has classified environmental tobacco smoke as a "group A" carcinogen, to which radon, asbestos, arsenic and benzene belong.

Of the estimated 53,000 annual deaths in the U.S. caused by passive smoking, 37,000 come from associated heart disease. A nonsmoker living with a smoker has a 30 percent higher risk of death from ischemic heart disease or myocardial infarction. Lung cancer risk also skyrockets. Any exposure from a spouse who smokes is associated with at least a 30 percent excess risk of lung cancer. Increasing daily amounts and the number of years of smoking significantly heighten the risk. The figure jumps to 80 percent if the spouse has been smoking four packs a day for 20 years. Another recent study points out that 17 percent of the cases of lung cancer among nonsmokers can be attributed to exposure to high levels of tobacco smoke during childhood and adolescence.

The health consequences of smoking among women are of special concern because of the deleterious effect on reproduction. Unfortunately, the fastest-growing segment of smokers in the U.S. is women younger than 23 years. Smoking reduces fertility, spurs the rate of spontaneous abortions and stillbirths, can cause excessive bleeding during pregnancy and results in lower birth weights in infants. Moreover, children of smokers do not grow as large or attain the same level of educational achievement as unexposed children.

Smoking is a significant cause of cardiovascular diseases and strokes in women, especially if they also use oral contraceptives. Lung cancer has now surpassed breast cancer as the primary cause of death from cancer among women. In 1993 lung cancer claimed an estimated 56,000 deaths, whereas breast cancer took 46,000 lives.

The elderly also face special harm from smoking. Among those older than 65, the rates of total mortality among current smokers are twice those among people who have never smoked. A 1992 *Time* magazine article noted that three life insurers owned by tobacco companies charge smokers nearly double for term insurance.

Smoking is associated with a variety of other ailments: cataracts, delayed healing of broken bones, periodontal maladies, predisposition to ulcer disease, hypertension, brain hemorrhages and skin wrinkles, to name just a few.

Recently some studies have suggested that cigarette smoking ameliorates symptoms of Alzheimer's disease. It is not surprising that with its powerful effect on the central nervous system, nicotine may influence the condition. Yet methodological flaws plague many of these studies. Moreover, other researchers suggest that smoking may increase the risk of Alzheimer's, in that it accelerates the natural consequences of aging. With its many and potent toxins, cigarettes would in any case be an inappropriate vehicle for delivering nicotine should the compound ever prove valuable in treating Alzheimer's.

There is much to be gained by those who kick the habit. After a year, mortality from heart disease drops halfway back to that of a nonsmoker; by five years, it drops to the rate of nonsmokers. A person's risk of lung cancer is cut in half in five years; by 10 years, it drops almost to the rate of nonsmokers. Such gains make sense, however, only if smokers quit in time, before they show any signs of tobacco's lethal effects.

Looking for a Market among Adolescents

In 1992, the most recent year for which data are available, the tobacco industry spent $5 billion on domestic marketing. That figure represents a huge increase from the approximate $250-million budget in 1971, when tobacco advertising was banned from television and radio. The current expenditure translates to about $75 for every adult smoker, or to $4,500 for every adolescent who became a smoker that year. This apparently high cost to attract a new smoker is very likely recouped over the average 25 years that this teen will smoke.

In the first half of this century, leaders of the tobacco companies boasted that innovative mass-marketing strategies built the industry. Recently, however, the tobacco business has maintained that its advertising is geared to draw established smokers to particular brands. But public health advocates insist that such advertising plays a role in generating new demand, with adolescents being the primary target. To explore the issue, we examined several marketing campaigns undertaken over the years and correlated them with the ages smokers say they began their habit. We find that, historically, there is considerable evidence that such campaigns led to an increase in cigarette smoking among adolescents of the targeted group.

National surveys collected the ages at which people started smoking. The 1955 Current Population Survey (CPS) was the first to query respondents for this information, although only summary data survive. Beginning in 1970, however, the National Health Interview Surveys (NHIS) included this question in some polls. Answers from all the surveys were combined to produce a sample of more than 165,000 individuals. Using a respondent's age at the time of the survey and the reported age of initiation, the year the person began smoking could be determined. Dividing the number of adolescents (defined as those 12 to 17 years old) who started smoking during a particular interval by the number who were "eligible" to begin at the start of the interval set the initiation rate for that group.

Mass-marketing campaigns began as early as the 1880s, which boosted tobacco consumption sixfold by 1900. Much of the rise was attributed to a greater number of people smoking cigarettes, as opposed to using cigars, pipes, snuff or chewing tobacco. Marketing strategies included painted billboards and an extensive distribution of coupons, which a recipient could redeem for free cigarettes and a variety of other premiums. Some brands included soft-porn pictures of women in the packages. Such tactics inspired outcry from educational leaders concerned about their corrupting influence on teenage boys. Thirteen percent of the males surveyed in 1955 who reached adolescence between 1890 and 1910 commenced smoking by 18 years of age, compared with almost no females.

The power of targeted advertising is more apparent if one considers the men born between 1890 and 1899. In 1912, when many of these men were teenagers, the R. J. Reynolds company launched the Camel brand of cigarettes with a revolutionary approach. In the months before the Camel debut, every city and hamlet in the country was bombarded with print advertising. According to the 1955 CPS, initiation by age 18 for males in this group jumped to 21.6 percent, a two thirds increase over those born before 1890. The NHIS initiation rate also reflected this change. For adolescent males, it went up from 2.9 percent between 1910 and 1912 to 4.9 percent between 1918 and 1921.

It was not until the mid-1920s that social mores permitted cigarette advertising to focus on women (public clamor forced a 1919 ad aimed at women by Lorillard Company to be with-

identifying themselves as "citizens against tax abuse" and spend millions of dollars to fight against tax increases. The industry probably feels that such large expenditures are necessary to fend off the perceived threat to profits. A 10 percent increase in cigarette prices reduces consumption by 4 percent, mostly by keeping new smokers away. The drop would probably be much larger in populations that are highly sensitive to price hikes, such as teenagers. Because the vast majority of smokers begin in their teens, a major drop in teenage smoking would seriously threaten the future of the tobacco industry.

The American Heart Association, the American Cancer Society and the American Lung Association have recommended a $2 increase in federal tax per pack of cigarettes. Their counsel is based in part on data from Canada and California, where cigarette tax hikes have significantly reduced consumption. The potential benefits of high tobacco taxes are many. Several states have earmarked the revenue for public health education campaigns, antitobacco advertising and health care for the poor. With the recent fall in average cigarette prices, tax increases have become particularly important to counteract a possible acceleration in consumption among teenagers. Many bills were introduced in Congress during the past two years to up the federal tax by $1 or less, both as independent proposals and as part of health care reform packages. None received sufficient support to pass.

Exports to Hook New Customers

Even if tighter marketing restrictions and higher excise taxes prove successful in decreasing tobacco smoking in the U.S., the industry has a means to counteract loss of revenue: exportation. Indeed, although total cigarette consumption in the U.S. has been declining for over a decade, domestic production has been buoyed by steadily increasing shipments overseas. Cigarette exportation climbed from 8 percent of production in 1984 to 30 percent in 1994. Unmanufactured tobacco leaf exportation now exceeds 34 percent of production. The U.S. currently leads the world in tobacco exports and has capitalized on the markets in underdeveloped countries, which have few if any restrictions on advertising or product labeling.

The six major transnational tobacco companies (three are based in the U.S., and the others in the U.K.) have experienced little resistance to gaining footholds in these developing countries. Often the only competition comes from government-run production companies, for which there is marginal advertising. Many have argued that the introduction of Western advertising in developing countries has done much more than shift the existing market share to the transnational companies. In Hong Kong, for instance, only 1 percent of the women now smoke, but the advertising by transnational companies has heavily targeted women—clearly indicating that the companies are making an effort to carve out a new market. This kind of exploitation equates with the disgraceful export of opium from England to China in the 1830s.

As assistant secretary of health under President George Bush, James Mason stated in 1990 at the Seventh World Conference on Tobacco and Health in Perth, Australia: "It is unconscionable for the mighty transnational tobacco companies to be peddling their poison abroad, particularly because their main targets are less developed countries. They play our free trade laws and exdrawn). In 1926 a poster depicted women imploring smokers of Chesterfield cigarettes to "Blow Some My Way." The most successful crusade, however, was for Lucky Strikes, which urged women to "Reach for a Lucky Instead of a Sweet." The 1955 CPS data showed that 7 percent of the women who were adolescents during the mid-1920s had started smoking by age 18, compared with only 2 percent in the preceding generation of female adolescents. Initiation rates from the NHIS data for adolescent girls were observed to increase threefold, from 0.6 percent between 1922 and 1925 to 1.8 percent between 1930 and 1933. In contrast, rates for males rose only slightly.

The next major boost in smoking initiation in adolescent females occurred in the late 1960s. In 1967 the tobacco industry launched "niche" brands aimed exclusively at women. The most popular was Virginia Slims. The visuals of this campaign emphasized a woman who was strong, independent and very thin. Consequently, and ironically in conjunction with the rise of the women's movement, initiation in female adolescents nearly doubled, from 3.7 percent between 1964 and 1967 to 6.2 percent between 1972 and 1975 (NHIS data). During the same period, rates for adolescent males remained stable.

Thus, in four distinct instances over the past 100 years, innovative and directed tobacco marketing campaigns were associated with marked surges in primary demand from adolescents only in the target group. The first two were directed at males and the second two at females. Of course, other factors helped to entrench smoking in society, such as the provision of free cigarettes during wartime and the romanticization of actors smoking in the movies. Yet it is clear from the data that advertising has been an overwhelming force in attracting new users.

JAMES NUBILE JB Pictures

OLD JOE CAMEL coolly surveys Times Square.

Despite subsequent regulations barring advertisement geared to minors, the tobacco industry nonetheless has retained its targeted approach. In 1988 R. J. Reynolds introduced another novel campaign featuring the ultracool cartoon character "Joe Camel." Recent data from California indicate a rising market share for Camels in youths and a turnaround in adolescent smoking prevalence, which had been declining during the 1980s. Future surveys of adults will most likely show a jump in adolescent smoking initiation rates coincident with the rolling out of this campaign. *—John P. Pierce and Elizabeth A. Gilpin*

The authors are researchers at the Cancer Prevention and Control Program at the University of California, San Diego.

port policies like a Stradivarius violin, pressuring our trade promotion agencies to keep open—and force open in some cases—other nations' markets for their products."

The U.S. government has remained remarkably unresponsive to such claims. One reason for this inaction may be that in 1990 the U.S. realized a $4.2-billion trade surplus from tobacco exports, accounting for 35 percent of the entire agricultural trade surplus. In that same year Vice President Dan Quayle stated during a North Carolina news conference: "I don't think it's news to North Carolina tobacco farmers that the American public as a whole is smoking less. We ought to think about the exports. We ought to think about opening up markets, breaking down the barriers rather than erecting new tariffs, new quotas and things of that sort."

Much of the aggressive trade behavior by the tobacco industry is sanctioned under section 301 of the 1974 Trade Act. Public health officials have repeatedly asked that Congress reevaluate this act and current tobacco trade practices, but the representatives have failed to take action. Moreover, many have questioned the difference in health standards applied to domestically consumed and exported tobacco. For example, there are no U.S.-imposed regulations on the labeling, tar content or advertisement of tobacco products exported to developing nations. It is truly ironic that the U.S. freely exports cigarettes to countries such as Colombia in the face of huge expenditures on both sides to restrict the trade of cocaine, which accounts for many fewer deaths.

The magnitude of tobacco-related diseases and deaths around the world cannot be overstated. Cigarette smoking is the number-one preventable cause of premature death in the U.S. Yet it enjoys remarkable tolerance among Americans. On an international level, trends in smoking prevalence suggest an even more profound rejection or ignorance of the health risks of tobacco use.

For these reasons, the obstacles facing the antitobacco campaign are formidable. From the standpoint of public health, it is clear that a battle plan must emphasize intervention programs that specifically target children and adolescents. These plans include increased government regulation of tobacco advertisements, restrictions on access to cigarettes by minors and higher tobacco taxes. Other possibilities are support for personal-injury litigation against the tobacco industry, government subsidies for the conversion of tobacco crops to other plants and comprehensive restrictions on workplace and public smoking. Concerned citizens, public health officials, government representatives and health care providers must join forces to adopt a multidisciplinary strategy to control this global epidemic.

FURTHER READING

SMOKING AND HEALTH: REPORT OF THE ADVISORY COMMITTEE TO THE SURGEON GENERAL OF THE PUBLIC HEALTH SERVICE. PHS Publication No. 1103, 1964.

THE WORLDWIDE SMOKING EPIDEMIC: TOBACCO TRADE, USE AND CONTROL. William C. Scott et al. in *Journal of the American Medical Association,* Vol. 263, No. 24, pages 3312–3318; June 27, 1990.

TOBACCO, BIOLOGY AND POLITICS. Stanton A. Glantz. Health Edco, Waco, Tex., 1992.

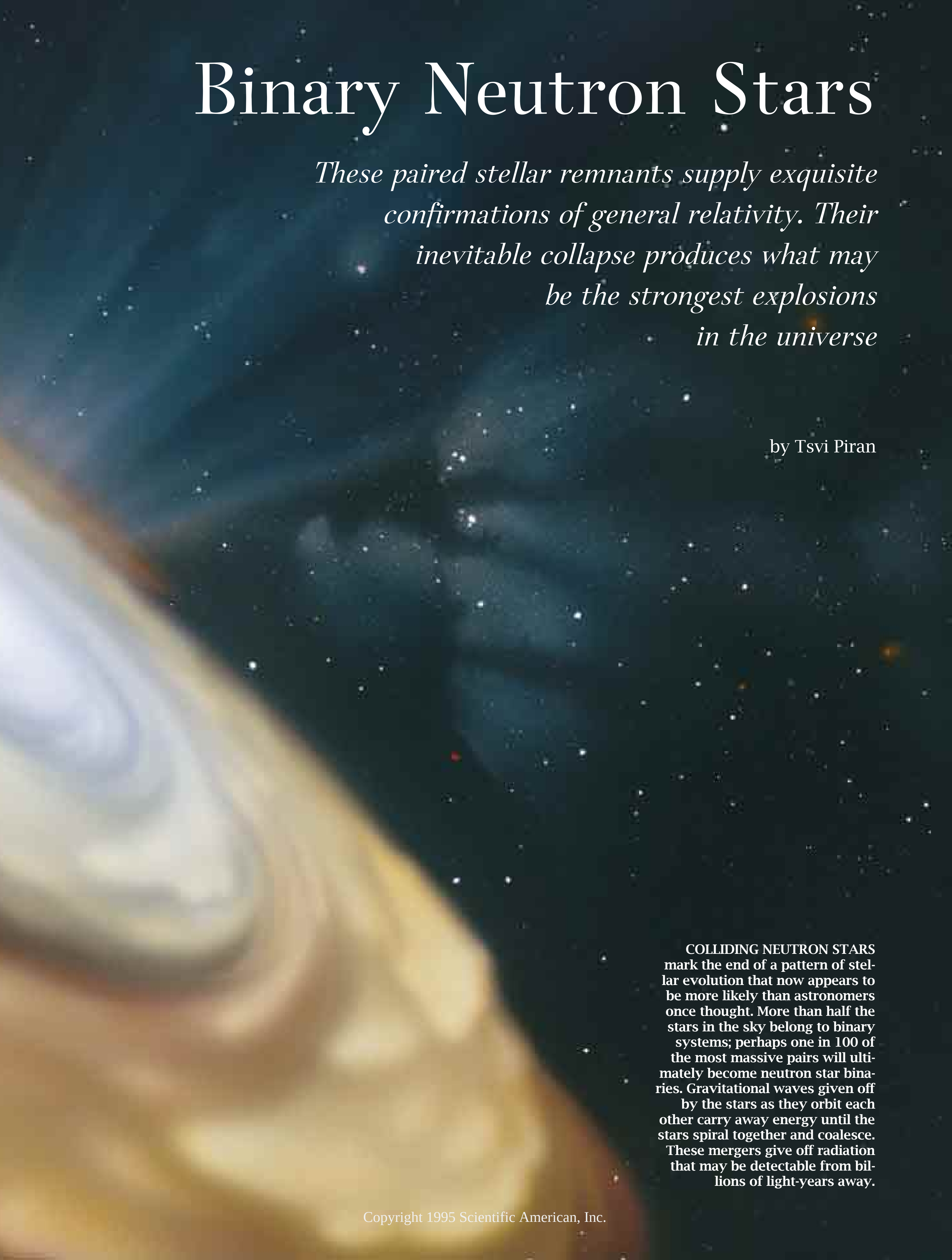
THE HUMAN COSTS OF TOBACCO USE. C. E. Bartecchi, T. D. MacKenzie and R. W. Schrier in *New England Journal of Medicine,* Vol. 330, No. 13, pages 907–912; March 31, 1994, and Vol. 330, No. 14, pages 975–980; April 7, 1994.

AL KAMAJIAN

Binary Neutron Stars

These paired stellar remnants supply exquisite confirmations of general relativity. Their inevitable collapse produces what may be the strongest explosions in the universe

by Tsvi Piran



COLLIDING NEUTRON STARS mark the end of a pattern of stellar evolution that now appears to be more likely than astronomers once thought. More than half the stars in the sky belong to binary systems; perhaps one in 100 of the most massive pairs will ultimately become neutron star binaries. Gravitational waves given off by the stars as they orbit each other carry away energy until the stars spiral together and coalesce. These mergers give off radiation that may be detectable from billions of light-years away.

In 1967 Jocelyn Bell and Antony Hewish found the first pulsar. Their radio telescope brought in signals from a source that emitted very regular pulses every 1.34 seconds. After eliminating terrestrial sources and provisionally discarding the notion that these signals might come from extraterrestrial intelligent beings, they were baffled. It was Thomas Gold of Cornell University who realized that the pulses originated from a rotating neutron star, beaming radio waves into space like a lighthouse. Researchers soon tuned in other pulsars.

Even as Bell and Hewish were making their discovery, military satellites orbiting the earth were detecting the signature of even more exotic signals: powerful gamma-ray bursts from outer space. The gamma rays triggered detectors intended to monitor illicit nuclear tests, but it was not until six years later that the observations were made public; even then, another 20 years passed before the bursts' origin was understood. Many people now think gamma-ray bursts are emitted by twin neutron stars in the throes of coalescence.

The discovery of binary neutron stars fell to Russell A. Hulse and Joseph H. Taylor, Jr., then at the University of Massachusetts at Amherst, who began a systematic pulsar survey in 1974. They used the Arecibo radio telescope in Puerto Rico, the largest in the world, and within a few months had found 40 previously unknown pulsars. Among their haul was a strange source named PSR 1913+16 (PSR denotes a pulsar, and the numbers stand for its position in the sky: 19 hours and 13 minutes longitude and a declination of 16 degrees). It emitted approximately 17 pulses per second, but the period of the pulses changed by as much as 80 microseconds from one day to the next. Pulsars are so regular that this small fluctuation stood out clearly.

Hulse and Taylor soon found that the timing of the signals varied in a regular pattern, repeating every seven hours and 45 minutes. This signature was not new; for many years astronomers have noted similar variations in the wavelength of light from binary stars (stars that are orbiting each other). The Doppler effect shortens the wavelength (and increases the frequency) of signals emitted when a source is moving toward the earth and increases wavelength (thus decreasing the frequency) when a source is moving away. Hulse and Taylor concluded that PSR 1913+16 was orbiting a companion star, even though available models of stellar evolution predicted only solitary pulsars.

The surprises did not end there. Analysis of the time delay indicated that the pulsar and its companion were separated by a mere 1.8 million kilometers. At that distance, a normal star (with a radius of roughly 600,000 kilometers) would almost certainly have blocked the pulsar's signal at some point during its orbit. The companion could also not be a white dwarf (radius of about 3,000 kilometers), because tidal interactions would have perturbed the orbit in a way that contradicted the observations. Hulse and Taylor concluded that the companion to PSR 1913+16 must be a neutron star.

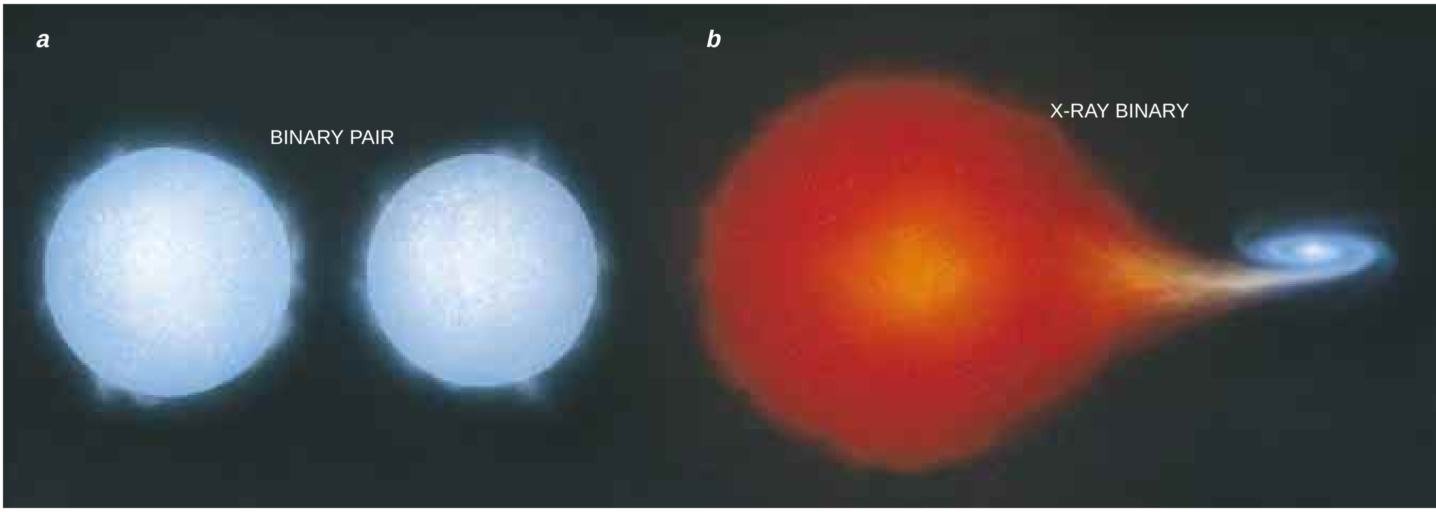
This finding earned the two a Nobel Prize in Physics in 1993. Astronomers have since mastered the challenge of understanding how binary neutron stars might exist at all, even as they have employed the signals these strange entities produce to conduct exceedingly fine tests of astrophysical models and of general relativity.

How a Binary Neutron Star Forms

By all the astrophysical theories that existed before 1974, binary neutron stars should not have existed. Astronomers believed that the repeated stellar catastrophes needed to create them would disrupt any gravitational binding between two stars.

Neutron stars are the remnants of massive stars, which perish in a supernova explosion after exhausting all their nuclear fuel. The death throes begin when a star of six solar masses or more consumes the hydrogen in its center, expands and becomes a red giant. At this stage, its core is already extremely dense: several solar masses within a radius of several thousand kilometers. An extended envelope more than 100 million kilometers across contains the rest of the mass. In the core, heavier elements such as silicon undergo nuclear fusion to become iron.

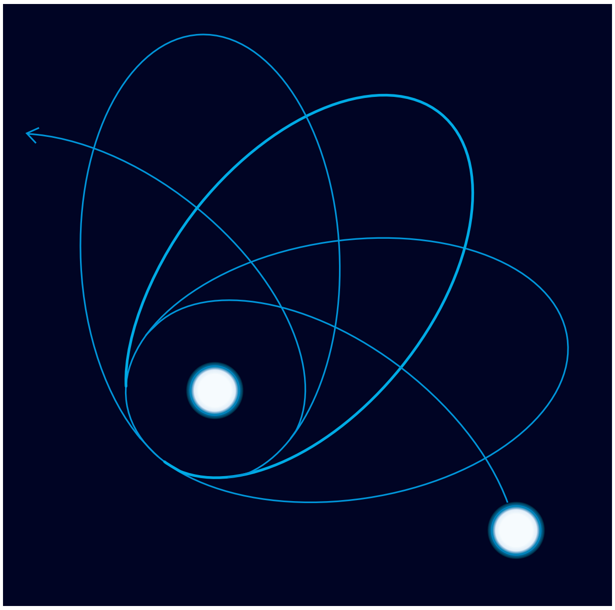
When the core reaches a temperature of several billion kelvins, the iron nuclei begin to break apart, absorbing heat from their surroundings and reducing the pressure in the core drastically. Unable to support itself against its own gravitational attraction, the core collapses. As its radius decreases from several thousand kilometers to 15, electrons and protons fuse into neutrons, leaving a very dense star of 1.4 solar masses in a volume no larger than an asteroid.



Meanwhile the energy released in the collapse heats the envelope of the star, which for a few weeks emits more light than an entire galaxy. Observations of old supernovae, such as the Crab Nebula's, whose light reached the earth in A.D. 1054, reveal a neutron star surrounded by a luminous cloud of gas, still moving out into interstellar space.

More than half the stars in the sky belong to binary systems. As a result, it is not surprising that at least a few massive pairs should remain bound together even after one of them undergoes a supernova explosion. The pair then becomes a massive x-ray binary, so named for the emission that the neutron star produces as it strips the outer atmosphere from its companion. Eventually the second star also explodes as a supernova and turns into a neutron star. The envelope ejected by the second supernova contains most of the mass of the binary (since the remaining neutron star contains a mere 1.4 solar masses). The ejection of such a large fraction of the total mass should therefore disrupt the binary and send the two neutron stars (the old one and the one that has just formed) flying into space with velocities of hundreds of kilometers per second.

Hulse and Taylor's discovery demonstrated, however, that some binaries survive the second supernova explosion. In retrospect, astronomers realized that the second supernova explosion might be asymmetrical, propelling the newly formed neutron star into a stable orbit rather than out into the void. The second supernova also may be less disruptive if the second star loses its envelope gradually during the massive x-ray binary phase. Since then, the discovery of three other neutron star binaries shows that other massive pairs have survived the second supernova.



ORBITAL PRECESSION, the rotation of the major axis of an elliptical orbit, results from relativistic perturbations of the motion of fast-moving bodies in intense gravitational fields. It is usually almost undetectable; Mercury's orbit precesses by less than 0.12 of a degree every century, but that of PSR 1913+16 changes by 4.2 degrees a year.

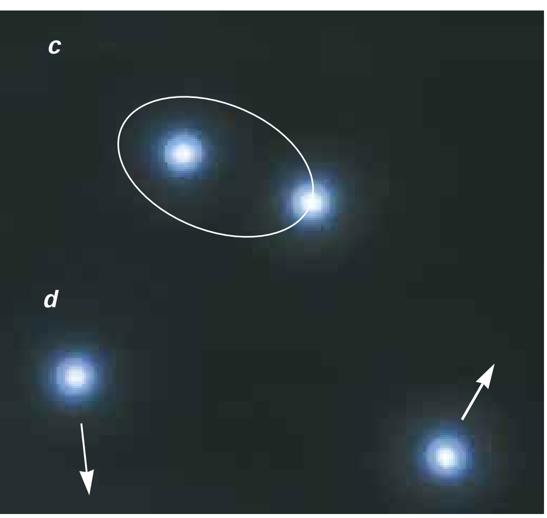
Several years ago Ramesh Narayan of Harvard University, Amotz Shemi of Tel Aviv University and I, along with E. Sterl Phinney of the California Institute of Technology, working independently, estimated that about 1 percent of massive x-ray binaries survive to form neutron star binaries. This figure implies that our galaxy contains a population of about 30,000 neutron star binaries. Following a similar line of argument, we also concluded that there should be a comparable number of binaries, yet unobserved, containing a neutron star and a black hole. Such a pair would form when one of the stars in a massive pair formed a supernova remnant containing more than about two solar masses and so collapsed to a singularity instead of a neutron star. Rarer, but still possible in theory, are black hole binaries, which start their lives as a pair of particularly massive stars; they should number about 300 in our galaxy.

Testing General Relativity

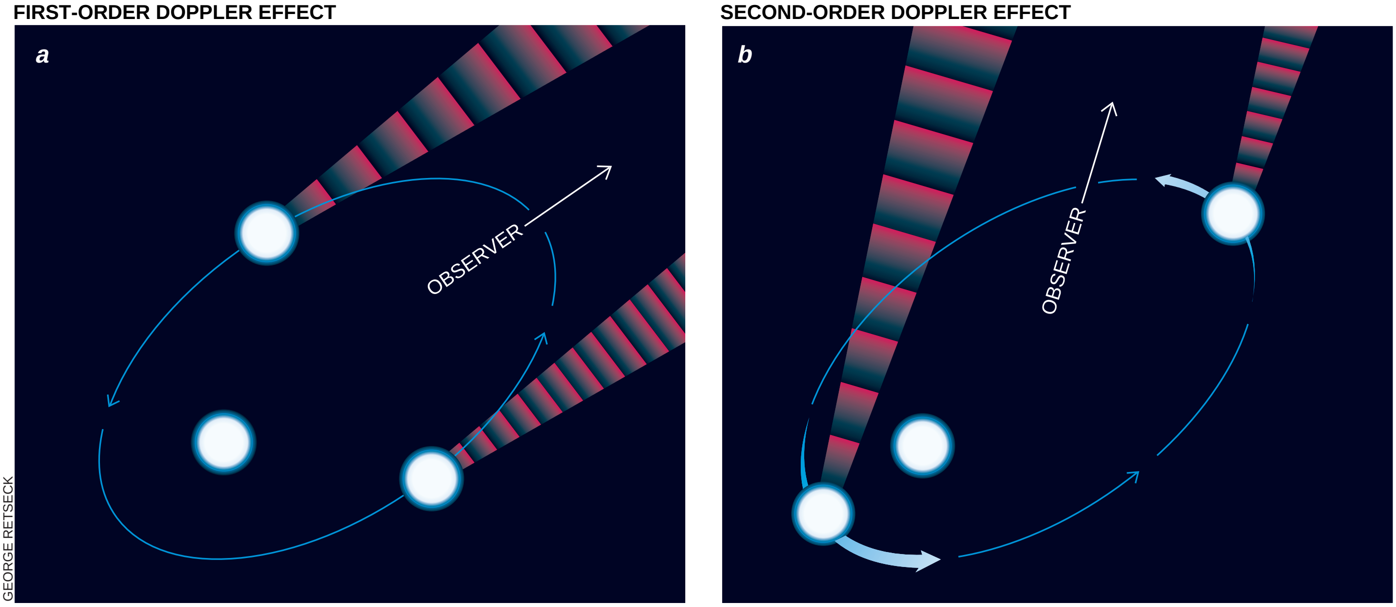
PSR 1913+16 has implications that reach far beyond the revision of theories of binary stellar evolution. Hulse and Taylor immediately realized that their discovery had provided an ideal site for testing Einstein's General Theory of Relativity.

Although this theory is accepted today as the only viable description of gravity, it has had only a few direct tests. Albert Einstein himself computed the precession of Mercury's orbit (the shift of the orbital axes and the point of Mercury's closest approach to the sun) and showed that observations agreed with his theory. Arthur Eddington detected the bending of light rays during a solar eclipse in 1919. In 1960 Robert V. Pound and Glen A. Rebka, Jr., then both at Harvard, first measured the gravitational redshift, the loss of energy by photons as they climb out of a powerful gravitational field. Finally, in 1964, Irwin I. Shapiro, also at Harvard, pointed out that light signals bent by a gravitational field should be delayed in comparison to those that take a straight path. He measured the delay by bouncing radar signals off other planets in the solar system. Although general relativity passed these tests with flying colors, they were all carried out in the (relativistically) weak gravitational field of the solar system. That fact left open the possibility that general relativity might break down in stronger gravitational fields.

Because a pulsar is effectively a clock orbiting in the strong gravitational field of its companion, relativity makes a range of clear predictions about how the ticks of that clock (the pulses) will



MASSIVE BINARY (*a*) evolves through a sequence of violent events. The heavier star in the pair burns its fuel faster and undergoes a supernova explosion; if the two stars stay bound together, the result is a massive x-ray binary (*b*) in which the neutron star remnant of the first star strips gas from its companion and emits x-radiation. Eventually the second star also exhausts its fuel. In roughly one of 100 cases, the resulting explosion leaves a pair of neutron stars orbiting each other (*c*); in the other 99, the two drift apart (*d*). There are enough binary star systems that a typical galaxy contains thousands of neutron star binaries.



GEORGE RETSECK

BINARY PULSAR SIGNALS are affected by relativistic phenomena. (Each illustration above shows one of the effects whose combination produces the timing of the pulses that radio astronomers observe.) The Doppler effect slows the rate at which pulses reach an observer when the pulsar is moving away from the earth in its orbit and increases the rate when the pulsar is moving toward the earth (*a*). The second-order Doppler effect and the gravitational redshift (*b*) impose a sim-

appear from the earth. First, the Doppler effect causes a periodic variation in the pulses' arrival time (the pattern that first alerted Taylor and Hulse).

A "second-order" Doppler effect, resulting from time dilation caused by the pulsar's rapid motion, leads to an additional (but much smaller) variation. This second-order effect can be distinguished because it depends on the square of velocity, which varies as the pulsar moves along its elliptical orbit. The second-order Doppler shift combines with the gravitational redshift, a slowing of the pulsar's clock when it is in the stronger gravitational field closer to its companion.

Like Mercury, PSR 1913+16 precesses in its orbit about its companion. The intense gravitational fields involved, however, mean that the periastron—the nadir of the orbit—rotates by 4.2 degrees a year, compared with Mercury's perihelion shift of a mere 42 arc seconds a century. The measured effects match the predictions of relativistic theory precisely. Remarkably, the precession and other orbital information supplied by the timing of the radio pulses make it possible to calculate the masses of the pulsar and its companion: 1.442 and 1.386 solar masses, respectively, with an uncertainty of 0.003 solar mass. This precision is impressive for a pair of objects 15,000 light-years away.

In 1991 Alexander Wolszczan of the Arecibo observatory found another binary pulsar that is almost a twin to PSR 1913+16. Each neutron star weighs between 1.27 and 1.41 solar masses. The Shapiro time delay, which was only marginally measured in PSR 1913+16, stands out clearly in signals from the pulsar that Wolszczan discovered.

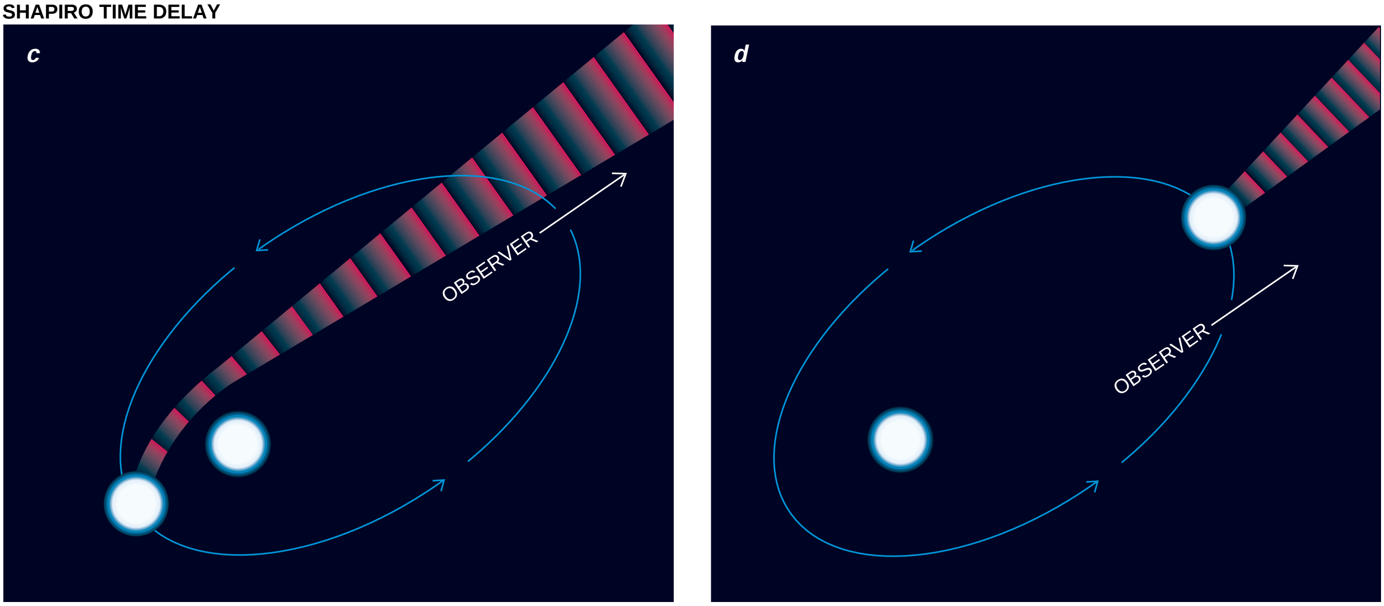
Measurements of PSR 1913+16 have also revealed a relativistic effect never seen before. In 1918, several years after the publication of his General Theory of Relativity, Einstein predicted the existence of gravitational radiation, an analogue to electromagnetic radiation. When electrically charged particles such as electrons and protons accelerate, they emit electromagnetic waves. Analogously, massive particles that move with varying acceleration emit gravitational waves, small ripples in the gravitational field that also propagate at the speed of light.

These ripples exert forces on other masses; if two objects are free to move, the distance between them will vary with the frequency of the wave. The size of the oscillation depends on the separation of the two objects and the strength of the waves. In principle, all objects whose acceleration varies emit gravitational radiation. Most objects are so small and move so slowly, however, that their gravitational radiation is utterly insignificant.

Binary pulsars are one of the few exceptions. The emission of gravitational waves produces a detectable effect on the binary system. In 1941, long before the discovery of the binary pulsar, the Russian physicists Lev D. Landau and Evgenii M. Lifshitz calculated the effect of this emission on the motion of a binary. Energy conservation requires that the energy carried away by the waves come from somewhere, in this case the orbital energy of the two stars. As a result, the distance between them must decrease.

PSR 1913+16 emits gravitational radiation at a rate of eight quadrillion gigawatts, about a fifth as much energy as the total radiation output of the sun. This luminosity is impressive as far as gravitational radiation sources are concerned but still too weak to be detected directly on the earth. Nevertheless, it has a noticeable effect on the pulsar's orbit. The distance between the two neutron stars decreases by a few meters a year, which suffices to produce a detectable variation in the timing of the radio pulses. By carefully monitoring the pulses from PSR 1913+16 over the years, Taylor and his collaborators have shown that the orbital separation decreases in exact agreement with the predictions of the General Theory of Relativity.

The reduction in the distance between the stars can be compared with the other general relativistic effects to arrive at a further confirmation. Just as measurements of the orbital decay produce a mathematical function relating the mass of the pulsar to the mass of its companion, so do the periastron shift



ilar variation because the pulsar's internal clock slows when it moves more rapidly in its orbit closest to its companion (*shown by longer arrow*). Most subtle is the so-called Shapiro time delay, which occurs as the gravitational field of the pulsar's companion bends signals passing near it (*c*). The signals travel farther than they would if they took a straight-line path (*d*) and so arrive later. This effect is undetectable in PSR 1913+16 but shows up clearly in a more recent discovery.

and the second-order Doppler effect. All three functions intersect at precisely the same point.

Undetectable Cataclysms

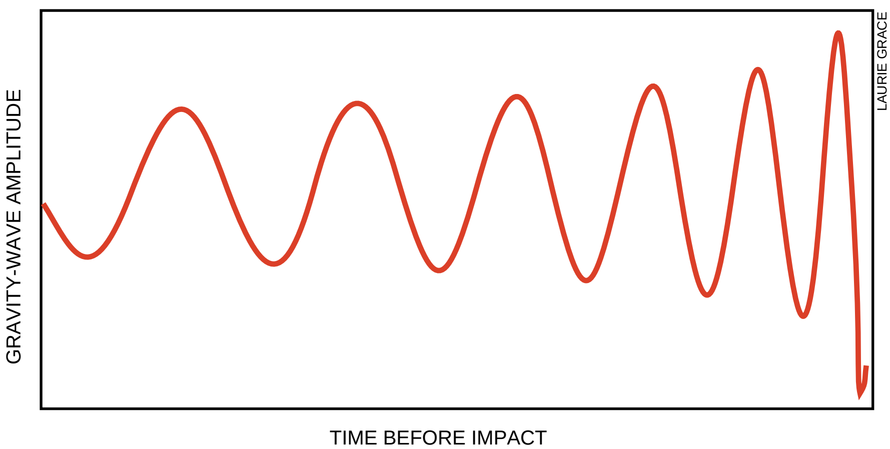
At present, the distance between PSR 1913+16 and its companion is decreasing only slowly. As the distance between the stars shrinks, the gravitational wave emission will increase, and the orbital decay will accelerate. Eventually the neutron stars will fall toward each other at a significant fraction of the speed of light, collide and merge. The 300 million years until PSR 1913+16 coalesces with its companion are long on a human scale but rather short on an astronomical one.

Given the number of neutron star binaries in the galaxy, one pair should merge roughly every 300,000 years, a cosmological blink of the eye. Extrapolating this rate to other galaxies implies that throughout the observable universe about one neutron star merger occurs every 20 minutes—frequently enough that astronomers should consider whether they can detect such collisions.

To figure out whether such occurrences are detectable requires a solid understanding of just what happens when two orbiting neutron stars collide. Shortly after the discovery of the first binary pulsar, Paul Clark and Douglas M. Eardley, then both at Yale University, concluded that the final outcome is a black hole. Current estimates of the maximum mass of a neutron star range between 1.4 and 2.0 solar masses. Rotation increases the maximal mass, but most models suggest that even a rapidly rotating neutron star cannot be significantly larger than 2.4 solar masses. Because the two stars together contain about 2.8 solar masses, collapse to a singularity is almost inevitable.

Melvyn B. Davies of Caltech, Willy Benz of the University of Arizona, Freidrich K. Thielemann of the Harvard-Smithsonian Center for Astrophysics and I have simulated the last moments of a neutron star binary in detail. The two objects are very dense and so behave effectively like point masses until they are quite close to each other. Tidal interaction between the stars becomes important only when they approach to within 30 kilometers, about twice the radius of a neutron star. At that stage, they begin to tear material from each other—about two tenths of a solar mass in total. Once the neutron stars touch, within a tiny fraction of a second they coalesce. The matter torn from the stars before the collision forms a disk around the central core and eventually spirals back into it.

What kinds of signals will this sequence of events generate? Clark and



GRAVITY WAVES from coalescing neutrons stars have a distinctive signature. As the two stars spiral together during the last minutes of their lives, they emit gravitational radiation whose frequency is related to their orbital period. The signal increases in amplitude and "chirps" from a few cycles to several hundred cycles per second.

Eardley realized that the colliding stars will warm up and reach temperatures of several billion kelvins. They figured that most of the thermal energy would be radiated as neutrinos and antineutrinos, much as it is in a supernova. Unfortunately, these weakly interacting, massless particles, which escape from the dense neutron star much more easily than do photons, are almost undetectable. When supernova 1987A exploded, the three detectors on the earth caught a total of 21 neutrinos out of the 5×10^{46} joules of radiation. Although the burst expected from a binary neutron star merger is slightly larger than that of a supernova, the typical event takes place much farther away than the mere 150,000 light-years of SN 1987A. To detect one merger a year would require picking up signals one sixteen-millionth the intensity of the 1987 event. Because current neutrino detectors must monitor interactions within thousands of tons of material, it is difficult to imagine the apparatus that would be required. Furthermore, supernovae are 1,000 times more frequent than are neutron star collisions. Even if we detected a neutrino burst from two neutron stars, it is unlikely we would be able to distinguish it among the far more numerous and far more intense supernova neutrino bursts.

Before it emits its neutrino burst, the neutron star binary sends out a similarly energetic (but not quite so undetectable) train of gravitational waves. During the 15 minutes before coalescence, the two stars cover the last 700 kilometers between them, and their orbital period shrinks from a fifth of a second to a few milliseconds. The resulting signal is just in the optimal range for terrestrial gravity-wave detectors.

An international network of such detectors is now being built in the U.S. and in Italy. The American Caltech-M.I.T. team is building detectors for the Laser Interferometer Gravitational-Wave Observatory (LIGO) near Hanford in Washington State and near Livingston in Louisiana. The French-Italian team is constructing its VIRGO facility near Pisa in Italy. The first detectors should be able to detect neutron star mergers up to 70 million light-years away; current estimates suggest that there is only one event per 100 years up to this distance. Researchers have proposed to improve their instruments dramatically over the subsequent few years; eventually they should be able to detect neutron star mergers as far away as three billion light-years—several hundred a year.

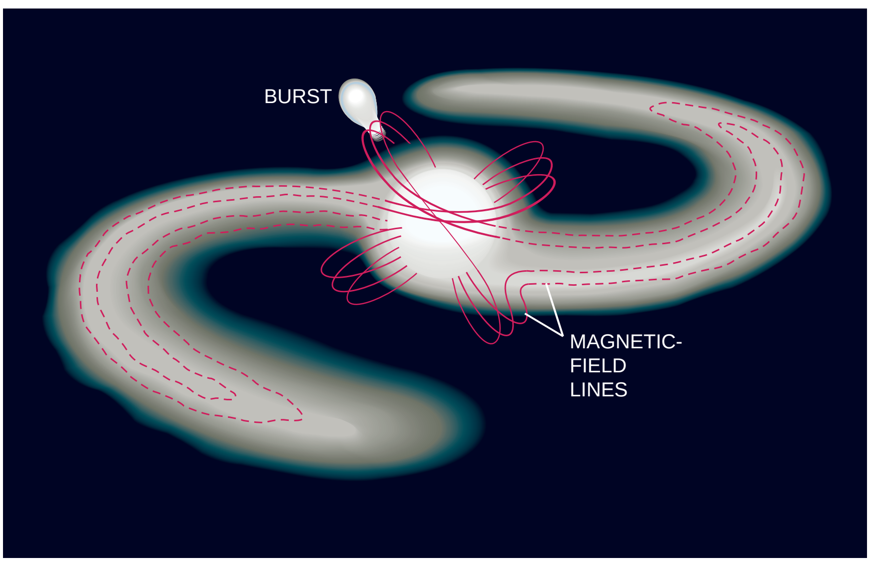
High-Energy Photons

For several years after the discovery of PSR 1913+16, I kept wondering whether there was a way to estimate what fraction of the coalescing stars' binding energy is emitted as electromagnetic radiation. Even if this fraction is tiny, the binding energy is so large that the resulting radiation would still be enormous. Furthermore, photons are much easier to detect than neutrinos or gravitational waves, and so mergers could be detected even from the most distant parts of the universe.

In 1987 J. Jeremy Goodman of Princeton University, Arnon Dar of the Technion and Shmuel Nussinov of Tel Aviv University noticed that about a tenth of a percent of the neutrinos and antineutrinos emitted by a collapsing supernova core collide with one other and annihilate to produce electron-positron pairs and gamma rays. In a supernova the absorption of these gamma rays by the star's envelope plays an important role in the explosion of the outer layers.

In 1989 David Eichler of Ben Gurion University of the Negev, Mario Livio of the Technion, David N. Schramm of the University of Chicago and I speculated that a similar fraction of the neutrinos released in a binary neutron star merger would also produce electron-positron pairs and gamma rays. The colliding neutron stars, however, have no envelope surrounding them, and so the gamma rays escape in a short, intense burst.

Gamma-ray bursts might arise from a more complex mechanism. The disk that forms during the neutron star merger falls back onto the central coalesced object within a few seconds, but during that time it, too, can trigger emissions. In 1992 Bohdan Paczyński of Princeton, Narayan of Harvard and I suggested that the rotation of the disk could intensify the neutron-star magnetic fields entangled in the disk's material, causing giant magnetic flares, a scaled-up version of the flares that rise from the surface of the sun. These short-lived magnetic disturbances could generate gamma-ray bursts in the same way that solar flares produce gamma rays and x-rays. The large variability in the observed bursts implies that both mechanisms may be at work.



GAMMA-RAY BURSTS may also result from a mechanism similar to the one that powers solar flares. The magnetic field from the coalesced neutron stars is amplified as it winds through the disk of material thrown out during the merger. The field accelerates charged particles until they emit gamma rays.

A Puzzle Unscrambled

Had it not been for the Limited Test Ban Treaty of 1963, we would not have known about these bursts until well into the next century. No one would have proposed a satellite to look for them, and had such a proposal been made it would surely have been turned down as too speculative. But the U.S. Department of Defense launched a series of satellites known as *Vela,* which carried omnidirectional x-ray and gamma-ray detectors to verify that no one was testing nuclear warheads in space.

These spacecraft never detected a nuclear explosion, but as soon as the first satellite was launched it began to detect entirely unexpected bursts of high-energy photons in the range of several hundred kiloelectron volts. Bursts lasted between a few dozen milliseconds and about 30 seconds. The lag between the arrival time of the bursts to different satellites indicated that the sources were outside the solar system. Still, the bursts were kept secret for several years, until in 1973 Ray W. Klebesadel, Ian B. Strong and Roy A. Olson of Los Alamos National Laboratory described them in a seminal paper. Theorists proposed more than 100 models in the next 20 years; in the late 1980s a consensus formed that the bursts originated on neutron stars in our own galaxy.

A minority led by Paczyński argued that the bursts originated at cosmological distances. In the spring of 1991 the *Compton Gamma Ray Observatory,* which was more sensitive than any previous gamma-ray satellite, was launched by the National Aeronautics and Space Administration. It revealed two unexpected facts. First, the distribution of burst intensities is not homogeneous in the way that it would be if the bursts were nearby. Second, the bursts came from all across the sky rather than being concentrated in the plane of the Milky Way, as they would be if they originated in the galactic disk. Together these facts demonstrate that the bursts do not originate from the disk of our galaxy. A lively debate still prevails over the possibility that the bursts might originate from the distant parts of the invisible halo of our galaxy, but as the *Compton Observatory* collects more data, this hypothesis seems less and less likely. It seems that the minority was right.

In the fall of 1991, I analyzed the distribution of burst intensities, as did Paczyński and his colleague Shude Mao. We concluded that the most distant bursts seen by the *Compton Observatory* came from several billion light-years away. Signals from such distances are redshifted (their wavelength is increased and energy decreased) by the expansion of the universe. As a result, we predicted that the cosmological redshift should lead to a correlation between the intensity of the bursts, their duration and their spectra. Fainter bursts, which tend to come from farther away, should last longer and contain a lower-energy distribution of gamma rays.

Recently a NASA team headed by Jay P. Norris of the Goddard Space Flight Center has found precisely such a correlation. The number of bursts that the *Compton Observatory* records also tallies quite well with our earlier estimates of the binary neutron star population. Roughly 30,000 mergers should occur every year throughout the observable universe, and the satellite's detectors can scan a sphere containing about 3 percent of that volume. Our rough estimates suggest 900 mergers a year in such a space; the *Compton Observatory* notes 1,000 bursts.

JOHN LONG

LIGO INTERFEROMETERS, one of which is shown here under construction, should one day be able to detect the gravitational radiation of colliding neutron stars from a distance of billions of light-years. If these signals arrive at the same time as gamma-ray bursts, a decades-old mystery may be solved.

Although the details of how colliding neutron stars give rise to gamma rays are still being worked out, the tantalizing agreement between these data from disparate sources implies that astronomers have been detecting neutron star mergers without knowing it for the past 25 years. Researchers have proposed a few other sources that might be capable of emitting the enormous amounts of energy needed for cosmological gamma-ray bursts. The merger model, however, is the only one based on an independently observed phenomenon, the spiraling in of a neutron star binary as a result of the emission of gravity waves.

It is the only model that makes a clear prediction that can be either confirmed or refuted. If, as I expect, LIGO and VIRGO detect the unique gravitational-wave signal of spiraling neutron stars in coincidence with a gamma-ray burst, astrophysicists will have opened a new window on the final stages of stellar evolution, one that no visible-light instruments can hope to match.

About the Author

TSVI PIRAN has studied general relativity and astrophysics for 20 years. He is a professor at the Hebrew University of Jerusalem, where he received his Ph.D. in 1976. He has also worked at the University of Oxford, Harvard University, Kyoto University and Fermilab. Piran, with Steven Weinberg of the University of Texas, established the Jerusalem Winter School for Theoretical Physics.

Further Reading

BLACK HOLES, WHITE DWARFS AND NEUTRON STARS: THE PHYSICS OF COMPACT OBJECTS. Stuart L. Shapiro and Saul A. Teukolsky. Wiley Interscience, 1983.

WAS EINSTEIN RIGHT? PUTTING GENERAL RELATIVITY TO THE TEST. Clifford M. Will. Basic Books, 1988.

LIGO: THE LASER INTERFEROMETER GRAVITATIONAL-WAVE OBSERVATORY. Alex Abramovici et al. in *Science,* Vol. 256, pages 325–333; April 17, 1992.

PROBING THE GAMMA-RAY SKY. Kevin Hurley in *Sky and Telescope,* Vol. 84, No. 6, pages 631–636; December 1992.

X-RAY BINARIES. Edward P. J. van den Heuvel and Jan van Paradijs in *Scientific American,* Vol. 269, No. 5, pages 64–71; November 1993.

Dendrimer Molecules

Chemists can now build fractal supermolecules. This new class of polymers promises to be valuable in biotechnology and environmental protection

by Donald A. Tomalia

In the center of Michigan, along the Chippewa River, some 130 miles southeast of Sleeping Bear Dunes National Lakeshore, the land is not productive enough for traditional agriculture, but it is adequate for growing trees. Thousands of trees of all types, with every branching pattern and shape imaginable, flourish there. Year after year young seedlings with single trunks emerge. Then their trunks elaborate branches, and those branches produce more branches in the same way, giving rise to the lush and varied forest.

As I pondered these trees near my home some 20 years ago with the eyes of a chemist, the systems of branches made me wonder whether one could design large, precisely defined molecules by adding branch after branch onto some original substance. The idea of gaining such control over the formation of a molecule appealed to me immediately on both theoretical and practical grounds, but it was not until the end of the 1970s that I found a way to put the concept into practice. Today my technique and other similar approaches are making it possible to construct treelike molecules that mimic a variety of biological structures, including proteins. There is good reason to believe that these synthetic constructions will prove valuable in medicine, the electronics industry and other fields.

Long ago nature devised exquisite strategies for manipulating the structures of the molecules necessary for initiating and sustaining life. Chemists have tried for years to achieve such mastery over the structures they create. Organic chemists have gained substantial command over the synthesis of small complex molecules. But the goal of constructing large well-defined molecules has been more elusive.

The idea of directing molecular growth to make these extremely large molecules with useful properties derived from experiments done in the 1930s, when Hermann Staudinger, then at the University of Freiburg, managed to link identical subunits, or monomers, into strings of spaghettilike molecules called random-coil polymers. Staudinger's work represented the first successful attempt at assembling large molecules from well-defined smaller components. But the investigator had little control over the lengths of the polymers, which spanned the continuum from microscopic (on the nanometer scale) to macroscopic (on the millimeter or centimeter scale). These polymers turned out to have interesting and valuable features; indeed, many familiar items are made of these types of random-coil polymers, including Styrofoam insulation, polyethylene milk cartons and Plexiglas.

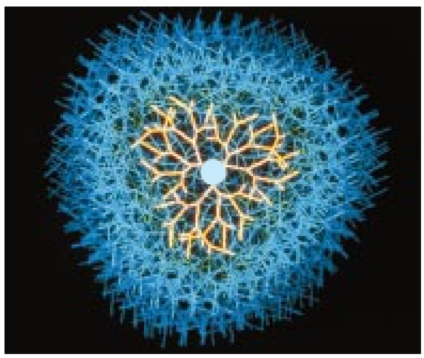
About 10 years later Paul J. Flory, then at Cornell University, and Walter H. Stockmayer, then at the Massachusetts Institute of Technology, took another step toward making large molecules. They developed a second class of polymer that is formed by bridging, or cross-linking, random-coil polymers at various sites on the long chains. The loops and connections give rise to inflexible three-dimensional structures referred to as cross-linked polymers. Their rigidity makes these substances insoluble in most liquids and therefore useful as coating materials on fiberglass hulls of boats or in urethane foams and epoxy systems.

In both random-coil and cross-linked polymers, monomers are joined into long, meandering chains of molecules having varying lengths and sizes, and the precise internal arrangements are impossible for chemists to predetermine. My idea in the mid-1970s was to gain such control. I finally figured out how to achieve that objective after making a surprising discovery in 1979.

DONALD A. TOMALIA is a research professor and director of nanoscopic chemistry and architecture at the Michigan Molecular Institute in Midland. He received his bachelor's degree in chemistry from the University of Michigan, Flint College. In 1968, while working as a synthetic polymer chemist in Dow Chemical Company Research Laboratories in Midland, Tomalia completed his Ph.D. in chemistry at Michigan State University. He is a cofounder, director and the chief scientist of Dendritech, Inc. His major areas of interest are in structure-controlled polymers, molecular recognition and nanoscopic chemistry.

Just Add Methanol and Stir

One spring day colleagues in my laboratory and I were following our standard procedure for making linear random-coil polymers called polyamidoamines. Although typically the synthesis did not require a solvent, on this particular day we added one—methanol (CH_3OH)—to the initial set of ingredients in order to facilitate stirring. We did not expect methanol to alter the substances in this reaction chemically.



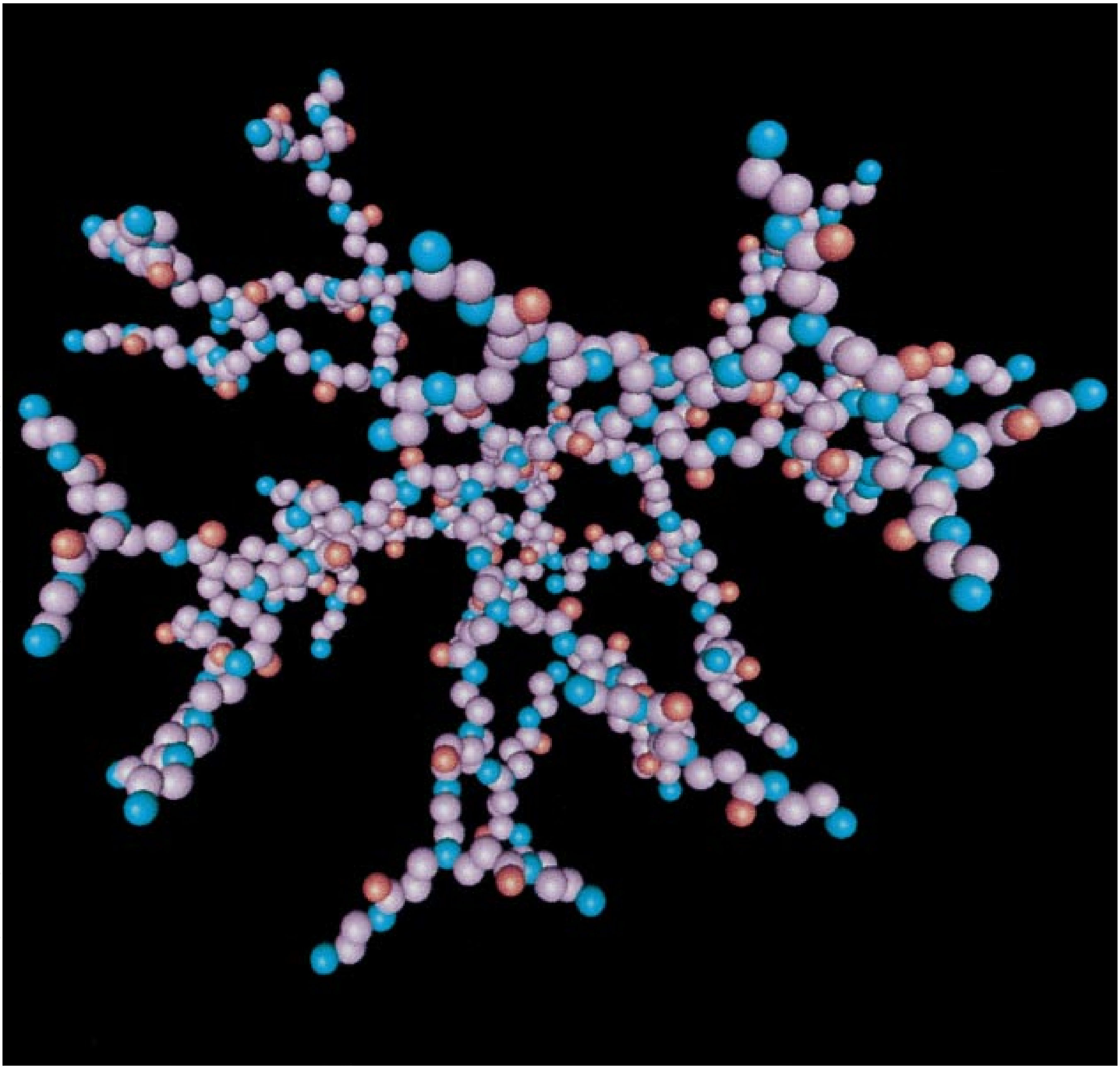
STARBURST DENDRIMERS, shown at the right and in cross section above, have an ammonia molecule at their core. These dendrimers consist of three branched "trees," parts of which are highlighted in the cross section. The trees were systematically built onto the core through an iterative process developed by the author and his colleagues.

The rest of the process went along as usual: we mixed the monomer methyl acrylate ($C_4H_6O_2$) with ethylene diamine ($C_2H_8N_2$).

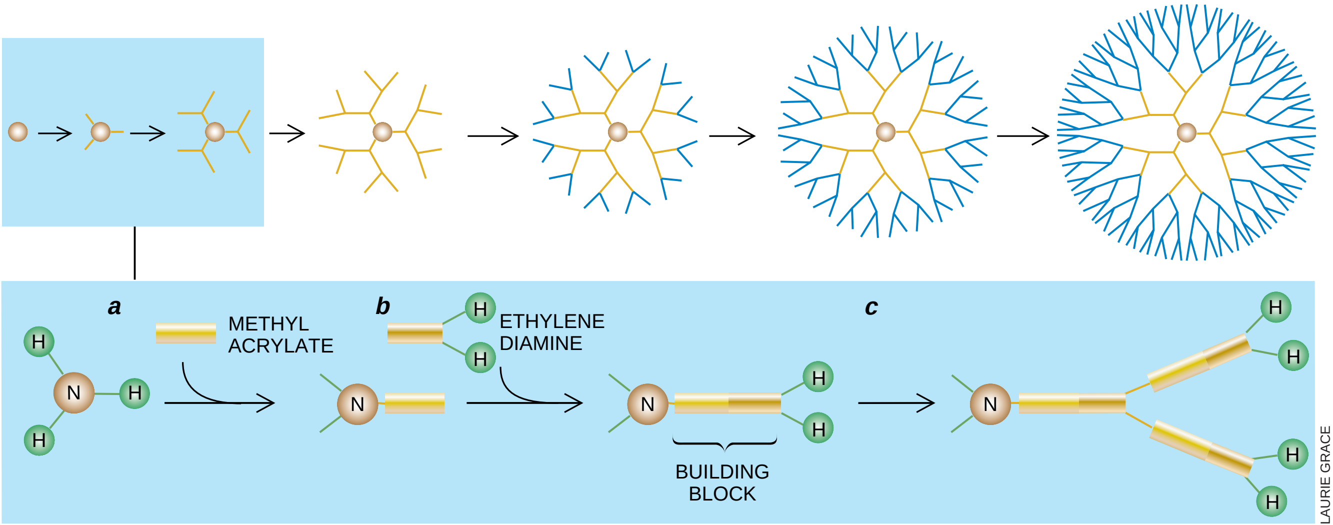
Much to everyone's amazement, we did not get the same random-coil polymer we usually made. Normally, when we mixed these two chemicals, they linked together in a one-to-one ratio, which resulted in a long continuous strand consisting of alternating methyl acrylate and ethylene diamine components. Instead, when we determined the structure of the product from this reaction, we discovered a remarkable arrangement: there were no long strands, only discrete units consisting of two methyl acrylate groups connected to each end of ethylene diamine. In this case the methanol did affect the reaction. It apparently facilitated the removal of hydrogen atoms from the nitrogens in ethylene diamine and enabled methyl acrylate to take their places.

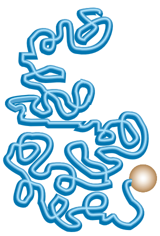
When we repeated the process—now adding ethylene diamine and methyl acrylate to the existing simple structures—we generated an even more intricate compound. This structure contained 12 methyl acrylate and five ethylene diamine monomers. I soon realized we might be able to adapt this technique to assemble the branching molecules I had envisioned back in the woods of Michigan. These molecules would be quite different from the long chains of monomers found in classical random-coil and cross-linked polymers.

Staring at the pattern of those trees, I had specifically imagined a system of molecular growth that resulted in large symmetrical structures built in stages, much as seasonal branching of young trees leads gradually to the development of a broad assembly of branches in older trees. Onto one reactive site of an initial core molecule, a second, linear molecule could be added, producing a stem. Two additional molecules, similar to the linear one, could be attached to the free end of the second molecule,



BRUCE EICHINGER *Biosym Technologies*; PETAR DVORNIC AND SRINIVAS UPPULURI *Michigan Molecular Institute (opposite page)*





GROWTH OF DENDRIMER proceeds by exponential addition to a core molecule (*top*). In this example, the core molecule contains one nitrogen atom and three hydrogen atoms (*detail*). First, the hydrogen atoms are replaced by a molecule called methyl acrylate (*a*), as is shown for just one of the three atoms. A second molecule—ethylene diamine—is added as well (*b*). Together the monomers constitute the basic building blocks of the dendrimer. The unattached end of this unit has two nitrogen-hydrogen bonds, and so the two steps can be repeated to affix two more sections to the growing structure (*c*). Now the unattached end of the structure has four nitrogen-hydrogen bonds, allowing four units to be attached subsequently. Repetition of the process gives rise to the intricate final structure. (Blue coloring highlights the fractallike growth.) The well-defined architecture contrasts with the irregular structure of a random-coil polymer (*left*).

producing a Y-shaped structure. Reiteration of this pattern, adding at least two more of these molecules to the tips of each added Y, would quickly lead to an ordered system of monomers.

This basic approach has resulted in the creation of an entirely new class of polymer architecture, in which concentric tiers of monomers—resembling the layers of an onion—are assembled around a single molecule (the initiator core molecule) at the center. As the layers build outward from this core molecule, the fractal, or dendritic, nature of the growing structure emerges: large regions resemble the smaller Ys formed by triplets of monomers. At the same time, the internal structure takes on a starlike appearance. We thus call the final products "Starburst dendrimers."

Adjustable Properties

We created our first dendrimers by repeatedly performing two simple operations that we still use to this day. When assembling dendrimers, we generally start with an ammonia molecule, which consists of a central nitrogen atom linked to three hydrogen atoms. To this molecule we add enough methanol to facilitate substitution of methyl acrylate for all three hydrogen atoms of the ammonia molecule. Next we add the second monomer, ethylene diamine, which attaches to the free end of each of the three methyl acrylate components. Because ethylene diamine has an amine (NH_2) group at its unbound terminal, all three branches on the central ammonia group now end in a nitrogen atom from which two hydrogen atoms protrude. In other words, the three outermost tips resemble the ammonia molecule at the core, except that six instead of three hydrogen atoms are available to react with additional methyl acrylate monomers.

Just as we are able to replace all three hydrogen atoms in ammonia with three methyl acrylate monomers, we are able to replace the six hydrogens from the three amine groups with six monomer units. To be more precise, in this second iteration of the two-step process, we attach six methyl acrylate molecules to the expanding structure and link ethylene diamine to each of the monomers. These six ethylene diamine–linked monomer groups—forming what we call the second generation of the dendrimer—provide 12 hydrogen atoms to start the third generation. Each successive generation exponentially increases the number of hydrogen atoms available to react.

We soon found that this approach, known as amplification chemistry, can be repeated through as many as nine or 10 generations, until the structure runs out of room to hold additional monomers in perfect branches. The procedure can yield enormous macromolecules, some of which have masses over a million times that of hydrogen (which has an atomic mass of one) and diameters more than 300 times that of hydrogen. And the outermost surface can carry hundreds or even thousands of reactive molecules, known as functional groups. These groups might be derived from the same monomers used to build the dendrimer, or they might be different functional groups, depending on the requirements of the intended application.

Because dendrimers have very regular and predictable patterns of growth, chemists can manipulate the characteristics not only of the interior of the molecule but also of the outer surface. Hence, they can additionally regulate the way the molecule chemically reacts with other molecules. In short, we can specify the size, shape and reactivity of the dendrimers, which allows us to manipulate the properties of the macromolecules we make to an extent that was not possible before.

Indeed, the adjustable physical and chemical properties of dendrimers are their most striking features. Chemists gain this control by carefully selecting the reactants used for making dendrimers. For example, the overall size of a dendrimer is determined by the number of generations included, the length of the monomers used in each generation and the angles between the monomers—features that depend on the chemical makeup of these molecules.

The final structure can also be shaped by the choice of the initiator core molecule. Although the first dendrimers were built from an ammonia core, we have since explored derivatives of ammonia as well as completely different families of compounds, including phosphorus-containing or silicone-containing molecules, benzene (C_6H_6) rings and carbon chains (which have hydrogen and sometimes oxygen atoms attached to them). Each different combination of core molecule and monomers results in a unique dendrimer structure with distinct properties. For instance, if we start with a molecule derived from ammonia but with only one available hydrogen atom, the dendrimer will resemble a mushroom cap. With two active hydrogens, the dendrimer looks more like a kidney-shaped molecule rather than the symmetrical spheroid that is obtained from a pure ammonia core.

As often happens in science, we were not the only researchers to discover a way to build branched molecules. Fritz Vögtle of the University of Bonn and his team were also investigating the possibility of constructing branched molecules around the same time. After we produced our first dendrimers, we learned that Vögtle's group had used a type of amplification chemistry similar to ours to synthesize small branched molecules. The other team's structures consisted mainly of derivatives of ammonia linked by the monomer acrylonitrile. The German researchers called their products "cascade molecules." Other German and Dutch investigators more recently showed that by using a different catalyst in the procedure, the Vögtle approach could be adapted for making large dendrimers as well.

Since 1979 many groups have used methods similar to ours and Vögtle's to synthesize these dendritic supermolecules. In 1985 George R. Newkome of the University of South Florida pioneered an alternative type of amplification chemistry to produce treelike molecules, which he referred to as arborols.

Another interesting approach to dendrimer synthesis was described in 1989 by Jean M. J. Fréchet and Craig J. Hawker of Cornell University and, separately, in 1990 by Timothy M. Miller and Thomas X. Neenan of AT&T Bell Laboratories. Instead of using the so-called divergent method—starting from the inside and building outward—as our laboratory does, these chemists used what is known as the convergent method of synthesis. They constructed individual branches first and then attached the units to a central core molecule.

Between 1980 and 1990 fewer than a dozen papers were published on this subject, but in the past several years there has been an explosion of activity. More than 20 types of dendrimer families with more than 100 different surfaces have now been reported. Apparently, a wide variety of monomers—including metals—may be used to create dendrimers. Moreover, it seems that virtually any functional group that can be found in an organic chemistry textbook can be attached to the surface of dendrimer molecules to carry out selected tasks.

Parallels in Nature

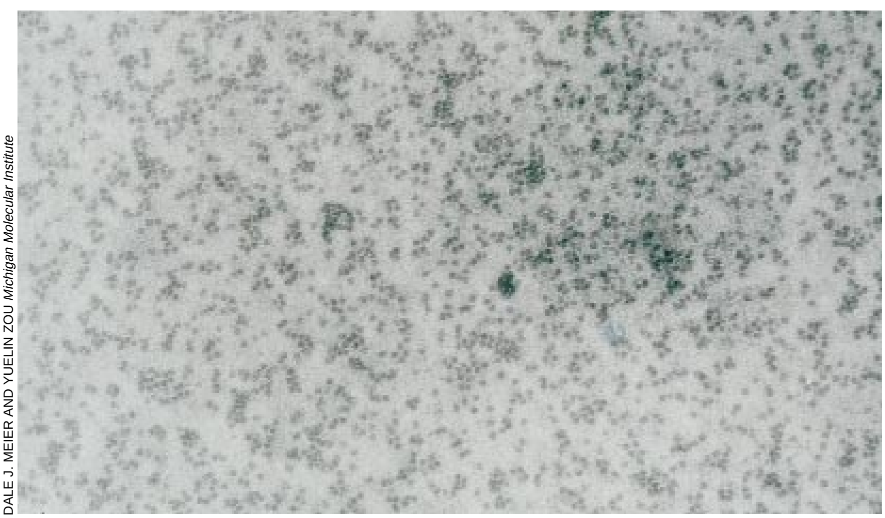
Chemistry is not just about connecting and rearranging atoms in various ways. It is a philosophy, a way of thinking about the dimensional hierarchy of the universe, from the simplest atoms to the most complex molecules and phenomena. The order in atoms and molecules is echoed everywhere in nature, from the branching schemes of trees and coral reefs to the dendritic networks of airways in the lungs and blood vessels in the circulatory system. The significance of these pervasive patterns is not entirely clear, but such connections are fascinating to contemplate.

For example, dendrimers have many intriguing organizational similarities to atoms. As dendrimers or atoms form, curious geometric or arithmetic patterns usually develop. These schemes may involve various repetitive structures or regular sequences of numbers. For example, the number of monomers in a dendrimer with an ammonia core increases in a well-defined sequence: 3, 6, 12, 24, 48 and so on. In this way, concentric dendrimer amplification resembles the accumulation of electrons in successive elements of the periodic table. The number of electrons within each orbital of an atom is also limited. The innermost electron cloud of an atom may not contain more than two electrons; the first generation in an ammonia-based dendrimer may not contain more than its saturation value of three monomers.

Additional parallels can be drawn between the branching network inside dendrimers and certain patterns of biological development. Dendritic growth in these molecules resembles the exponential growth seen during cell mitosis, the process in which one cell becomes two, two cells become four and so on. Furthermore, the original core of the dendrimer determines its final structure, just as the characteristics of the dividing cells derive their unique features from the genetic makeup of the original cell.

The similarities between dendrimers and atoms (the fundamental building blocks of nature) and the parallels between dendritic and biological development may well have practical implications. They suggest that individual or even larger assemblages of the structures might eventually be exploited as synthetic replicas of biological molecules. This notion is reinforced by the fact that dendrimers are about the same size as some of the most important molecules in nature, including enzymes, antibodies, DNA, RNA and viruses.

Scientists have frequently marveled at the way living systems outdo our most advanced techniques for manipulating and combining atoms into large biological molecules. Although dendrimers represent only one of several approaches to building molecules in this



DALE J. MEIER AND YUELIN ZOU *Michigan Molecular Institute*

INDIVIDUAL DENDRIMERS (*black dots*), each about 11 nanometers across, can be seen with an electron microscope. They are similar in size to many large biological molecules. Dendrimers can also be assembled into much bigger clusters.

size range, it seems clear that these structures will function as the basis for what I like to call a new, nanoscopic chemistry set for constructing complex molecules that mimic biological compounds in size, shape and function.

One of the many exciting indications that dendrimers can have practical uses has recently come from my laboratory and that of James R. Baker, Jr., of the University of Michigan Hospital and, separately, from the laboratory of Francis C. Szoka, Jr., of the University of California at San Francisco. We have found evidence that dendrimers might one day be valuable in gene therapy, as vehicles for bringing DNA sequences into cells.

Dendrimers Deliver DNA

The DNA-transporting structures we fashioned resemble clusters of proteins called histones. In the human body, nuclear DNA is found wrapped around such clusters. Our dendrimers are so close in shape and size to a histone cluster that DNA wraps around them just as it does around the natural protein complex. The DNA we studied contained a genetic sequence that codes for the protein luciferase, the substance that gives fireflies their luminescence. This gene is rather easy to track: when it is successfully transferred to a new cell and remains functional, the cell begins to glow.

In petri dish experiments, we combined histonelike dendrimers and the luciferase gene with close to 30 different types of cells from various species, including humans. In nearly all cases, the dendrimers transported genetic material into the cell and gave rise to the luciferase protein. (We do not yet completely understand how the dendrimer-DNA unit makes its way into cells.)

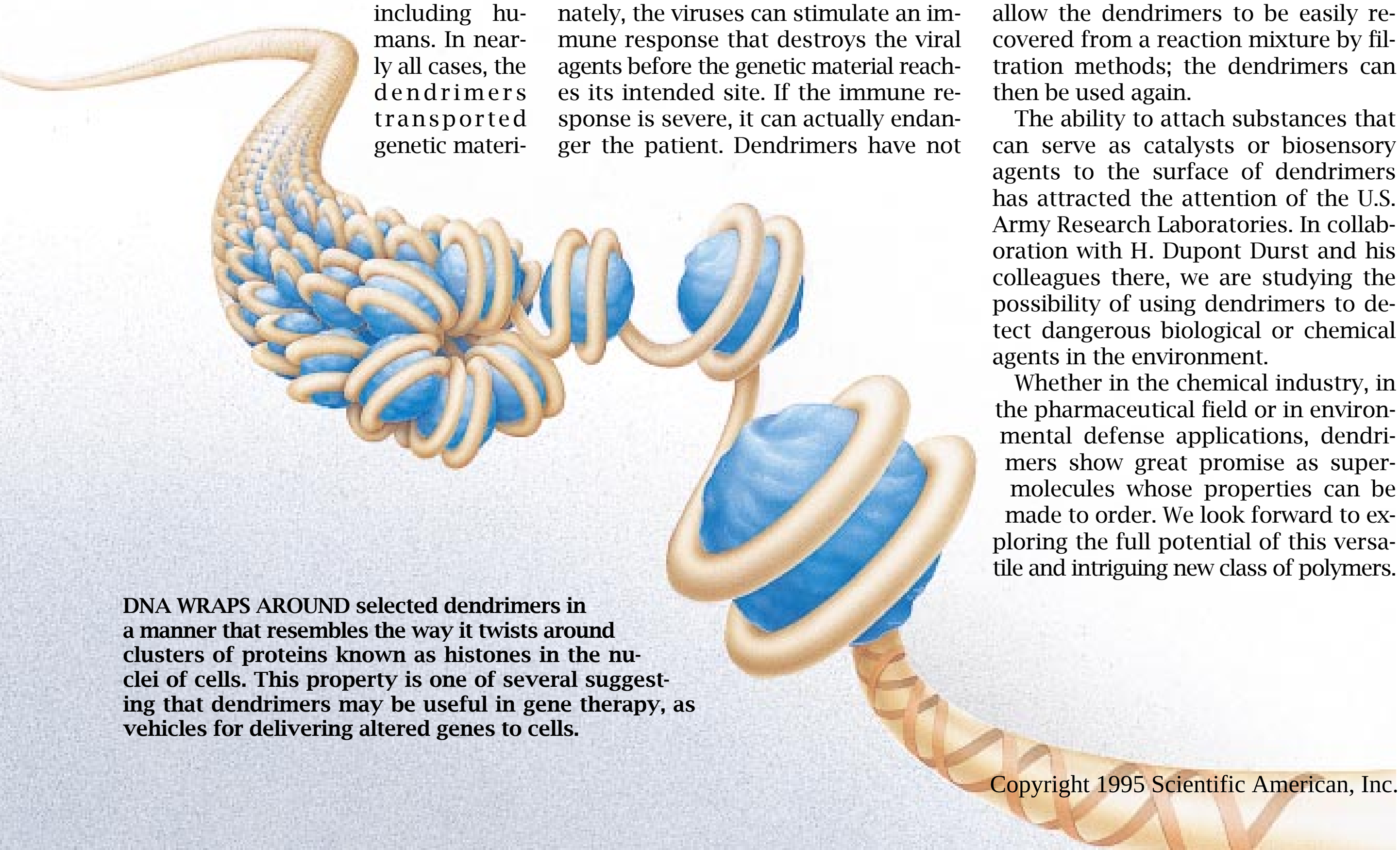
We have other reasons for suspecting that dendrimers might be of service in gene therapy. Notably, the structures can be designed to home in on specific target cells. For example, attaching certain substances, particularly sugar and protein groups, to dendrimer surfaces causes these polymers to adhere more favorably with some cell membranes than with others. By carefully selecting the components we add, we can direct a dendrimer-DNA combination to specific types of cells.

Furthermore, dendrimers may have an advantage over the current method of delivering genetic material to cells. Today scientists often use modified viruses to bring genes to cells. Unfortunately, the viruses can stimulate an immune response that destroys the viral agents before the genetic material reaches its intended site. If the immune response is severe, it can actually endanger the patient. Dendrimers have not triggered such an extreme response in any of our laboratory studies. Many research teams are also developing techniques for transporting other kinds of molecules, such as drugs, to selected targets in the body.

Last December, Gerard van Koten and his group at Utrecht University in the Netherlands identified another application for dendrimers that could have great significance for the chemical industry. In many manufacturing processes, chemical plants must use catalysts to enhance the efficiency of certain reactions. The outside of a dendrimer can be covered with many catalytic sites, so that one dendrimer can induce a large number of catalytic reactions. These dendrimers typically dissolve in the reaction mixture readily, which further facilitates catalysis. Additionally, the large size of the structures should allow the dendrimers to be easily recovered from a reaction mixture by filtration methods; the dendrimers can then be used again.

The ability to attach substances that can serve as catalysts or biosensory agents to the surface of dendrimers has attracted the attention of the U.S. Army Research Laboratories. In collaboration with H. Dupont Durst and his colleagues there, we are studying the possibility of using dendrimers to detect dangerous biological or chemical agents in the environment.

Whether in the chemical industry, in the pharmaceutical field or in environmental defense applications, dendrimers show great promise as supermolecules whose properties can be made to order. We look forward to exploring the full potential of this versatile and intriguing new class of polymers.



DNA WRAPS AROUND selected dendrimers in a manner that resembles the way it twists around clusters of proteins known as histones in the nuclei of cells. This property is one of several suggesting that dendrimers may be useful in gene therapy, as vehicles for delivering altered genes to cells.

TOMO NARASHIMA

FURTHER READING

STARBURST DENDRIMERS: MOLECULAR-LEVEL CONTROL OF SIZE, SHAPE, SURFACE CHEMISTRY, TOPOLOGY, AND FLEXIBILITY FROM ATOMS TO MACROSCOPIC MATTER. Donald A. Tomalia, Adel M. Naylor and William A. Goddard III in *Angewandte Chemie, International Edition in English,* Vol. 29, No. 2, pages 138–175; February 1990.

RISING CHEMICAL "STARS" COULD PLAY MANY ROLES. Joe Alper in *Science,* Vol. 251, pages 1562–1564; March 29, 1991.

DENDRIMERS, ARBOROLS, AND CASCADE MOLECULES: BREAKTHROUGH INTO GENERATIONS OF NEW MATERIALS. Hans-Bernhard Mekelburger, Wilfried Jaworek and Fritz Vögtle in *Angewandte Chemie, International Edition in English,* Vol. 31, No. 12, pages 1571–1576; December 1992.

FUNCTIONAL POLYMERS AND DENDRIMERS: REACTIVITY, MOLECULAR ARCHITECTURE, AND INTERFACIAL ENERGY. Jean M. J. Fréchet in *Science,* Vol. 263, pages 1710–1715; March 25, 1994.

STARBURST/CASCADE DENDRIMERS: FUNDAMENTAL BUILDING BLOCKS FOR A NEW NANOSCOPIC CHEMISTRY SET. D. A. Tomalia in *Advanced Materials,* Vol. 6, Nos. 7–8, pages 529–539; July–August 1994.

The Ocean's Salt Fingers

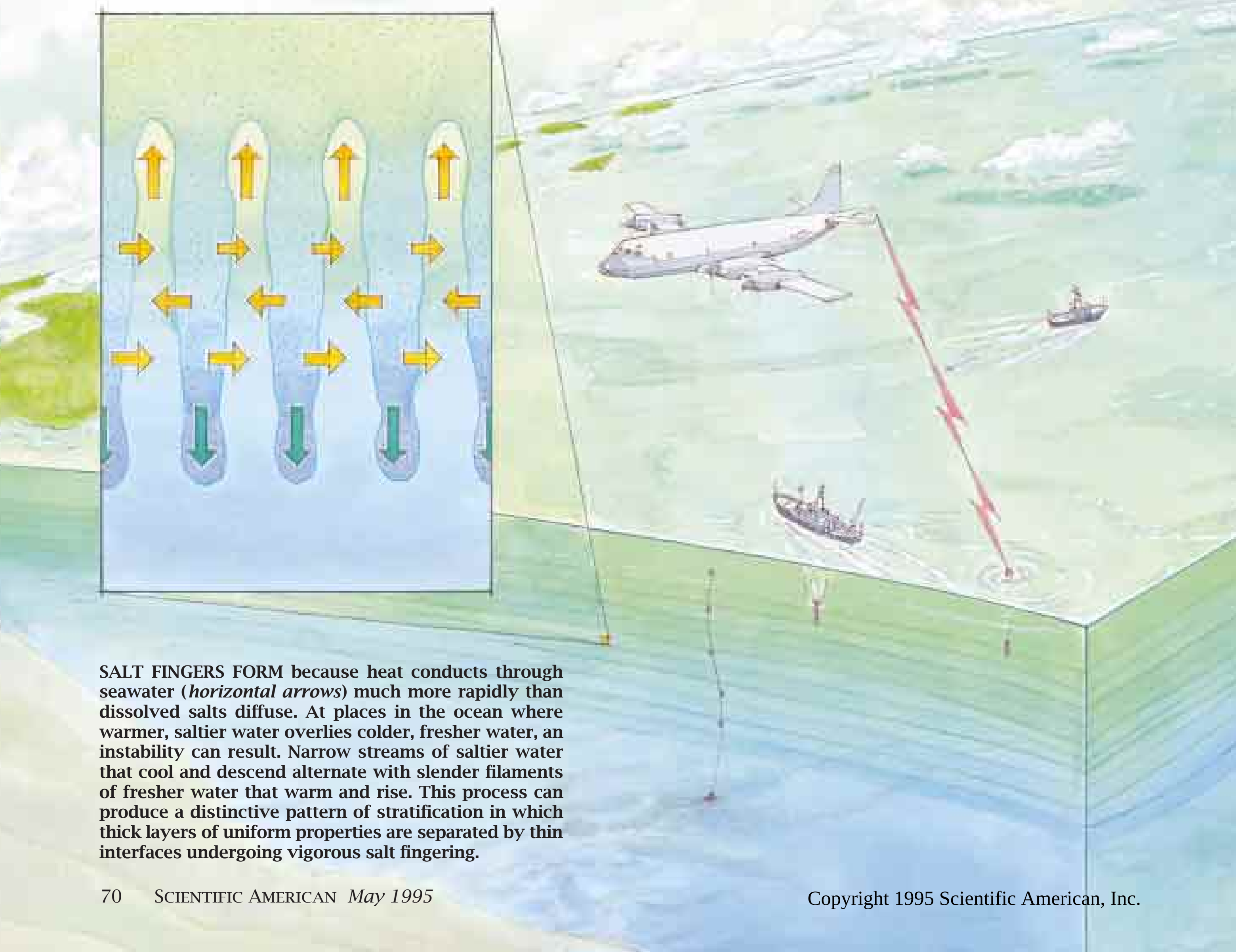
A small-scale oddity in the way seawater mixes can have large-scale consequences for the structure of the ocean

by Raymond W. Schmitt, Jr.

In the mid-1950s Henry M. Stommel and Arnold B. Arons, researchers at the Woods Hole Oceanographic Institution, stood staring at a chalkboard, groping for a way to measure the pressure at the bottom of the sea. In desperation they were considering constructing a three-mile-long tube from the surface to the bottom and drawing the ocean's deepest water up into the pipe. Because water is less salty near the seafloor than at the surface, the column of water within the tube would, after coming to thermal equilibrium, be lighter and so stand higher than the ocean surface around it.

Stommel and Arons imagined that they could then easily measure changes in pressure at the base of this ungainly instrument by watching the water level in the tube go up and down. As they gazed at their sketch, focused entirely on the problem of pressure measurement, Arons suddenly got a strange idea. He added a spigot to the top of the pipe in the diagram and said, "Hank, if we open the faucet, it will run forever."

Stommel dubbed this astonishing mental construction "the perpetual salt fountain." He and his colleagues immediately convinced themselves of the veracity of their idea by setting up a model salt fountain in the laboratory. Sub-



SALT FINGERS FORM because heat conducts through seawater (*horizontal arrows*) much more rapidly than dissolved salts diffuse. At places in the ocean where warmer, saltier water overlies colder, fresher water, an instability can result. Narrow streams of saltier water that cool and descend alternate with slender filaments of fresher water that warm and rise. This process can produce a distinctive pattern of stratification in which thick layers of uniform properties are separated by thin interfaces undergoing vigorous salt fingering.

sequently, they attempted to fabricate an actual salt fountain at sea, although their modest efforts toward a full-scale demonstration of the concept gave ambiguous results.

Stommel and Arons's salt fountain might have remained a mere oceanographic curiosity had not their theoretically minded colleague Melvin E. Stern realized a few years later that the ocean was fully capable of producing salt fountains on its own. He pointed out that seawater conducts heat about 100 times faster than it allows the diffusion of dissolved salts. As a result, a "parcel" of water in the ocean can reach thermal equilibrium with its surroundings far sooner than it can achieve chemical equilibrium by sharing dissolved salts. Thus, adjacent water parcels can differ significantly in salinity even in the absence of a physical barrier (like the wall of a pipe), and those differences can drive motion in the fluid. Because this phenomenon occurs only at scales of a few centimeters (the range of effective heat conduction), the small, elongate streams that form by this process have come to be known to oceanographers as salt fingers.

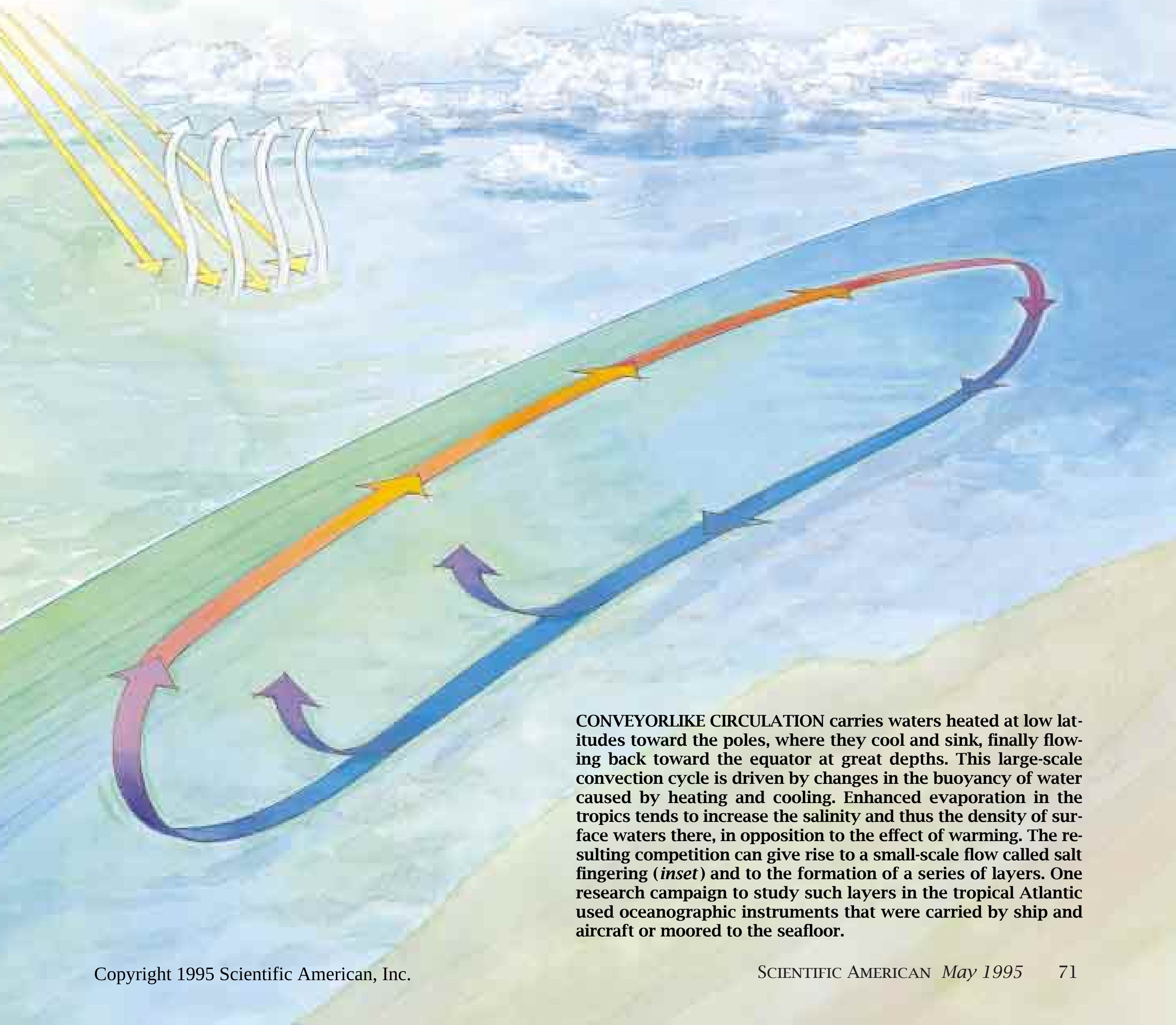
My interest in studying salt fingers was sparked by fascination, but it has been maintained over many years by a growing appreciation that these miniature oddities can exert considerable control on the large-scale structure of the upper ocean. Oceanographers now realize that a fundamental understanding of salt fingers and the extent of their global influence is needed to model the ocean's temperature and salinity accurately. Such advances should help determine many critical but as yet poorly known quantities, such as the rate at which the mixing in the ocean can redistribute carbon dioxide, pollutants and heat.

Nature's Salt Fountains

The sun warms the earth more intensely near the equator than at the poles, but the ocean works vigorously to correct the uneven distribution of heat. Warm waters flow poleward from the tropics along the surface of the ocean while frigid waters at high latitudes sink and flow back toward the equator at great depths.

This immense convective system arises because cold water is denser than warm water, but in the ocean a complicating factor—the presence of dissolved salt—can produce similarly important density changes. The salinity of the sea varies from place to place because of

BARRY ROSS



CONVEYORLIKE CIRCULATION carries waters heated at low latitudes toward the poles, where they cool and sink, finally flowing back toward the equator at great depths. This large-scale convection cycle is driven by changes in the buoyancy of water caused by heating and cooling. Enhanced evaporation in the tropics tends to increase the salinity and thus the density of surface waters there, in opposition to the effect of warming. The resulting competition can give rise to a small-scale flow called salt fingering (*inset*) and to the formation of a series of layers. One research campaign to study such layers in the tropical Atlantic used oceanographic instruments that were carried by ship and aircraft or moored to the seafloor.

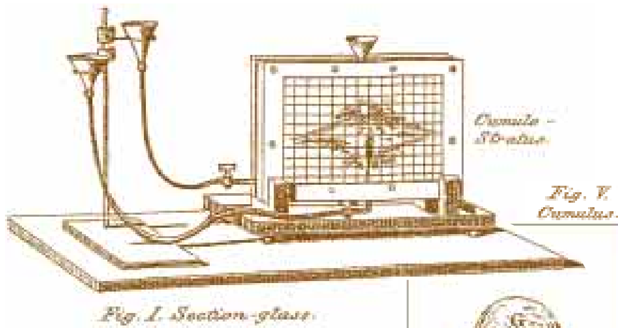
Missed Opportunity for Discovery during the 19th Century

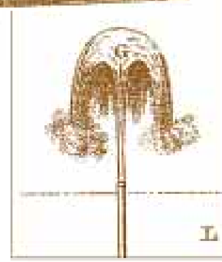
I have lately become fascinated by two researchers who came close to understanding the physics of salt fingers well before their discovery by Melvin E. Stern of Woods Hole in the late 1950s. First to approach the problem was English economist W. Stanley Jevons. During his youth, Jevons spent five years in Australia, where he pursued a number of scientific topics, including meteorology.

PRINCETON UNIVERSITY PRESS

W. STANLEY JEVONS in Australia at the age of 22.

Curiosity about the formation of clouds led Jevons to carry out an experiment in which he introduced a layer of warm sugar water above cold freshwater. In one of his first publications, an 1857 report in *Philosophical Magazine,* he stated that he had observed an "infiltration of minute, thread-like streams." Jevons had created sugar fingers, and he came close to realizing how they formed. He wrote, "The parts of these strata, however, which are immediately in contact, soon communicate their heat and tend to assume a mean temperature; and it is evident that whenever this is the case, the portions of liquid containing sugar must always be slightly denser than those that are pure, and must consequently sink below and displace the latter." This statement, with its implicit understanding that heat diffuses faster than dissolved substances, is a reasonable description of salt fingers. He then went off course, however, by oversimplifying the physics involved and assuming





SYDNEY MAGAZINE OF SCIENCE AND ART

SECTION GLASS APPARATUS that was constructed by Jevons to simulate the formation of clouds in the laboratory.

two closely related phenomena. Evaporation removes freshwater from the surface of the ocean and leaves a higher concentration of salts, which increases the density of the remaining water. Conversely, precipitation dilutes the seawater and decreases its density.

Typically the changes in local seawater density arising from evaporation and precipitation tend to oppose those from the global pattern of heating or cooling. That is, greater evaporation at low latitudes increases near-surface salinity and density, whereas enhanced precipitation at high latitudes decreases both these properties.

This common pattern leads to something of a competition between the effects of temperature and salinity in the upper ocean. At most locations, the ocean is gravitationally stable, with lighter, warm water overlying heavier, cool water. But the ocean's salt gradient (considered in isolation from its temperature gradient) is decidedly top-heavy. The precarious salinity distribution survives in most places in the ocean by virtue of the stabilizing influence of temperature.

Nevertheless, according to Stern's theory, at those places where the differences in salinity between shallow and deep layers become large enough, fountainlike flow between adjacent levels should begin spontaneously. Calculations and laboratory experiments show that these flows, when fully developed, form a regular pattern of long, thin vertical filaments that appear as small square cells in horizontal cross section.

With theory making this clear prediction, seagoing oceanographers began searching in the mid-1970s for direct evidence of these salt fingers. One site studied for this purpose was in the eastern North Atlantic Ocean, where the salty outflow of the Mediterranean Sea can be detected to depths of more than one kilometer. (These waters are highly saline because the Mediterranean experiences a large excess of evaporation over precipitation.) As this water flows westward from the Strait of Gibraltar, it sinks and spreads into the Atlantic.

On an expedition in the 1970s to detect salt fingers, Bruce A. Magnell, then a graduate student at the Massachusetts Institute of Technology, towed a special instrument through the Mediterranean outflow in the Atlantic. The device contained unique, rapid-response sensors that could track small-scale temperature and salinity changes across salt fingers. From the same ship, Albert J. Williams of Woods Hole deployed a free-falling optical apparatus that could image the variation in refractive index caused by fingers. Both methods produced evidence that the expected centimeter-scale fluctuations occurred at the transition between Mediterranean and Atlantic waters. Salt fountains, albeit tiny ones, did in fact exist in nature.

The success of the initial fieldwork provided impetus for many further studies of salt fingers. For example, Thomas R. Osborn of Johns Hopkins University recently observed from a submarine that an asymmetrical form of salt fingering could take place: he found narrow, downward-falling plumes of warm, salty water that were surrounded by broad, upward flow. I have been able to show that such fingers are al-

RAYMOND W. SCHMITT, JR., after earning a B.S. in physics from Carnegie Mellon University, left his landlocked home state of Pennsylvania and went on to the University of Rhode Island. There he completed a Ph.D. in physical oceanography on the shores of Narragansett Bay. In 1978 he began a postdoctoral fellowship at the Woods Hole Oceanographic Institution overlooking Vineyard Sound in Massachusetts. He has remained there since and is currently a senior scientist. Schmitt's research interests include small-scale mixing in the ocean from both salt fingering and mechanical turbulence. He has published numerous articles on salt-finger mixing and the history of its scientific investigation. He also studies the role of the ocean in the global water cycle.

that all convection should occur as "minute streamlets."

A few years later Jevons returned to England and went on to become quite distinguished in the fields of economics and logic. His career was cut short by his untimely death by drowning at age 47.

Another opportunity to uncover the physical basis of salt fingers came around the time of Jevons's death, in 1882. In the following year the renowned English physicist Lord Rayleigh (John William Strutt) published a theoretical analysis of Jevons's experiments in a paper entitled "Investigation of the Character of the Equilibrium of an Incompressible Heavy Fluid of Variable Density." In it he provided a mathematical treatment of motion in stratified fluids. Rayleigh did not, however, consider the role of diffusion and so missed the opportunity to understand how salt fingers operate. He did formulate the conditions for the simpler Rayleigh-Taylor instability, now known to be important in plasma dynamics and supernova explosions.


UNIVERSITY OF WISCONSIN PRESS

JOHN W. STRUTT at age 28, three years before succeeding his father as the third Baron Rayleigh.

Rayleigh's published acknowledgment of Jevons's earlier work is limited to a brief footnote in his 1883 paper, in which he also states that he had arrived at his theory in 1880. Intrigued about the timing and motivation of Rayleigh's work, I recently examined his original notebooks, which are archived at Hanscom Air Force Base in Massachusetts. A laboratory notebook in the hand of Rayleigh's assistant, his sister-in-law Eleanor Sidgwick, reveals that they, too, performed a sugar-finger experiment, at the Cavendish Laboratory in Cambridge, England, in April 1880. She wrote: "We repeated several times the experiment of W. S. Jevons [see Phil. Mag. for July 1857] on the formation of cirrous clouds.... The effects obtained resembled those described by him.... In all cases moreover the extremities of the filaments were expanded in a mushroom like form." Strangely, no mention of these experiments is made in Rayleigh's 1883 paper.

Rayleigh, who later received a Nobel Prize for the discovery of argon, was an astute theoretician and talented experimentalist. Yet despite having read Jevons's hint and duplicated his experiments, Rayleigh failed to recognize the role of heat conduction in the formation of salt fingers. Perhaps he could have reached this understanding through discussions with Jevons (both were Fellows of the Royal Society) had Jevons not died prematurely. We can only speculate that he delayed publication in order to have such discussions but never got the chance. Discovery of the physics of salt fingers had to wait for almost a century. *—R.W.S.*

lowed by theory; however, no one has yet identified the specific mechanism that causes the asymmetrical geometry to form.

Undersea Staircases

My own experience with salt fingers has come from laboratory and theoretical work as well as a number of oceanographic expeditions. The most ambitious of these voyages took place thousands of miles away from Gibraltar, on the other side of the Atlantic near the Caribbean Sea. The vertical salt gradient in that region is especially favorable to salt fingers, and like the Mediterranean outflow in the eastern Atlantic, the ocean there displays a very curious structure in the way in which temperature and salinity change with depth. Ordinarily, the upper ocean displays a continuous variation in temperature, linking the warm, sun-drenched surface layer to the cold, dark water of the abyss. Oceanographers call this gradual temperature transition the main thermocline.

Yet for at least as long as oceanographers have been able to make detailed measurements in this part of the North Atlantic (the past 25 years), the thermocline has been far from smooth. Instead temperature declines with increasing depths through a series of distinct steps, each between five and 40 meters thick. Oceanographers have come to refer to such structures as thermohaline staircases, a name that reflects the steplike changes in both temperature and the concentration of halide salts. Like vast geologic strata, individual layers of uniform temperature and salinity can be traced for hundreds of kilometers. It is difficult to imagine how such regular features can survive in a continually churning ocean. Should not the ever present internal waves and eddies (the ocean's turbulent "weather") quickly destroy the staircase pattern? What strange process can maintain such a subsea ziggurat?

The answer to this oceanographic puzzle lies in salt fingers. The motion of water in the fingers transports salt downward. Heat, too, travels downward with the salty filaments, even as it conducts sideways into the surrounding water. The motion thereby reduces the normal vertical gradient in the temperature and salinity. But remarkably enough, the overall density of seawater is affected oppositely: the ocean develops *greater* density differences when it forms salt fingers. This is because fingers transport more salt than heat; this action further reduces the density of the already lighter water on top and increases the density of the heavier water below. The result appears counterintuitive: ordinary mixing caused by turbulence would act to decrease the density gradient, not to increase it.

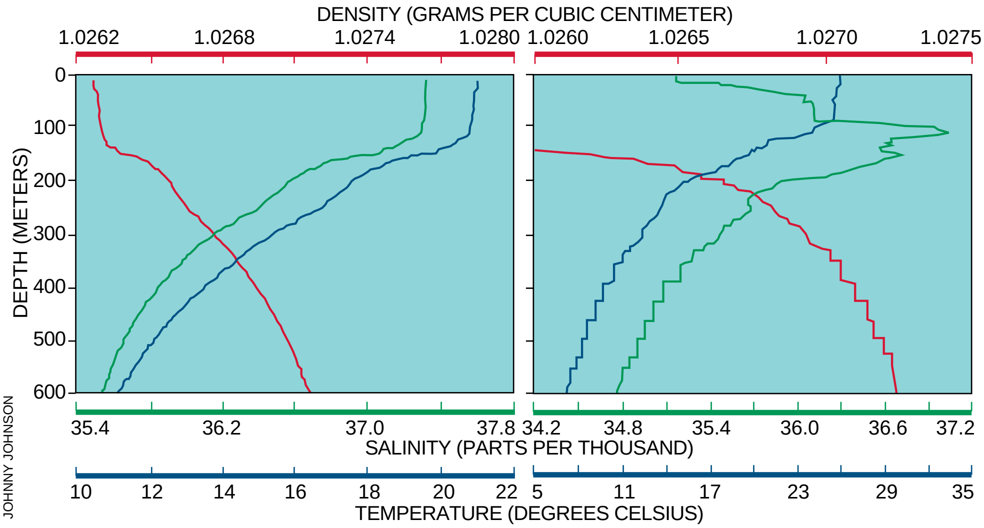
The peculiar density flux of salt fingering has a surprising consequence that neatly explains the staircase structure observed in the tropical Atlantic and other areas. Whereas ordinary mixing through random turbulence would be expected to smooth out any initial irregularities in a density profile, mixing by salt fingers serves to enhance them. In zones where the density gradient already changes sharply, the shift can become even more extreme. These places experience the highest density flux, which becomes large enough to reduce the density variation in adjacent regions. As a result, the upper ocean organizes itself into layers of fairly uniform temperature and salinity, bounded by "sheets," or interfaces, of high gradients. The layers contain large-scale convective flows; the interfaces are laced with salt fingers.

Thermohaline staircases were first observed in the laboratory at Woods Hole by Stern and J. Stewart Turner of the University of Cambridge in the late 1960s. At about the same time, other researchers began finding staircases in the ocean with electronic instruments that could register temperature and salinity continuously as they were lowered into the sea. Only recently have oceanographers developed a detailed picture of the internal workings of such remarkable structures.

Operation "C-SALT"

In 1985 my colleagues and I had the opportunity to study a large area where staircases form in the vicinity of the Caribbean. Our program was funded by the National Science Foundation and the Office of Naval Research, and good navy style demanded that we develop a suitable acronym. We adopted C-SALT, which stands for "Caribbean Sheets and Layers Transects."

Prior to our fieldwork, Janice D. Boyd of the Naval Research Laboratory determined from existing collections of oceanographic data that staircaselike temperature and salinity structures were commonly observed over a large area just east of Barbados. To probe this region, we assembled an observational armada that included aircraft and ships as well as stationary instruments moored to the sea bottom. The first rather pleasant surprise was that the system we sought covered a vast expanse. An area of over one million square kilometers—approximately equal to the extent of California and Texas combined—showed staircase-shaped profiles in our measurements. Typically 10 distinct layers could be seen. That such relatively thin features (each about 30 meters thick) could stretch for hundreds of kilometers over a large area of the ocean was entirely unexpected.



OCEANOGRAPHIC PROFILES typically reveal a gradual change between the warm, saltier surface layer and cold, fresher waters below (*left*). In the transition zone salt fingering can create steplike changes in temperature and salinity that appear remarkably regular (*right*).

Another discovery about the staircase system was the manner in which temperature and salinity varied within individual layers. We found that these undersea strata became cooler and fresher to the south, warmer and saltier to the north. This pattern matches what can be observed in the laboratory when such layers are allowed to decay over time. From our measurements we concluded that the lateral changes in physical properties represent a balance between horizontal flow within the layers and differences in the vertical fluxes across the top and bottom interfaces.

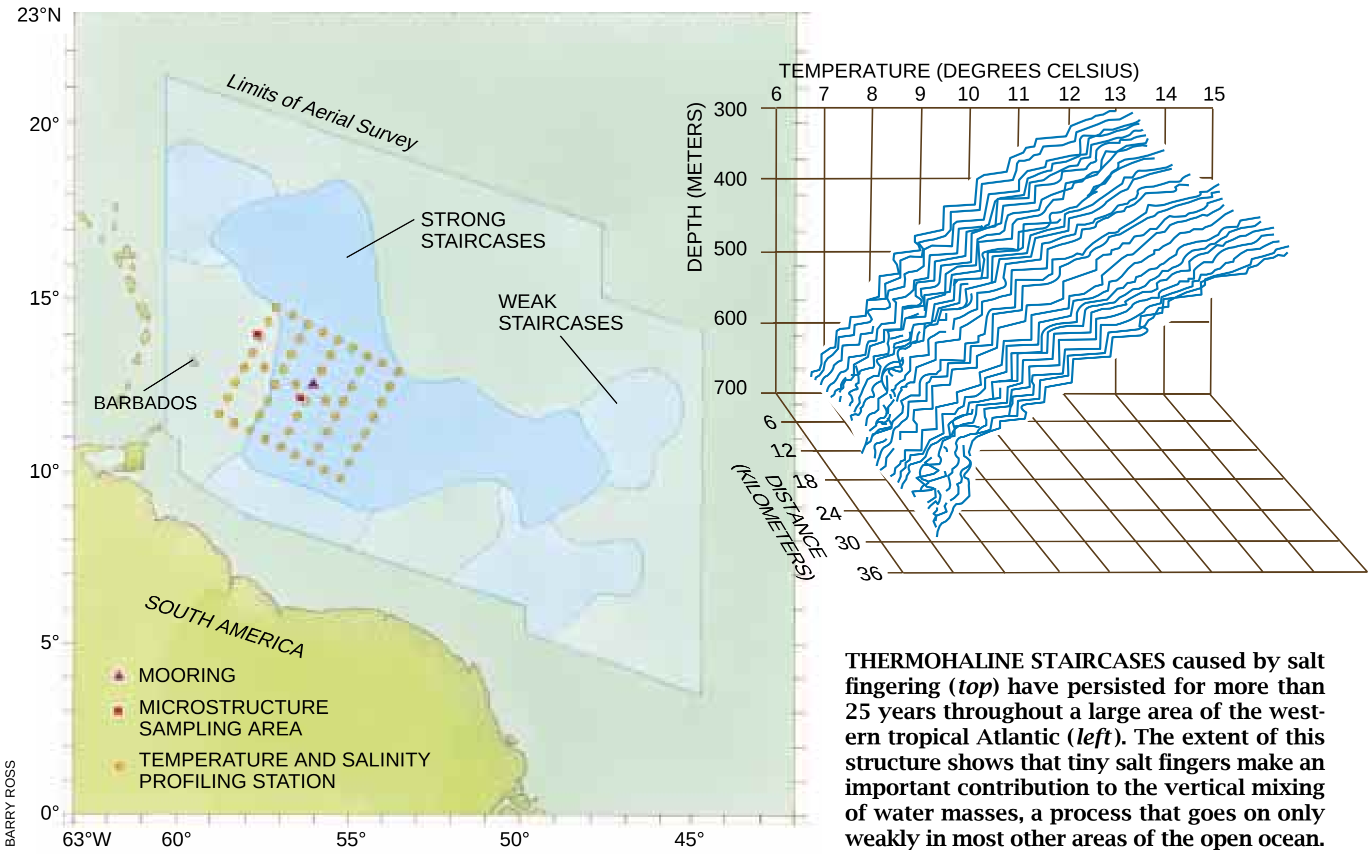
Our documentation of the remarkably consistent way these properties varied with location provided an important clue that salt fingers were in fact maintaining the staircases. That is, we observed that within individual layers the changes in salinity were larger than those in temperature. No known mixing process apart from salt fingers could have caused such an imbalance. In fact, the measured ratio between salinity and temperature variations was very close to that expected from laboratory experiments and from theoretical predictions. Our study of staircase structures within the thermocline had confirmed that salt fingers were indeed the dominant form of mixing in the region, and we were quite excited by the discovery that the details of small-scale mixing had an observable effect on the large-scale temperature and salinity structure of the ocean.

PURPLE FINGERS develop when warm, saline water containing potassium permanganate (which acts as dissolved salt and colorful tracer) floats on cold freshwater.

Fingers at the Interfaces

During the final stages of C-SALT, we set out to measure mixing rates using two different ships. On one vessel, Michael C. Gregg and Thomas B. Sanford (both at the University of Washington), along with Williams and me, deployed instruments that could record temperature and salinity profiles as they fell freely through the water. Rolf G. Lueck of the Chesapeake Bay Institute towed a streamlined device that measured turbulence. From the other ship George O. Marmorino and his colleagues from the Naval Research Laboratory deployed a string of sensors capable of mapping the temperature structure of the staircase and also detecting the small-scale effects of salt fingers.

With each instrument we found a similar pattern of relatively regular temperature variations in the interfaces between layers that was quite distinct from the spiky signature of turbulence. The scale of the undulations that we observed corresponded to a three-centimeter cell size, in excellent agreement with salt-finger theory. The data obtained from the navy's towed tempera-



THERMOHALINE STAIRCASES caused by salt fingering (*top*) have persisted for more than 25 years throughout a large area of the western tropical Atlantic (*left*). The extent of this structure shows that tiny salt fingers make an important contribution to the vertical mixing of water masses, a process that goes on only weakly in most other areas of the open ocean.

ture sensors further revealed convective plumes operating within the mixed layers. The ocean's workings seemed to be following the patterns we had observed in the laboratory.

Still, the overall vigor of salt fingering my colleagues and I encountered during C-SALT was somewhat lower than we had anticipated. Also surprising to us was that Williams's optical device showed the fingers to be tilted far from vertical by the changes in velocity (shear) between layers. The weaker mixing seems to have resulted because the interfaces proved to be a few meters thick rather than a few tens of centimeters, as we had expected from laboratory experiments. More sophisticated physical models recently developed by Eric Kunze of the University of Washington account for the thickness variations and the effects of the vertical shear on the ocean's salt fingers.

Measurements made during C-SALT also highlighted an important thermodynamic difference between salt fingers and ordinary ocean turbulence. When turbulence from waves and eddies mixes the upper ocean, it distorts the normal temperature distribution as it converts some of the kinetic energy of the surface movements into potential energy in the thermal structure (by lifting dense parcels of water and depressing light ones). Some of this kinetic energy is dissipated, however, as heat. Salt fingers also mix parcels of water upward and downward. But compared with turbulence, salt fingering causes relatively little heat dissipation and so may allow five to 20 times as much vertical mixing for the equivalent energy loss.

The staircases encountered during the C-SALT expeditions cover an enormous expanse (about 25 percent of the area of the Atlantic between 10 and 15 degrees north latitude). Our estimates suggest that the mixing rate for salt is 10 times larger in this region than outside it, where only weak mixing occurs. Thus, staircases like those encountered during the C-SALT expeditions may provide a critical path for the ocean's vertical transfer of salt, oxygen and nutrients, as well as for many recently introduced pollutants.

Fingering in Other Realms

Progress in understanding the mixing from salt fingers has application far beyond calculating the ocean's response to environmental change. It has been suggested that the tall, narrow structures found in basaltic rock formations (termed columnar jointing) result from a type of "basalt fingering" within the cooling magma. Recent experiments on molten glasses by Yan Liang and Frank M. Richter of the University of Chicago and E. Bruce Watson of Rensselaer Polytechnic Institute confirm that fingers can evolve in such viscous fluids. Fingers also occur in metal alloys, leading to speckling and strength defects in castings that are cooled from below.

Salt-finger mixing can in theory also arise within stars and in the atmospheres of gaseous planets. Will space probes of the next century discover extraterrestrial analogues to the C-SALT staircases when they explore the atmospheres of Jupiter and Saturn? The answer must await advances in planetary exploration, but further insight into this fascinating physical phenomenon can meanwhile be obtained from the continued study of the earth's own magnificent fluid laboratory, the ocean.

FURTHER READING

THE MICROSTRUCTURE OF THE OCEAN. Michael C. Gregg in *Scientific American,* Vol. 228, No. 2, pages 65–77; February 1973.

OCEAN CIRCULATION PHYSICS. Melvin E. Stern. Academic Press, 1975.

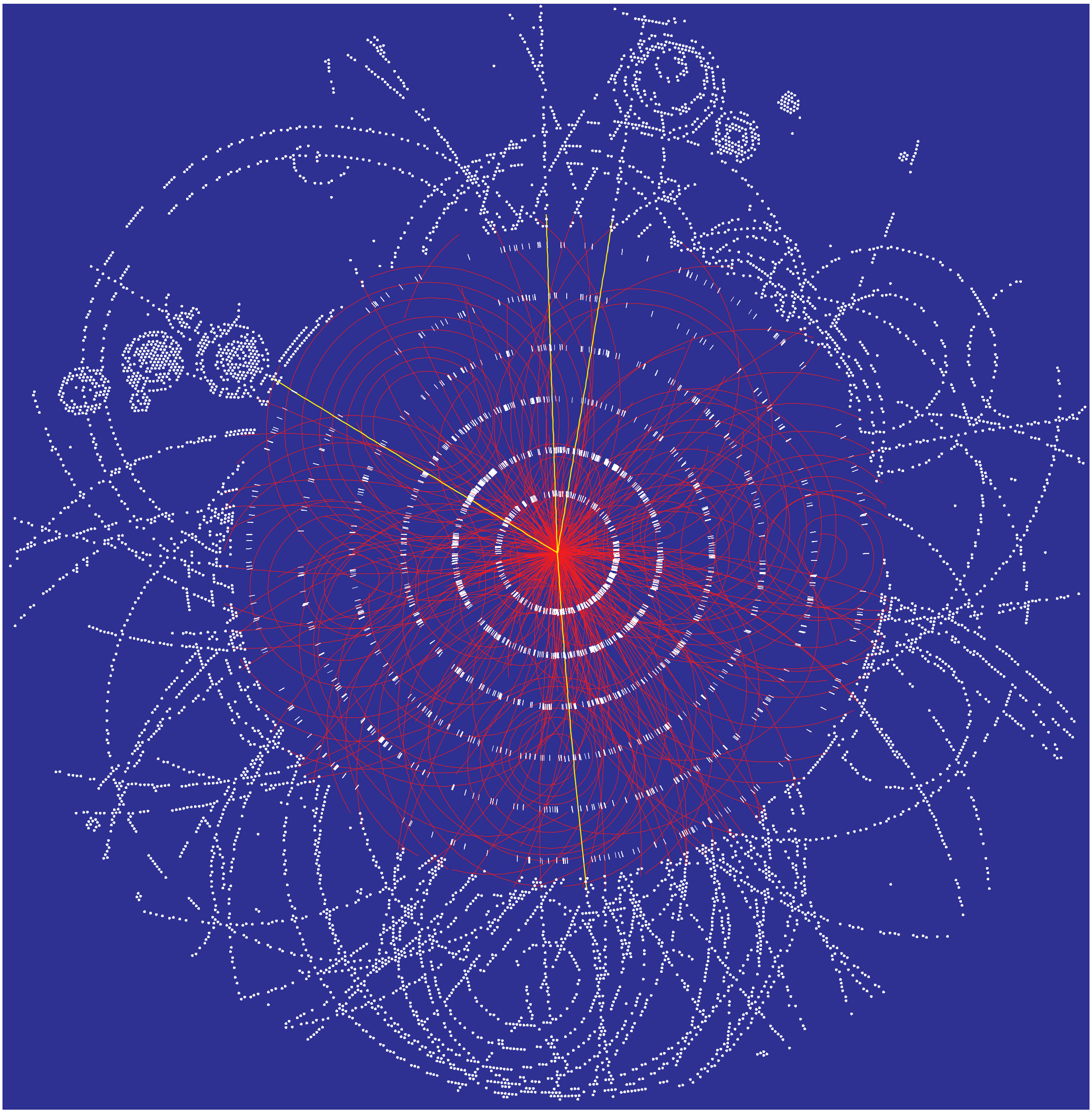
BUOYANCY EFFECTS IN FLUIDS. J. S. Turner. Cambridge University Press, 1979.

DOUBLE DIFFUSION IN OCEANOGRAPHY. R. W. Schmitt in *Annual Review of Fluid Mechanics,* Vol. 26, pages 255–285; 1994.

The Silicon Microstrip Detector

Produced with the same tools used to create integrated circuits, these detectors recently helped to find the top quark and are central to other crucial experiments

by Alan M. Litke and Andreas S. Schwarz


CERN

Silicon, the elixir of the electronics revolution, has long been known as a material of uncommon adaptability. Now, roughly half a century after the discoveries that eventually ushered in today's world of ubiquitous computers, communications technologies and photovoltaic cells, this versatility appears headed toward remarkable heights. Coming years promise chips that will let electronics communicate with human nerves and tiny machines able to perform microsurgical tasks.

Yet another fascinating application is in detecting charged particles, in which a kind of silicon-based microscope is used to peer into the very heart of matter. Fifteen years in the making, the silicon microstrip detector allows physicists to measure the positions of particles to within 10 microns or less. Such precision—at least 10 times better than what has been readily achievable with previous electronic particle-detection technologies—has opened new experimental possibilities in particle physics and other scientific fields.

At the intersection of high technology and high-energy physics, these detectors have already helped researchers confirm the existence of the long-sought top quark. The recently completed effort to establish that particle's existence conclusively is just one of several projects that use these detectors to address fundamental questions and that would be far more difficult without them. Notable among these endeavors will be the hunt for the Higgs boson, the mysterious and elusive particle that physicists believe imbues matter with mass.

Moreover, our story extends far beyond particle physics. Thanks to the versatility and richness of the underlying silicon technology, new types of detectors are poised for wide scientific application. Possible uses ranging from x-ray and gamma-ray astronomy to medical imaging are being actively pursued. Indeed, the microstrip detector is part of a broad revolution in scientific imaging, in which modern semiconductor technology is being brought to bear on the detection of light, x-rays, charged particles and even neurophysiological signals. As such, the detectors nicely illustrate the flourishing symbiosis between science and technology.

DETECTION OF HIGGS BOSON, the elusive particle believed to endow matter with mass, may look like this computer simulation of an event at the planned Large Hadron Collider at CERN. Protons smashed against protons may produce a Higgs particle that decays into four other particles, called muons (*yellow lines*). Hundreds of background particles (*red*), which must be separated from the few tracks of interest, will also be produced. Six cylindrical layers (*white lines*) of silicon microstrip and pixel detectors, concentric with the collision region, will record where the charged particles pass through these detector layers. These positions, combined with those from an outer array of gaseous detectors (*white dots*), will permit precise reconstruction of the trajectories.

How Microstrip Detectors Work

In a silicon strip detector, particles are detected with diodes, electronic "one-way streets" that allow current to flow in one direction but not the other. This ability to block the flow of current is the whole point of using a diode as a detecting element; otherwise statistical fluctuations in a large current flow would mask the relatively faint signals caused by the particle being detected.

Using the techniques of integrated-circuit fabrication, the diodes are deposited on wafers of silicon. They can be set down in almost any geometric pattern but are usually fabricated as parallel strips, each a diode several centimeters long. Strip-to-strip spacing is typically 25 microns, accurate to a small fraction of a micron.

A charged particle passing through a wafer, or an x-ray being absorbed into it, generates an electrical signal on a small number of strips. The signal reveals, to within a fraction of the strip spacing, where the particle intercepted the wafer along one dimension. For the typical spacing mentioned, the resolution would be about five to 10 microns. In most major experiments, researchers are now upgrading their instruments by using detectors that have strips on both sides of the wafer, with the strips on one side at an angle to the strips on the other (alternatively, two wafers are glued together in the same basic configuration). Such double-sided microstrip detectors give two independent coordinates for a hit, thus specifying a point in space.

Transforming a piece of raw silicon into such a precision instrument takes some special processing. A pinch of phosphorus here, a dash of boron there and a few other choice ingredients subtly but remarkably alter the electrical properties of silicon. The starting point is a thin wafer of crystalline silicon, typically 300 microns thick and 10 centimeters in diameter. The silicon crystal has a diamondlike structure: every atom has four equidistant nearest neighbors and shares one of its four outer electrons with each neighbor.

This very pure silicon is then lightly doped with an impurity atom, such as phosphorus, that has five outer electrons. (We do mean "lightly": there is only one impurity atom for every 50 billion or so silicon atoms.) Like the silicon atoms, each phosphorus atom nestling in the crystalline lattice shares four of its outer electrons with its four neighbors. The phosphorus atom's fifth electron is then easily liberated to carry an electric current. This kind of doping is therefore known as *n*-type; the *n* refers to the negative charge carried by the free electron.

Once an appropriately doped wafer has been prepared, standard techniques of the integrated-circuit trade, such as photolithography and ion implantation, place hundreds or even thousands of strips on it. Each strip is formed by implanting into the wafer surface a line of impurity atoms five to 10 microns wide and approximately 0.05 micron thick. The impurity atoms, such as boron, used for this implant have three outer electrons, which is significant because when a boron atom replaces a silicon atom in the crystalline lattice, it readily acquires a fourth electron by bonding to its four neighbors. Acquisition of this extra electron creates the absence of an electron, or "hole," somewhere else in the crystal. Silicon doped in this manner—with impurity atoms that have three outer electrons—is called *p*-type, because the available charge carrier, the hole, is in effect a positively charged particle.

Thus, a diode consists of a *p*-type material, an *n*-type material and an interface, or "junction," between the two. In the region of this junction, electrons from the *n*-type material diffuse into the *p*-type material and are neutralized by the holes, leaving behind ionized impurity atoms on both sides of the junction. (The diffusion of holes from the *p*-type region to the *n*-type region has

ALAN M. LITKE and ANDREAS S. SCHWARZ collaborated in the mid-1980s on the first silicon microstrip–based system for detecting high-energy particles in colliding-beam experiments. That project was carried out at the Stanford Linear Accelerator Center. Litke is research physicist at the Santa Cruz Institute for Particle Physics and adjunct professor at the University of California, Santa Cruz. He now works on the Aleph experiment at CERN, the European laboratory for particle physics near Geneva. Schwarz is senior scientist at DESY, the German particle physics laboratory in Hamburg. He is participating in an experiment that will study the subtle difference in the behavior of matter and antimatter.

the same result.) These fixed, charged atoms give rise to an electric field that acts like a barrier that tends to keep holes on the *p*-side and electrons on the *n*-side. The net outcome is a thin “depletion” region at the junction, free of charge carriers but containing a strong electric field. Suitably enlarged, this depletion region is where particles are detected when the diode is part of a microstrip detector.

For efficient detection of both charged particles and x-rays, this depletion region must be as deep as is practical. More depth means a longer path over which a charged particle generates a signal as it traverses the wafer, thereby making for a stronger signal. In the case of x-rays, the efficiency of detection increases because there is more material for absorption.

Applying a voltage with the proper polarity to the diode increases the depletion depth. In a microstrip detector, a negative terminal is attached to the *p*-type strips and a positive terminal to the *n*-type substrate. Under the influence of the voltage, the charge carriers are driven away from the junction, causing the depleted region to expand.

The depleted region of the diode is now ready to function as a particle detector. When a charged particle passes through, it interacts electrically with the lattice, depositing energy and liberating electron-hole pairs in a narrow column along its path. Under the influence of the electric field established by the ionized impurity atoms, the electrons and holes drift in opposite directions, with the holes moving toward the strips. The motion of these charge carriers generates an electrical signal on one or a few nearby strips. Moreover, because these carriers freed by the intruding particle drift only a very short distance, the signal is generated in only 20 billionths of a second or so.

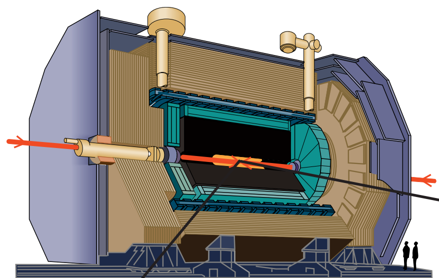
Once generated, these signals must be amplified and recorded before they can be used to determine a point on the particle’s trajectory. These crucial functions are carried out by special electronic readout chips equipped with an array of amplifier circuits that is comparable, in density, to the strips on the detector wafer. Through a wire-bonding technique, each strip on the detector is ultrasonically stitched to its corresponding amplifier on a readout chip with an aluminum wire one quarter the width of a human hair.

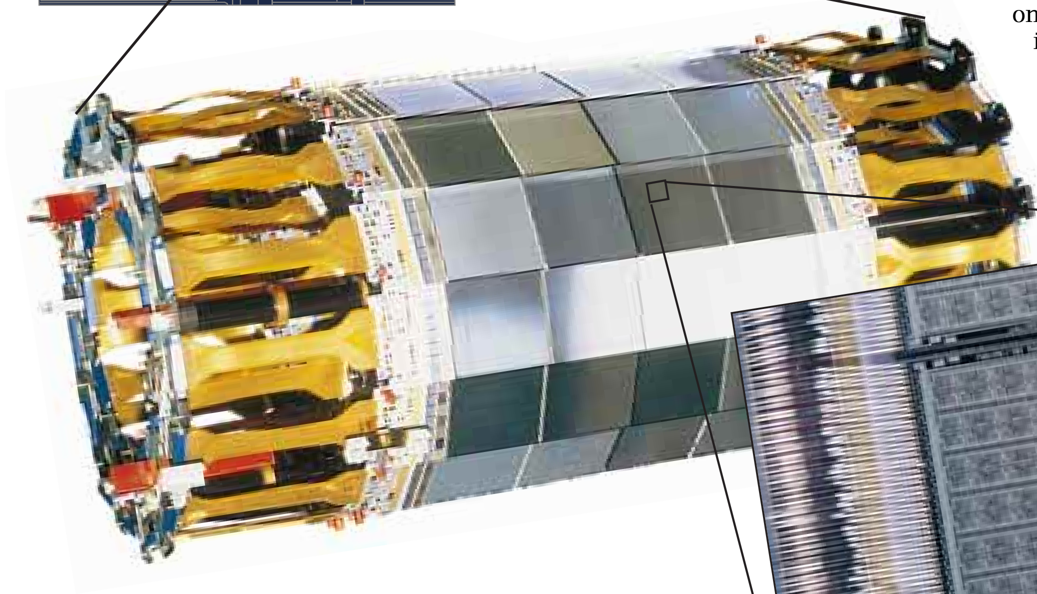
Out of such details comes the superior spatial resolution of microstrip detectors—their very raison d’être. Before these detectors became available in the 1980s, particle tracking was accomplished electronically through the ionization of a gas, with signals detected on a conducting wire. Developed some 27 years ago, gaseous wire chambers remain an important part of all current and planned high-energy physics experiments, and their invention was a major breakthrough that won Georges Charpak the 1992 Nobel Prize in Physics. Microstrip detectors, however, are about 10 times more accurate.

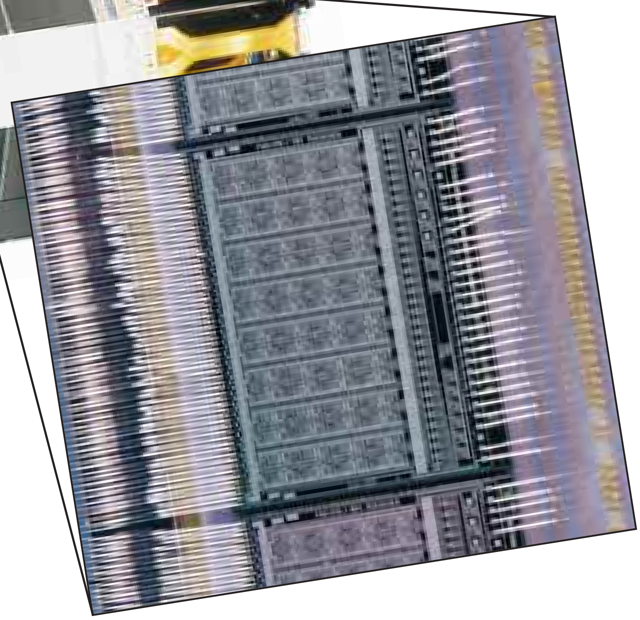
Applications in Particle Physics

Although their applications have flourished in initially unimagined ways, silicon microstrip detectors were born of a specific scientific need: to detect and study particles with “charm”—that is, ones containing the charm quark. This quark is one of the six enshrined in physicists’ evolving Standard Model of the basic constituents of matter and their interactions. In this model, all matter is built from six quarks (up, down, charm, strange, top and bottom), along with six particles known as leptons and the 12 corresponding antiparticles of the quarks and leptons.

Charmed particles have lifetimes on the order of one trillionth of a second (one picosecond). After being shaken loose in a high-energy collision, a charmed par-







SILICON VERTEX DETECTOR is at the heart of a much larger system of particle detectors, called Aleph (*top*). Beams hit head-on at the center of the vertex detector (*bottom*), inside its two concentric, cylindrically shaped arrangements of microstrip detector modules. Each module consists of four particle-sensing silicon wafers and several amplifying chips, bonded together by myriad wires (*right*). Not visible in this photograph are the individual detecting strips on the silicon wafer.

CERN (*top* and *bottom*); DESY (*right*) ticle will generally travel a few millimeters, depending on its energy and type, before decaying into several other particles. The point where the charmed particle is produced and liberated is known as the primary vertex; the point where it decays is called the secondary vertex. Distinguishing the secondary vertex from the primary is one of the key requirements of detecting charmed particles and measuring their properties.

These vertices are found by tracking the trajectories of the particles emerging from a collision with a set of detectors arranged near the point where the collisions take place. In a colliding-beam experiment, this set typically consists of several concentric layers of silicon microstrip detectors, outside of which are some gaseous wire chambers. Tracks recorded by these detectors are extrapolated back to the vicinity of the primary vertex. The trajectories of particles arising from the decay of a charmed particle converge at a secondary vertex that, if all has gone well, is distinct from the primary vertex.

As it turns out, the precision needed to extrapolate the tracks in order to keep the different vertices in focus should be much less than the particle lifetime multiplied by the speed of light. For charmed particles, the required precision is a few tens of microns, a value well within the capabilities of silicon microstrip detectors.

Quite fortuitously, two other types of particles have lifetimes comparable to those of charmed particles. The tau lepton, the heaviest known member of the lepton family, has a lifetime of 0.3 picosecond. Bottom particles, which contain the bottom quark, have lifetimes on the order of 1.5 picoseconds. Silicon microstrip detectors are well suited for detecting and studying particles such as these.

Telltale Bottom Particles

Detection of bottom particles is particularly important for high-energy physics; it is one of the central aims of many of the current and planned microstrip detection systems. For one, these particles are expected to behave differently from their antiparticle counterparts, so their study should shed light on the subtle imperfection in the symmetry between matter and antimatter (physicists refer to this as "CP violation"). In addition, particles containing the top quark ("top particles") decay nearly all the time into bottom ones. Moreover, the Higgs, if it is not too massive, is expected to decay most of the time into a pair of particles, one a bottom and the other an antibottom. Thus, three of modern physics' most pressing experimental quests are intimately tied to bottom particles and therefore to microstrip detectors.

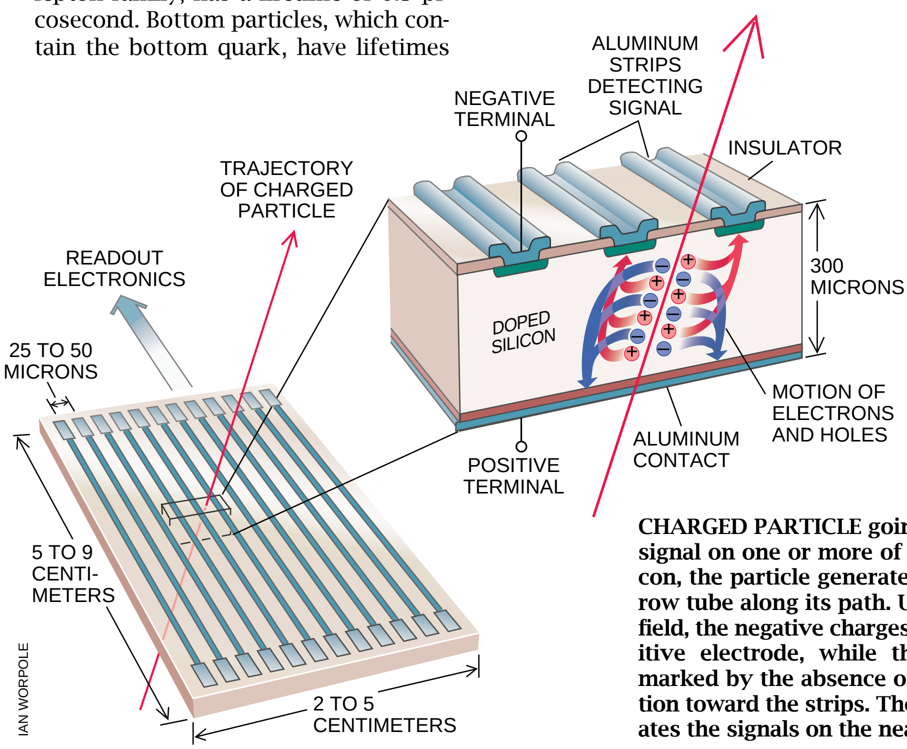
One effective way of producing and studying bottom particles is by colliding energetic beams of electrons and antielectrons (positrons) head-on to produce particles called Z^0, which then decay into bottom and antibottom ones. (The Z^0 is one of three "gauge bosons" that mediate the weak force that governs radioactive decay; the other two are designated W^+ and W^-.) One colliding-beam machine for this purpose is the Large Electron-Positron Collider (LEP), located in a 27-kilometer circular tunnel that straddles the French-Swiss border near Geneva. The LEP is part of CERN, the European laboratory for particle physics. Four experiments gather data at the LEP, including one called Aleph, on which we work with 400 other physicists.

At Aleph's center is a vertex detector consisting of a cylindrical arrangement of double-sided silicon microstrip detectors, built in two concentric layers, placed around the beam pipe containing the collision region [*see illustration on opposite page*]. Surrounding this vertex detector is a large array of particle detectors, including two gaseous tracking chambers enclosed in a superconducting magnet. This complex detecting system is used to identify, and measure the momenta of, particles that fly out from the collision region whenever a Z^0 decays.

About 70 percent of the time, the Z^0 decays into a few (usually two) narrow showers, or "jets," of particles traveling out from the collision region. In about one fifth of these events, one of the jets will contain a bottom particle, and a second jet will include an antibottom one. Each bottom particle travels a few millimeters on average before decaying, in most cases into a charmed particle and one or more others [*see illustration on next two pages*].

With its silicon microstrip devices, Aleph's vertex detector can reveal a magnified view of the event with extraordinary sharpness. Brought into focus are the primary vertex, where the Z^0 decays and the bottom particle is born, and the secondary vertex, where it decays into a charmed particle. Even a tertiary vertex, where the charmed particle decays, is obvious.

This kind of detailed information, collected from a large number of events, discloses some fundamental properties of bottom particles. Their lifetimes, which tell us about the weak interactions of the bottom quark, are measured from the decay length—the distance between the primary and secondary vertices. This same length, furthermore, is a factor in measurements of the frequency with which the neutral bottom particle



CHARGED PARTICLE going through a microstrip detector causes a signal on one or more of its parallel strips. As it traverses the silicon, the particle generates positive-negative charge pairs in a narrow tube along its path. Under the influence of an internal electric field, the negative charges—electrons—drift rapidly toward the positive electrode, while the positive charges, or "holes"—places marked by the absence of an electron—drift in the opposite direction toward the strips. The motion of the electrons and holes generates the signals on the nearby strips.

becomes its antiparticle—a fascinating example of the transformation of matter into antimatter.

The Top and the Higgs

Silicon microstrip detectors played an important role in the successful search for the long-sought top quark (the most massive of the six) and will be central to future searches for the "most wanted" particle in high-energy physics, the Higgs boson. Confirmation of the top quark's existence emerged in early March from the Fermi National Accelerator Laboratory in Batavia, Ill., where the Tevatron collider, currently the world's highest-energy colliding-particle machine, smashes protons at 900 billion electron volts against antiprotons of the same energy.

In a tiny fraction of the proton-antiproton collisions, particles containing the top quark are produced. Almost all these top particles decay into a bottom particle and a charged gauge boson (a W^+ or a W^-). A silicon vertex detector was used to single out events containing bottom decays, thereby separating top events from background ones and greatly emphasizing the top signal.

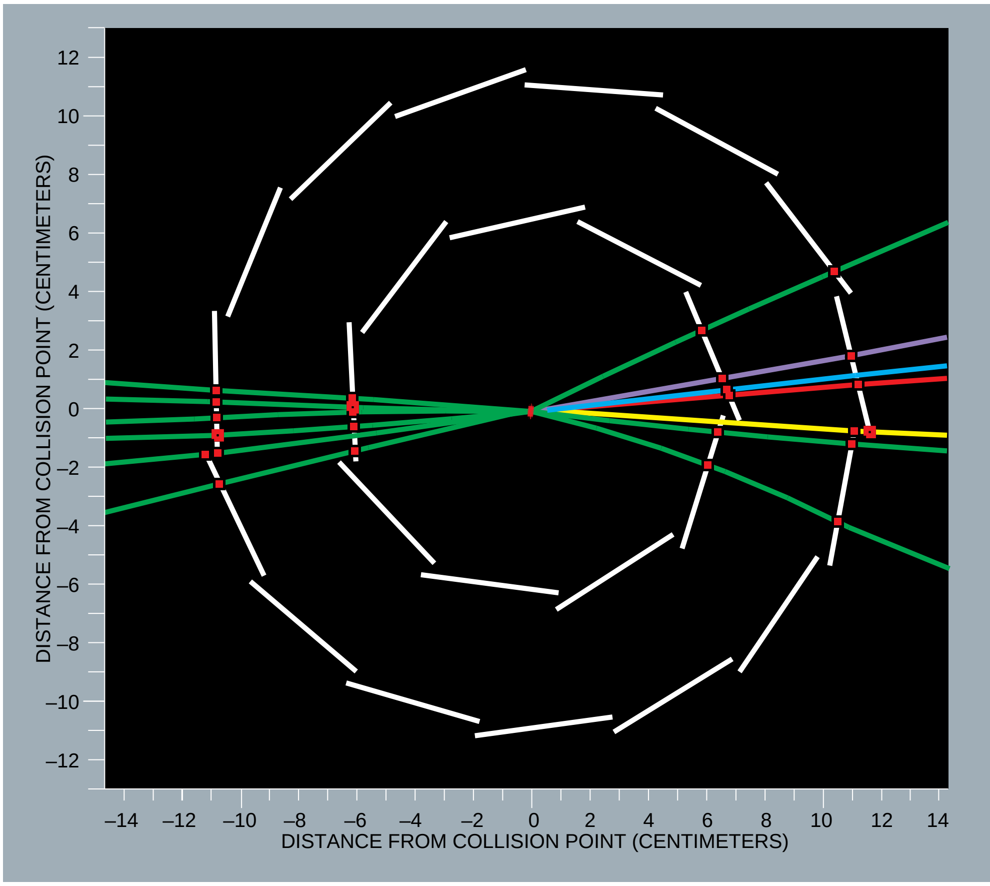
More elusive still is the Higgs boson, which after more than three decades of theoretical and experimental work, remains the Holy Grail of particle physics. It is hypothesized to imbue quarks and the charged leptons, as well as the Z^0, W^+ and W^-, with mass [see "The Higgs Boson," by Martinus J. G. Veltman; SCIENTIFIC AMERICAN, November 1986].

Searches for the Higgs at the LEP have come up empty-handed, indicating that it is too massive to be produced at a detectable rate at the present LEP collision energy. The search, though, will soon continue with renewed vigor, when the LEP's energy is approximately doubled and researchers use improved silicon vertex detectors to identify the bottom decays characteristic of their prey.

In the end, it may take a still more energetic machine. The one that is perhaps most likely to find the Higgs, and even new physics beyond the Standard Model, is the Large Hadron Collider (LHC), to be built by CERN. About eight times as energetic as the Tevatron, it will collide protons against protons inside the LEP tunnel and extend the search for the Higgs boson to masses several times greater than anything detectable with today's instruments.

Even in the energetic LHC, Higgs bosons will be uncommon indeed. Investigators will compensate by setting the rate for interactions extraordinarily high—almost one billion every second—but such a frequency will contribute to difficult experimental conditions. The very rare events of interest will have to be disentangled from a fantastic quantity of debris, and the detectors and electronics will be harshly irradiated.

One of the key technologies in this forbidding environment will be the silicon microstrip detector. Its high speed and close spacing of detecting elements will enable researchers to separate events from different beam collisions and recognize the few important tracks among them. Its high precision will permit accurate measurement of the momenta of energetic charged particles from the slight bending of their trajectories in a magnetic field.



BRIEF LIFE OF PARTICLE is recorded by concentric rings of silicon microstrip detectors in the Aleph system, shown as white lines in the image above. The red squares indicate where charged particles (*colored lines*) traversed the detectors and left a detected signal. These particles were created by the violent annihilation

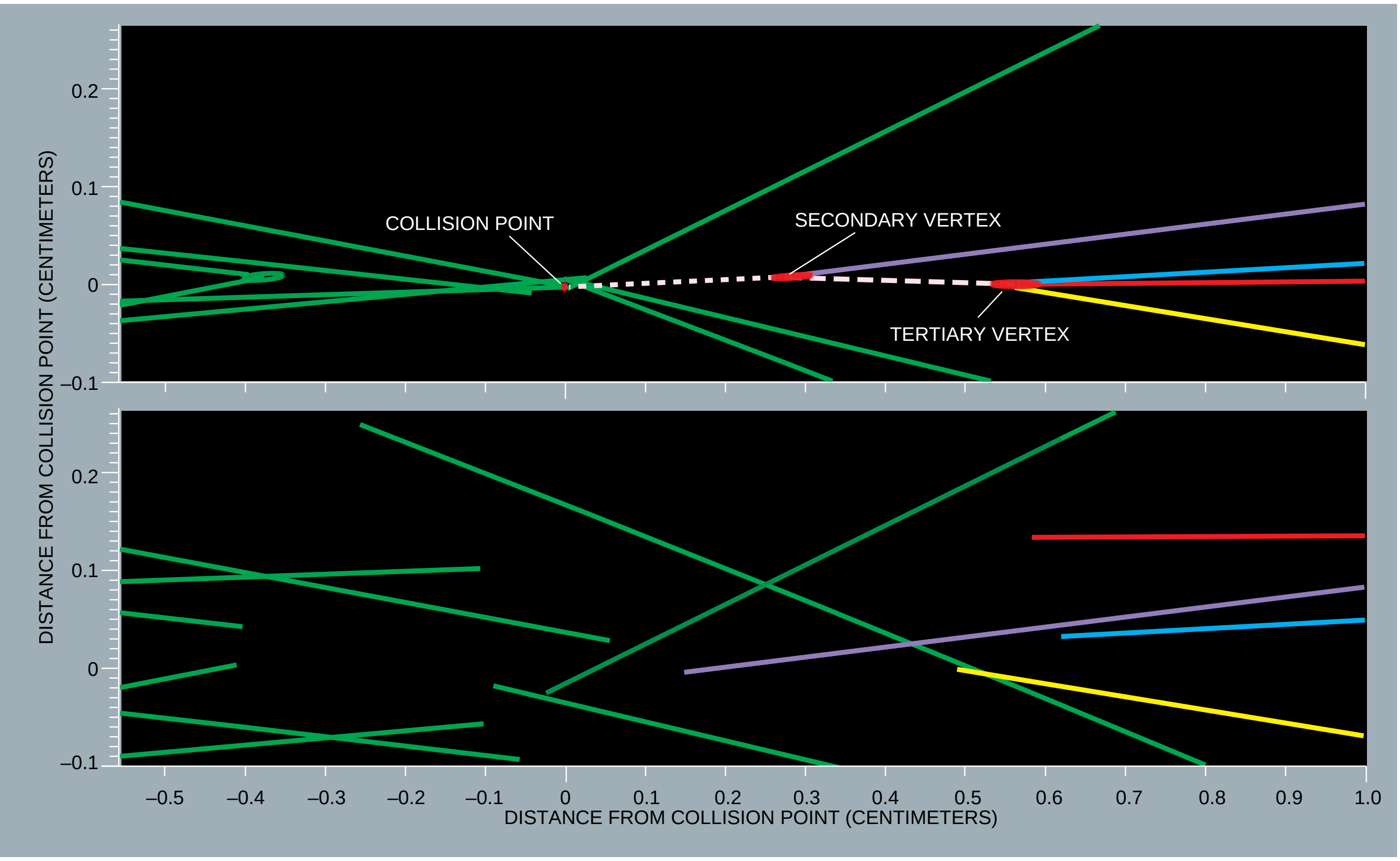
Seeing Stars

The revolution in precision tracking fostered by silicon microstrip detectors is starting to spill over to other areas of research, notably astrophysics and x-ray imaging. One exciting possibility is in satellite-based gamma-ray astronomy. Groups from Stanford University and other institutions are working on a conceptual design of a Gamma (ray) Large Area Silicon Telescope able to map out galactic and extragalactic gamma-ray sources with unprecedented precision and sensitivity.

In this telescope, silicon microstrip detectors will be interwoven with thin metal plates that will, in general, convert an incoming high-energy gamma ray into an electron-positron pair. This pair will be tracked in the layers of microstrip detector, thereby giving information on the direction of the incident gamma ray.

Useful as it is, the microstrip detector is not perfect. A double-sided detector, with its independent, perpendicular sets of detecting strips, can by itself pinpoint no more than one hit at a time. A single particle generates one signal on each side—these can be unambiguously associated to give a two-dimensional point. Multiple particles, however, generate multiple signals, causing ambiguity—a signal on one side can be associated with any of several signals on the other side.

Many applications, such as x-ray imaging, require an unambiguous location for every hit, including simultaneous



of an electron and an antielectron at the center of the rings. As revealed by the magnified image (*top*), one of the products of the annihilation is a particle that disintegrates into other particles at a secondary vertex. One of the particles in turn decays into still more at a tertiary vertex. Without the data from the microstrip detectors (*bottom*), the tracks cannot be reconstructed with enough precision to discern the initial production and subsequent decay points.

ones. Creation of so-called pixel detectors with this capability is a very active area. One variety, under development at CERN and elsewhere, is the hybrid pixel detector. The detector chip in such instruments is fabricated in much the same way as a strip detector, but the strips are replaced with a two-dimensional array of small squares or rectangles (the pixels), typically 50 to 500 microns on a side. This chip is sandwiched to another one carrying a corresponding array of amplifier/signal processing elements such that each pixel is connected to one of its elements via a set of conducting bumps.

Pixel detectors are planned for the innermost layers of the charged-particle tracking systems of the big LHC experiments. Close to the collision region, with its extreme density of particles, track recognition and reconstruction will be best accomplished with the fine-granularity, space-point measurements possible with pixel devices.

Another possible use of pixel detectors, still at an early stage of development, is in x-ray radiography for medical imaging, especially mammography. X-rays absorbed in a pixel detector directly generate signals that can be converted into images in an electronic, digital format for immediate computer display, analysis and storage; the delays and storage concerns associated with photographic film could be eliminated.

Symbiosis in Action

Silicon microstrip detectors are a fine example of an underappreciated partnership. The scientific quest to fathom nature often spurs the development of new technology and new instruments and sometimes leads to the creation of entirely novel industries. These innovative instruments then lead to further progress in science.

For instance, the desire to understand the atom led, in the 1920s, to the exposition of quantum mechanics. This, in turn, brought about a deeper understanding of matter in the solid state and to the invention of the transistor in 1947. The transistor anchored the still unfolding electronics and computer revolutions; its development was the cornerstone of the semiconductor industry, whose technology is now being used to fabricate microstrip detectors. These devices are helping to answer some of the most urgent questions of fundamental physics: the nature of the top quark and the Higgs boson, to give just two examples.

Science and technology are inextricably woven together in a tapestry of great beauty. The whole is much greater than the sum of the parts.

FURTHER READING

SEMICONDUCTOR DEVICES: PHYSICS AND TECHNOLOGY. Simon M. Sze. John Wiley and Sons, 1985.

GAMMA LARGE AREA SILICON TELESCOPE (GLAST): APPLYING SILICON STRIP DETECTOR TECHNOLOGY TO THE DETECTION OF GAMMA RAYS IN SPACE. W. B. Atwood in *Nuclear Instruments and Methods in Physics Research A*, Vol. 342, No. 1, pages 302–307; March 15, 1994.

HEAVY FLAVOUR PHYSICS AT COLLIDERS WITH SILICON STRIP VERTEX DETECTORS. Andreas S. Schwarz in *Physics Reports*, Vol. 238, Nos. 1–2, pages 1–133; March 1994.

The Atomic Intrigues of Niels Bohr


NIELS BOHR INSTITUTE *AIP Niels Bohr Library*

Bohr, 1935

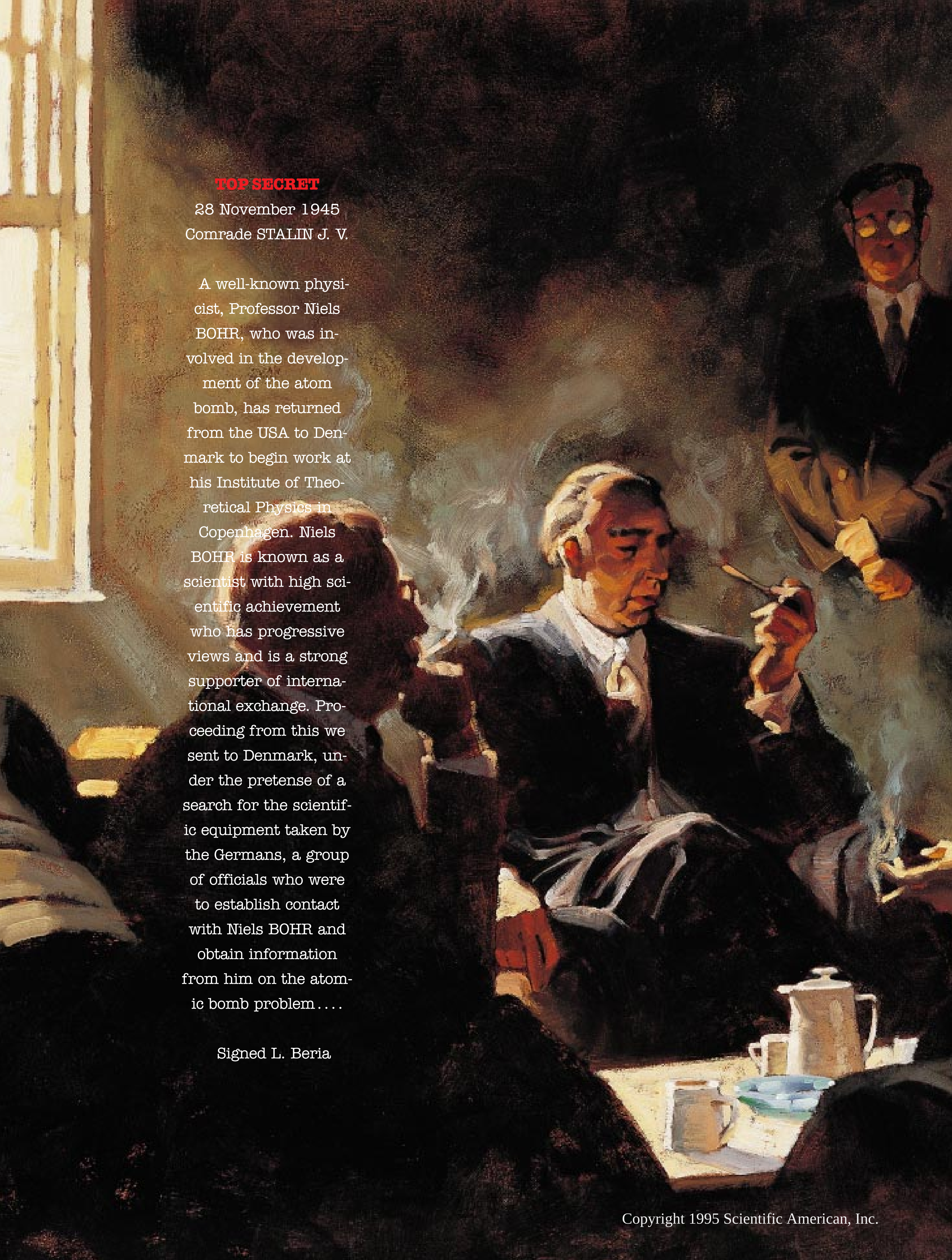
Scandal is not usually linked with the name Niels Bohr; genius and character are. The Nobel Prize–winning physicist who first described the atom through quantum mechanics was a titan of 20th-century science. During World War II, he opposed the Nazi occupation of his native Denmark and as a refugee participated in the Manhattan Project, which gave nuclear might to the U.S. Yet Bohr was also an outspoken advocate of international cooperation, urging Franklin D. Roosevelt and Winston Churchill to provide other countries with knowledge of the bomb's workings in the interest of world peace.

Would Bohr have acted unilaterally on those convictions? Last year some seemingly well-placed sources publicly alleged that Bohr had given classified information about the U.S. bomb program to Soviet agents. Suspicions centered on a meeting between Bohr and representatives of the secret police that took place in 1945.

Those misgivings are unfounded, as the first of the following articles makes clear. The authors, renowned physicists in their own right, review the circumstances leading up to the encounter and show that Bohr made no attempt to conceal it from the authorities. Moreover, citing a memorandum to Joseph Stalin that includes a transcript of the conversation, they argue that Bohr remained deliberately vague about what he knew. Their scholarship should help erase the stain on Bohr's good reputation.

The second article illuminates a less well-known but in many ways equally fascinating episode of Bohr's life—one in which he brought insights about the Nazi nuclear weapons program to the Allies. The author, a physicist-turned-sleuth, follows a 50-year-old trail to discover whether Bohr received a drawing of a nuclear device from the German scientist Werner Heisenberg. The story is an important reminder of a time not so long ago when the balance of power and the cloak-and-dagger of espionage hung over the work of every physicist studying the secrets of the atom.

—The Editors



TOP SECRET

28 November 1945

Comrade STALIN J. V.

A well-known physicist, Professor Niels BOHR, who was involved in the development of the atom bomb, has returned from the USA to Denmark to begin work at his Institute of Theoretical Physics in Copenhagen. Niels BOHR is known as a scientist with high scientific achievement who has progressive views and is a strong supporter of international exchange. Proceeding from this we sent to Denmark, under the pretense of a search for the scientific equipment taken by the Germans, a group of officials who were to establish contact with Niels BOHR and obtain information from him on the atomic bomb problem....

Signed L. Beria

GREGORY MANCHESS

Did Bohr Share Nuclear Secrets?

Niels Bohr met with a Soviet agent in late 1945. Although some have accused Bohr of divulging nuclear secrets, a recently disclosed memo offers evidence to the contrary

by Hans A. Bethe, Kurt Gottfried and Roald Z. Sagdeev

"The most vital information for developing the first Soviet atomic bomb came from scientists designing the American atomic bomb at Los Alamos.... They agreed to share information on nuclear weapons with Soviet scientists...."

This sensational allegation from the 1994 book *Special Tasks* goes on to accuse Niels Bohr, as well as other famous Manhattan Project scientists, of providing Soviet spies with confidential information about the American effort to develop nuclear bombs. The book, based on the recollections of Pavel Sudoplatov, one of Joseph Stalin's most senior secret police officers, does not document its grave charges with any evidence. Nevertheless, the book's publication triggered an intense controversy about the truth of the claims.

Recently a memorandum to Stalin that describes the contact with Bohr was released from the KGB archives. The document contains a supposedly verbatim transcript of the very meeting, held in November 1945, in which Bohr allegedly supplied the Soviets with classified information. We have now reviewed this document thoroughly and find it is evidence that the charges against Bohr are false.

The controversy and our analysis of the memo can be understood only if one knows something about the development of nuclear physics before the war and the efforts to apply the resulting scientific knowledge to military purposes during World War II. By the early part of this century scientists had recognized that an atom consists of electrons orbiting a nucleus, which contains positively charged protons. They also knew that the energy inside the nucleus is millions of times greater than that discharged by chemical explosives. But rather little was known about the structure of the nucleus until the discovery of the neutron in 1932 by James Chadwick of the University of Cambridge. His work demonstrated that nuclei are composed of protons and neutrons.

The neutron quickly became an incisive probe for revealing properties of the nucleus. Having no electrical charge, neutrons are not repelled by the positively charged nuclei and can readily enter the nuclei of many elements in a process known as neutron capture. Many nuclear reactions involving neutron capture were studied, but none produced dramatic changes in nuclei.

In December 1938, however, Otto Hahn and Fritz Strassmann, working at the Kaiser Wilhelm Institute in Berlin, observed an astonishing reaction involving uranium, the heaviest element then known. The two were amazed to find that when uranium was bombarded with neutrons, barium, a relatively light element, was among the reaction products. Otto Frisch and Lise Meitner, two refugees from Nazi Germany, promptly interpreted Hahn and Strassmann's observation as indicating

MEETING BETWEEN NIELS BOHR AND SOVIET AGENT Iakov Terletskii (*right*) is shown in this artist's conception. The visit took place in November 1945 at Bohr's office in Copenhagen. Terletskii brought a translator (*left*); Bohr (*center*) was accompanied by two of his sons, Ernest (*far right*), who stood guard in another room with a loaded pistol, and Aage (*standing*), who remained in the room. A memo to Joseph Stalin (*section translated at left*) introduced a transcript of the conversation between Bohr and Terletskii. Excerpts of the transcript begin on the next page.

A Conversation in Copenhagen

The following excerpts of Terletskii and Bohr's discussion are taken from a translation of the transcript sent to Stalin in 1945. The entire transcript can be obtained from Scientific American *on America Online.*

TERLETSKII: By what practical method was uranium 235 obtained in large quantities, and what method is considered to be the most promising (diffusion, magnetic or some other)?

BOHR: The theory of obtaining uranium 235 is well known to scientists of all countries; it was developed before the war and presents no secret. The war has introduced nothing basically new. It should be pointed out, however, that the problem of the uranium reactor and of plutonium resulting from it were both solved during the war, although they were not new in principle either. They were practically implemented. The main development is in separating uranium 235 from the natural mixture of isotopes. After the required amount of uranium 235 is obtained, no theoretical difficulty stands in the way of a bomb. Separation of uranium 235 is achieved by the diffusion method, which is well known, and also by mass spectroscopy. No new method has been used. The American success is due to practical implementation of well-known designs on an incredibly huge scale. I should warn you that, while in the USA, I did not take part in the technological development of the project, and therefore I am not familiar with its design features or

the breakup of the uranium nucleus into two nuclei of roughly equal size. Frisch and Meitner dubbed the process "fission." They reasoned that the electrical repulsion among the many protons in the heaviest nuclei made those nuclei nearly unstable, so that capture of an additional neutron could cause the nuclei to fall apart. They also calculated that if uranium split in two, vastly more energy was liberated than in all nuclear reactions previously observed.

Possibility of a Chain Reaction

Frenzied activity ensued in the small international community of nuclear physicists. At first, their dominant motive was scientific curiosity, and all new results were published openly. But two discoveries in early 1939 were to be of capital importance to the military programs that began after Adolf Hitler's invasion of Poland in September 1939.

Bohr, while visiting Princeton University from Copenhagen in the winter of 1939, surmised that in natural uranium ore only a very tiny fraction of the nuclei was responsible for the observed fission. The second crucial discovery, made by two independent groups, was that several neutrons were emitted by the uranium nucleus during the process of fission. These secondary neutrons could, therefore, induce fission in other nuclei and thus initiate a chain reaction.

With the American physicist John Wheeler, Bohr developed a detailed theory of fission that enabled the two scientists to predict whether the nucleus of a given element would undergo fission after neutron capture. They also estimated probabilities for nuclei to undergo fission "spontaneously," that is, without first being destabilized by a captured neutron.

Using their new theory, Bohr and Wheeler affirmed Bohr's idea that just a small fraction of the nuclei in natural uranium was susceptible to neutron-induced fission. Specifically, they showed that such nuclei were in a rare form known as uranium 235. Like most elements, uranium comes in various forms, called isotopes, that differ from one another only in the number of neutrons in the nucleus. For example, uranium 235, which makes up 0.7 percent of natural uranium ore, has 143 neutrons and 92 protons; the mass number 235 is the sum of the number of neutrons and protons. The more abundant isotope, uranium 238, has 146 neutrons in addition to 92 protons.

Bohr's prediction that only uranium 235 underwent neutron-induced fission was soon confirmed by several experiments. The finding meant that technologies based on fission of uranium would probably require isolation of uranium 235. Yet achieving such isolation would be extremely difficult. The way in which an atom behaves in all chemical processes is determined by the electrical charge of the nucleus and therefore does not depend on the number of neutrons. Hence, uranium 235 cannot be separated from uranium 238 in the ore by any chemical process. Separation must rely only on the very small (1 percent) difference in weight between the rare and abundant isotopes. No remotely affordable method for large-scale separation was then known.

This difficulty initially engendered great skepticism that nuclear fission would find practical application. Many physicists in all the major nations that would soon be at war understood that a chain reaction in uranium could have two distinct but closely related functions of profound importance. First, a controlled, self-sustaining chain reaction would make a nuclear power plant possible. Second, if the chain reaction multiplied indefinitely, it might generate an enormous explosion. But because the prospect of isolating enough uranium 235 to make a bomb seemed far-fetched, hardly anyone believed that such a device could be fabricated in the foreseeable future, much less influence the outcome of the war.

Despite this unlikely prospect, the fear that Germany—which had played so outstanding a role in modern physics—might develop nuclear weapons led physicists in Britain and the U.S. to advocate exploration of possible military uses of fission. The American nuclear weapons program began modestly in 1940 with a $6,000 government grant to Enrico Fermi and other scientists, many of them refugees from Europe. Fermi had fled Italy and moved to Columbia University in January 1939. His first project was to produce a self-sustaining chain reaction in natural urani-

HANS A. BETHE, KURT GOTTFRIED and ROALD Z. SAGDEEV have long shared an interest in the political implications of nuclear weapons. Bethe, now an emeritus professor at Cornell University, was head of the theoretical physics division at Los Alamos Laboratory during World War II. In 1967 he won the Nobel Prize for his research in astrophysics; his recent work has been in the area of supernova explosions. Gottfried is professor of physics at Cornell, where he specializes in nuclear and elementary particle physics and served as chairman of the department. He is a founding member of the Union of Concerned Scientists. Sagdeev is Distinguished Professor of Physics and director of the East-West Center for Space Science at the University of Maryland. Before coming to the U.S., he was director of the Soviet Space Research Institute; he also served as a science adviser to former president Mikhail S. Gorbachev.

the size of the apparatuses or even any part of them....

Authors' commentary: If this was really Bohr's answer, he was resorting to uncharacteristic intellectual arrogance to evade a substantive answer. The separation methods he mentioned could isolate only minuscule amounts of specific isotopes, not kilograms. This problem was widely believed (especially by Bohr) to pose an insurmountable hurdle to building nuclear bombs. Ingenious innovations were needed to make large-scale isotope separation feasible. Bohr's answer is as revealing as if Isaac Newton, on being asked how to put a person on the moon, were to reply that apart from some technical details everything needed was in the *Principia.*

TERLETSKII: What is the number of spontaneous fissions per unit time for all the above-mentioned substances [uranium 235, uranium 238, plutonium 239, plutonium 240]?

BOHR: Spontaneous fission is rather insignificant and should be disregarded for calculations. The lifetime for spontaneous fission is about 7,000 years. I cannot give you a more accurate figure, but you understand that with this value for the lifetime of spontaneous fission one should not expect it to affect the process significantly.

Authors' commentary: This comes after a string of questions about reactors and follows one on the number of neutrons released in fission, which is relevant to both reactors and bombs. Therefore, it is ambiguous whether the question refers to bombs or reactors. Bohr's answer is correct for reactor design, but not for bomb design. As explained in the text, spontaneous fission is a vital consideration in a

um despite its very small content of fissionable uranium 235; Fermi finally succeeded in December 1942 in Chicago.

A Second Path to the Bomb

Before this time, physicists in Germany, Russia and the U.S. had become aware of another possible route to a nuclear bomb. They had predicted independently, and in secret, that when the abundant isotope uranium 238 captured a neutron, it could transform into a new element—now called plutonium—and that this element would undergo neutron-induced fission as readily as the rare uranium 235. Because it is a distinct element, plutonium can be separated by chemical means from the uranium in which it has been produced, thereby sidestepping the vexing isotope separation problem. Thus, it turned out that Fermi's uranium reactor also made plutonium. This reactor became the prototype for the plutonium production plant built at Hanford, Wash.

Back in 1940, however, when the American nuclear research effort was still concentrated on whether a nuclear reactor could be made operational, the British played a decisive part in showing that a nuclear weapon might be a realistic objective. In March, Frisch and another refugee from Germany, Rudolf Peierls, working clandestinely and without government support at the University of Birmingham in England, calculated the minimum amount of pure uranium 235 that would lead to an explosive chain reaction. They arrived at the startling result that this "critical mass" would be as small as several kilograms.

Their calculation and their promising ideas for isotope separation were conveyed in a memo to Winston Churchill's cabinet, which decided to pursue in earnest research toward a nuclear bomb. The British decision helped to convince President Franklin D. Roosevelt in 1942 that he should create the Manhattan Project. This U.S. effort to develop nuclear weapons soon employed thousands at the weapons design laboratory in Los Alamos, N.M., as well as at many other large facilities spread across the nation.

Wanting to leave no stone unturned, the U.S. pursued both the plutonium and uranium routes to the bomb. The first bomb concept considered at Los Alamos Laboratory—the gun design—involved shooting two subcritical masses of uranium 235 into each other to form a critical mass suddenly. Almost pure uranium 235, separated from the ore at prodigious expense, was used in the gun-design bomb that destroyed Hiroshima, Japan, on August 6, 1945.

The same approach would not work with plutonium, however. When the first plutonium sample arrived at Los Alamos in 1944, it brought a big surprise. Plutonium was found to have an appreciable tendency to fission spontaneously. If the gun design were used to assemble the critical mass, the neutrons emitted during spontaneous fission would initiate a chain reaction producing an enormous amount of heat. That heat would disable the device and allow only a small explosion to occur.

An entirely new and far more difficult arrangement—the implosion design—was devised to detonate a plutonium bomb. The technique used a subcritical sphere of plutonium surrounded by ordinary explosives. A chemical explosion compressed the plutonium to a much higher density than it has in a normal environment. Because of this high density, the chain reaction grew with great speed. The compression was so rapid that the spontaneous fission of plutonium had no time to disable the device. The implosion design was used against Nagasaki on August 9, 1945.


AIP EMILIO SEGRÈ VISUAL ARCHIVES


AIP EMILIO SEGRÈ VISUAL ARCHIVES

BOHR'S PHYSICIST FRIENDS Peter L. Kapitsa (*top*) and Lev D. Landau (*bottom*) are thought to have provided motivation for Bohr to speak with the Soviets. Before the war, Landau had been imprisoned, and Kapitsa had not been allowed to return to England after a visit to his relatives. Bohr apparently hoped the encounter would give him a chance to convey his support and concern for the two scientists.

plutonium bomb, as it necessitated an entirely new design.

TERLETSKII: From what materials were atomic bombs made?

BOHR: I don't know exactly what material was used for the bombs dropped on Japan. I don't think any theoretical scientist will answer that question. Only the military know the answer. As a scientist, I can say that they were made either of plutonium or uranium 235.

Authors' commentary: Many senior scientists at Los Alamos (such as Bethe) knew which materials were used, but we do not know whether Bohr did.

TERLETSKII: Is the superdense matter used before the explosion of the bomb or under the impact of the explosion?

BOHR: There is no need for that. The point is that during the explosion uranium particles are moving with a velocity equal to that of the neutrons. If this were not so, the bomb would have disintegrated in just a small blast. But with an explosion, due to the equal velocity, the process of uranium fission continues even after the explosion.

Authors' commentary: The question is garbled. Presumably it was supposed to ask whether the high density of the fissile material produced by the initial chemical explosion is able to enhance the chain reaction, and an answer thereto would have been very valuable. What Bohr is reported to have said is incomprehensible to us and must have been so to the Soviets.

Two days after this second bombing, the U.S. government made public its official report on the Manhattan Project. Written by Princeton physicist Henry DeWolf Smyth, this book divulged a great deal of information but also served an important security purpose. Thousands of people had been engaged in the project, so an unambiguous definition of what was secret was crucial. As Major General Leslie R. Groves, the military commander of the Manhattan Project, wrote in the preface of the document, "No requests for additional information should be made."

Soviet Scientists and Spies

The Soviets, of course, were also interested in a nuclear weapon during the war. Thanks to David Holloway's new book *Stalin and the Bomb,* the interplay between science, espionage and politics that marked this effort have now become known in the West.

Starting in the mid-1930s, a band of young experimenters at the Leningrad Physicotechnical Institute, led by Igor V. Kurchatov, had become adept at nuclear research. For example, Kurchatov's student Georgii N. Flerov was the first to confirm the Bohr-Wheeler prediction that uranium fissioned spontaneously; he described this result in the American journal *Physical Review* in July 1940.

Soviet theorists were world-class. At the Institute of Chemical Physics in Leningrad, Iulii B. Khariton and Iakov B. Zel'dovich (who, with Andrei Sakharov, were to lead the postwar Soviet hydrogen bomb program) carried out and published pioneering studies of the fission chain reaction in natural uranium. In 1941, after secrecy had been imposed, Khariton and Zel'dovich, like Frisch and Peierls before them, correctly estimated the critical mass of uranium 235. (Surprisingly, the German physicists never did this calculation during the war, even though it was the reputation of German science that had convinced Churchill and Roosevelt that they should pursue the bomb at any cost.)

Soviet physicists brought the military potential of fission to their government's attention. But the U.S.S.R. had to contend with Hitler's invasion after June 1941 and could not devote major resources to the uncertain goal of a nuclear bomb until Germany's defeat.

Soviet espionage was a great success, however. In September 1941 Soviet agents in London acquired a report to the British cabinet based on the Frisch-Peierls memo. Apparently divulged by a senior British civil servant, this information was invaluable because it also revealed that the cabinet had decided to collaborate on bomb research with the U.S. A few months later the Soviets' crucial source became the physicist Klaus Fuchs, a communist who had fled Nazi Germany and settled in Britain, where he was working with Peierls. Fuchs volunteered a stream of technical reports first from England and then from the U.S., when he was a member of the British team participating in the Manhattan Project. Among these documents were descriptions of isotope separation and nuclear reactors. Most significant, in June 1945 detailed specifications for the implosion bomb used later at Nagasaki were transmitted to the Soviets.

Thus, when President Harry S Truman, two weeks before the bombing of Hiroshima, made his famous veiled hint to Stalin that "we have a new weapon of unusual destructive force," he was wrong to conclude that Stalin did not understand his allusion. On the contrary, Stalin quickly ordered an acceleration of the Soviet effort to build nuclear weapons. He also appointed Lavrentii P. Beria, his notorious chief of secret police, to be overseer of this project.

Beria instilled fear in all who came into his orbit. He trusted no one—neither his scientists nor his spies. Kurchatov, who was head of scientific research for the program from its inception, was the only scientist fully aware of the nuclear espionage. In 1945, however, Beria hired a small group of scientists to edit the ever growing espionage files. This team was led by Iakov Terletskii, a young physicist of a caliber well below that of Kurchatov's team but who would be fully under Beria's control.

A Visitor from Moscow

In November 1945 Terletskii, at the direction of Beria, went to Copenhagen to meet Bohr, who had recently returned home after participating in the Manhattan Project. According to the memo to Stalin, Beria devised the mission in the hope that Bohr, a known advocate of international cooperation, might say something useful about Western nuclear research. Undoubtedly, Beria also thought that if Bohr did divulge any secrets, whether intentionally or not, the information might verify other Soviet intelligence and research and perhaps make Bohr vulnerable to blackmail.

On November 2 a communist member of the Danish parliament asked Bohr to meet secretly with Terletskii, who was carrying a letter to Bohr from his old friend Peter L. Kapitsa. Kapitsa, the outstanding Soviet experimental physicist of his generation, was, in effect, being held captive by the Soviet government.

According to Bohr's son Aage, who also worked at Los Alamos and is now a Nobel Prize–winning physicist himself at the University of Copenhagen, Niels Bohr was aghast. He called the approach "a regrettable mistake" and insisted to the Danish parliament member that any conversation would have to be open and that he could discuss only publicly available information. Bohr notified Western authorities, and the British reported the suspicious approach to Groves before the encounter. The authorities expressed concern that an attempt to kidnap Bohr might be afoot.

Nevertheless, a meeting was held on November 14, with the Danish government providing for Bohr's protection. Bohr's son Ernest, then age 21, was stationed in the next room and armed with a pistol. Western authorities surely were curious to see what questions the Soviet agent would ask. Newly declassified papers from the British cabinet office show that Bohr kept in contact with the British embassy in the weeks before the talks. A cable to the top official in the British Foreign Office reported that Bohr paid a long visit to the embassy on the very day he first met Terletskii.

Two independent eyewitness accounts of the Bohr-Terletskii exchange exist, and they agree remarkably well. Terletskii's memoirs appeared in Russia shortly after his death in 1993. Aage Bohr, who was 23 at the time, remained in the room for the entire meeting at his father's insistence; he recalls clearly how the events unfolded.

Communication between the parties was poor. A Soviet expert in international trade who knew no physics acted as translator. As two of the authors (Bethe and Gottfried) can testify, Bohr tended to lower his soft voice to emphasize a crucial point and was not easy to understand in English or German even when one knew the subject well, which Terletskii did not. In his memoirs, Terletskii admits that he had only a flimsy grasp of what Bohr said and that notes were not taken until he and the interpreter tried to reconstruct the conversation afterward.

According to both accounts, Bohr spent much of the appointment speaking at length about Kapitsa and Lev D. Landau. Kapitsa had worked in Cambridge with Ernest Rutherford until 1934, when Stalin would not let him return after a visit to his relatives in the U.S.S.R. Landau, a brilliant theorist, had once worked at Bohr's Institute of Theoretical Physics in Copenhagen and had been imprisoned for one year by Stalin before the war. Bohr sought to convey his support for them to the top of the Soviet government through its emissary, Terletskii.

Terletskii also relates that—much to his dismay—when the chance to ask his questions finally came, Aage was still present. This is confirmed by Aage, who recalls "a somewhat desperate Terletskii who, after a long conversation about Kapitsa ... was very anxious to present a number of questions. They came rather fast and through the interpreter, and we did not grasp the detailed content of the questions."

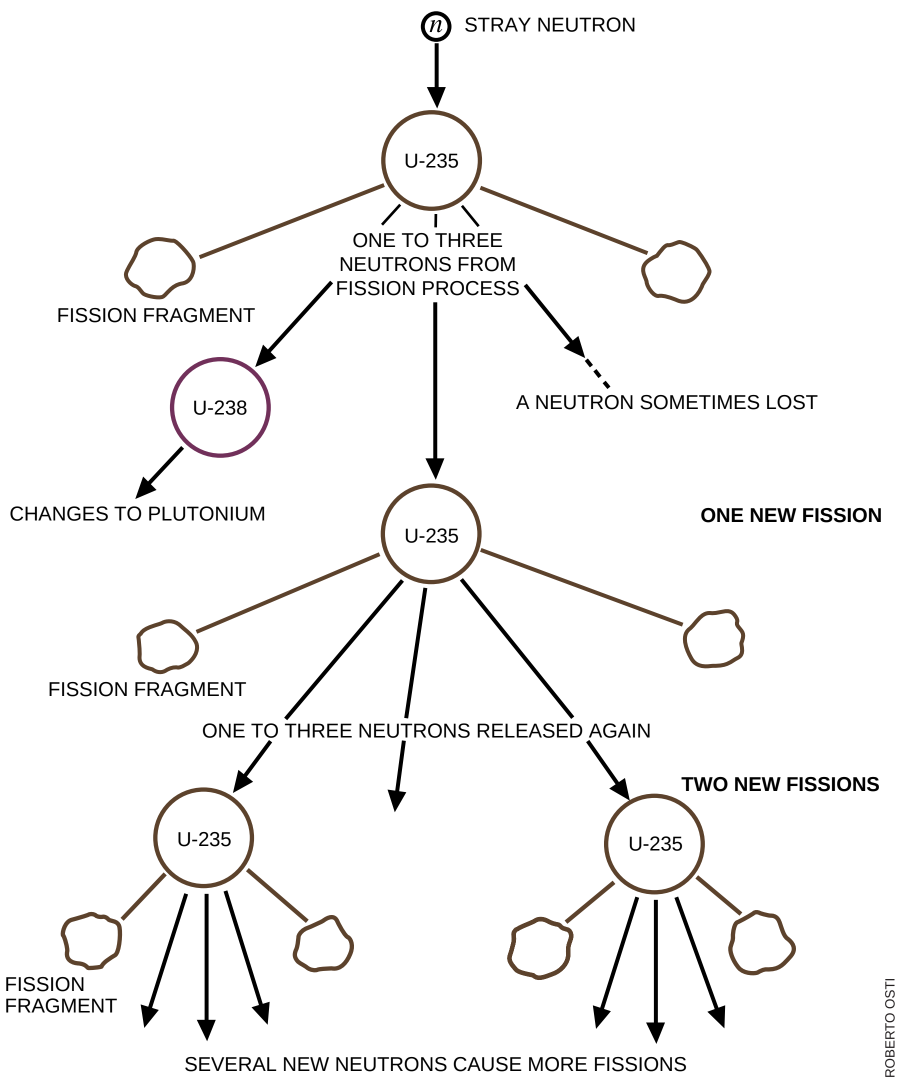
At a brief second visit on November 16, Bohr handed Terletskii the Smyth report. By that time 100,000 copies had been sold, and the Soviet government's translation was nearly finished.

The Transcript

When Beria transmitted the memo to Stalin on Terletskii's mission, he attached a purportedly verbatim transcript of Terletskii's questions and Bohr's answers [*see box on preceding three pages*]. We now turn to an analysis of this document and leave for later the question of whether it should be taken at face value.

More than half of Terletskii's 22 questions (which were prepared by Kurchatov's team) were about nuclear reactors and isotope separation techniques, and several were on fission itself and bomb detonation mechanisms. One query asked whether defense against nuclear weapons was possible.

The transcript confirms Terletskii's recollection that Bohr's "answers were very general. Each time, he said that in Los Alamos he was not told of the details ... [and] that he had never visited



CHAIN REACTION triggered by fission of a uranium nucleus in a nuclear reactor made with uranium ore was described in the Smyth report, the government's public account of the Manhattan Project. As shown in this illustration adapted from the document, a stray neutron (*top*)—from fission of a uranium 235 nucleus that is not shown—is captured by another uranium 235 nucleus, which breaks into two fragments and several neutrons. Some of these neutrons are, in turn, captured by other uranium 235 nuclei that undergo fission themselves (*center*), releasing neutrons that induce additional fission reactions (*bottom*). If, instead, uranium 238 absorbs neutrons, it can subsequently transform into plutonium (*left*). When Bohr gave the Smyth report to Terletskii, the Soviets had almost finished their translation.

the East Coast laboratories" where the Manhattan Project's isotope separation facilities were located. Indeed, on nuclear reactors and isotope separation techniques, Bohr's answers provided only information known before the war, even though subsequent developments were covered in detail in the Smyth report. (Close to half the text was devoted to these topics.) When asked about fission itself, Bohr usually referred to the prewar literature, especially to his famous paper with Wheeler.

AP WIDE WORLD PHOTOS

NUCLEAR BLAST was produced during experiments conducted by the U.S. in July 1946. The U.S.S.R. conducted its first test of nuclear weapons in August 1949.

Bohr supposedly gave an incorrect response, however, in stating that no U.S. reactor employed heavy water as the "moderator" for slowing down neutrons. In fact, the reactor at Argonne Laboratory near Chicago did use heavy water, as stated in the Smyth report. Whether this minor error was Bohr's or Terletskii's we do not know.

Terletskii's questions relating to bomb design merit close scrutiny. When asked about the number of neutrons emitted by various isotopes of uranium and plutonium, Bohr simply replied, "More than two." Terletskii then asked, "Can you provide a more accurate figure?" and Bohr responded, "No, I cannot.... The exact number is not of great significance...." But the amount of uranium 235 or plutonium needed for a bomb actually does depend quite sensitively on the number of neutrons per fission.

The second question seemingly dealing with bomb design concerned spontaneous fission, which, as we have seen, is crucial to consider when devising the plutonium weapon. It is unclear, however, whether Terletskii's inquiry referred to reactor or bomb design, and Bohr's answer was correct only for reactor design, in which spontaneous fission is unimportant. The last question concerned how a chain reaction evolves after the chemical explosion compresses the fissionable material. Bohr's answer was a model of opaque irrelevance.

Because Bohr had not been involved with isotope separation or reactor design during the Manhattan Project, he was in no position to give detailed responses to questions on these topics. During his time at Los Alamos, however, he was briefed by Richard Feynman about chain reactions in bombs and participated in devising the implosion design. Had Bohr wanted to convey vital information, these questions offered the opportunity to do so. Instead he gave the Soviets either partial or incomprehensible answers that divulged nothing.

In his response to whether defense against the bomb was feasible, Bohr went on at length about the need for international control, a position he consistently took in public. But his supposed answer also contained the absurdity that "the great Oppenheimer has retired in protest and stopped his work on the project." While it is true that J. Robert Oppenheimer had left Los Alamos to return to the University of California at Berkeley, he became the Truman administration's most influential technical adviser on nuclear weapons policy.

Bohr's purported statement about Oppenheimer, quite apart from the skepticism with which any document signed by Beria should be treated, convinces us that Bohr's replies were "massaged" as they made their way toward Stalin's desk. Indeed, in his memoirs, Terletskii recounts that the secret police colonel in charge of the mission to Copenhagen coached him on how best to relay what he had heard. Despite these efforts, Beria was disappointed with the results and "became uncontrolled and interrupted with obscenities directed at Bohr and Americans" when Terletskii reported on his conversation with Bohr.

Nevertheless, in the memo to Stalin, Beria implied that Terletskii's mission had been a success—which is hardly surprising as the operation had been Beria's brainchild. We find it interesting that the memo never mentions the Smyth report and apparently was not checked against it. Probably Beria and his people suspected that so informative a document was a masterpiece of disinformation. Therefore, we cannot dismiss Terletskii's claim that Bohr, in conveying the Smyth report, gave the document a credibility that it would not have had otherwise.

Regardless of how much Bohr's words were altered by Terletskii and others working for Beria, the memo to Stalin offers the most authoritative version of the meeting available. The purpose of any alterations would have been to exaggerate the importance of the information provided by Bohr. Yet the transcript has Bohr saying nothing about matters of technical or military significance not covered in the Smyth report. Thus, the allegation that Bohr shared nuclear secrets with the Soviets is refuted by Beria's own account of the encounter between his agent and Bohr.

FURTHER READING

THE MAKING OF THE ATOMIC BOMB. Richard Rhodes. Simon & Schuster, 1986.

ATOMIC ENERGY FOR MILITARY PURPOSES: THE OFFICIAL REPORT ON THE DEVELOPMENT OF THE ATOMIC BOMB UNDER THE AUSPICES OF THE UNITED STATES GOVERNMENT, 1940–1945. Henry D. Smyth. Reissued by Stanford University Press, 1989.

SPECIAL TASKS: THE MEMOIRS OF AN UNWANTED WITNESS: A SOVIET SPY-MASTER. Pavel and Anatoli Sudoplatov, with Jerrold L. and Leona P. Schecter. Little, Brown and Company, 1994.

STALIN AND THE BOMB. David Holloway. Yale University Press, 1994.

WERE THE ATOMIC SCIENTISTS SPIES? Thomas Powers in *New York Review of Books,* Vol. 41, No. 11, pages 10–17; June 9, 1994.

What Did Heisenberg Tell Bohr about the Bomb?

In 1941 Werner Heisenberg and Niels Bohr met privately in Copenhagen. Almost two years later at Los Alamos, Bohr showed a sketch of what he believed was Heisenberg's design for a nuclear weapon

by Jeremy Bernstein

In September 1943 Niels Bohr learned that the gestapo in Copenhagen intended to arrest him. A few weeks later, on the 29th, he, his wife and several others hoping to escape from Denmark crawled in complete darkness to a beach outside Carlsberg. There they boarded a boat and crossed the Øresund in secret to Sweden. On October 6 the British flew Bohr alone from Sweden to Scotland. Later that same day he traveled to London and in the evening met with Sir John Anderson, the physical chemist in charge of the nascent British atomic bomb project. Anderson gave the Danish physicist a briefing on the Anglo-American program. According to Bohr's son Aage, who followed his father to England a week later and was his assistant throughout the war, Bohr was deeply surprised—shocked may be a better description—by how far the Anglo-American program had already progressed.

Bohr's alarm very likely had two sources. First, during the 1930s, when nuclear physics was developing, Bohr had said on several occasions that he thought any practical use of nuclear energy was all but impossible. That view was reinforced in the spring of 1939, when he realized an important detail concerning the fission of uranium. In December 1938 the German physical chemists Otto Hahn and Fritz Strassmann had discovered that uranium could be fissioned if it was bombarded with neutrons. (Hahn's former assistant Lise Meitner and her nephew Otto Frisch conjectured that the uranium nucleus had actually been split in the experiments and so coined the name "fission" for the process.) The experiments used natural uranium, 99 percent of which is in the isotope uranium 238. About seven tenths of a percent is in the isotope uranium 235, whose nucleus contains three fewer neutrons.

Chemically, the isotopes are indistinguishable. What Bohr realized was that because of their structural differences, only the very rare isotope uranium 235 had fissioned in the Hahn-Strassmann experiments. He concluded, then, that making a nuclear weapon would be almost impossible because it would require separating these isotopes—a daunting task. In December 1939 he said in a lecture, "With present technical means it is, however, impossible to purify the rare uranium isotope in sufficient quantity to realize the chain reaction." One can therefore well understand why Bohr was shocked to learn four years later that that was just what the Allies intended to do.

The second reason for Bohr's alarm can be traced back to a meeting he had had with the German physicist Werner Heisenberg in mid-September 1941, almost two years before his escape to Britain. By 1941 the Germans had occupied Denmark for more than a year. During that period, they established a so-called German Cultural Institute in Copenhagen to generate German cultural propaganda. Among its activities, the institute organized scientific meetings. Heisenberg was one of several German scientists who came under its auspices to Copenhagen, in this case to a meeting of astronomers. He had known Bohr since 1922 and had spent a good deal of time at Bohr's institute in Copenhagen, where Bohr had acted as a kind of muse for the creation of quantum theory. Now Heisenberg had returned as a representative of a despised occupying power, touting the certainty of its victory, according to some accounts.

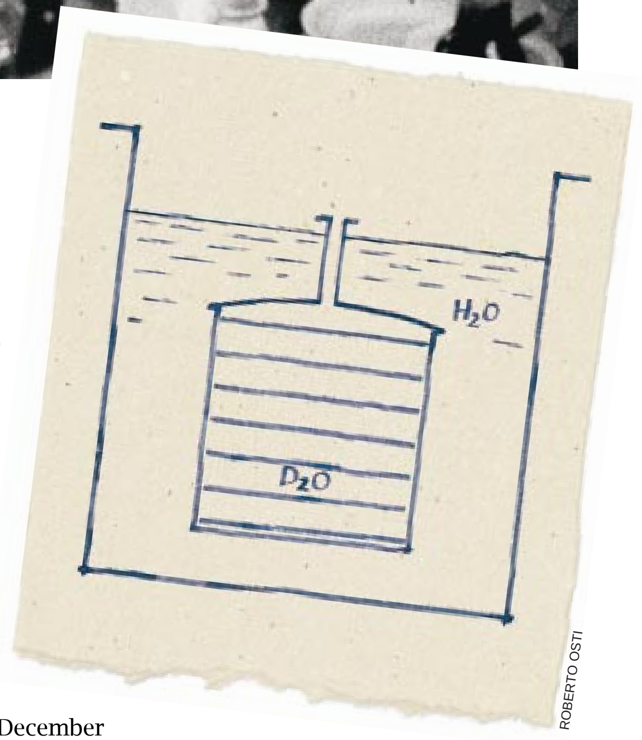
Heisenberg's Visit

Heisenberg spent a week in Copenhagen and visited Bohr's institute on several occasions. During one of these visits, he and Bohr talked privately. Neither man seems to have made any notes, so one cannot be entirely sure what was said. Also, Bohr was a poor listener, so the two may well have talked past each other. Nevertheless, Bohr came away from the discussion with the distinct impression that Heisenberg was working on nuclear weapons. As Aage Bohr later recalled, "Heisenberg brought up the question of the military applications of atomic energy. My father was very reticent and expressed his skepticism because of the great technical difficulties that had to be overcome, but he had the impression that Heisenberg thought that the new possibilities could decide the outcome of the war if the war dragged on." Now, two years later, Bohr was learning for the first time of the Allied nuclear weapons program. What had the Germans done during those two years? No wonder Bohr was alarmed.

It would be fascinating to know in detail what was meant by "new possibilities," but one can make an educated guess. By the mid-1940s physicists on both sides of the conflict realized that aside from fissioning uranium, there was an entirely separate route to mak-

JEREMY BERNSTEIN is professor of physics at the Stevens Institute of Technology and an adjunct professor at the Rockefeller University. He also serves as a vice president of the board of trustees of the Aspen Center for Physics. He has written some 50 technical papers, 12 books and numerous magazine articles. He has worked as a staff writer at the *New Yorker* magazine, taught nonfiction writing at Princeton University and won several science writing awards. He is a fellow of the American Physical Society, a Benjamin Franklin Fellow of the Royal Society of the Arts and a member of the French and American Alpine Clubs.

PAUL EHRENFEST, JR. *AIP Emilio Segrè Visual Archives*



ROBERTO OSTI

WERNER HEISENBERG AND NIELS BOHR, shown above in 1934, had a curious meeting seven years later. Some speculate that the two physicists—then from opposite sides of the war—discussed nuclear weapons. When Bohr arrived at Los Alamos Laboratory in late 1943, he reportedly had a sketch of what he thought to be Heisenberg's plan to build a bomb. If so, the drawing most likely resembled the reactor design shown at the right.

ing a nuclear weapon—the use of what later came to be known as plutonium. That element is somewhat heavier than uranium and has a different chemistry, but given its nuclear structure, it is at least as fissionable. Unlike uranium, though, plutonium does not exist naturally and must be manufactured in a nuclear reactor by bombarding the reactor's uranium fuel rods with neutrons. Once made, the plutonium can be separated from its uranium matrix by chemical means.

From the moment this process was understood, any reactor became, in a certain sense, a component of a nuclear weapon. There is no doubt whatsoever that Heisenberg knew this fact well when he visited Bohr. He even gave lectures, whose texts have been preserved, describing such a possibility to highly placed German officials. Is this what he was trying to tell Bohr and, if so, why? There was such a lack of agreement between the two men as to what exactly was said that we will probably never know for sure.

As a corollary to this larger puzzle there is a smaller one. There is evidence that during the course of the Copenhagen meeting, Heisenberg gave Bohr a drawing. It is not clear whether Heisenberg made the drawing at the meeting or beforehand. Being familiar with how theoretical physicists communicate, I would imagine he drew the sketch on the spot to help convey an idea. In any case, under circumstances I will shortly describe, this drawing, or a replica, found its way to Los Alamos Laboratory in December 1943, where it created a considerable stir: it appeared to contain direct information about how the Germans were planning to make nuclear weapons. Before I describe how the drawing got to Los Alamos, let me tell how I learned of its existence. There is a relation.

The Mysterious Sketch

Beginning in November 1977, I conducted a series of interviews with the physicist Hans Bethe. Those sessions lasted on and off for two years and resulted in a three-part profile for the *New Yorker* magazine and a subsequent book.

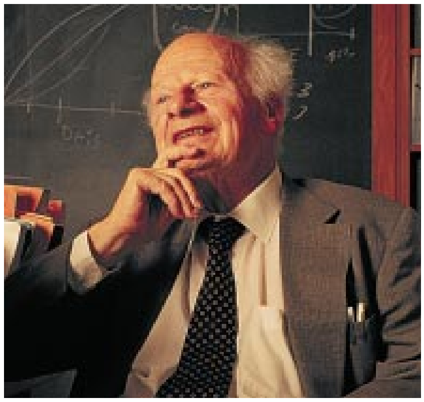
The interviews, which I taped, followed the chronology of Bethe's life. Bethe, who was born in Strasbourg in 1906, emigrated to the U.S. in 1935 and has been at Cornell University ever since. He became an American citizen in 1941, by which time, as he recalled, he was "desperate to do something—to make some contribution to the war effort." He, like Bohr, was at first certain that nuclear weapons were entirely impractical and went to work on the development of radar at the Massachusetts Institute of Technology.

In the summer of 1942 J. Robert Oppenheimer convened a study group at the University of California at Berkeley to investigate nuclear weapons. By this time Bethe was acknowledged as one of the leading nuclear theorists in the world, so Oppenheimer naturally asked him to participate. On the way to California by train, Bethe stopped in Chicago to pick up Edward Teller. There Bethe got the chance to see Enrico Fermi's developing nuclear reactor and, in his words, "became convinced that the atomic bomb project was real, and that it would probably work." He spent that summer working on the theory of nuclear weapons and in April 1943 went to Los Alamos, which had just opened as a laboratory. Eventually he became head of its theory division.

Now to the drawing. On November 29, 1943, Bohr and his son Aage sailed from Glasgow on the *Aquitania* for New York City. They arrived on December 6. Bohr was assigned the code name of Nicholas Baker, and Aage became James Baker; they were also given bodyguards. On December 28, after having had high-level meetings in Washington, D.C., with many officials—including Major General Leslie R. Groves, the commanding officer in charge of the Manhattan Project—Bohr departed for Los Alamos. On the 31st, presumably just after arriving at the laboratory, he met with a select group of physicists. The principal purpose of this meeting was for Bohr to tell the attending physicists what he knew about the German effort to make a nuclear weapon—in particular what he had learned from Heisenberg.

During one of my interviews with Bethe, he described this meeting, though not in any detail, and told me about the drawing. This is what he said to me (I have it on my tapes): "Heisenberg gave Bohr a drawing. This drawing was

"I am quite positive there was a drawing.... [It] made a great impression on me."
—Bethe



ROBERT PROCHNOW

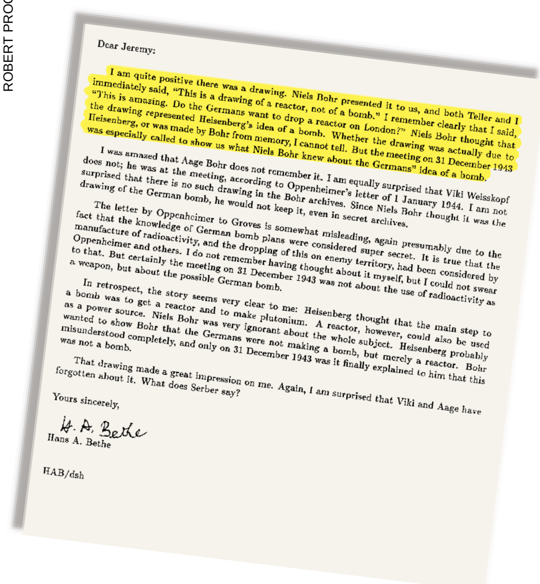
LOS ALAMOS NATIONAL LABORATORY

HANS BETHE clearly recalls seeing a design credited to Heisenberg at Los Alamos in 1943. He wrote to the author about this recollection last year (*right*). At the left, Bethe sits in front of Enrico Fermi, who designed the first working nuclear reactor in the U.S., in Chicago. Fermi's success helped to convince Bethe that nuclear weapons could be made.

transmitted by Bohr later on to us at Los Alamos. This drawing was clearly the drawing of a reactor. But our conclusion was, when seeing it, these Germans are totally crazy. Do they want to throw a reactor down on London?" Only after the war did the Los Alamos scientists learn that the Germans knew perfectly well, at least in principle, what to do with a reactor—use it to make plutonium. But Bohr was concerned that one could actually use this reactor as some sort of weapon.

As far as I know, until I described this matter in the *New Yorker,* no one had ever mentioned such a drawing in print. In fact, my article on Bethe was frequently cited as the source for this odd sidelight on the Bohr-Heisenberg relationship. Hence, I found myself as a kind of a footnote to a footnote to history. My authority was shaken, though, at the start of 1994, during one of my periodic visits to the Rockefeller University in New York City, where I am an adjunct professor. Abraham Pais, a biographer of both Einstein and Bohr and a professor of physics emeritus at the university, called me into his office. I have known Pais for 40 years but had not seen him in a while. This visit, then, was his first opportunity to tell me about a call he had received several months earlier.

It was from Thomas Powers, who at that time was writing *Heisenberg's War.* Powers had learned about the drawing from my book on Bethe. He was struck by the fact that at first glance it seemed as if Heisenberg had given to Bohr, in the middle of a war, a drawing of a highly classified German military project. That was such an extraordinary thing for Heisenberg to have done, if he did do it, that Powers wanted to check the matter out. He therefore got in touch with Aage Bohr in Copenhagen (his father had died in 1962). In a letter dated November 16, 1989, Aage Bohr wrote, "Heisenberg certainly drew no sketch of a reactor during his visit in 1941. The operation of a reactor was not discussed at all."



Dear Jeremy:

I am quite positive there was a drawing. Niels Bohr presented it to us, and both Teller and I immediately said, "This is a drawing of a reactor, not of a bomb." I remember clearly that I said, "This is amazing. Do the Germans want to drop a reactor on London?" Niels Bohr thought that the drawing represented Heisenberg's idea of a bomb. Whether the drawing was actually due to Heisenberg, or was made by Bohr from memory, I cannot tell. But the meeting on 31 December 1943 was especially called to show us what Niels Bohr knew about the Germans" idea of a bomb.

I was amazed that Aage Bohr does not remember it. I am equally surprised that Viki Weisskopf does not; he was at the meeting, according to Oppenheimer's letter of 1 January 1944. I am not surprised that there is no such drawing in the Bohr archives. Since Niels Bohr thought it was the drawing of the German bomb, he would not keep it, even in secret archives.

The letter by Oppenheimer to Groves is somewhat misleading, again presumably due to the fact that the knowledge of German bomb plans were considered super secret. It is true that the manufacture of radioactivity, and the dropping of this on enemy territory, had been considered by Oppenheimer and others. I do not remember having thought about it myself, but I could not swear to that. But certainly the meeting on 31 December 1943 was not about the use of radioactivity as a weapon, but about the possible German bomb.

In retrospect, the story seems very clear to me: Heisenberg thought that the main step to a bomb was to get a reactor and to make plutonium. A reactor, however, could also be used as a power source. Niels Bohr was very ignorant about the whole subject. Heisenberg probably wanted to show Bohr that the Germans were not making a bomb, but merely a reactor. Bohr misunderstood completely, and only on 31 December 1943 was it finally explained to him that this was not a bomb.

That drawing made a great impression on me. Again, I am surprised that Viki and Aage have forgotten about it. What does Serber say?

Yours sincerely,

H. A. Bethe
Hans A. Bethe

HAB/dsh

Stunned, Powers next contacted Bethe, who repeated to him exactly what he had told me 10 years earlier. In a quandary, Powers had called Pais, and now Pais was asking me. But Pais had done his own investigation. He had spoken with Aage Bohr, who once again insisted that there had never been any such drawing. Pais had also checked the archives in Copenhagen where all Bohr's private papers and journals are stored. Nowhere, he told me, did he find any mention of this drawing. Now it was my turn to be stunned. It is one thing to be a footnote to a footnote to history, but it is quite another to be a footnote to a footnote to incorrect history.

I promised Pais I would look into the matter myself, although, in truth, when I left his office I did not have the foggiest idea of how I would go about it. Obviously, contacting Bethe again would not get me much further. Nothing could be more direct than what he had told me and repeated to Powers. I would need witnesses independent of Bethe and Aage Bohr. That much was clear. But who? Oppenheimer was dead. Niels Bohr was dead. Groves was dead. Who else could have seen that drawing?

The Investigation

I began, in fact, with less information than I have so far given the reader. All Bethe had told me was that Bohr had "transmitted" a drawing to Los Alamos. He had not related any specific details about the December 31 meeting, so initially I had no idea who might have been there. Indeed, I did not even have the specific date. All that I learned subsequently. But I did know physicists who were at Los Alamos at the time and who might have seen or heard about the drawing. Two came to mind. One was

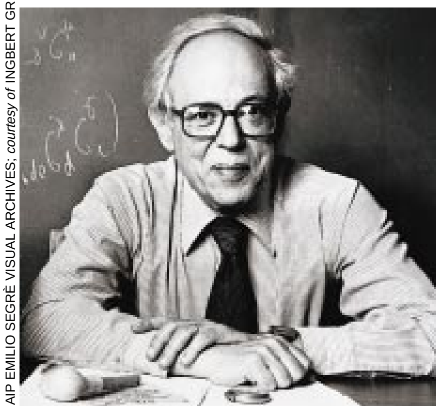
AP/WIDE WORLD PHOTOS

Victor Weisskopf, an old friend, who had been close to Oppenheimer.

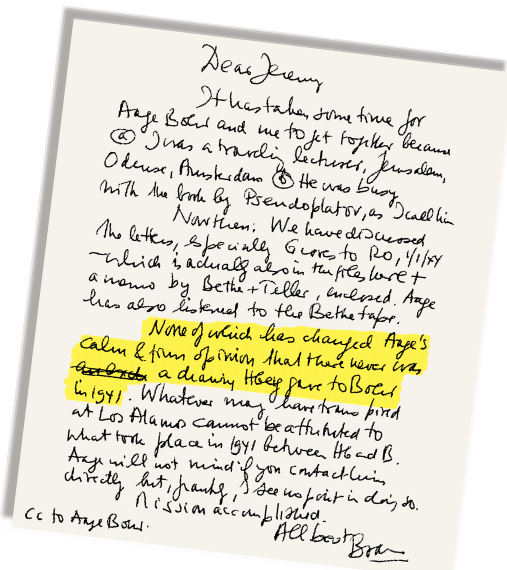
The other was Rudolf Peierls. Peierls and Otto Frisch had in March 1940 made the first correct calculation—in principle—to determine the amount of uranium 235, or the critical mass, needed to make a bomb. (The fact that this mass turned out to be pounds rather than tons is what really prompted the Allied effort.) Peierls, along with Frisch, was at Los Alamos as of early 1944. I have also known Peierls for many years and have frequently discussed with him the history of nuclear weapons. So it was quite natural for me to write him as well. This I did in early February, and soon after, both men answered.

Peierls replied that he had never seen the "famous sketch" yet did not think that either Bethe or Aage Bohr had deliberately lied. He proposed that perhaps Niels Bohr had kept knowledge of the sensitive document from his family or

"[Nothing] has changed Aage's calm and firm opinion that there never was a drawing."
—Pais



AIP EMILIO SEGRÈ VISUAL ARCHIVES; *courtesy of* INGBERT GRÜTTNER



Dear Jeremy
It has taken some time for Aage Bohr and me to get together because (a) I was a traveling lecturer, Jerusalem, Odense, Amsterdam (b) He was busy with the book by Pseudoplator, as I call him.
Now then: We have discussed the letters, specially Groves to RO, 1/1/44 —which is actually also in the files here + a memo by Bethe + Teller, enclosed. Aage has also listened to the Bethe tape.
None of which has changed Aage's calm & firm opinion that there never was [illegible] a drawing Heisenberg gave to Bohr in 1941. Whatever may have transpired at Los Alamos cannot be attributed to what took place in 1941 between Hb and B.
Aage will not mind if you contact him directly but, frankly, I see no point in doing so.
Mission accomplished.
All best
Bram
cc to Aage Bohr.

ABRAHAM PAIS asked Aage Bohr, one of Bohr's sons, whether Heisenberg might have given his father a drawing in 1941. Aage Bohr (*left*) knew his father's activities well, having served as his assistant during the war. Pais wrote the author last summer (*right*) that Aage Bohr was certain Heisenberg had not made a sketch for his father.

that perhaps Heisenberg had only shown the sketch to Bohr, who might then have redrawn it. He suggested I contact Bethe about this possibility. Weisskopf also wrote proposing I contact Bethe once more, because he, too, had never seen or heard about the drawing.

Neither of these letters was what I had hoped to receive. Clearly, I had to write Bethe to tell him what I had learned and to see if he could shed any further light on the situation. But then I had an inspiration. I would call Robert Serber. Serber, a professor of physics emeritus at Columbia University who lives in New York City, is also an old friend. After receiving his Ph.D. in 1934 from the University of Wisconsin, he had won one of five National Research Council Fellowships in physics and chose to go work with Oppenheimer at Berkeley. During the next few years, he had become very close to Oppenheimer.

After a brief interlude at the University of Illinois from 1938 until 1942, Serber returned to Berkeley to work on the bomb with Oppenheimer. He was there in the summer of 1942 when Bethe and Teller arrived. By March 1943 he had moved, with the first batch of scientists, to Los Alamos. One of his early tasks was to give a series of introductory lectures on bomb physics to the arriving scientists. These lessons were collected into what came to be called *The Los Alamos Primer,* declassified in 1965 and first published in its entirety in 1992. If anyone knew about the drawing, it would be Serber because he was in constant contact with Oppenheimer throughout this period.

I called Serber, and immediately I knew I had struck a gold mine. Not only did he remember the drawing vividly, but he also remembered the precise circumstances under which he had seen it. He had been summoned to Oppenheimer's office on December 31, where a meeting was already in progress. Oppenheimer showed him a drawing with no explanation and asked him to identify it. This was the kind of intellectual game Oppenheimer liked to play. Serber looked at it and said it was clearly the drawing of a reactor. Oppenheimer replied that in fact it was a drawing of Heisenberg's reactor and had been given to the assembled group by Bohr. Bohr, who was, as Serber recalled, standing next to Oppenheimer, did not disagree.

That is what Serber told me. But he also said he had some written material related to this meeting. A few days later copies of two documents arrived: a letter from Oppenheimer to General Groves sent the day after the meeting and a two-page memorandum written by Bethe and Teller on the explosive possibilities of the reactor. Unfortunately, although these documents were very suggestive, they did not, at least when I first read them, settle the issue completely. The Bethe-Teller memorandum did hold significant clues, but I will return to them later. Oppenheimer's letter made no mention of the drawing or of Heisenberg or of the Germans. But the last sentence clearly implied that Bohr had spoken to Groves in Washington about these matters. Perhaps something in Groves's own archives might prove enlightening.

Meanwhile I had at last written to Bethe, and on March 2, I received his answer. It begins, "I am quite positive there was a drawing. Niels Bohr presented it to us, and both Teller and I immediately said, 'This is a drawing of a reactor, not of a bomb.'... Whether the drawing was actually due Heisenberg, or was made by Bohr from memory, I cannot tell. But the meeting on 31 December 1943 was especially called to show us what Niels Bohr knew about the Germans' idea of a bomb."

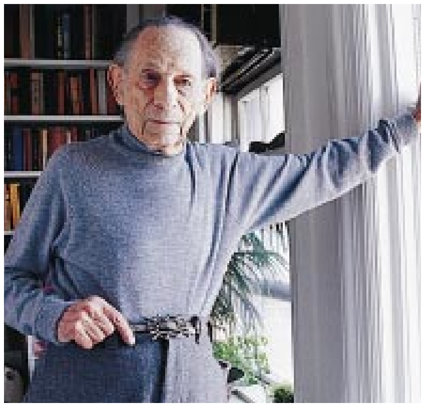
Bethe offered a theory to explain the mystery: "Heisenberg thought that the main step to a bomb was to get a reactor and to make plutonium. A reactor,

UPI/BETTMANN

however, could also be used as a power source. Niels Bohr was very ignorant about the whole subject. Heisenberg probably wanted to show Bohr that the Germans were not making a bomb but merely a reactor. Bohr misunderstood completely, and only on 31 December 1943 was it finally explained to him that this was not a bomb. That drawing made a great impression on me. Again, I am surprised that Viki [Weisskopf] and Aage have forgotten about it. What does Serber say?"

I was able to write Bethe and tell him what Serber had said. I also wrote Teller to ask for his recollections of the meeting. I was not sure I would get an answer and never have. But I had also written again to Weisskopf, sending him copies of the memorandums from Serber. On February 23, I received a typically gracious Weisskopf letter, acknowledging that he had indeed seen the sketch but later forgotten about it.

"...the [device] suggested...by Baker would be a quite useless military weapon." –Oppenheimer



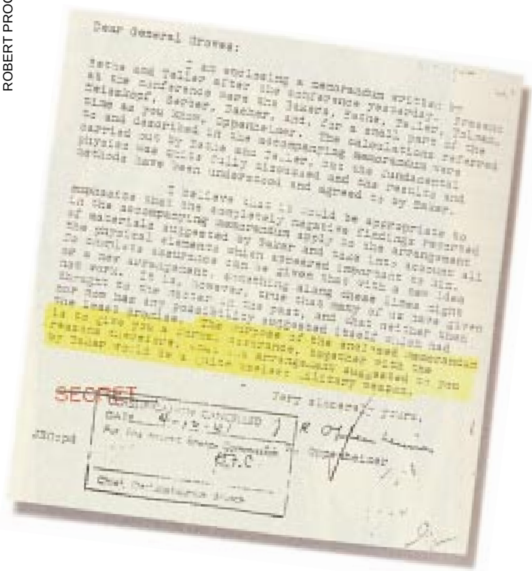
ROBERT PROCHNOW

ROBERT SERBER not only remembered seeing a drawing on December 31, 1943, but had kept several revealing documents. He sent the author a letter (*shown at right*) from J. Robert Oppenheimer to Major General Leslie R. Groves (*both shown at left*). In it, Oppenheimer assured Groves that the "arrangement of materials" Baker (Bohr's code name) had described could not be used as a weapon.

I now had, I thought, enough material to return to Pais. I played for him my Bethe tape and gave him copies of all the documents. He was about to return to Copenhagen, where he spends about half the year with his Danish wife. He promised me that he would speak to Aage Bohr at an opportune moment. That happened late in June. By the 30th Pais had written to tell me what had happened. He and Aage Bohr had met, discussed the letters and reviewed the tapes. Still, Aage Bohr felt certain that Heisenberg never gave any such drawing to his father. So I wrote to Aage Bohr directly. In February of this year his assistant, Finn Aaserud, wrote, "Aage Bohr maintains that it is entirely impossible that Bohr brought with him to the U.S. a drawing from the 1941 meeting with Heisenberg and indeed that the discussion at Los Alamos you refer to had anything to do with the 1941 encounter at all."

Where does this leave us? I have asked myself this question many times since receiving Pais's letter last June. I was at a loss until recently, when I took another look at the memorandum that Bethe and Teller prepared for Oppenheimer and Bohr and eventually for Groves. It suddenly struck me that in the first sentence of the second paragraph of this report Heisenberg's imprint stands out like a sore thumb. It reads, "The proposed pile [reactor] consists of uranium sheets immersed into heavy water." In other words, Bethe and Teller were not considering any old reactor design but rather a very particular one that Bohr had described to them. This design is actually the faulty reactor Heisenberg invented in late 1939 and early 1940, which he clung to until nearly the end of the war!

It is almost unthinkable that in the few short weeks from when Bohr learned about the Allied project to when he arrived at Los Alamos he would have produced his own design possessing the same flaws as did Heisenberg's. He must have gotten this idea from Heisenberg,



Dear General Groves:

SECRET

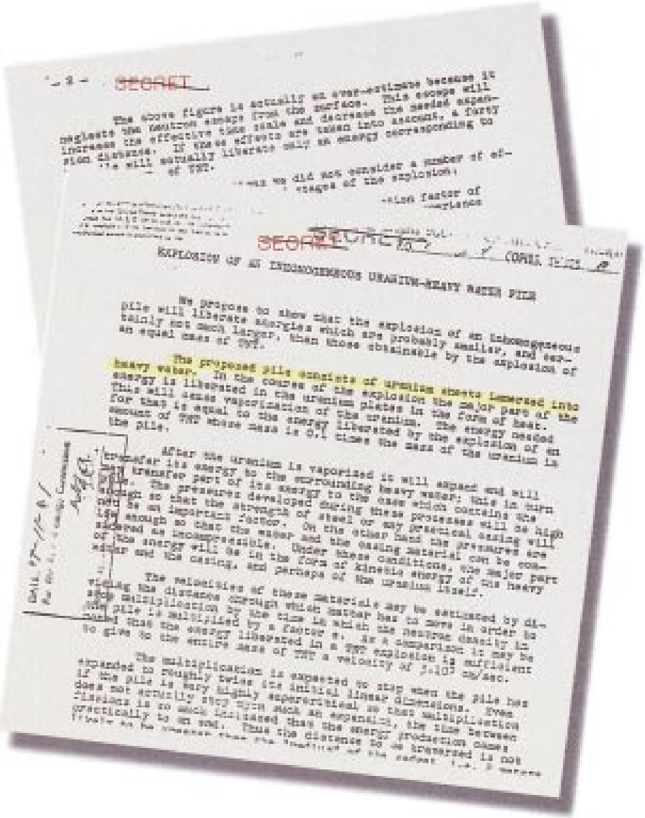
either verbally or in the form of a drawing. Where else could it have come from?

The Evidence

Let me explain. Any reactor requires fuel elements, the uranium, and what is known as a moderator, a device that slows the speed of neutrons hitting the fuel. Neutrons traveling near the speed of sound are vastly more effective in causing fission than are the rapidly moving neutrons produced by the fissioning itself. So the fuel elements in a reactor are embedded in the moderator. But a designer must carefully choose from which material the moderator should be made and also how the fuel elements should be placed in it. The latter involves both art and science.

The trick is that the uranium itself can absorb neutrons without producing fission. This absorption becomes stronger as the neutrons are slowing down. If the geometry of the fuel elements is not well thought out, the uranium will absorb so many neutrons that a self-sustaining chain reaction will never take place. In fact, the most efficient design involves separated lumps of uranium embedded in a lattice within the moderator. How big these lumps should be, and how they should be arranged, involves art. But the worst possible solution is placing the uranium in sheets, or layers.

To return to the matter at hand, note that Bethe and Teller wrote, "The proposed pile consists of uranium sheets." Heisenberg chose just such a design because it involved easier calculations than did other schemes. Then there is the question of the moderator. Bethe and Teller stated that the sheets were to be "immersed into heavy water." This specification, once explained, also has Heisenberg written all over it. The role of the moderator is, as I have mentioned, to slow down the fissioned neutrons. The best materials for this purpose are the lightest because a collision between a neutron and an object having a similar mass results in the greatest energy loss. If the neutron collides with a heavier object it will bounce off and change its direction but not its speed.


SECRET

EXPLOSION OF AN INHOMOGENEOUS URANIUM-HEAVY WATER PILE

We propose to show that the explosion of an inhomogeneous pile will liberate energies which are probably smaller, and certainly not much larger, than those obtainable by the explosion of an equal mass of TNT.

The proposed pile consists of uranium sheets immersed into heavy water. In the course of the explosion the major part of the energy is liberated in the uranium plates in the form of heat. This will cause vaporization of the uranium. The energy needed for that is equal to the energy liberated by the explosion of an amount of TNT whose mass is 0.1 times the mass of the uranium in the pile.

MEMORANDUM from Bethe and Edward Teller discusses the explosive possibilities of the nuclear reactor Bohr described to them when he arrived at Los Alamos. The document does not mention Heisenberg, yet the design it analyzes matches one that the German invented in 1939 and supported throughout the war.

If mass were the only consideration, the ideal moderator would be hydrogen, whose nucleus is a single proton having a mass sensibly the same as the neutron's. But, in reality, ordinary hydrogen fails as a moderator because it absorbs neutrons. In contrast, "heavy hydrogen," which has an extra neutron in its nucleus, does not absorb neutrons. Heavy hydrogen is found in "heavy water." But in seawater, say, this heavy water is only about one part in 5,000. So to use it as a moderator, it must be separated from ordinary water—an expensive and difficult process.

Carbon, on the other hand, is abundant and cheap, although somewhat less effective as a moderator. By late 1940 Heisenberg had concluded that only carbon and heavy hydrogen should be used as moderators. But in January 1941, Walther Bothe, who was the leading experimental nuclear physicist left in Germany, began working with graphite, the form of carbon commonly used in pencils. His experiments seemed to show that graphite absorbed neutrons too strongly to serve as an effective moderator. What Bothe did not realize was that unless the graphite is purified far beyond any ordinary industrial requirement, it will contain boron impurities. Boron soaks up neutrons like a sponge. One part boron in 500,000 of graphite can ruin that graphite as a moderator. All the same, because of Bothe's experiment, Heisenberg and other German physicists decided that heavy water was the only practical choice.

Needless to say, physicists who were responsible for the successful reactor program here made the same kinds of calculations. Like Heisenberg, they decided that a carbon reactor would need more natural uranium than a heavy-water reactor. Fermi and his colleague Leo Szilard had also done experiments on neutron absorption by carbon. But Szilard was a fanatic about the purity of the graphite, and so their graphite, unlike Bothe's, worked well as a moderator. Because carbon was so cheap compared with heavy water, they decided that it was the best moderator. Fermi's reactor, which first operated on December 2, 1942, had a lattice of uranium lumps embedded in carbon. All the German experimental reactors—none of which ever operated—used heavy-water moderators. Look again at the sentence in the Bethe-Teller memorandum: "The proposed pile consists of uranium sheets immersed into heavy water." It is as if someone had written "Made in Germany" on this design.

Putting everything together, there seems to be little doubt that Heisenberg attempted to describe a nuclear device to Bohr. It seems that this device was his version of a reactor. He may, or may not, have given Bohr a drawing, but Bohr clearly retained a visual memory of the design. Bohr, however, did not understand the difference between a reactor and a bomb at the time and assumed that Heisenberg was describing a bomb.

So Aage Bohr may be quite right when he says, as far as his father was concerned, there was no discussion of a reactor. He may also be right that Heisenberg never gave Bohr a drawing. None of the individuals I have contacted are sure that the drawing they saw was in Heisenberg's hand—only that it was a drawing of Heisenberg's reactor. This I think solves the puzzle, but it does not solve the mystery. What was the purpose of Heisenberg's visit in the first place? Those who admire Heisenberg have argued that it was to show Bohr that the Germans were working only on a "peaceful" reactor.

It also must be noted that when Heisenberg visited Bohr, he clearly knew that reactors could be used to manufacture plutonium and that plutonium could fuel a nuclear weapon. Why, then, did he visit Bohr? What message was he trying to convey? What was he trying to persuade Bohr to do, or not to do? What was he trying to learn? That is the real mystery, one we may never solve.

FURTHER READING

HANS BETHE: PROPHET OF ENERGY. Jeremy Bernstein. Basic Books, 1980.

NIELS BOHR'S TIMES: IN PHYSICS, PHILOSOPHY AND POLITY. Abraham Pais. Oxford University Press, 1991.

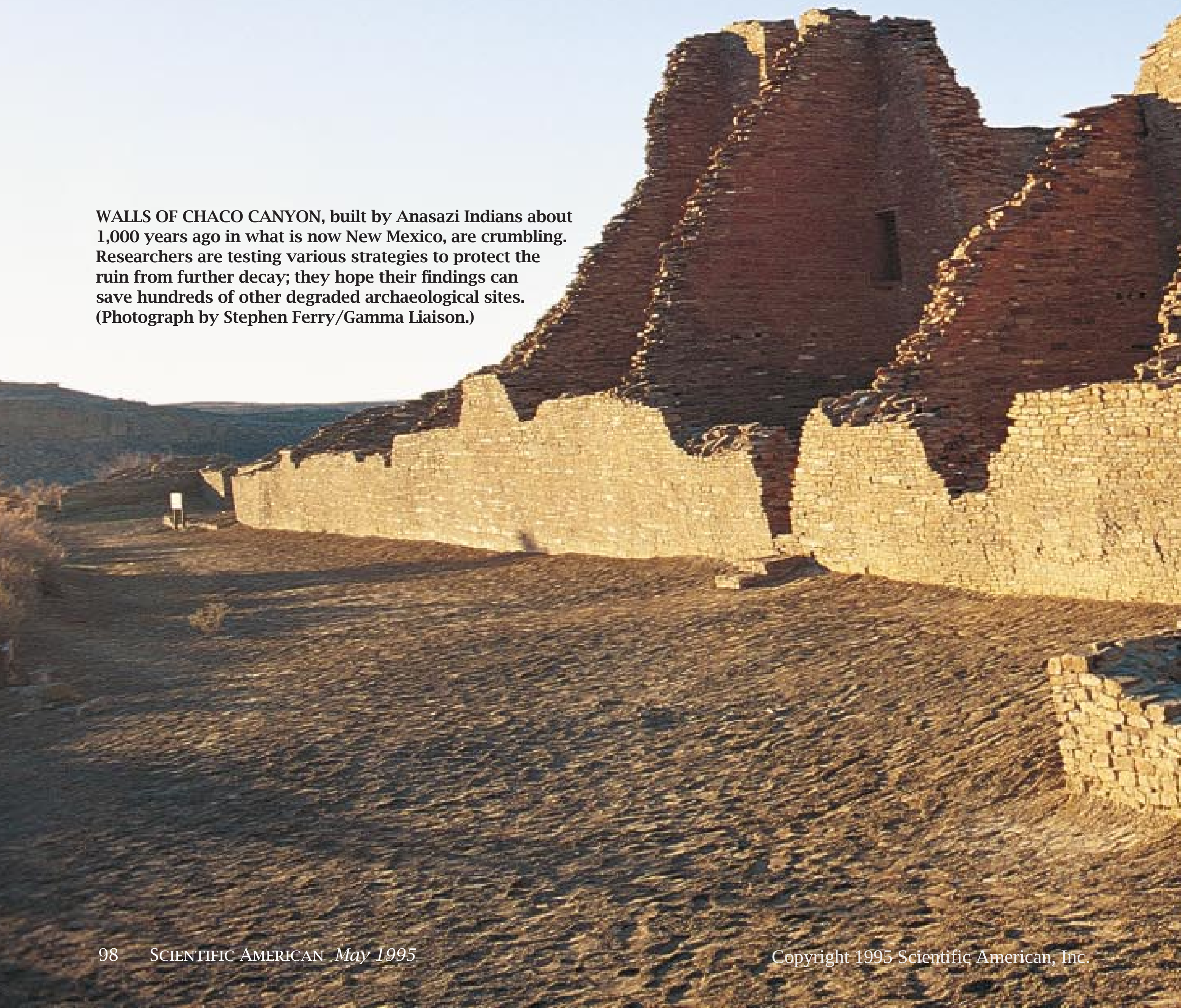
THE LOS ALAMOS PRIMER: THE FIRST LECTURES ON HOW TO BUILD AN ATOMIC BOMB. Robert Serber. Edited by Richard Rhodes. University of California Press, 1992.

HEISENBERG'S WAR: THE SECRET HISTORY OF THE GERMAN BOMB. Thomas Powers. Alfred A. Knopf, 1993.

The Preservation of Past

by Marguerite Holloway, *staff writer*

Conservators are racing to save monuments threatened by development, pollution, looting and neglect. In the process, they are transforming the field of archaeology into a new science



WALLS OF CHACO CANYON, built by Anasazi Indians about 1,000 years ago in what is now New Mexico, are crumbling. Researchers are testing various strategies to protect the ruin from further decay; they hope their findings can save hundreds of other degraded archaeological sites. (Photograph by Stephen Ferry/Gamma Liaison.)

Just before sunset, the stones of Pueblo Bonito in Chaco Canyon, New Mexico, glow golden, and their varied patterns stand out like friezes. The ruin—an elaborate city built and inhabited by Anasazi Indians between A.D. 850 and 1150—becomes empty of the day's tourists, and the cold of the desert night begins to settle in the shadowy rooms and kivas, circular ceremonial centers sunken in the ground. The small vents and tapered doors grow black.

One of these dark entrances—a high, squat door sitting at the junction of two walls—captures the attention of Dabney Ford, an archaeologist with the National Park Service who has worked at and studied Chaco Canyon for more than 10 years. She explains that the opening is aligned to catch sunlight only on the winter solstice. On that day, rays of light reflect on a specially built wall corner on the opposite side of the room. Although archaeologists understand relatively little about Anasazi culture, some have recently come to believe that the design of Pueblo Bonito—like that of Teotihuacán in Mexico, Machu Picchu in Peru and many other ancient cities—reflects astronomical cycles. The door is just one example of an architecture oriented by solar and lunar fluxes—one that perhaps embodies Anasazi political and social hierarchy.

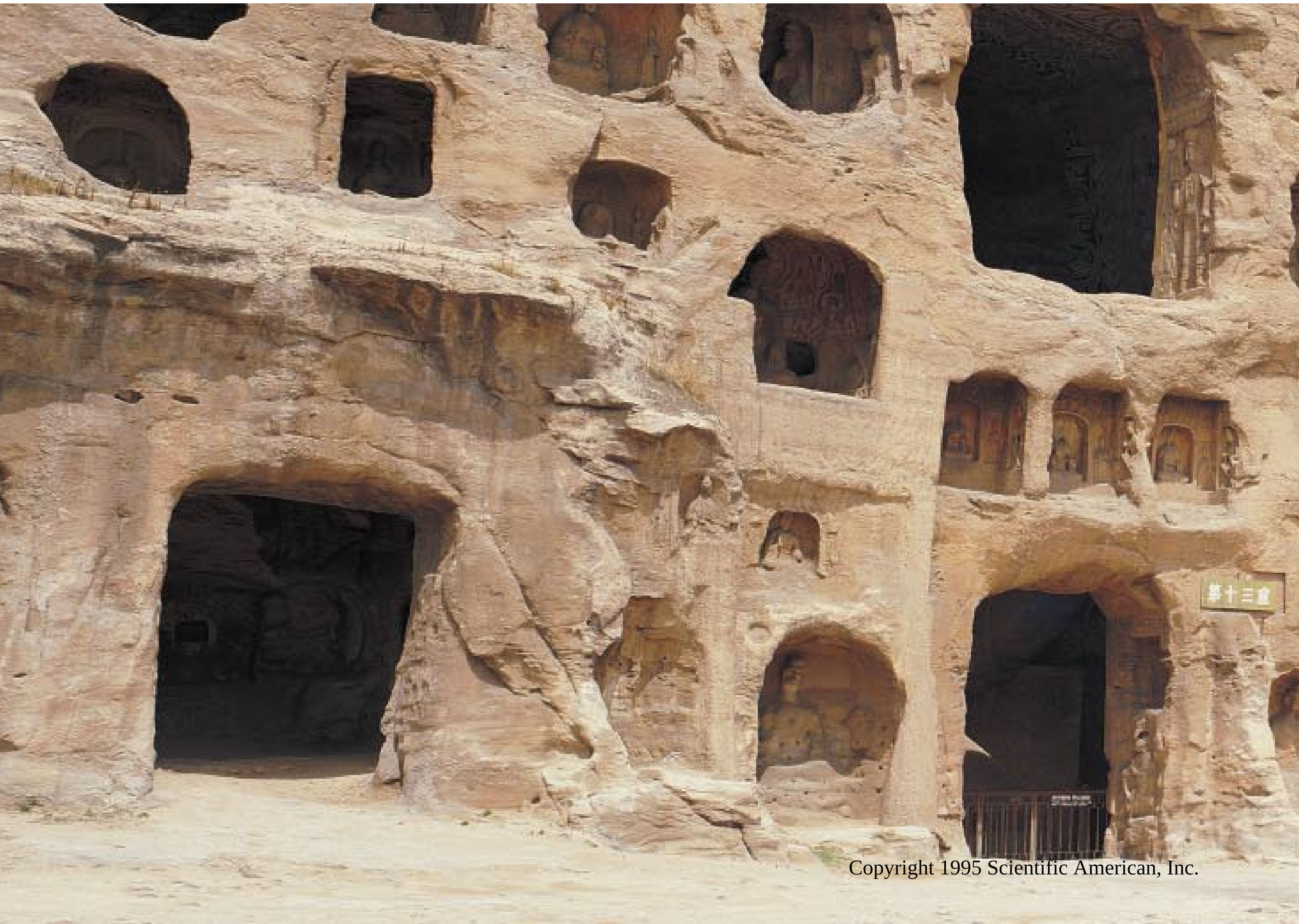
"Architects and archaeologists have had a hard time perceiving the complexity of this society," Ford notes, as she moves through tiny doorways, pointing out wood beams that have survived intact for hundreds of years. Chaco Canyon "is pretty humbling."

It is also falling apart. Being able to reinterpret Anasazi culture (if one is an archaeologist), visit the site (if one is a tourist) or pray in the sacred kivas (if one is a tribal descendant) depends on the walls' not buckling or collapsing as water seeps into the aging rocks and mortar. It depends on preventing salt crystals from growing in the stones and breaking them apart, vandals from digging up artifacts and visitors from climbing on walls or wood beams, splintering in minutes what endured for a millennium. Preserving Chaco Canyon entails solving a series of problems posed by archaeological sites around the world.

These concerns have, in the minds of many archaeologists, art historians and preservationists, reached crisis proportions. All 440 places designated by UNESCO as World Heritage Sites (including Chaco Canyon) are threatened, explains Bernd von Droste, director of the World Heritage Center in Paris. Angkor in Cambodia, the Pyramids in Egypt, the Parthenon in Athens, the old city of Dubrovnik in Croatia, the earthen bas-reliefs of the Royal Palaces of Abomey in Benin and the hominid fossil footprints found by archaeologist Mary Leakey in Tanzania are among hundreds of cultural and anthropological icons in danger of destruction. "In another 20 years, most of the archaeology of Egypt is going to be in museums," remarks Fred Wendorf of Southern Methodist University, who has been working in Egypt since 1962. "Everything else is going to be covered by water or plowed up in fields or coated with asphalt."

The principal threats to archaeologi-

DUSAN STULIK *Getty Conservation Institute*



cal sites are the same ones cited by environmentalists as endangering biodiversity: development, population growth, tourism, illegal traffic, air pollution, war, neglect and, in some cases, botched efforts at conservation. Nevertheless, the dangers stalking the world's cultural patrimony have not catalyzed any grassroots campaign to rival that of the green movement. (Some archaeologists speculate that, despite the public's fascination with Indiana Jones, roofless buildings do not have the same appeal as baby seals. Degradation also can be too slow to engage general interest: the fading looks of a sculpture may be seen only by comparing a contemporary photograph with one taken 50 years ago.)

Within the archaeological community, however, the growing perils are fostering a reevaluation of the entire field. "The whole attitude of archaeologists has been to plunder, then leave," notes Pamela Jerome, a professor at Columbia University, who is teaching one of the first courses exclusively on site conservation given at a graduate school. Typically archaeologists have excavated and then moved on—sometimes covering ruins over again to preserve them, a process called backfilling. Their training has not included work in conservation. That effort has instead been the province of museum-trained conservators, who restore only portable artifacts and collected objects. The two groups have often been at odds, Jerome says: one trying to push ahead with a dig, the other saying slow down, save this.

Wasting Away

But field archaeologists are now facing many of the challenges that have always confronted object conservators: disintegration and loss. They "have begun to realize that the rate of site destruction is pretty sizable and that archaeology is a nonrenewable resource. They have become more interested in preserving sites in situ," describes Francis P. McManamon, an archaeologist at the National Park Service. "That awareness has been reinforced by the recognition that you still have to take care of all the material that was excavated." Park Service collections alone hold an estimated 24.6 million objects, of which 16.8 million have not yet been catalogued. (Renovating facilities to care for these artifacts properly, something required under 1990 regulations, could cost more than $59.8 million.) The staggering amount of material that has been amassed worldwide has led some researchers to call for moratoriums on excavations until collections have been thoroughly researched; in many cases, recovered objects waste away in dusty vaults, unstudied and ill cared for.

The artifact crisis and the site crisis together have led to a new ethic and science. Borrowing from the wealth of knowledge about conserving artifacts and from materials science, scientists at several institutions—including the Getty Conservation Institute (GCI) in Marina del Rey, Calif., the International Centre for the Study of the Preservation and the Restoration of Cultural Property in Rome, the International Council on Monuments and Sites in Paris, UNESCO, the World Monuments Fund in New York City and various laboratories and universities—are formalizing the discipline of site conservation. Symposiums and courses such as Jerome's are beginning to appear. The International Institute for Conservation of Historic and Artistic Works will hold its 1996 meet-

ing on archaeological preservation. A journal, *Conservation and Management of Archaeological Sites,* to be published in London, is slated to start this year.

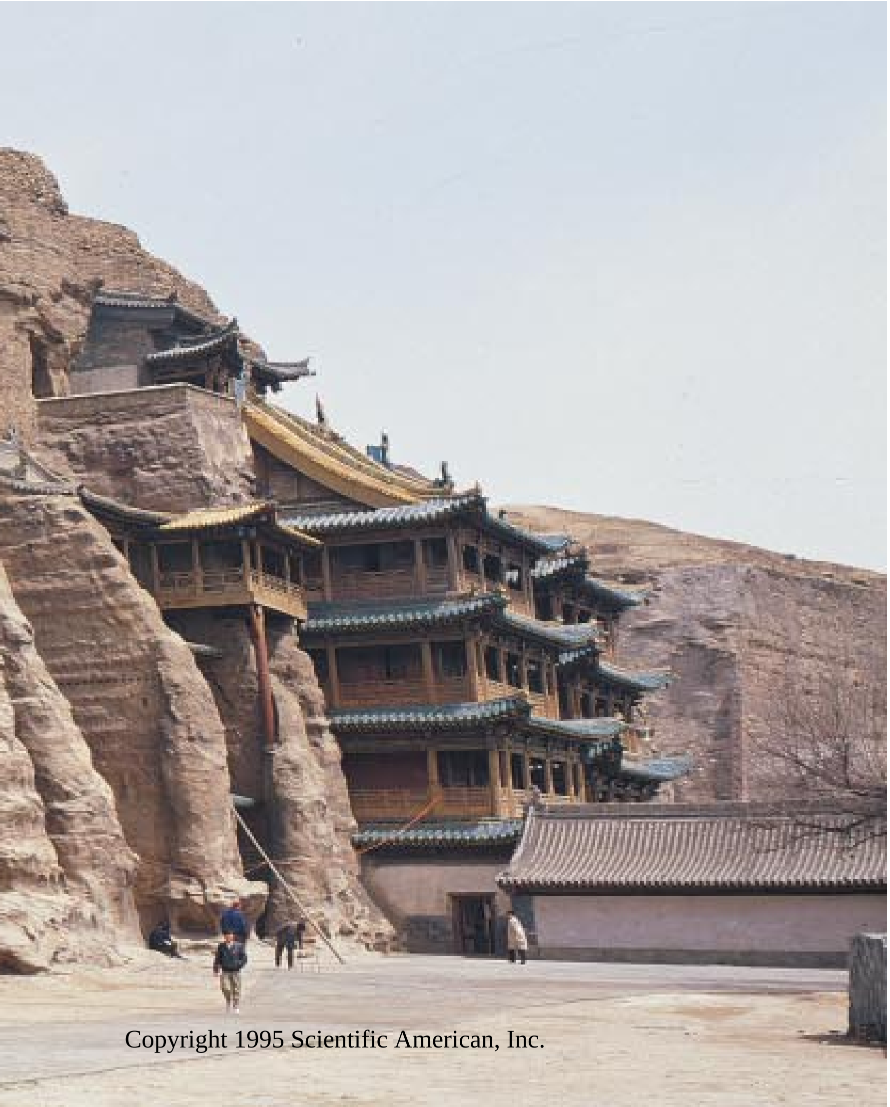
The primary goal of this burgeoning area of research is to slow deterioration. Buried ruins can survive for eons because they reach equilibrium with their stable surroundings. When dug out of sand or a mat of jungle growth or opened to air, their components begin to fall apart. "You are exposing it to sun, to rain, to dewfall, to oxygen, to all these things," describes Neville H. Agnew, associate director of programs at the GCI. "You catalyze a whole assemblage of degeneration that just springs into action—sometimes even within hours of excavation you can see deterioration starting." Some protective strategies being investigated include determining how to backfill ruins correctly, how to monitor caves or tombs for changes in humidity that could cause wall paintings to flake off, how to treat stones of tropical ruins with biocides and how to protect adobe, or earthen, buildings from the ravages of earthquakes.

Although backfilling has been practiced casually by archaeologists for years, it has never been tested rigorously. Conservationists view reburial as the best solution for many places, but the procedure remains controversial because researchers want their sites left uncovered so they, or the public, can wander around. This concern is evident at Chaco Canyon, where the GCI is collaborating with the Park Service on a backfilling experiment. Every year the park receives 80,000 visitors, who, presumably, want to see most of the site.

Even in Chetro Ketl, one of the great houses of Chaco, it is clear that some effort is needed to stanch further deterioration. The effects of wind, rain and a failed attempt at stabilization are all apparent. Heavy cement applied to the top of walls in the 1960s to keep out moisture has broken the lighter, underlying stone. Erosion around the base of many walls has destabilized them to the verge of collapse. Water has entered the crowns of other walls, frozen and subsequently forced apart the stones below. And water has caused salt to effloresce, or leach out, from the structure. Once such saline crystals form, they grow or shrink depending on the availability of moisture—and when they grow, they, too, push apart the stones, ultimately crumbling them.

Ford points out the white salt stains along the walls as she walks through the building, explaining the setup. Only 10 of 60 targeted rooms at Chetro Ketl have been backfilled, because reburying the rest would cost about $150,000—the annual maintenance budget for the entire 10-mile-long section of canyon and its thousands of ruins. Some rooms have been refilled only a few feet, with drains installed to pull moisture away from the walls and out of the site into drainage ditches. In other places, troughs buried in the center of the chambers collect and evaporate water.

In one room with wood beams, backfilling has been quite extensive and includes what are called geosynthetic materials—in this case, a sheet of black knotted plastic, which holds the soil firmly in place. Keeping the soil stable and getting the damp away are crucial so that the reburied wood does not rot, Ford says, and this test is being carefully monitored. She scrambles over to a nearby kiva and gestures toward a lunar probe look-alike standing, ungainly, in the middle of the sacred space,



YUNGANG GROTTOES in China are threatened by pollution, arid climate and water seepage (*left*). These factors are causing the 1,500-year-old sculptures to deteriorate, as is evident in a comparison of the same statue in 1933 (*below, left*) and in 1991 (*below, right*). Conservators are working to drain water from the site and to construct shelters to protect certain caves.

HEDDA MORRISON, *courtesy of Getty Conservation Institute*

MARGARET G.H. MAC LEAN *Getty Conservation Institute*

GETTY CONSERVATION INSTITUTE

recording fluctuations in moisture and temperature.

Although the experiments are not finished, in some chambers the system seems to be stabilizing the walls. In other places, the attempts are not so successful, and the right strategy remains elusive. "It may be a real mess," Ford shrugs, "but, hey, that's science."

Unforeseen results are inevitable because every location is different and because site conservation technology is just emerging. Agnew of the GCI recalls being called in to help preserve fossilized dinosaur tracks in Lark Quarry, a remote region of Australia. A shelter designed to protect the prints had also protected kangaroos from the hot sun, and so they would congregate there, cooling off and defecating, which did little to help preserve the tracks. Agnew concludes that the best approach if one does not have the resources needed to maintain a monitoring crew—which Lark Quarry management did not—is just "to rebury the site until that time when it can be opened for public display."

That lesson is being applied to fossil footprints in Tanzania that are more than 3.5 million years old. The GCI is evaluating how best to cover the famous Laetoli tracks that established bipedalism in hominids of that era. Although Leakey covered the site with dirt in 1979, tree roots have destroyed part of the trackway. Researchers are now considering how water flows in and around the site, whether they should use biocides to control the growth of weeds and trees, how they should signal future excavators that these levels of dirt and other materials are not original, how deep the prints should be buried and what kinds of geosynthetic materials should be incorporated. Once covered, it is unclear when, or if, the prints will be unearthed—or how it would be determined that experts should reinterpret the site or that the public should view the prints.

"There are challenges at many levels—they are not just scientific and technological," Agnew remarks. "We have to understand the mechanisms of deterioration of material and sites from a scientific point of view; that is critical. But it goes well beyond that as well." The issue of access, for instance, is pivotal. Archaeologists have the privilege of excavating places that belong to all of humanity. But what if strangers love the site to death?

The Egyptian Antiquities Organization is grappling with that concern at Nefertari's tomb in the Valley of the Queens. Originally opened in 1904, the tomb of the favorite wife of Ramses II houses magnificent wall paintings depicting Nefertari's voyage to the afterlife. The luxurious red, yellow, white and green murals were destabilized during discovery and after the introduction of desert air: plaster was damaged, and salts crystallized, sometimes an inch thick, pushing off the mud plaster and growing through the pigment. Before the 1980s it was clear that the paintings would not survive much longer. Between 1988 and 1992 some of the paintings were painstakingly removed and cleaned, and flakes of paint were reattached. The paintings were then put back in place.

The past two years have been devoted to monitoring humidity and changes in temperature to see how stable the internal environment is and whether it could support visitors. Shin Maekawa, head of environmental sciences at the GCI, tried to determine what the original state of the tomb was before it was opened because that condition had preserved the paintings perfectly for about

NEFERTARI'S TOMB in Egypt houses wall paintings that were damaged when the site was entered in 1904 (*left*). To preserve the recently cleaned and protected murals, researchers argue that tourism should be limited. A virtual-reality guide could offer visitors an alternative (*below*).

INFOBYTE; GETTY CONSERVATION INSTITUTE

MONITORING STATION sits in a kiva at Chetro Ketl. The device, designed by the Getty Conservation Institute, records changes in soil temperature and moisture. The data may help researchers determine how best to rebury ruins.

3,200 years. Through a series of experiments he found that the tomb had been quite humid through the centuries. In fact, humidity probably reached 50 percent in the summer months.

"Until that time, people thought that in order to conserve wall paintings it had to be drier, but that is not true," Maekawa explains. What turns out to be most critical is keeping such levels stable, something that can be achieved only if the number of visitors is strictly limited. "It took the tomb three days to recover from having 20 people inside for half an hour in the summer," Maekawa reports. Other popular tombs—such as Tutankhamen's, which already shows extensive deterioration—may have more than 3,000 visitors a day. "Despite all that effort we put in, if unregulated numbers of tourists are allowed into the tomb again, the wall paintings will be placed in danger," Maekawa warns.

Nevertheless, the government is under great pressure to open it. Tourism brings Egypt about $3 billion annually, and the reopening of Nefertari's tomb could boost the number of visitors. Indeed, although it can be beneficial, tourism is cited as one of the gravest threats to sites everywhere. According to von Droste of UNESCO, there were 500 million international tourists in 1994; about 40 percent of them traveled to see monuments such as the Sphinx, the Mayan temples of Tikal in Guatemala or the city of Petra in Jordan. Unfortunately, von Droste says, tourism often produces money for operators outside the countries being visited, so "revenues rarely go to the protection of the site." To address this problem, the GCI and others are trying to devise sustainable site management plans—as well as plans for historic centers within cities—that involve the local community, scientists, the government and the tourist industry.

In some cases, replicas can serve as proxies, protecting originals from the touch of tourists. The cave paintings of Lascaux in France were nearly destroyed by visitors brushing up against the walls. So the government built a replica nearby. The original caves are now viewed only by scholars, and the facsimile has become so popular that the number of tourists has to be limited. The GCI and Infobyte in Rome developed a virtual-reality guide to Nefertari's tomb—an alternative to visiting the real thing that may become common. But such solutions do not convey the power of a place. Nor do they resolve the question of who has the right to visit a site, whether sacred places should be left untouched or whether something should be preserved if no one can see it.

Limited Access

"It is worth saving sites, even if no one can get in there other than a few select scholars and other people," Agnew states energetically. "Of course, one does not like to be elitist and exclude the visiting public, but in fact, 90 percent of most tourists are not interested in art history. They come ill prepared: they wander in, and they wander out. Those people can go see other tombs in Egypt or other sites that are not threatened. I do not believe everyone has an inalienable right to see everything they want to see, any more than everyone can get into a Pavarotti concert—there are a limited number of seats. That is the reality of life."

Most preferable to conservators would be to extract archaeological information without touching the site at all. Consequently, researchers are increasingly using remote sensing—that is, recording information with cameras, radar, magnetometers—in short, any technology that will reveal aspects about a site from afar. "We are trying to develop the tools for conservation," says Farouk El-Baz, director of the Center for Remote Sensing at Boston University. El-Baz and his colleagues have used one of the techniques to determine how moisture was entering Nefertari's tomb as well as to extract air from a sealed chamber at the base of the Great Pyramid of Giza and to photograph the room's contents.

Remote sensing by satellite was recently used to map Angkor in Cambodia, and conservators are attempting to figure out how to apply the information. Angkor was the capital of the Khmer Empire from the ninth until the 15th century. Because each king would build a new capital, Angkor has more than 60 major monuments. The site is overgrown, stones have fallen out of place and the details of many sculptures have faded. Teams from several countries are working on various temples, each following its own distinct approach.

Preserving monuments in tropical environments is an unruly task because the jungle keeps taking over. Conservators have to decide whether to use biocides to kill vegetation and whether to shelter sites. Angkor gets 12 feet of water a year. "That much water is going to support plant growth, no matter what you do," says John Stubbs of the World Monuments Fund, which is preserving the compound of Preah Khan.

Stubbs has decided not to use biocides. Instead the team is removing heavy plant growth by hand, documenting many of the statues and performing anastylosis—replacing stones from where they have landed on the ground to what is assumed to be their original position in the ruin. On the whole, the World Monuments Fund approach is hands-off and reflects the growing consensus among archaeologists that sites should not be reconstructed, because that activity entails a potentially incorrect reinterpretation of the past.

Nevertheless, Stubbs does favor a certain kind of reconstruction and hopes to use the satellite data from the National Aeronautics and Space Administration to that end. "I would like to restore water features because I am quite convinced that the architecture was supposed to be seen in reflection," he explains. "The whole culture is about capturing and manipulating water." The image shows the reaches of the estuary, which covers the area from the Mekong River up to Angkor during the rainy season. The information could help conservators understand the region's hydrology, knowledge that could, in turn, help them determine how and whether to drain parts of the site to preserve it.

Stubbs's biggest fear about Angkor—and he is not alone—is that looting will leave it barren. "The popularity of the site, the new access, the hunger for Khmer art and the lawlessness of the place are all a recipe for pillaging," he notes. Stubbs says he has visited shops in Bangkok, inquired about buying Khmer antiquities and been shown pictures of objects taken from Preah Khan—and even some photographs of statues still in place.

Many archaeologists single out looting as the major threat to ancient sites and objects. It is impossible to get good numbers on the value of the stolen objects, but estimates run from $2 billion to $6 billion annually. From 80 to 90 percent of antiquities, including those sitting in museums, are considered "hot." The scale of seizures intensified in the mid-1980s, when people began to look at antiquities as investments, explains Ricardo J. Elia of Boston University. All the more horrible, he fumes, is that an artifact obtained as a spoil of plunder, at the expense of a destroyed site, comes to be seen as an art object when it is bought by a wealthy collector and ends up in a museum: "It starts out real dirty and ends up real clean."

For conservators, stopping illicit traffic and using science and technology to protect monuments and ruins are, in many ways, easier tasks than answering questions that are raised about saving sites. Given limited resources and conflicts of interest, what should be saved? "It used to be that time made the choice," reflects Miguel Angel Corzo, director of the GCI. "War came, and people used houses and monuments as quarries. It was just the natural evolution of things that brought us to where we are today. But now we have an extraordinary capacity to save whatever we want; the technology is there. What are we going to leave behind? Are we going to take the same attitude, 'Time will tell'? Or are we going to be people of the 21st century and say, 'This is who we are going to be remembered as'?"

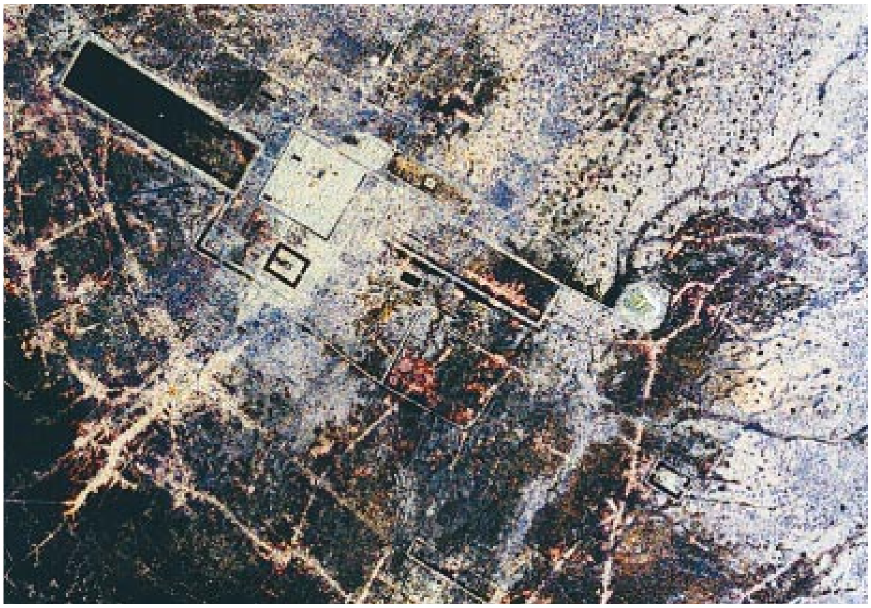
It is tragically clear that wiping out a site can wipe out cultural identity—something that happens when ethnic groups clash, each trying to obliterate the culture of the other. China's razing of the monasteries of Tibet is but one example. The interpretation of archaeological remains powerfully shapes the outlook of an entire people, as can be seen in the evolving view of the Great Zimbabwe, a series of majestic granite structures constructed between A.D. 1400 and 1600. British colonialists claimed that the circular compound could not have been built by native Africans: northern invaders from either Greece, Egypt or the Middle East must have swept down and built the sophisticated enclosure. It has been established, however, that it was the ancestors of regional Shona people who built the Great Zimbabwe. The cultural identity and independence of Zimbabwe (formerly Rhodesia) are closely linked to this monument.

African countries, in particular, could lose a profound part of their cultural identity if their sites are not conserved. The continent has a dearth of native archaeologists, and many places have not yet been excavated. According to a 1992 World Bank conference report, *Culture and Development in Africa,* there is only one archaeology program in 11 central African countries, and the total annual budget for the entire region is $20,000.

Despite the vast challenges and many unresolved ethical questions, the archaeological community is developing a philosophy of conservation. This shift could be seen in the response to the recent discovery of some 300 Cro-Magnon paintings in a cave near Avignon, France. Immediately after its discovery, the cave and its 20,000-year-old masterpieces were sealed off. "The cave should be left as it is. It must be studied, of course, but bearing in mind that the collecting of the information must not be destructive," says Jean Clottes, a prehistorian and adviser to the French Culture Ministry. "It is out of the question to start excavating all over the place." As Clottes points out, ever smaller samples of materials are needed to date paintings, and so the investigations of archaeologists present may seem like ravages to those to come. Furthermore, "theories change," Clottes notes. "Everyone believes in one for 10 or 15 years, then there is another theory."

It is for the sake of these theories and for the awe of place and time that people want to hold on to their past. Multiculturalism—the celebration of any culture, no matter how small—as well as the movement toward the next millennium may be contributing to a growing nostalgia or, as some archaeologists put it, a need to reconnect with history. For others, though, the attempt to hold on to everything makes little sense. "We Westerners tend to feel that it is massively important to preserve," comments Marian A. Kaminitz, a conservator at the National Museum of the American Indian. Others "have a strong sense of allowing things to return, of not holding on to them."

In the gathering dark of Chaco Canyon, Pueblo Bonito returns to the desert. An absolute stillness settles around the ruins. "It is beneficial that they are protecting it," says Petuuche Gilbert, a descendant of the Anasazi. But we could also let it go, he adds: "It would be part of the natural cycle."



JET PROPULSION LABORATORY/NASA

ANGKOR BY SATELLITE offers a fresh perspective for archaeologists studying the ancient Khmer city in Cambodia. Conservators hope data on vegetation growth and hydrology can aid them in protecting the site from further degradation and, perhaps, in restoring some of the original canals and reservoirs.

Fibonacci Forgeries

"A toast, ladies and gentlemen: To the Queen!" We stood, dutifully raised our glasses and murmured the sovereign's name. It should have been a poignant moment, but it was spoiled by the person next to me, who flopped back into his chair and muttered, "Thank God, now I can smoke!" (It is a quaint British custom that at a formal dinner no one may smoke before the party drinks to the queen's health.)

"I'd rather you didn't," I said. "I'm a nonsmoker."

"You had the smoked salmon," he said and roared with laughter as he lit a cigarette. I wrinkled my nose and glanced at his badge: Richard Byrd. He had penciled in, "Call me Dicky."

"The salmon, by the way, was terrible," I told him. "In fact, I think it was really dogfish dyed orange." He did not look surprised. The annual dinner of CAT-DOG (the Charitable Association for Tax-Deductible Offerings of Generosity) was always a disaster.

"I wondered why the fish was wearing a flea collar," said the lady on my left. I had met her at other gatherings: Amanda Bander-Gander, a leading light of the local Animal Protection Society.

"No, dear, you just dropped your napkin ring on it," her husband yelled from five places down on the opposite side of the table. Alexander Bander-Gander, a lawyer, was sandwiched between Athanasius Fell, a doctor, and Dennis Racket, an old friend of mine from the tennis club.

"So what business are you in?" Dicky asked me.

I leaned closer to him and whispered, "I'm a mathematician."

As usual, though I had spoken below the threshold of aural perception, it was a party-stopper. The entire table went silent.

"I was never—" he began.

"Any good at math at school," I finished. "That's what they all say."

"Hated it then," Athanasius mumbled. "Still do."

"I must be the exception!" a voice boomed at my right. "Absolutely loved it! Name's Adam Smasher. I'm a nuclear physicist. I have a little puzzle I'll ask all of you. What's the next number in the sequence 1, 1, 2, 3, 5, 8, 13, 21?"

"Nineteen," I grunted automatically, while battling with a bread roll seemingly baked with cement.

"You're not supposed to answer," he said. "Anyway, you're wrong—it's 34. What made you think it was 19?"

I drained my glass. "According to Carl E. Linderholm's great classic *Mathematics Made Difficult,* the next term is always 19, whatever the sequence: 1, 2, 3, 4, 5—19 and 1, 2, 4, 8, 16, 32—19. Even 2, 3, 5, 7, 11, 13, 17—19."

"That's ridiculous."

"No, it's simple and general and universally applicable and thus superior to any other solution. The Laplace interpolation formula can fit a polynomial to any sequence whatsoever, so you can choose whichever number you want to come next, having a perfectly valid reason. For simplicity, you always choose the same number."

"Why 19?" Dennis asked.

"It's supposed to be one more than your favorite number," I said, "to fool anyone present who likes to psychoanalyze people based on their favorite number."

"Nonsense. I'll tell you the real answer," Adam said. "Each number is the sum of the previous two. So the next is 13 + 21, or 34, then 55, then 89, then 144 and so on. It's the—"

"Fibonacci sequence," I interrupted. "God, I'm so fed up with the blasted Fibonacci sequence! Even the name's phony! 'Leonardo Fibonacci, son of Bonacci!' That's a nickname invented by Guil-

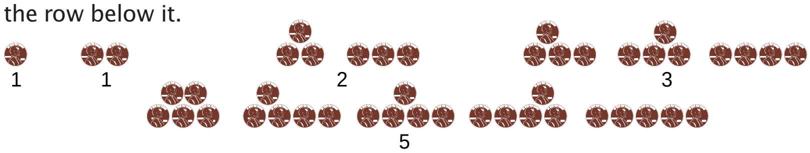
DIANNE GASPAS ETTL

Fibonacci or Forgery?

Sequence 1 is 2, 3, 5, 8, 13 . . . a_n, where a_n equals 1 plus the sum of the first n terms of the Fibonacci sequence. For example, $a_1 = 1 + 1 = 2$; $a_2 = 1 + (1+1) = 3$; $a_3 = 1 + (1+1+2) = 5$ and so on.

Sequence 2 is 1, 3, 8, 21, 55 . . . a_n, where a_n equals the sum of the first n terms of the sequence formed by removing every other Fibonacci number. So $a_1 = 1$; $a_2 = 1 + 2 = 3$; $a_3 = 1 + 2 + 5 = 8$ and so on.

Sequence 3 is 1, 1, 2, 3, 5, 8 . . . a_n, where a_n is the number of ways in which you can arrange n coins in horizontal rows such that all the coins in each row touch and every coin above the bottom row touches two coins in the row below it.



Sequence 4 is 1, 2, 5, 13 . . . a_n. It is the same as sequence 3, except that now n is the number of coins in the bottom row.

Sequence 5 is 1, 2, 5, 13 . . . a_n, where $a_n = (n-1)2^{n-2}+1$.

Sequence 6 is 1, 2, 5, 13 . . . a_n, as defined by the number of disconnected graphs with $n+1$ vertices.

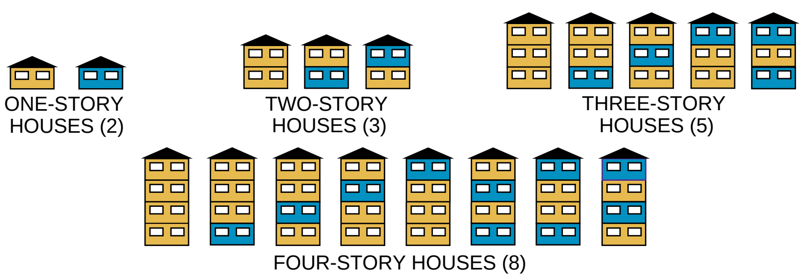
Sequence 7 is 1, 1, 2, 3, 5, 8 . . . a_n, where a_n equals the integer nearest to $((1+\sqrt{5})/2)^n/\sqrt{5}$.

Sequence 8 is 1, 2, 5, 13 . . . a_n, defined by the number of connected graphs with $n+2$ vertices having just one cycle.

Sequence 9 is 1, 2, 5, 13 . . . a_n, where a_n is the coefficient of $x^{n+2}/(n+2)!$ in the power series solution to the differential equation $y'' = e^x y$, which starts $y = 1 + x^2/2! + x^3/3! + 2x^4/4! + 5x^5/5! + 13x^6/6! \ldots$

Sequence 10 is 1, 2, 5, 13 . . . a_n, where a_n equals $a_{n-1} + na_{n-2}$, when $a_{-1} = a_0 = 1/2$.

Sequence 11 is 2, 3, 5, 8, 13 . . . a_n. In this sequence, apartment blocks of n floors are to be painted blue and yellow, with the rule that no two adjacent floors can be blue. (They can, however, be yellow.) Let a_n be the number of ways to paint a block with n floors.



laume Libri in 1838, long after Fibonacci died. The famed Fibonacci was in fact named Leonardo Pisano Bigollo. Pisano means that he lived in Pisa; no one knows what Bigollo means. At any rate, his sequence ought to be called the Leonardo Pisano Bigollo sequence, except that's too long."

"You mathematicians have a lot of attitude," Dicky remarked. "At least, an attitude that differs from most other people's."

"Proof," Adam declared. "Mathematicians always want to prove things. That's certainly a strange attitude. Never understood why myself. If you keep trying it, and it keeps working, it's got to be right! So why waste time getting into all sorts of logical tangles proving the silly thing?"

"Well, why do you physicists bother doing experiments? If a theory tells you what you want to hear, why not just assume it's true?"

"Because you can't just go around believing theories without testing them!"

"And mathematicians don't think you can go around believing theorems without proving them. Alexander, why do lawyers insist on cases being tried in court? Why not just let the judge look at the evidence and decide whether the defendant has committed the crime?"

"You can't do that! There could be a miscarriage of justice!"

"Right. That's what mathematicians worry about when they insist on proofs. They don't want to find out later they were wrong. It might be embarrassing."

Adam shook his head sadly. "You know full well it's not like that. Mathematics is basically simple. If you can see an obvious pattern, it can't be coincidence. Why bother to prove it?"

I thought for a few seconds. "I'll give you an example. Here's a sequence, and I want you to tell me the next number."

"I'll do my best."

"1, 1, 2, 3, 5, 8, 13, 21, 34, 55," I said.

He looked puzzled. "Don't be silly. I just asked you that one. It's the Fibonacci sequence."

"Is it? Then what's next?"

"89," Adam replied.

"Wrong. It's 91."

"But it looks just like the—"

"You're leaping to conclusions, Adam, based on previous prejudices. Most injudicious. Your sequence was the Fibonacci sequence, but the nth term in my sequence is the least integer not less than $\sqrt{e^{n-2}}$, where $e = 2.71828$, the base of natural logarithms. I want the 11th term, which is the least integer not less than $\sqrt{e^9}$, or 90.017. That's 91."

"Humph. Well, that's an accident, a rare exception. I'll believe you if you can show me some more examples of mis-

Solutions to the Misleading Sequences

Sequence 1: Fibonacci. For example, to form a_6, you add 8, the sixth term in the Fibonacci sequence. The sum is 8 + 13, a sum of consecutive Fibonacci numbers. By definition, then, the answer, 21, is itself a Fibonacci number. This pattern continues.

Sequence 2: Fibonacci. For example, to go from a_5 to a_6, you add 89. The sum is 55 + 89, a sum of consecutive Fibonacci numbers. By definition, then, the answer, 144, is itself a Fibonacci number. This pattern continues.

Sequence 3: Forgery. The sequence continues 12, 18, 26.

Sequence 4: Fibonacci. The proof depends on the identity $f_{2n-1} = f_{2n-3} + 2f_{2n-5} + 3f_{2n-7} + \ldots + (n-1)f_1 + 1$ for Fibonacci numbers f_n.

Sequence 5: Forgery. The sequence continues 33, 81, 193.

Sequence 6: Forgery. The sequence continues 44, 191, 1,229, 13,588.

Sequence 7: Fibonacci. This pattern follows from Binet's formula $f_n = [((1+\sqrt{5})/2)^n + ((1-\sqrt{5})/2)^n]/\sqrt{5}$.

Sequence 8: Forgery. The sequence continues 33, 89, 240, 657, 1,806.

Sequence 9: Forgery. The sequence continues 36, 109, 359, 1,266.

Sequence 10: Forgery. The sequence continues 38, 116, 382.

Sequence 11: Fibonacci. Consider, for example, a five-floor building. The fifth floor can be either yellow or blue. If it is yellow, the rest of the building can be painted in any possible way for a four-floor building. If it is blue, the fourth floor must be yellow, and the rest of the building can be painted in any possible way for a three-floor building. So $a_5 = a_4 + a_3$, as it is for Fibonacci numbers. The pattern is general.

Odd Prime Sequence: Forgery. The sequence 3, 5, 7, 11...331, 337 and so on consists of all the numbers n that divide $2^{n-1} - 1$ exactly. The next term is 341, which is not prime (11 × 31 = 341) but divides $2^{340} - 1$.

leading sequences," Alexander interjected. "Can you? Or have you exhausted your repertoire?"

"There are hundreds," I said. "It's easier than you think. Richard K. Guy, a mathematician at the University of Calgary, collects them. He refers to them all as the Strong Law of Small Numbers. There aren't enough small numbers to meet the many demands placed on them, so what look like patterns involving small numbers might just be coincidence. And often are."

I showed them 11 sequences that looked like the Fibonacci numbers, or the alternate Fibonacci numbers 1, 2, 5, 13 and so on [*see box on opposite page for the sequences*]. I asked them to decide which were really Fibonacci sequences and which were Fibonacci forgeries [*see box above for the answers*].

My companions began to argue about whose solution was right when I proposed one last puzzle. "What is the next term in the sequence 3, 5, 7, 11, 13, 17, 19..."—I carried on for some time listing the odd primes—"...331, 337?"

"Those are the odd primes," Adam said. "You can't generate that many primes by accident. The next term must be—uh—347."

"Are you sure?" I asked quietly.

FURTHER READING

MATHEMATICAL GAMES: PATTERNS IN PRIMES ARE A CLUE TO THE STRONG LAW OF SMALL NUMBERS. Martin Gardner in *Scientific American,* Vol. 243, No. 6, pages 18–28; December 1980.

A DIARY ON INFORMATION THEORY. Alfréd Rényi. Wiley, 1987.

THE STRONG LAW OF SMALL NUMBERS. Richard K. Guy in *American Mathematical Monthly,* Vol. 95, No. 8, pages 697–712; October 1988.

THE SECOND STRONG LAW OF SMALL NUMBERS. Richard K. Guy in *Mathematics Magazine,* Vol. 63, No. 1, pages 3–20; February 1990.

SCIENTIFIC AMERICAN

CORRESPONDENCE

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

Back issues: $6 each ($7 outside U.S.) prepaid. Most numbers available. Credit card (Mastercard/Visa) orders for two or more issues accepted. To order, fax (212) 355-0408.

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

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

Photocopying rights are hereby granted by Scientific American, Inc., to libraries and others registered with the Copyright Clearance Center (CCC) to photocopy articles in this issue of *Scientific American* for the fee of $3.00 per copy of each article plus $0.50 per page. Such clearance does not extend to the photocopying of articles for promotion or other commercial purposes. Correspondence and payment should be addressed to Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923. Specify CCC Reference Number ISSN 0036-8733/95. $3.00 + 0.50.

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

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

Subscription correspondence should be addressed to Subscription Manager, Scientific American, P.O. Box 3187, Harlan, IA 51537. The date of the last issue of your subscription appears on each month's mailing label. For change of address notify us at least four weeks in advance. Please send your old address (mailing label, if possible) and your new address. E-mail: SCAinquiry@aol.com.

Candid Camera

Review by Philip Morrison

PALE BLUE DOT: A VISION OF THE HUMAN FUTURE IN SPACE, by Carl Sagan. Random House, 1994 ($35).

This book opens with a generous gift to us all. It was made early in 1990, when the space probe *Voyager* completed its scripted dozen-year tour of duty. Well beyond Neptune and far north of the plane of the solar system, the craft received a final set of new commands, no part of the original mission. Look back, *Voyager,* to the now distant inner planets! Carl Sagan and a few others had argued and waited years for the Jet Propulsion Laboratory and NASA to schedule a shoestring effort to snap just one candid portrait of the earth among the planets. They succeeded brilliantly.

The spacecraft recorded a mosaic of 60 frames, planets against the sky. It sent the bits to far-off earth; that slow video transmission from past its design range took three months. The four planets in clear view would at best be mere dots of light. (The glare of the sun only six light-hours away from the camera masked nearby Mercury. Pluto and Mars happened to be poorly placed. Uranus and Neptune were so faint that they required long exposure; they could not help themselves from moving, so their images are streaked.) Our earth entire is seen as one bright bluish dot in the image that faces the title page.

The cloud-and-sea-marked blue marble we know well from closer space views dwindled to a featureless dot once the telescopic camera was carried a few thousand times farther away. The earth was for the first time put into the same perspective as the other planets in the sky. In 1957 a metal ball from a human workshop had been set into the heavens; mere speed transformed it into something celestial, "shining, circular, and perpetual" in orbit. Aristotle was denied and Newton confirmed. One after another the planets became close-up scenes, real places, a red Mars landscape looking not unlike the Arizona desert. Here they are at last as wandering stars: the earth, too.

Anyone can now *see* at once that Copernicus was right. The earth is only one planet among planets; blue, yes, but then icy Neptune is blue as well. Sagan eloquently draws an inference about this our home. "On it everyone you love... every human being who ever was, lived out their lives.... Every hunter and forager, every hero and coward, every creator and destroyer of civilization... every saint and sinner in the history of our species lived there—on a mote of dust suspended in a sunbeam."

Around us stretches a vast cosmic darkness, studded with countless pools of myriads of suns. Only the abstract laws of physics seem to bring any order; no Designer is visible, no Parent seen to care for us. Yet the universe of science is no petty place; it is dauntingly large, old, diverse, even humbling. A religion to match our awesome data, Sagan thinks, may someday emerge. We are not center stage. "We have not been given the lead in the cosmic drama. Perhaps someone else has. Perhaps no one else has... a good reason for humility." The self-serving conclusion is what the author mistrusts above all.

After this high thinking, Sagan catches his readers up in a dozen exciting, personal, even audacious reviews of what we have found and may yet seek in space. Colorful and fresh space images are plentiful in these pages, along with paintings, diagrams, maps and some graphs, but no equations at all. Many of these treat issues Sagan has been close to. Here he is more than an author; he is an original researcher, articulate and engaged.

A sample of three out of many rich topics will have to suffice. One is a witty little essay. The French use a mild expletive, *sacre-bleu.* Translate it: Good heavens! Like our "geez," it is a euphemism for invoking Deity. The earth's distant signature is that sacred blue. Yet just outside the air, astronauts see a black sky with stars, even in sunlight. All airless moons and asteroids share black skies.

Molecular light scattering from the density fluctuations of any molecule is preferentially blue; what scatters

EARTH: a pale blue dot in a sunbeam. Photographed by* Voyager *in 1990.

is removed from the direct path, so that the long low light of any sunset is rendered in red hues. Venus has an atmosphere so thick that down at ground level it is always sunset, as the sky color reported by the *Venera* landers confirms. The meager sky above the rusty deserts of Mars looks pale salmon, colored by the long-lasting residues of finest dust, stirred up into the high air by windstorms. The two giant planets are certainly dark at the prodigious depths of their atmospheres. All we see there from without is the cloud layers that lie highest. The unearthly blue skies of Neptune and Uranus are puzzling; easy answers do not appear quite to work. Does some unknown blue stuff lie there inside? *Sacre-jeune, sacre-rouge, sacre-noir!* For this reader, awe of plenitude arises from this game of words.

Are we alone in the Milky Way? The most ambitious of searches for radio signals from putative distant sharers of our own scientific curiosity has been pressed for seven years by Paul Horowitz of Harvard University; it was financed by private gift, largely from Steven Spielberg of *E.T.* The modest radio dish in use is surplus, but the student-built multichannel receiver is world class, automatically monitoring eight million simultaneous narrow-frequency channels.

Signals that rise far above the noise are the candidates for our attention and are tested for origin from interference or for electronic fault. A few are culled from 60 trillion observations over the entire sky. Not one of the finds repeats. One fact, says Sagan, co-author of the analysis, "sends a chill down my spine." Of the 11 best candidates, eight lie within or close to the band of the Milky Way, where most stars collect. A fair bet against that happening by chance should pay off at better than 100 to 1. Probably these are rare, nonrepeating chip errors; without repeats or other structure we remain ignorant and very properly withhold belief. We will keep on listening, now without federal support. "Science offers little in the way of cheap thrills. The standards of evidence are strict. But…they allow us to see far, illuminating even a great darkness."

This book glowingly communicates current wonders and large issues still ahead, like future "interplanetary violence" against the earth by collisions both random and intentional. It displays openly the hopes and judgments of one gifted, adept and devoted human being, surely sometimes right and sometimes wrong. No recent book has done better at making plain the subtle nature and the fascination of scientific investigation.

GREAT PAINTINGS ***opens directly into a no-frills database of artworks.***

The Gallery in the Machine

Review by Ben Davis

GREAT PAINTINGS, RENAISSANCE TO IMPRESSIONISM: THE FRICK COLLECTION. Digital Collections, Inc., 1994 ($79.95, Macintosh or Windows). **ART GALLERY: THE COLLECTION OF THE NATIONAL GALLERY, LONDON.** Microsoft Corporation, 1993 ($59.95, Macintosh or Windows). **AMERICAN VISIONS: 20TH CENTURY ART FROM THE ROY R. NEUBERGER COLLECTION.** Eden Interactive, 1994 ($59.99, hybrid Macintosh/Windows).

There is something mysterious about visiting an art museum, especially the enigmatic Frick Collection in New York City. What is El Greco's famous portrait *St. Jerome* doing here—shouldn't it be in the Prado? And how has Johannes Vermeer's *Mistress and Maid* found its way into this converted private mansion that was completed just in 1914? The Frick is a quiet island of art treasures located on Manhattan's Upper East Side. But like a number of other museums, it has also moved into cyberspace, popping up in digital form on a CD-ROM.

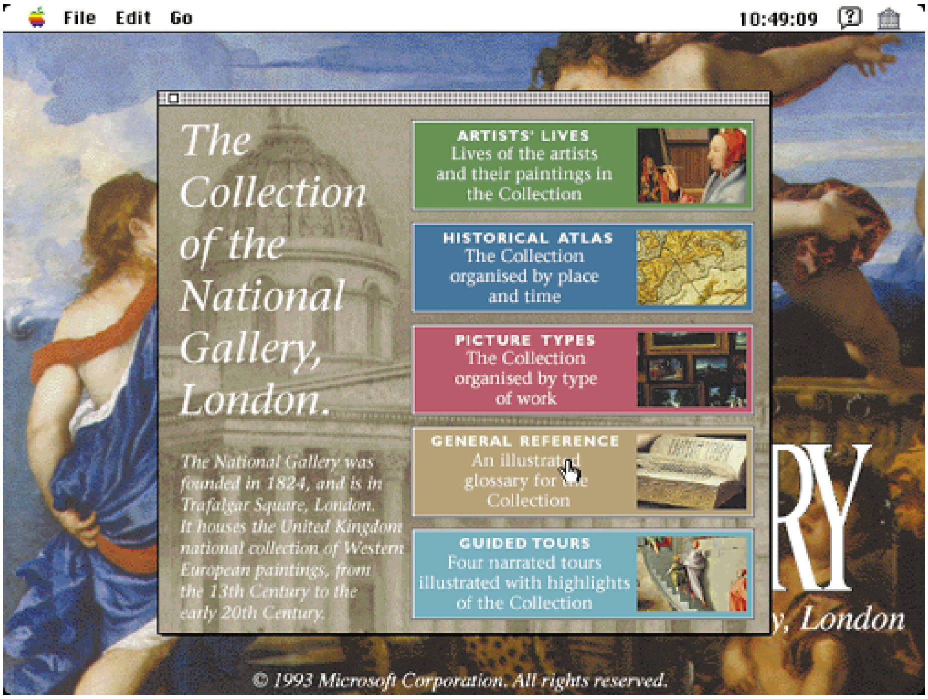
Museums have always relied on technology to aid in the presentation, preservation, storage and retrieval of collections. But the latest innovation—the digital museum collection—marks a radical turn in the relation between art and technology. For the first time, the diverse functions of a traditional art museum have been converted into a single form, electronic bits; artworks have been transformed into information objects.

Publishers and museums are grappling to find appropriate applications for this new way to view art. Are the disks a research tool, a form of catalogue, an "edutainment" product? Should we view them as high-tech coffee-table books or as primitive versions of virtual museums that will someday reside on the Information Highway? Do they function as interactive previews of the real museum collections or actually substitute for a visit? Can the mystery of great imagery survive when seen on a cathode-ray tube?

The three CD-ROMs reviewed here take tellingly diverse approaches. *Great Paintings, Renaissance to Impressionism* is a database containing some (but not all) of the works in the Frick Collection. *Art Gallery* is a database as well as a rudimentary art encyclopedia and educational tutorial. *American Visions* consists of a database as well as a point-of-view presentation. The disks require four to eight megabytes of memory and, of course, a CD-ROM drive.

All three CD-ROMs share a peculiar omission: none of them tells you how to get to the real museum. Nor do any of them provide floor plans so that you can locate the works that you have seen in digital form. The disks never encourage you to experience the real works, which is the point of looking at digital art collections—isn't it?

Where these disks differ is in their intentions, which are most acutely revealed in the way each one uses motion. Curiously, even though museums are



ART GALLERY ***offers several ways to view or tour the museum collection.***

not notably dynamic settings, it is the moving images and audio that create in the CD-ROMs a sense of place, a feeling of connection to the actual museums. Whether through clever animations that explicate painting techniques or digital-video testimonials, the motion in the multimedia beckons the user into the virtual museum.

Yet the Frick Collection CD-ROM is constructed as an image database that uses essentially no motion at all. This disk seems to be designed more for an art researcher than for a casual visitor. The data interface resembles Windows, workmanlike and devoid of frills. The disk provides several ways to explore the database, including key words, style and type of work, themes or topics, and date of execution. Images can be sorted according to these criteria to make slide show–style presentations. You can enlarge areas of the works, although you quickly end up with a screenful of oversized pixels.

Unfortunately, most consumers buying an art CD-ROM do not know what a researcher would be looking for. How did Velázquez paint the eyes of his noble subjects to make them appear a bit strange? No ready answer awaits in the rather minimalist digital documentation available on the disk. Even a simple timeline or an example of how to conduct a historical search would make the browsing much more rewarding.

The lack of context is particularly frustrating because of the eclectic nature of the Frick Collection and the haunting ambience of the building that houses it. On a recent visit I rediscovered the richly furnished rooms and the startling diversity of works that had left me with such vivid memories. Henry Clay Frick himself insisted that the works he had collected be seen in the setting of the "elegance and distinction" of his home.

The corresponding CD-ROM completely lacks this aura. Nowhere do you see the marvelous architecture in which the works reside. It takes a bit of persistent exploring to find the one chunk of background on the collection: a brief but excellent text by Charles Ryskamp, the director of the Frick Collection, which is available under one of the "info" options. Perhaps the positive side of the sparse background material is that when you visit the Frick you will be able to discover what a jewel it is without any preparation from the CD-ROM.

Microsoft's *Art Gallery* embraces a warmer, more accessible approach—that of the art encyclopedia. This disk aims to expose the user to a broad range of experiences while maintaining a familiar pattern of access. Closely patterned on workhorse teaching references (such as H. W. Janson's *History of Art*), this style still gives the browser the feeling of researching through a database, but this time with the assistance of a kinder, gentler interface created by an art historian.

You can search the collection by the name of the artist or look through groupings labeled "historical atlas," "picture types" and "general reference." You also have the option of jumping into four user-friendly tutorials. This educational approach is appropriate to a museum such as London's National Gallery. Here the collection reflects the activity of an institution rather than the viewpoint of a single art lover.

The tutorials in *Art Gallery* offer wonderful if somewhat brief formal explanations of how perspective works, how composition works, how color affects subject matter and how a work of art is restored. Each tutorial includes narration, descriptive text and animated analyses of about half a dozen paintings. An unraveling of Jan Van Eyck's *The Arnolfini Wedding,* for example, includes views of the painting, a detail of the artist reflected in the mirror in the painting and a short but thorough discourse on the effect that Van Eyck was trying to achieve.

When you search through the main collection, small thumbnail images of works appear alongside black text on a white screen that seems intended to simulate a page, a card or a white wall. A click of the mouse expands the image. Viewing the computer screen in this setup feels something like looking into a lightbulb; a neutral gray background might be less jarring. Most of the paintings are rendered in color. Why some of them are shown in black and white is not obvious. An art novice might mistakenly assume these are monochromatic paintings.

The disk incorporates a refreshingly simple alphabetical index to aid you in finding a specific work. Clicking on a letter reveals all items from that part of the alphabet. Selecting a specific name then brings up information about that artist or artwork. Some words in the descriptions of the works are highlighted; when clicked, they take you to other images and related text, a kind of connection known as a hyperlink. Text marked with an asterisk can be clicked on to call up a box containing a more in-depth explanation of the topic.

Hyperlinks between artists and historical movements create an expanding network of pathways designed so that the user can always find a little more information on any specific artist, work or theme. This interactive mode of presentation offers multiple ways to look at a painting and helps to create a sense that the works are all related in some way. Another useful and distinctively multimedia feature in *Art Gallery* is the audio icon that, when clicked, gives the pronunciation of the artist's name. It is a little like having your history teacher on hand. The electronic voice is welcome but inconsistent: it tells you how to say only the last name of Jan Van

Eyck (is the first name pronounced JAN or YAHN?) but gives both names in other cases (Harmen Steenwyck).

Art Gallery also suffers some more serious pedagogic gaps. For instance, Frans Hals, who painted in a surprising, abstract style in some of his very representational portraits, is passed over in the guided tours without any comment. And if you searched under surrealism, you would find works by Bosch, Picasso and Goya that have little or no connection to surrealist imagery. In general, the material on the disk is probably too conceptually thin to enable a serious-minded user to discover and prove a hypothesis about art history. Nevertheless, the disk manages to cover a lot of artistic territory.

Where *Art Gallery* aims for breadth, *American Visions* shoots for specificity, to good effect. Like the Frick Collection, the Roy R. Neuberger Collection (located at the State University of New York at Purchase) came together through the will of a single man. But unlike the Frick CD-ROM, *American Visions* explains why the works were collected and tells stories about them. The Neuberger disk uses an interactive museum approach that feels much friendlier than the Frick's image-database model. This disk also imparts more of a sense of place than the other two because it includes anecdotal video clips of artists and critics.

Starting up *American Visions* takes the user to a theatrical opening, replete with a jazz soundtrack and a collage of paintings, faces, dates and names that gradually covers the screen. Once in, you can take an audio tour of how to use the disk, watch a brief movie about Roy Neuberger (you learn that his focus was on "buying the work of living artists"—particularly in Paris in the 1920s because "good times cause things to happen in the arts") or enter the database and explore the collection. For users dazed by the amount of material available, the most reassuring philosophical statement in the audio tour comes at the end when the narrator says, "And don't forget to turn your computer off at least once a week."

The disk's graphic interface is attractive but somewhat confusing. Rows of similar-looking thumbnail buttons represent many different things: the artworks, photographs of the artists, anecdotal movies, text about the works, artists and social events of the period. The user can uncover more buttons (and hence more choices) by clicking on arrows at the end of the rows. Once you have chosen a painting, for instance, you can request to see links, which calls up a vertical column of new buttons next to the work. This setup of hyperlinks creates a great many ways to circulate through the collection, but the material on the CD-ROM is finite, so you will soon run into the same material from a number of different routes.

The generous use of sound and video motion in *American Visions* quickly pulls the viewer inside the collection.

AMERICAN VISIONS *provides direct access to paintings, video and text.*

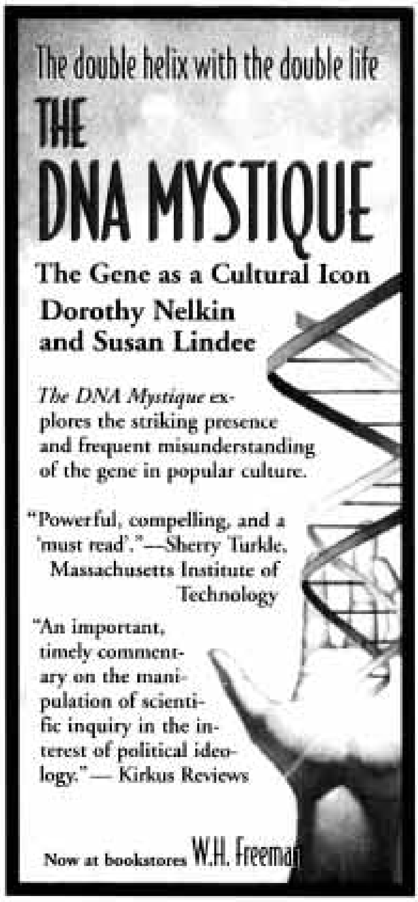
The disk's most novel feature is the instant accessibility of movies about the artists and their works—a resource that would not be available in a real museum and one not provided by the other disks. The movie clips, sometimes accompanied by snippets of jazz, range from color interviews with Neuberger to old black-and-white films, such as one of Jackson Pollack painting. The CD-ROM contains many more movies about artists than about individual pieces, so it gives a rather personality-driven sense of art history.

A set of icons stretched across the bottom of the screen permits the user to reorient the virtual pathways connecting the information in the disk. These icons provide access to a timeline, a viewer for looking at individual artworks or an alphabetical listing of artists. The options designed into *American Visions* allow you to gain a genuine sense of discovery and a feeling for the ways in which Neuberger's style of collecting was influenced by the artists' sensibilities, the mood of the times and various social events.

The structural approach taken by the Neuberger CD-ROM is very much in tune with the present state of multimedia technology. The fairly formal presentations in the Frick and the National Gallery disks proffer the standard view of art as a composition- and technique-centered activity. The more dynamic Neuberger disk brings the work into direct contact with its social context. That intellectual stance is reinforced by the articles by Peter Samis of the San Francisco Museum of Modern Art included in *American Visions.* Samis's articles clearly describe contemporary structural criticism, which holds that art can be fully understood only when it is linked to other forms of expression. CD-ROM multimedia, which explicitly synchronize art and technology, offer an apt means for elucidating the structuralist line of thought.

The publication of these three virtual museums denotes an important moment in the history of art. The medium is clearly still in its infancy; these CD-ROMs miss much of the inner essence of the artworks. But that very failure may help the user appreciate the mysterious power of art that the technology cannot yet replicate. For now, the magic of the museum experience can be explored only by going to the actual paintings—being there, breathing the same air as the art.

BEN DAVIS is a research associate at the Center for Educational Computing Initiatives at the Massachusetts Institute of Technology.



ESSAY by William J. Mitchell

The Parable of the Pizza Parlor

Question: Why didn't the software agent show up at the party?
Answer: Because it had no body to go with.

This nerdy variant on the old Halloween joke about the skeleton neatly encapsulates some commonly expressed hopes and flip-side fears about the Information Superhighway—that it will inexorably replace transportation with telecommunication, face-to-face meetings with transactions in cyberspace and human secretaries and assistants with disembodied software agents. In reality, a robust infobahn is likely to produce a considerably more complex and subtle redistribution of functions among buildings, transportation systems and computer networks. Let me illustrate this point with a homely story I call "The Parable of the Pizza Parlor."

Not so long ago pizza parlors were mostly found on Main Street. They had advertising signs out front to pull in customers, counters where those customers placed orders and handed over cash, kitchens where pizza was baked and an eating space containing tables and seats. All these components were wrapped up in one small building.

In the era of the automobile, a competing configuration emerged. The pizza parlor (by now perhaps part of a chain) did not rely just on its sign; it also advertised in the Yellow Pages and in the mass media. It moved from Main Street to a location beside the highway, and it acquired a parking lot. Many customers now telephoned in their orders and had them delivered by car to their homes and offices. Transportation and telecommunication systems began to play significant roles in the pizza parlor's workings, and its architectural unity fragmented as consumption shifted from a single seating area to the many different locations where customers ate their delivered pizzas.

Main Street began to die as the pizza parlor and other businesses left for more attractive sites. Soon the old, familiar Main Street was no longer the place where people went to hang out.

Sometime in the mid-1990s the pizza parlor went on-line on the World Wide Web. The street address turned into a network address, and the counter became a screen display that allowed a customer at any computer terminal to design a virtual pizza and pay using some form of digital cash. The kitchen transmuted into a nationwide collection of food preparation centers at locations carefully selected to provide maximum coverage of the market. Each order was automatically routed to the preparation point nearest to the customer. There the local kitchen produced and packaged the pizza, which was then delivered via a radio-controlled vehicle.

ADAM FERNANDEZ

The old advertising signs gradually disappeared, and soon there was no newspaper or television advertising either. Instead customers were attracted through graphic "storefronts" in on-line virtual malls and through network "Yellow Pages" listings.

The electronic pizza parlor, some future observer might note, was a big hit. Pizza suppliers reached a much larger market than ever before, and because customers now consumed the product at home, the suppliers did not have to build and maintain restaurant facilities in expensive locations.

The customers also liked the new setup; they could always get exactly what they wanted, quickly, reliably and inexpensively. But they sometimes missed the atmosphere of the ramshackle old parlors, the conversations that unfolded there and the opportunities that the pizzerias afforded to get out of the house and feel like part of a local community.

The pizza parlor may seem like a minor institution, but this story illustrates a more general trend. Using sophisticated telecommunications, for example, office workers may now telecommute from home or simply rely on personal electronic devices to set up virtual workplaces anywhere. Retailers that combine on-line interactive catalogues with direct delivery from the warehouse can now compete with pedestrian shopping streets and automobile-oriented malls. And increasingly effective distance-learning and telemedicine systems are reducing the need to go to school or visit the doctor's office.

In short, cities will be transformed as the Information Superhighway develops. We will have to rethink the spatial relationships, transportation connections and telecommunication linkages among homes, workplaces and service providers. Housing will have to be reconfigured as a wide range of transactions that once took place elsewhere moves back into domestic space.

The weakening or disappearance of traditional gathering places will require the creation of different foci for community life—both physical places and on-line, virtual locations. Offices, hospitals, schools and shopping centers will fragment and recombine in surprising ways as virtual transactions and telepresence relax traditional requirements of proximity. Perhaps reinvigorated local communities will cluster around restaurants, parks and health clubs while also benefiting from strong electronic connections to the wider world.

This restructuring will take place on a massive scale. Depending on the design and policy choices we make in the coming years, it could produce more equitable access to services and economic opportunities, or it could yield electronically serviced islands of privilege surrounded by zones of disinvestment, unemployment and poverty. The stakes are huge.

Cities will certainly not disappear as an increasing amount of human interaction shifts into cyberspace; they will evolve into complex hybrids of physical spaces and on-line locations. They will have places where you need a body to go and places where you don't.

WILLIAM J. MITCHELL is dean of the School of Architecture and Planning at the Massachusetts Institute of Technology. He is author of City of Bits: Space, Place, and the Infobahn *(MIT Press), also available on-line at http://www-mitpress.mit.edu/City_of_Bits/*