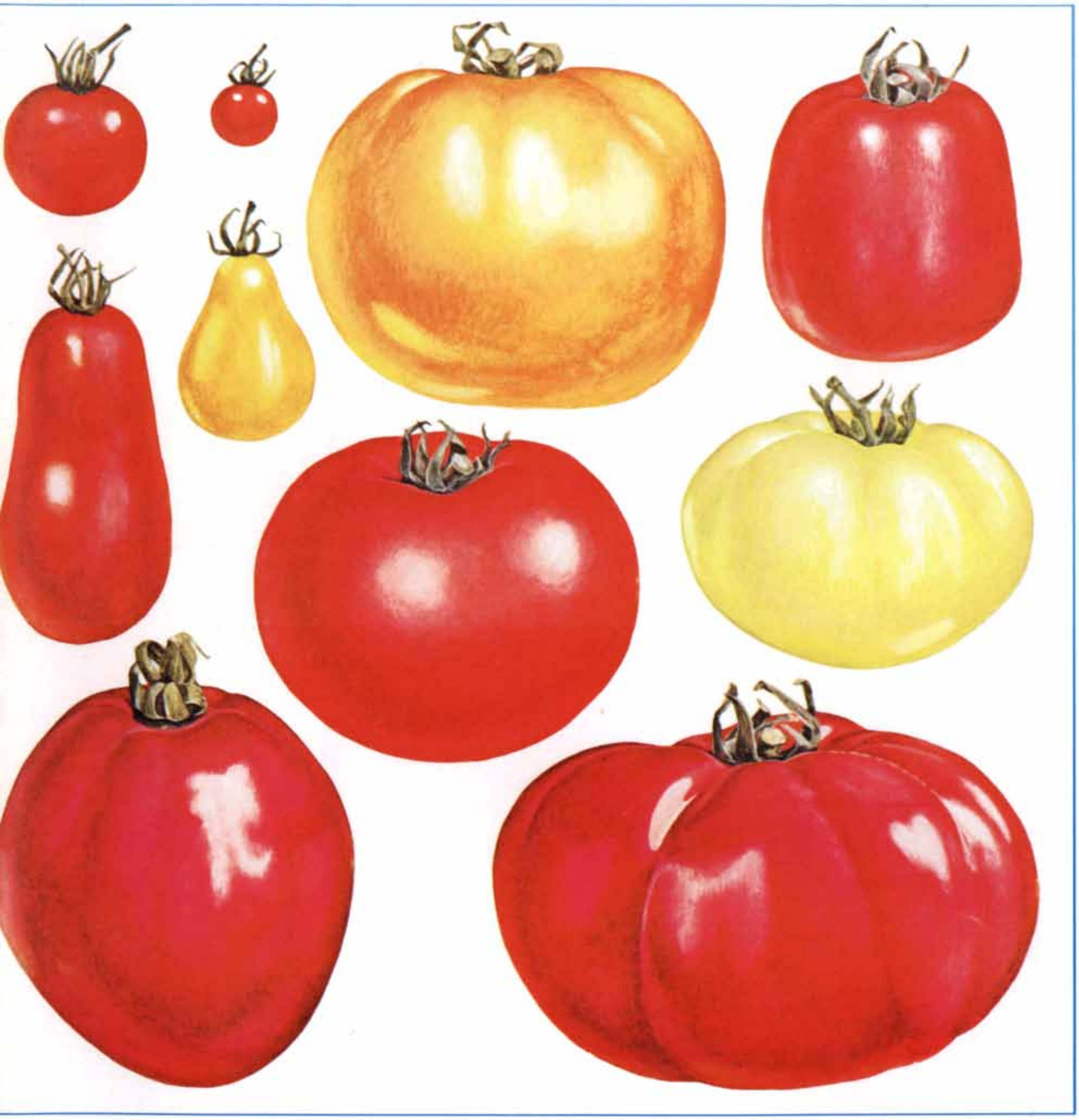
SCIENTIFIC AMERICAN



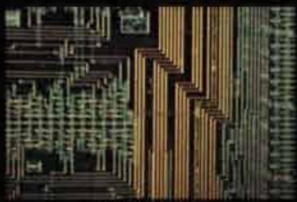
THE TOMATO

$1.50

August 1978

hexa·photo·cybernetic

Technology this advanced deserves its own word.



The incredible new Canon A-1. It's on the cutting edge of technology, with state-of-the-art electronics and computerized operation.

Yet, although the way we build it is complex, the way you use it is the ultimate in simplicity. If pushbutton-easy fine photography is what you're after, the A-1 can give it to you six ways.

1 Shutter Priority automation: You pick the speed to freeze or blur action or prevent camera shake, the A-1 picks the perfect aperture automatically.

2 Aperture Priority automation: Select the aperture you want for a soft background or total overall sharpness, the A-1 will give you the speed that's correct.

3 Programmed automation: On the A-1, this unique exposure mode picks both aperture and speed in ideal combinations depending on the light. You don't have to set *anything*.

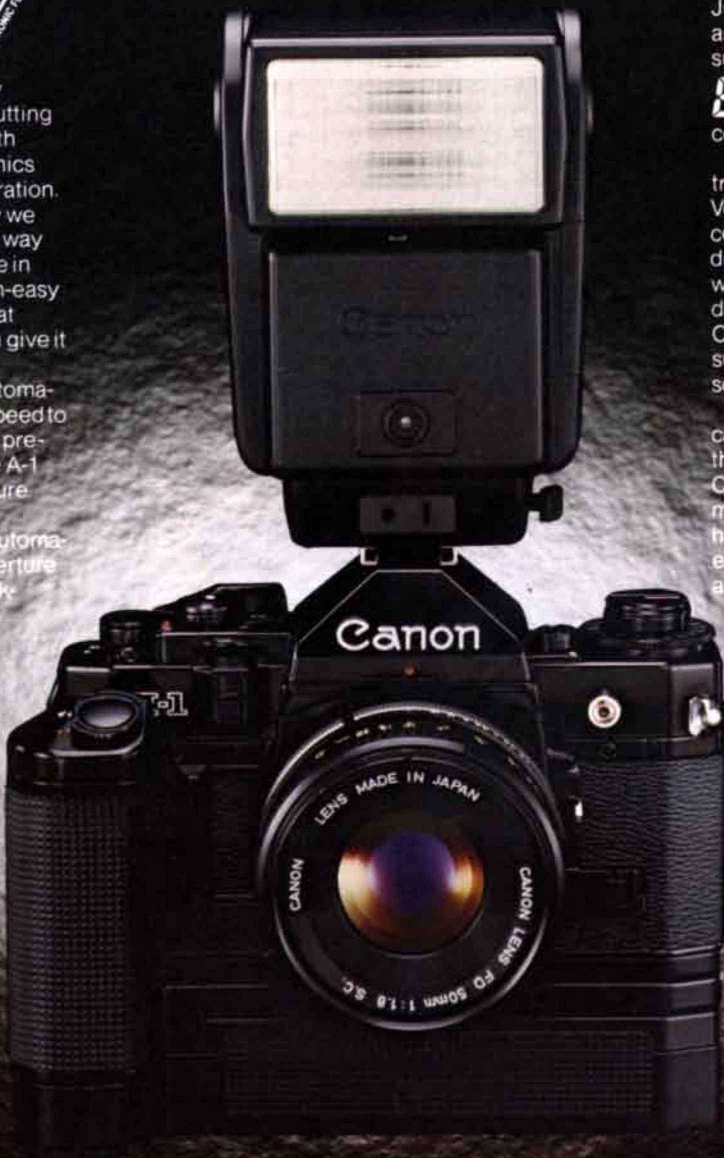
4 Stopped-Down automation: This mode lets you get automatic exposure with any lens or accessory you can mount on the camera. Great for astrophotography, photomicrography, close-ups.

5 Flash automation: The A-1 has the easiest electronic flash of any camera of its type. With special Canon flash units, your aperture *and* speed are set automatically. Just mount the flash, turn it on and shoot for perfect exposures every time.

6 For any type of photography that's not already covered, there's a manual.

There's more. Digital control and viewfinder display. Versatile control features. Accessories like a rugged motor drive, compact motorized film winder and interchangeable date/data back. Plus forty Canon FD lenses for unsurpassed image quality and scope.

You won't find another camera like it on the face of the earth. Because only Canon has the technology to make it. Luckily, you don't have to go to the ends of the earth to get an A-1. Only as far as your local Canon dealer.



THE OFFICIAL 35MM CAMERA OF THE 1980 OLYMPIC WINTER GAMES

Canon A-1

So advanced, it's simple.

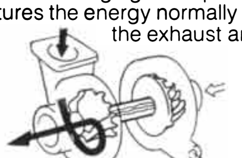
Canon USA, Inc., 10 Nevada Drive, Lake Success, New York 11040
140 Industrial Drive, Elmhurst, Illinois 60126 • 123 Paularino Avenue East, Costa Mesa, California 92626
Bldg. B-2 1050 Ala Moana Blvd., Honolulu, Hawaii 96814 • Canon Optics & Business Machines Canada, Ltd., Ontario

Your passion for power matched to our passion for engineering.

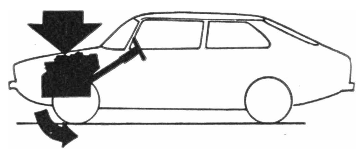
Driving the Saab Turbo puts the power of the future in your hands. The experience is both visceral and cerebral. Even a pro like Car and Driver's Steve Thompson felt it. He said, "...Saab Turbo...a genuine Fourth-of-July driving experience, full of sudden pleasure and high spirit..."

The Turbo is the culmination of years of our engineering passion. We'll highlight just a few of the many engineering details of this amazing machine.

The Saab Turbo-Charged System:
Turbo-charging is the process that captures the energy normally lost through the exhaust and puts the energy back into the work process to obtain extra power.

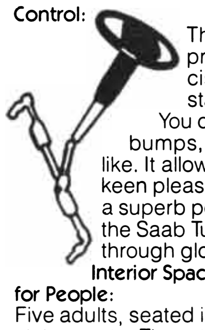


Our breakthrough was designing a turbo-charging system able to function at low speeds. That means our 2-litre, overhead cam fuel-injected engine has extra power for you to command in daily driving situations. To pass, a surge of turbo power shoots you ahead. Or when you're moving on to a busy highway, count on turbo power to supply the big thrust.



We Pioneered Front Wheel Drive:
We realized that front wheel drive supplied the superior traction needed for all sorts of road surfaces; snow, mud, etc. It's been a standard feature on all our cars. And an integral feature of the Saab Turbo.

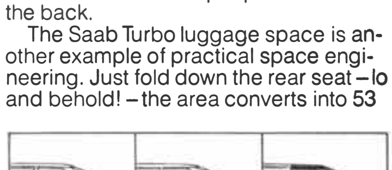
Rack and Pinion Steering for Confident Control:



This steering system provides direct, precise control with instant wheel response. You can react instantly to bumps, potholes and the like. It allows you to feel the keen pleasure of maneuvering a superb power machine like the Saab Turbo, confidently, through glove-tight turns.

Interior Space Honestly Designed for People:
Five adults, seated in the Saab Turbo, are sitting pretty. The orthopedically-created seats and innovative interior design are proof we know that people need plenty of room for their feet, shoulders, elbows and heads. Even the people who sit in the back.

The Saab Turbo luggage space is another example of practical space engineering. Just fold down the rear seat – lo and behold! – the area converts into 53



cubic feet of cargo space. That's practically like having a station wagon at your command.

We manifested our passion by creating the finest performance production turbo for every day driving. It's your turn. Unleash your passion. At least, test the Saab Turbo.

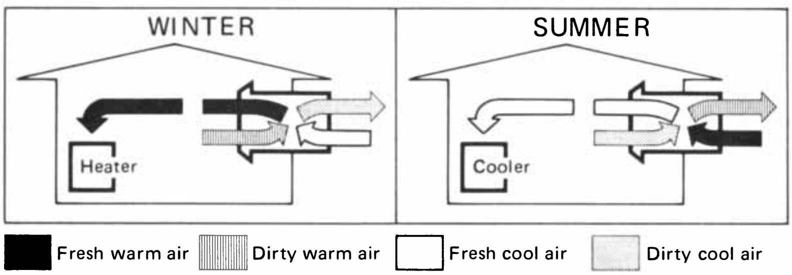
The Saab Turbo: $10,498*

*The manufacturers suggested retail P.O.E. price for the SAAB Turbo includes dealer preparation: taxes, title, destination charges and options are additional.

All SAAB models are available through our Tourist Delivery Plan. Ask your local SAAB dealer for more information.

SAAB IS THE COMMAND PERFORMANCE CAR

Announcing an inexpensive practical method for conserving energy in summer and winter.



Mitsubishi Electric has come up with a series of elegantly simple total-heat exchangers that the average home, building, or factory owner can easily afford.

Capable of exchanging both sensible heat (temperature) and latent heat (humidity), these Lossnays effectively use the permeating properties of specially treated paper.

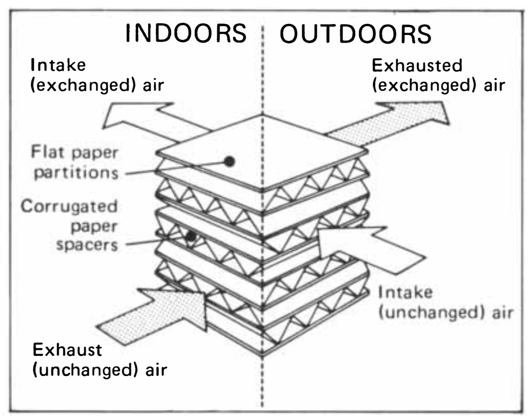
The basic element is constructed entirely of paper, is totally static, and remarkably easy to maintain.

No mixing of intake and exhaust air occurs in a Lossnay heat exchanger. And yet the average rate of total-heat exchange is a high 70%. With the price of energy today, it's definitely worth getting the exact figures for your building.

Mitsubishi makes dozens of models to accommodate any size facility, with air-flow capacities ranging from 60 to 60,000CFM.

Moreover, they're sold at prices you can afford.

For more information, please write to us at the address below. Lossnays will be also shown at the 6th International Heat Transfer Conference and Industrial Exposition, Sheraton Centre, Toronto, August 7—11, 1978.



Mitsubishi Lossnay Heat exchangers efficiently conserve heating and cooling energy normally lost through ventilation. The Lossnay system is a simple, practical way to reduce energy costs with a low initial investment.

MITSUBISHI ELECTRIC

For more information, contact MITSUBISHI ELECTRIC CORPORATION, 2-3, Marunouchi 2-chome, Chiyoda-ku, Tokyo 100. TELEX: J24532 CABLE: MELCO TOKYO

7806-X01-A9-38

Established 1845 SCIENTIFIC AMERICAN August 1978 Volume 239 Number 2

ARTICLES

DEPARTMENTS

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y 10017 COPYRIGHT © 1978 BY SCIENTIFIC AMERICAN, INC. ALL RIGHTS RESERVED. PRINTED IN THE U.S.A. 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. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $18 PER YEAR, U.S., ITS POSSESSIONS AND CANADA; $22 PER YEAR, ALL OTHER COUNTRIES.

executive health

the report that briefs you on what to watch

Volume XIV, Number 10 • July, 1978 • Pickfair Bldg., Rancho Santa Fe, Calif. 92067 • Area 714:756-2600

New research findings . . .

ON WALKING . . . nature's own amazing *"anti-age antibiotic"!*

Not running, not jogging, but *walking* is your most *efficient* form of exercise and the *only* one you can *safely* follow *all* the years of your life!

Today, after some 2000 years, modern medical researchers are demonstrating (somewhat to their surprise) that Hippocrates, "the father of medicine" was right . . . walking is man's best *"medicine"!*

Take one example. At a three-day conference on *Exercise in Aging—Its Role in Prevention of Physical Decline* (held October 27-29, 1977, at the National Institute of Health, Bethesda, Maryland) researchers from across the United States, Canada and Western Europe presented papers on this (until now) largely neglected area of research. As their papers were presented, these important points of agreement emerged:

(1) Walking is the most *efficient* form of exercise . . . and the only one you can safely follow all the years of your life.

(2) Exercise can enable your body to maintain a *vital reserve* which has a protective effect during stress.

(3) Exercised *bones* do not demineralize. As a result they are far less likely to break or lose their range of motion.

(4) Exercised *lungs* still exhibit the emphysema-like changes of age, but are far less diminished in their capacity compared to the lungs of sedentary people.

(5) Exercised *cardiovascular systems* show a similar maximum preservation of function.

(6) The benefits of exercise in preventing or correcting obesity are striking.

(7) Late-onset diabetes is almost entirely reversible by exercise if you are overweight.

(8) Daily exercise permits greater food intake and better blood circulation, thus improving each body cell's nourishment while preventing obesity.

(9) The physically impaired, particularly the arthritic, can perhaps benefit the most from exercise . . .

"HOW LONG WILL I LIVE?"

Discover what research scientists now know about how to take the best care you can of the *only* body you will ever have!

The errors of our younger years are drafts upon our older years, payable (with interest) some 30 years from date.

The so-called "diseases of old age" are essentially the diseases of 50 to 70 . . . **"the dangerous years!"**

Research scientists find that people who survive these "dangerous years" **successfully** (without acquiring some "chronic" disease such as cancer or heart trouble) are likely to live on for another **healthy** quarter of a century. They seem to have developed what might almost be called an **immunity** to these killers. Why? To find out, some careful studies have now been made of these extraordinarily long-lived individuals.

May we suggest that you get the benefit of these new research studies for yourself? Do as so many thousands of executives do. Subscribe to Executive Health Report. The members of our **Editorial Board** are among the world's most distinguished authorities on preventive medicine. Their wise advice can help you not only live longer but enjoy those extra years! (Note their high qualifications as shown on opposite page.)

Executive Health Report is not sold on newsstands but only by private subscription at $18 a year in the U.S.A. and its possessions. $20 a year in Canada and Mexico. Individual reports (back issues) $1.50 per copy. All other countries $26 by surface mail, $30 by air mail. (Only International Money Order or check cashable on U.S. bank will be acceptable.)

Subscribe now under this unusual introductory offer:

(1) Your choice of any three of the reports listed below ($1.50 each) FREE!

(2) If you are not *completely* satisfied with your first issue, your money will be *promptly* refunded.

Never forget. **"Men's lives are chains of chances,"** but as Euripides saw clearly so long ago, **"Chance fights ever on the side of the prudent."** Your only insurance against "tomorrow" is what you do today. You have but one life . . . doesn't it make sense to find out how to take the best care of it you can?

Please study the reports listed here and circle your three choices:

New research findings On Walking . . . nature's own amazing "anti-age anti-biotic"!

Albert Szent Gyorgyi, M.D., Ph.D.: How new understandings about the biological function of ascorbic acid may profoundly affect our lives.

Linus Pauling, Ph.D.: For the best of health, how much vitamin C do you need? People who take the optimum amount of vitamin C may well have, at each age, only one quarter as much illness and chance of dying as those who do not take extra vitamin C.

George C. Griffith, M.D.: On those irregular heart beats (Arrhythmias). Some mean little or nothing, but others warn your heart is in trouble.

Miles H. Robinson, M.D.: On sugar and white flour . . . The dangerous twins. How, with the best or intentions, we have managed to process natural foods into appetite-tempting, disease-breeding trouble-makers . . .

Roger J. Williams, Ph.D., D.Sc.: On your startling biochemical individuality. Some amazing facts about your body you need to know if you want to understand yourself (and other people) better.

Alexander Leaf, M.D.: On the physical fitness of men who live to a great age.

Linus Pauling, Ph.D.: On vitamin C and heart disease. Can vitamin C protect you, and how much should you take?

James F. Toole, M.D.: On Strokes and "Little Strokes" . . . their causes and what you should know about them to help protect yourself!

Sir Hans Krebs, M.D.: On the Overuse and Misuse of Medication.

Alton Ochsner, M.D.: "On the Role of Vitamins C and E in Medicine." A world-famous surgeon tells you how he uses these two essential vitamins.

John K. Lattimer, M.D., Sc.D.: On That Treacherous Gland . . . Your Prostate.

John Yudkin, M.D.: On "This Sliming Business" . . . The truth about the prevention and cure of overweight!

Linus Pauling, Ph.D.: What About Vitamin E? Eminent investigators now suspect it may be one of the key factors to help resist disease and slow the aging process.

Linus Pauling, Ph.D.: On Vitamin C and Cancer. Recent studies show that vitamin C has a large life-extending effect for patients with advanced cancer and suggest a similar large effect for earlier stages of the disease.

Mark D. Altschule, M.D.: Is It True What They Say About Cholesterol?

Mark D. Altschule, M.D.: What Causes Your Arteries to Harden?

Alton Ochsner, M.D.: On "The Chair Disease" . . . Why blood clots in your veins are a little-realized occupational hazard of desk-bound executives.

Hans Selye, M.D.: On Stress Without Distress. Your mind can make or break you!

Hardin B. Jones, Ph.D.: On The Problems Executives Must Anticipate With the Growth of Marijuana Smoking. The Senior Scientist at the world-renowned Donner Laboratory of Medical Research, University of California, Berkeley, explains some little-understood facts that executives should know as the growth of marijuana use continues among employees.

The B Vitamins . . . Part II. On Vitamin B_6 (Pyridoxine) "The Sleeping Giant of Nutrition."

Stephen R. Elek, M.D.: On "The Hurry-Up Disease." Why it may be a key factor that triggers heart attacks as early as age 40.

Please use the coupon below under our special money-back guarantee.

EXECUTIVE HEALTH, Pickfair Bldg., Rancho Santa Fe, CA 92067

Gentlemen: Enclosed is my check for $.......................for a year's subscription to Executive Health to start with this month's issue. I have circled the three $1.50 reports I am to receive free. It is understood that if I am not *completely* satisfied with my first issue, my money will be *promptly* refunded. In addition, I would appreciate your sending me a complete list of your 101 other reports because among them may be some from which I might greatly benefit and would otherwise miss.

NAME (please print)________________________________

ADDRESS____

STATE/COUNTRY____________________ZIP_________ SA-12

THE COVER

The fruits assembled in the painting on the cover are the mature berries of 10 varieties of the tomato plant (*Lycopersicon esculentum,* except for the tiny currant tomato near the top left, which is *L. pimpinellifolium*). Innovative breeding has helped to make the tomato a major commercial food crop in the U.S. (see "The Tomato," by Charles M. Rick, page 76), but most of the tomatoes on the cover are grown by home gardeners. Cultivars in top row are (*left to right*) cherry and currant tomatoes, the orange Jubilee and the "square round" processing tomato, which has been developed recently for mechanized harvesting. In the second row are the San Marzano paste variety, Yellow Pear, a typical commercial or home-garden F_1 hybrid and the pale yellow Snowball. In the bottom row are Oxheart and Ponderosa, two so-called beefsteak cultivars. Colors that differ from the normal red color are controlled by specific recessive genes: *r,* for yellow flesh, in Yellow Pear and Snowball; *t,* for orange flesh, in Jubilee, and *y,* for colorless skin, in Snowball, Oxheart and Ponderosa.

THE ILLUSTRATIONS

Cover painting by Enid Kotschnig

Page	Source
18	James Young
19–22	Ilil Arbel
24	Ilil Arbel (*top*), Ikuo Sakurai (*bottom*)
25	Ilil Arbel
28	Jerome Kuhl
38	U.S. Department of Energy
39–43	Allen Beechel
45	C. Roger Lynds, Leo Goldberg and Jack Harvey
46	Armin J. Deutsch
47–48	Dan Todd
49	Ray J. Weymann
50	Dan Todd
51	Ray J. Weymann
52	Dan Todd
55	Koobi Fora Research Project
56	National Aeronautics and Space Administration
57–58	Bunji Tagawa
59–61	Tom Prentiss
62–63	Bunji Tagawa
64–66	Tom Prentiss
72	James E. Harris
76	Enid Kotschnig
78–79	Ilil Arbel
80	Charles M. Rick (*top*), Ilil Arbel (*bottom*)
81–82	Ilil Arbel
84	The John Henry Company (*top*), Ilil Arbel (*bottom*)
86	Ilil Arbel
90–99	Gabor Kiss
103	Patricia J. Wynne
104	Patricia J. Wynne (*top*), Albert Miller (*bottom*)
105–112	Albert Miller
117	John A. O'Keefe
118	Adolph E. Brotman
119	Institute of Geology, U.S.S.R. Academy of Sciences (*top*); Robert F. Fudali (*bottom*)
120	NERC ©, Institute of Geological Sciences, London (*top*); Smithsonian Institution (*middle*); Billy P. Glass (*bottom*)
121	Adolph E. Brotman
122	Dean R. Chapman
123–125	Adolph E. Brotman
129–130	Lorelle M. Raboni
131	Albert H. Markhart III (*top*), Lorelle M. Raboni (*bottom*)
132–134	Lorelle M. Raboni
135	Ben Rose
139	Lorelle M. Raboni
140–143	Michael Goodman

LETTERS

Sirs:

In their article "Deinstitutionalization and Mental Health Services" [SCIENTIFIC AMERICAN, February] Dr. Ellen L. Bassuk and Mr. Samuel Gerson discuss a very important problem and indicate quite accurately that this issue is very complex, involving the interaction of social, political and economic forces. Some of their assumptions, interpretations and explanations, however, are questionable and some of their data are misleading.

There is the implication that the Community Mental Health Centers (CMHC) Construction Act of 1963 and the subsequent development of comprehensive CMHC's to implement it are responsible for the large numbers of patients discharged from state mental hospitals. Furthermore, the authors seem to imply that CMHC's were created to take care of these discharges in the community and have failed to do so. In actuality, as the authors indicate, the trend toward increased discharges from state mental hospitals started in 1955. This was approximately eight years before legislation was enacted and 11 years before the first CMHC was operational.

The authors at times refer to community mental health care as a philosophy and at other times as a CMHC. They indicate that the philosophy of community-based mental health care is responsible for the deinstitutionalization that produced the increase in discharges from state mental hospitals. The assumption is questionable. The primary reason for the increased discharges was probably financial. With the advent of Medicare and Medicaid, and recently Supplemental Security Income (Title XX), the cost of caring for the mental patient could be shifted from state support to Federal support. The cause of deinstitutionalization is therefore probably more economic than philosophical.

The objective of community-based services is to serve persons where they live. CMHC's serve those persons who reside within specifically designated areas (catchment areas), and for those persons CMHC's have proved to be a viable alternative to state mental hospital care. A number of studies have shown that the rate of first admissions and readmissions to state mental hospitals from catchment areas with CMHC's is lower than it is from other areas in the same state. Thus CMHC's, where they exist, are reducing the flow of patients into state mental hospitals. The authors of the article obscure this in their discussion of admissions by not differentiating between areas served by CMHC's and other areas.

Federal, state and local funds for CMHC's support mental health treatment, not housing, welfare and other community services. CMHC's are not expected to care primarily for the person being returned from the state mental hospital to the community. Their mandate is much broader and is based on the catchment-area concept. They must attend to the full range of mental health needs of all ages and conditions of persons residing in their service area. Indeed, one of the reasons the National Institute of Mental Health developed the Community Support Program, referred to by the authors, is the recognition that CMHC's could not become comprehensive psychosocial rehabilitation centers. The support system in the community that former mental hospital patients require includes shelter and social and economic assistance as well as intensive clinical service.

The main criticism of the article must be directed at its pessimism about community-based mental health care that is not supported by data and the authors' failure to differentiate between the Community Mental Health Centers program and community socioeconomic-rehabilitative supports that are beyond the mental health system's authority and responsibility.

If any of your readers are interested in a more detailed critique of the Bassuk and Gerson article, they may obtain it from Fred E. Spaner, Ph.D., Community Mental Health Services Support Branch, Division of Mental Health Service Programs, National Institute of

Scientific American, August, 1978; Vol. 239, No. 2. Published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, president; Dennis Flanagan, vice-president; Donald H. Miller, Jr., vice-president and secretary; George S. Conn, treasurer; Arlene Wright, assistant treasurer.

Editorial correspondence should be addressed to The Editors, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Manuscripts are submitted at the author's risk and will not be returned unless accompanied by postage.

Advertising correspondence should be addressed to C. John Kirby, Advertising Director, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

Offprint correspondence and orders should be addressed to W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104. For each offprint ordered please enclose 40 cents.

Subscription correspondence should be addressed to Subscription Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. For change of address, notify us at least four weeks in advance. Send both old and new addresses and enclose an address imprint from a recent issue. (Date of last issue on your subscription is shown at upper right-hand corner of each month's mailing label.)

Name ____________________

New Address ____________________

____

____

Old Address ____________________

____

____

Mental Health, Room 11C-06, 5600 Fishers Lane, Rockville, Md. 20852.

FRED E. SPANER, PH.D.

STEVEN S. SHARFSTEIN, M.D.

CARL A. TAUBE

National Institute of Mental Health
Rockville, Md.

Sirs:

In the few months since our article "Deinstitutionalization and Mental Health Services" appeared we have received a broad range of comments about the successes and failures of community-based care for formerly hospitalized mental patients. Although the responses reflect diverse opinions and ideologies, all of them share a passionate commitment to one or another treatment approach. Increasingly these positions are being argued in the political arena and in the national media of communication. Under the glare of this attention it is not surprising that discussions about deinstitutionalization and community-based care tend to be polarized into positions of advocacy and opposition. In such a process any recognition of substantial problems is perceived as a condemnation of an entire approach by those whose focus is limited to the positive elements. Furthermore, there seems to be a growing tendency for all parties involved in the delivery of mental health services to maintain that areas of neglect are not their responsibility. The letter from Dr. Spaner, Dr. Sharfstein and Mr. Taube appears to us to be a reflection of these trends. Their comments are shaped throughout by disclaimers of responsibility for the growth of deinstitutionalization and for the care of discharged mental patients. They begin their critique by noting that since the census of hospitalized patients began to decrease eight years prior to the passage of the Community Mental Health Centers Construction Act in 1963, deinstitutionalization proceeded independently of the community mental health movement. They further state that deinstitutionalization was a function of economic realities rather than philosophical changes in the approach to treatment. To us both of these arguments appear specious. First, the bulk of deinstitutionalization (85 percent) occurred after 1963. Second, we believe that massive releases of patients from mental hospitals could not have been effected without a receptive theoretical framework among mental health care planners. This framework, in its broadest outlines, is that processes of chronicity are environmentally based and could therefore be ameliorated by means of a shift in the locus of treatment and alterations in the patients' social network. This view was, and remains, a fundamental aspect of the community mental health movement, and as such it continues to offer a theoretical basis for deinstitutionalization.

As we stated in our article and in our reply to the letter from Dr. John C. Wolfe ["Letters," SCIENTIFIC AMERICAN, June], it is in the translation of this theoretical perspective into an operational system of community-based services that major difficulties have occurred. Dr. Spaner, Dr. Sharfstein and Mr. Taube, in failing to acknowledge any problem areas, would have us believe all is well. They refer to a number of studies showing how well-functioning CMHC's have reduced the number of state-hospital admissions, and they take deserved credit for these successes. They do not, however, mention the relatively low number of such facilities in relation to the overall need, nor do they speak of their shortcomings in community mental health services for severely disturbed individuals. Instead they argue that the basic life-support needs of the chronically disabled patients go beyond their domain. Although we are probably in general agreement on this point, it should be noted that the current problems facing deinstitutionalized and chronically disabled patients exist in part because the theoretical basis of the community mental health movement has led mental health planners to believe improvement could occur solely through a new approach to treatment. It is therefore ironic that Dr. Spaner, Dr. Sharfstein and Mr. Taube decry our "pessimism" about this assumption when they concur in the view that the mental health care system alone is unable to adequately provide for the needs of chronically disabled mental patients.

It appears to us that Dr. Spaner, Dr. Sharfstein and Mr. Taube perceive our discussion as an attack on the entire community mental health network of services rather than as a call for a recognition of the fact that theories and polemics have not met, and cannot adequately meet, the myriad needs of severely disturbed individuals. As Dr. Wolfe stated in his letter, the time has come for all involved in the care of the chronically disabled to assume a corporate responsibility. We believe this can best be realized by an acknowledgment of current and previous areas of neglect rather than by statements that seek to obscure these difficulties and take any mention of them as "pessimistic."

ELLEN L. BASSUK

Harvard Medical School
Boston

SAMUEL GERSON

Cambridge Hospital
Cambridge, Mass.

Technics
by Panasonic

Introducing the Technics SA-1000. With more power and less distortion than any other receiver we've made: 330 watts per channel minimum RMS into eight ohms from 20 Hz-20 kHz with no more than 0.03% total harmonic distortion.

But that's only one reason to buy the SA-1000. Dynamic range is another. To capture the volume, clarity and sheer dynamics of a live symphony, you need an equally dynamic amplifier section. Like 72,000 μF worth of high-capacitance filtering, separate DC rectifiers, current-mirror loading and direct coupling. The results are impressive: tremendous reserve power, negligible transient crosstalk distortion and excellent stability.

And just for the record, the SA-1000's phono equalizer gives you everything from a super-high S/N ratio of 97 dB (10 mV, IHF A). To a phono input that can handle a 300 mV signal at 1 kHz.

On FM you'll get outstanding specs plus two RF stages with low-noise, 4-pole, dual-gate MOS FETs, Technics-developed flat group delay filters and a Phase Locked Loop IC in the MPX section.

FM Sensitivity		FM Selectivity	Stereo Separation at 1 kHz
IHF '58	Stereo-50 dB*		
0.9 μV	36.2 dBf	85 dB	50 dB

*IHF '75 standard

As good as all that sounds, Technics Acoustic Control makes it sound even better, because it adds low and high range boost and filter switches which vary the way each tone control performs at a particular setting. There's also a midrange control with a variable center frequency. And 24 LED peak-power indicators that let you keep an eye on what your ears will hear.

The Technics SA-1000. In the world of receivers, it bats 1000.

A few receivers give you 0.03% THD. Only Technics gives it to you with 330 watts per channel.

makes even those numbers pale. That old Diesel stigma of feeble torque and leisurely pickup has vanished in a burst of turbocharged energy. This is one Diesel that can break away from stoplights and tollbooths *with* the traffic, not behind it. That extra thrust you need on freeway entrance ramps can be found. You can sustain a normal driving pace on the Interstates, hour after hour.

Mercedes-Benz test data show that although 400 pounds heavier than its lively 300 D Sedan stablemate, the Turbodiesel can zip from zero to 55 mph in 2.6 seconds *less* time – placing it among the quickest Diesel cars in history.

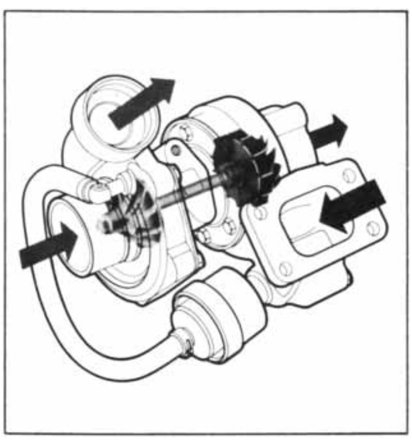
No minor feat for an engine of only three liters or 183.0 cubic inches in capacity; yet no great surprise for an engine so efficient that it produces *.601 horsepower* per cubic inch of displacement – the best ratio of power to engine size of any Diesel passenger car power plant in the world.

Fuel appetite down

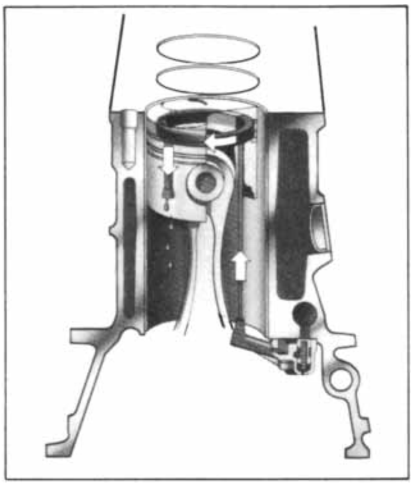
Startling as it may seem, this extra performance helped reduce the Turbodiesel engine's normal appetite for fuel.

Turbocharging so handily solved the Diesel need for power that it freed the engineers to specify a more economical rear-axle ratio – in effect, gearing the car to go further on the power produced by a given gallon of fuel.

Consider the Turbodiesel's performance. Then consider that EPA



Amazing turbocharger device weighs only 17 pounds but boosts horsepower by 43 percent.



Each piston in the 5-cylinder Turbodiesel engine is cooled by a fine spray of oil injected from below.

estimates show 29 mpg in highway driving and 24 mpg in the city. Naturally, your mileage will depend on the condition and equipment of your car and on where and how you drive.

Record-breaking reliability

This breakthrough has hardly been rushed to the market. Mercedes-Benz placed the 300 SD Turbodiesel engine in production only after 5 years of testing – and only after placing it in the reliability record books.

Fitted with a modified version of this engine, a C-111-3 research car went out on April 30, 1978 and set nine world records – including one stint of 2,345 miles in twelve hours at an average speed of 195.39 mph. For the entire record run, the engineers report a fuel mileage figure of 14.7 mpg.

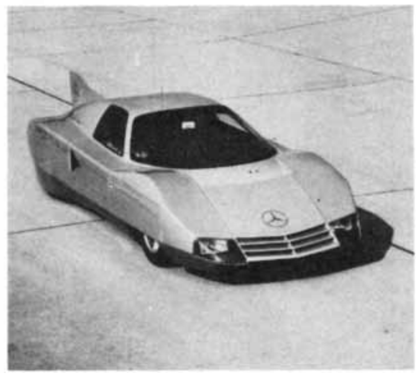
A car apart

The Diesel turbocharging trail blazed by Mercedes-Benz with the Turbodiesel may some day be followed by others. But no Diesel engine will ever share the privilege of propelling a comparable car. It is the unique combination of that engine and this automobile that truly sets the Turbodiesel apart.

As befits the most elegant Diesel Mercedes-Benz has ever built, the Turbodiesel sits on the longest wheelbase of any Diesel car produced by the company in modern times. Yet its turning circle is a tight 38 feet and its crisp handling lets you nip through traffic.

Fastidious engineering is everywhere. You glide along on a suspension that is neither spongy nor harsh, but designed to provide both ride comfort *and* roadholding. The key is the independent suspension of all four wheels, allowing each wheel to individually react to the road surface.

Each shock absorber is gas pressurized to help cushion even minor ripples. The automatic transmission provides not three but four speeds, and the option of shifting for yourself if you prefer. From a monocoque body shell to 11-inch disc brakes at all four wheels, nothing from the vast store of Mercedes-Benz technical expertise has been held back.



This turbocharged five-cylinder Mercedes-Benz Diesel research car just shattered nine world speed records – including one lap at an average speed of 203.37 mph.

Inside, you and your passengers are a coddled group: surrounded by thick padding, velour carpet underfoot, cradled in spacious seats front and rear. Bi-level climate control, electric windows, automatic cruise control, central vacuum locking system and AM/FM stereo radio are all built in.

Also built into your 300 SD Turbodiesel: *120 safety features,* perhaps the most attractive fact of them all.

Summing it up

In 1886, the gasoline-powered automobile was pioneered by the two men who founded Mercedes-Benz. In 1936 came the world's first production Diesel passenger car, again from Mercedes-Benz. And now in 1978 begins the era of the turbocharged Diesel passenger car.

And once again, the innovator is Mercedes-Benz.


THE SEAGRAM'S GIN
"WIMBLEDON WHISTLE."
Seagram's Gin and Schweppes Bitter Lemon...A Great Match!
Schweppes
Schweppes
Bitter Lemon
Seagram's
Extra Dry
Gin
Pour 1½ oz. Seagram's Extra Dry Gin
over cubes, fill with Schweppes Bitter Lemon, garnish
with orange slice, and you're ready to serve an ace.
Seagram's Extra Dry. The Perfect Martini Gin. Perfect all ways.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

AUGUST, 1928: "Two trends are likely to have robbed the railroads of all but a fraction of their passenger traffic before another decade has rolled around. Since 1920, although the nation has increased 13 per cent in population, there has been a decline in railroad passenger traffic of about a third. This is attributable to the private automobile and more recently the motor coach. The trend is now toward the highway for local travel, while the long journeys continue to be made by rail. But do we now espy on the distant horizon the competitor that is destined also to take away a share of the remaining long-distance passenger hauling? What of the airplane now coming on apace? Overland journeys are soon to be made alternately by rail (at night) and by airplane (in daylight), the entire coast-to-coast journey thus being cut to 48 hours. How much longer will it be at the present rate of progress in commercial aviation before the entire journey will regularly be accomplished in a single leap?"

"According to recent press notices, interplanetary transportation must remain impossible until atomic energy can be obtained and controlled. This attitude is much like that of the scientists of 30 years ago, who declared that an airplane could not operate unless the force of gravity could be neutralized. If atomic energy were available, it would be a very convenient means of propelling an interplanetary rocket. Atomic energy is not, however, necessary, as an interplanetary flight is possible with means even now at our disposal. If a propellant of high-energy content, such as hydrogen together with oxygen, is used with high efficiency and in the proper way, an interplanetary flight is possible in a rocket that is neither tremendously bulky nor unwieldy."

"The death within one year, at Akkra in South Africa, of three workers in yellow fever adds to the long roll of the martyrs of medicine, to the list of physicians and investigators who have died as a result of disease incurred during research for methods to relieve suffering and to prevent death. Last September, Dr. Adrian Stokes died of yellow fever at only 40 years of age. He had already contributed greatly to medicine, particularly in the field of bacteriology, and he had shown the possibility of inoculating rhesus monkeys with yellow fever. On May 21 Dr. Hideyo Noguchi, a noted Japanese investigator, died of yellow fever in the same place while in the midst of researches attempting to identify the South American form of the disease. On May 30 Dr. William Alexander Young, a third member of the investigating party, also succumbed to yellow fever."

"Within recent years great strides have been made in the perfection of the chromium-plating process. Almost 75 years have elapsed since Robert Wilhelm Bunsen first obtained metallic chromium by the electrolysis of chromic acid, but it is only since the World War that the conditions necessary for obtaining a satisfactory commercial coating of the metal have been worked out. Metallic chromium ranks next to diamond in hardness, and it is many times harder than the best steel. It presents a high resistance to corrosive and tarnishing influences such as all alkalies, oxygen, chlorine up to 300 degrees centigrade, sulfur, superheated steam and concentrated nitric and sulfuric acids. These qualities, together with its platinum-like finish and high polish, make it a valuable plating metal."

"Evidently the 'talking movie' is with us to stay. A research branch of the Western Electric Company has recently announced that contracts have been signed indicating that the major motion-picture producers in the country will adopt one form or another of talking movie. But deliver us from the speaking voice of some of the present-day motion-picture stars! What a sad disillusionment it is to see the sweet face of a popular star, only to hear when she speaks the coarse, harsh voice of the forewoman of the local hat factory. Will the advent of the talking movie mean that a new set of stars will rise in the firmament of moviedom? Or will 'ghost talkers' spring up in the industry?"

SCIENTIFIC AMERICAN

AUGUST, 1878: "There are two principal systems proposed for aërial navigation: first, devices lighter than air, or aërostats, which rise and maintain themselves at a height by reason of their very low density, so that it is merely necessary to guide and move them, and second, apparatus heavier than air, which must be raised, moved and guided mechanically. This latter class has two subdivisions: the aëroplane, consisting essentially of a plane moved by some kind of motor (such as a screw or some other device) and sustained by the resistance of the air to its movement, and apparatus endeavoring to imitate exactly the flight of a bird. We shall now suppose an aëroplane weighing 1,000 kilograms and moving horizontally with a velocity of 10 meters per second. To have the least dead weight it must work at an angle of 54°44′, and then the work of translation would be 14,088.69 kilograms, or about 187 horse power. Such a system would demand a propelling force much greater than is possible with known or probable motors. What can American inventive genius and practical skill do in this important direction?"

"The fifth publication of Behm and Wagner's well-known *Population of the Earth* makes the number of the earth's human inhabitants for the current year 1,439,145,300, an increase of 15 million over the estimate of last year. The increase is attributed partly to natural growth and partly to exacter knowledge resulting from recent censuses. The distribution of the population among the grand geographical divisions is as follows: Europe, 312,398,480; Asia, 831,000,000; Africa, 205,219,500; Australia and Polynesia, 4,411,300; America, 86,116,000."

"An argument for the Texas Pacific Railroad draws at great length upon the vast extent of territory to be developed by the road. All of western Texas, all of New Mexico, Arizona, southern Nevada, southern Utah and a large portion of northwestern Mexico would be tributary to the road, an area of about half a million square miles in extent. The part within our own national bounds is equal in size to Germany and France combined, or enough to make 10 states the size of New York, in all of which there is not a single mile of railway. The assertion that there was not enough arable land left in the region to make a good-sized county in Wisconsin is resented as libel; a large part is unquestionably sterile, yet there remains a very considerable area of the highest fertility. The valleys of New Mexico, Arizona and the bordering states of Mexico are exceedingly productive when irrigated or where the supply of moisture is sufficient. Areas along the Union Pacific, previously supposed to be beyond redemption, now bear abdundant crops; a like effect would follow the building of the southern road."

"Great consternation has been caused in the oil regions of Pennsylvania by the recent decline in the price of petroleum to less than a dollar per barrel. It is generally admitted that these ruinous prices are a natural and unavoidable result of the immense overproduction, and in some quarters the belief is expressed that if there is no other way of curbing the desire to sink new oil wells more rapidly than new markets can be found, a still further decline will in the end prove beneficial by warning all whom it may concern of the folly of glutting the market with excessive quantities of a product such as petroleum."

THE AUTHORS

DONALD R. OLANDER ("The Gas Centrifuge") is professor of nuclear engineering at the University of California at Berkeley. He did his undergraduate work at Columbia University, where he obtained his bachelor's degrees in chemistry and chemical engineering, the latter in 1954. He then spent a year at the University of Paris studying chemical physics on a Fulbright fellowship. On his return he entered the doctoral program in chemical engineering at the Massachusetts Institute of Technology and did thesis research on the reprocessing of nuclear fuels. After receiving his Ph.D. in 1958, he joined the chemical engineering department at Berkeley, moving to the newly created nuclear engineering department in 1961.

RAY J. WEYMANN ("Stellar Winds") is professor of astronomy at the University of Arizona. He received his bachelor's degree from the California Institute of Technology in 1956 and his Ph.D. in astronomy from Princeton University in 1959. Having done his thesis research on the theory of mass loss from red-giant stars, he pursued the subject observationally for the next two years as a postdoctoral fellow at the Hale Observatories. He then joined the Steward Observatory of the University of Arizona, where he served as director and head of the department of astronomy from 1970 to 1975. His past research interests have included intergalactic matter and Seyfert galaxies (about which he wrote an article for SCIENTIFIC AMERICAN), and he is currently interested in the origins and properties of the absorption lines in the spectra of quasars. Weymann has also been heavily involved in the joint Smithsonian-Arizona multiple-mirror-telescope project.

ALAN WALKER and RICHARD E. F. LEAKEY ("The Hominids of East Turkana") have collaborated on efforts to reconstruct and interpret the fossil skulls of ancient hominids. Walker is professor of cell biology and anatomy at the Johns Hopkins School of Medicine. Born in England, he studied geology and zoology at the University of Cambridge and obtained his Ph.D. from the University of London. In 1965 he went to East Africa to teach anatomy at the Makerere Medical School in Uganda. From there he moved to the University of Nairobi Medical School, where he began his collaboration with Leakey. Since 1973 Walker has taught anatomy in the U.S., but he continues to make frequent visits to Kenya to work on the East Turkana fossils. Leakey, the son of Louis and Mary Leakey, is director of the National Museums of Kenya and chairman of the Foundation for Research into the Origins of Man. His first expedition to the East Turkana fossil localities was in 1968. Since then the Koobi Fora Research Project has expanded into an international team of investigators.

CHARLES M. RICK ("The Tomato") is professor in the department of vegetable crops at the College of Agricultural and Environmental Sciences of the University of California at Davis. He was graduated from Pennsylvania State University, and after a summer working as an assistant plant breeder for the W. Atlee Burpee Company in California he went on to do graduate work in plant genetics at Harvard University. After receiving his Ph.D. in 1940, he joined the Davis faculty, where he has remained except for trips to the Andes to collect tomato varieties and do research on them. Rick writes: "I was bitten by the gardening bug in my grade school days and have not recovered since. I remember the excitement I felt when placing my first mail order for seeds and how I did everything wrong in sowing and handling the seedlings. Later, while still at home, I spent every available hour working in the orchard my father managed. It was clear even then that my career would deal with plants."

WARREN G. PROCTOR ("Negative Absolute Temperatures") is guest scientist at the Center for Nuclear Studies at Grenoble. He did his undergraduate work in electrical engineering at the California Institute of Technology and obtained his Ph.D. in physics from Stanford University in 1950. After a few years on the faculty of the University of Washington he joined Varian Associates in Palo Alto, where over the next 20 years his research focused on investigations of nuclear magnetic resonance in solids. This latter period was interspersed with leaves of absence at various institutions, and it was during a sojourn at the Saclay Nuclear Research Center that the work described in his article was done in collaboration with Anatole Abragam. "In recent years," Proctor writes, "I can best be described as a scientific vagabond, having just finished a two-year appointment as guest professor at the University of Helsinki, where I became very interested in the lifetimes of excited states of impurities in semiconductors. I am currently working on this problem here in Grenoble."

H. CRAIG HELLER, LARRY I. CRAWSHAW and HAROLD T. HAMMEL ("The Thermostat of Vertebrate Animals") met when Heller and Crawshaw were postdoctoral fellows studying temperature regulation in Hammel's laboratory at the Scripps Institution of Oceanography of the University of California at San Diego. Heller is associate professor of biology at Stanford University. He was graduated from Ursinus College in Pennsylvania and received his Ph.D. from Yale University in 1970. Crawshaw is associate professor of biology at Portland State University in Oregon. He did his undergraduate work at the University of California at Los Angeles and obtained his Ph.D. from the University of California at Santa Barbara in 1970. Hammel is professor of physiology at the Scripps Institution and associate director of its Physiological Research Laboratory. He received his bachelor's degree in physics from Purdue University in 1943 and then joined the Los Alamos Scientific Laboratory, where from 1944 to 1946 he was junior physicist in the Manhattan Engineer District. He then did graduate work at Cornell University, obtaining his master's degree in physics and his Ph.D. in zoology, the latter in 1953. After eight years on the faculty of the University of Pennsylvania School of Medicine and six years at the Yale School of Medicine, he joined the Scripps Institution.

JOHN A. O'KEEFE ("The Tektite Problem") is a geophysicist at the Goddard Space Flight Center of the National Aeronautics and Space Administration. He did his undergraduate work at Harvard College and received his Ph.D. in astronomy from the University of Chicago in 1941 with a thesis on dark clouds in our galaxy. After two years as professor of physical sciences at Brenau College in Gainesville, Ga., he joined the Army Corps of Engineers as a mathematician specializing in geodesy (the study of the size and shape of the earth) and served as chief of the Research and Analysis Branch of the U.S. Army Map Service. In 1958 he joined NASA as assistant chief of the Theoretical Division, and soon afterward he discovered the pear-shaped component (the third zonal harmonic) of the earth's shape. O'Keefe has been fascinated with tektites since 1954 and has edited one book and written another on their nature and possible origin.

STEVEN VOGEL ("Organisms That Capture Currents") is associate professor of zoology at Duke University. He obtained his bachelor's degree from Tufts University and his Ph.D. from Harvard University in 1966 with a thesis on insect flight; he then joined the Duke faculty. Vogel writes: "My interest in flight expanded to a general concern with the adaptations of organisms to the moving fluids around them and to my current fascination with induced flow and ventilatory processes in sponges, leaves, burrows and mounds."

MATHEMATICAL GAMES

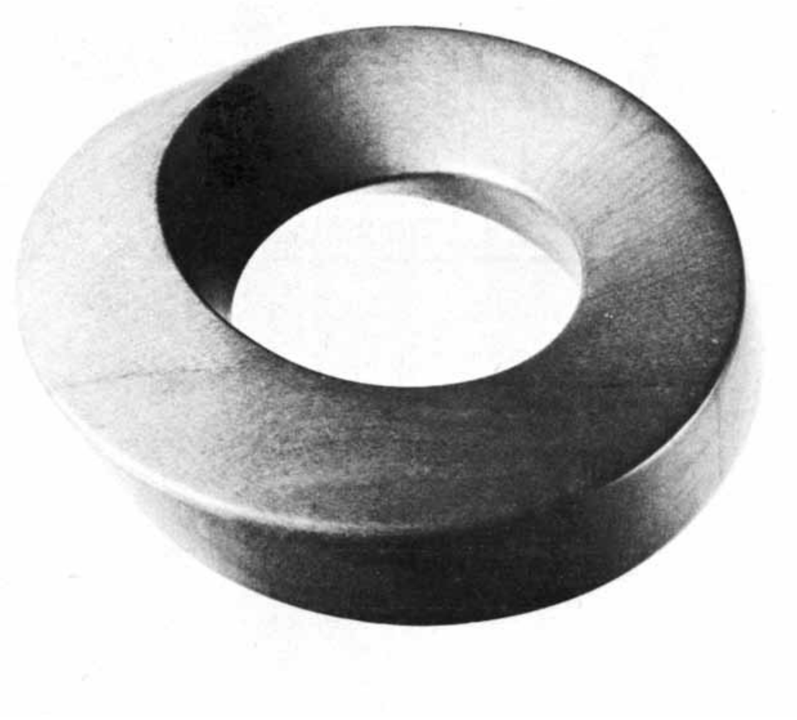
A Möbius band has a finite thickness, and so it is actually a twisted prism

by Martin Gardner

The well-known Möbius band, formed by giving a paper strip a half twist and joining the ends, is a model of an abstract surface of zero thickness. No paper, however, is zero thick. The cross section of a Möbius band actually is a rectangle, very much longer than it is wide. The band itself may be regarded as a four-sided prism, twisted so that one end has rotated 180 degrees before the two ends are joined. Viewed in this way it is a solid ring with two distinct "faces." One face is the flat surface of the band, which circles the ring twice. The other is the band's narrow but also flat edge, which also circles the ring twice.

Over the past 15 years I have received dozens of letters from readers who independently noticed that models of Möbius bands actually are twisted prisms and who generalized such rings by considering prisms with cross sections that are regular polygons with any number of sides. Although such structures have many strange properties, surprisingly little is known about them. Indeed, the only reference I know in English is the 1972 article "Some Novel Möbius Strips," listed in the bibliography of this issue [*page 148*].

There is no agreed-on name for these structures, and so let us call them prismatic rings. Let n be the number of sides of the polygon cross section, and let k be the number of $1/n$ turns the prism is given before its ends are joined. If the prism is not twisted, then $k = 0$. If it is twisted (in either direction) so that each side joins an adjacent side, then $k = 1$. If each side joins the next side but one, then $k = 2$, and so on.



A twisted prismatic ring with one face and one edge

The easiest prismatic ring to visualize is the ring with a square cross section. If $k = 0$ (no twists), the ring obviously has four sides and four edges. If $k = 1$, we get the beautiful solid in the illustration on this page. (It is a photograph of a wood carving by Roger I. Canfield, who sent it to me after reading my December 1968 column on Möbius bands.) Like the Möbius surface, it has only one "face" and only one edge.

I first encountered the $n = 4$ ring about 25 years ago in a science-fiction story. I have forgotten the author, but as I remember the episode it involved a man who found himself walking inside a prismatic square ring that had been given one twist. There was no gravity, but the man wore shoes that attached him magnetically to the floor. He noticed a mark on the floor, and he was puzzled when he returned to the same place and found the mark on one wall. Another circuit put the mark on the ceiling, a third put it on the other wall and a fourth brought it back to the floor again.

When the square prism is twisted twice, it becomes topologically equivalent to the familiar Möbius band viewed as a solid: two-faced and two-edged. Three twists (like one twist the other way) produce one face and one edge, and four twists bring the structure back to four faces and four edges. On all prismatic rings the faces equal the edges in number, and so henceforth we shall consider faces only. The sequence of numbers for the faces repeats periodically when the twists exceed n. Thus a square prism with five twists has the same number of faces as one with a single twist. Note also that all twisted rings are mirror-asymmetric and therefore have mirror-image counterparts.

Let us generalize to prismatic rings with cross sections that are regular polygons of n sides. Given k twists it is easy to predict the number of faces. It is the GCD (greatest common divisor) of n and k. From this fact several interesting properties follow. If the cross section has a prime number of sides, the number of faces is n only when k (the number of twists) is 0 or any multiple of n. Otherwise the ring has only one face. If n is not prime, the ring has one face only when n and k are prime to each other (have no common divisor). The table at the top of the opposite page (sent to me in 1964 by John Steefel) gives the number of faces on prismatic rings with cross sections that have two through 15 sides and that have been given zero through 15 twists.

Note that the Möbius strip appears here as a degenerate prismatic ring with a two-sided cross section.

Now the fun begins. We all know that crazy things happen to twisted bands when they are cut down the middle. Equally crazy things happen when twisted prismatic rings are cut in various ways. The middle illustration on this page shows eight ways a ring with a square cross section can be cut.

Consider the first way, a simple bisection down the middle. If $k = 0$, the result obviously is two separate rings, each with four faces and no twists. If $k = 1$, the cutting goes twice around the ring, and the result is the same as if two perpendicular cuts are made as shown in the second diagram in the middle illustration. This operation forms a single ring four times as long as the original, with four faces and 16 twists of 90 degrees each.

When $k = 2$, the cutting forms a single ring with four faces and eight twists, but now the ring is only twice the size of the original. When $k = 3$, the cutting goes around a second time to produce a knotted single ring four times as long as the original. It has four faces and 24 twists. Two interlocked rings result when $k = 4$, each the same size as the original and each with four faces and four twists.

The cases of $k = 2$ and $k = 4$ can be modeled with paper strips, viewing them as solids. As we have seen, when $k = 2$, we have the familiar Möbius band, and $k = 4$ is a paper band with two half twists. Simply cut the strips down the middle, then examine the results, remembering that the edges are considered faces. To experiment with $k = 1$ or $k = 3$ an actual solid model is helpful. As Charles J. Matthews suggests in the article mentioned above, the simplest way to make such a model is with "salt ceramic": a mixture of one cup of table salt, half a cup of cornstarch and three-fourths of a cup of cold water. Put it in a double boiler, heat it and stir it until it thickens and follows the spoon. Let it cool on wax paper. Knead out the lumps and shape the substance into strips about eight inches long, with four faces about half an inch wide. Form into the desired ring, smooth out the cracks with water and let dry. If you like, Matthews writes, you can paint each face a different color before you start cutting.

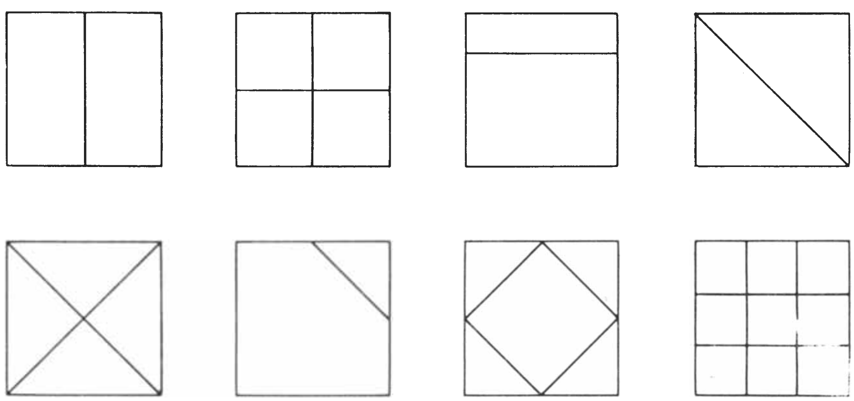
As we have seen, when $k = 1$ or $k = 3$, the methods of cutting in the first two diagrams are the same. When $k = 2$ or $k = 4$, they are not the same. Here again we can model the double cut of the second diagram easily with paper strips. Simply put one strip on another and give the double strip either one half twist or two before joining the edges. Such strips, viewed as solids, are the same as rings with two or four quarter twists that have been bisected. Cutting these rings down the middle is then the same as adding a second cut perpendicular to the first one. In this way we see that if $k = 2$, and the ring is bisected both ways, the result is two interlocked rings (one ring is twisted twice around the other), each four-faced, twice as long as the original and twisted eight times. If $k = 4$, the result is four rings, all interlocked. Each is the same size as the original, with four faces and four twists.

Interested readers may wish to exper-

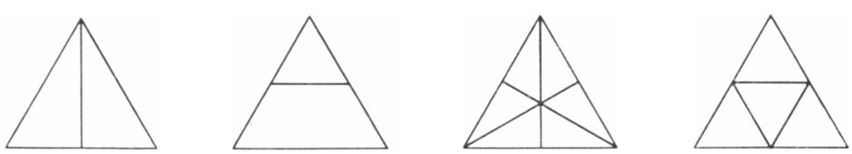
k = NUMBER OF TWISTS; *n* = SIDES OF CROSS SECTION

n \ k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1
3	3	1	1	3	1	1	3	1	1	3	1	1	3	1	1	3
4	4	1	2	1	4	1	2	1	4	1	2	1	4	1	2	1
5	5	1	1	1	1	5	1	1	1	1	5	1	1	1	1	5
6	6	1	2	3	2	1	6	1	2	3	2	1	6	1	2	3
7	7	1	1	1	1	1	1	7	1	1	1	1	1	1	7	1
8	8	1	2	1	4	1	2	1	8	1	2	1	4	1	2	1
9	9	1	1	3	1	1	3	1	1	9	1	1	3	1	1	3
10	10	1	2	1	2	5	2	1	2	1	10	1	2	1	2	5
11	11	1	1	1	1	1	1	1	1	1	1	11	1	1	1	1
12	12	1	2	3	4	1	6	1	4	3	2	1	12	1	2	3
13	13	1	1	1	1	1	1	1	1	1	1	1	1	13	1	1
14	14	1	2	1	2	1	2	7	2	1	2	1	2	1	14	1
15	15	1	1	3	1	5	3	1	1	3	5	1	3	1	1	5

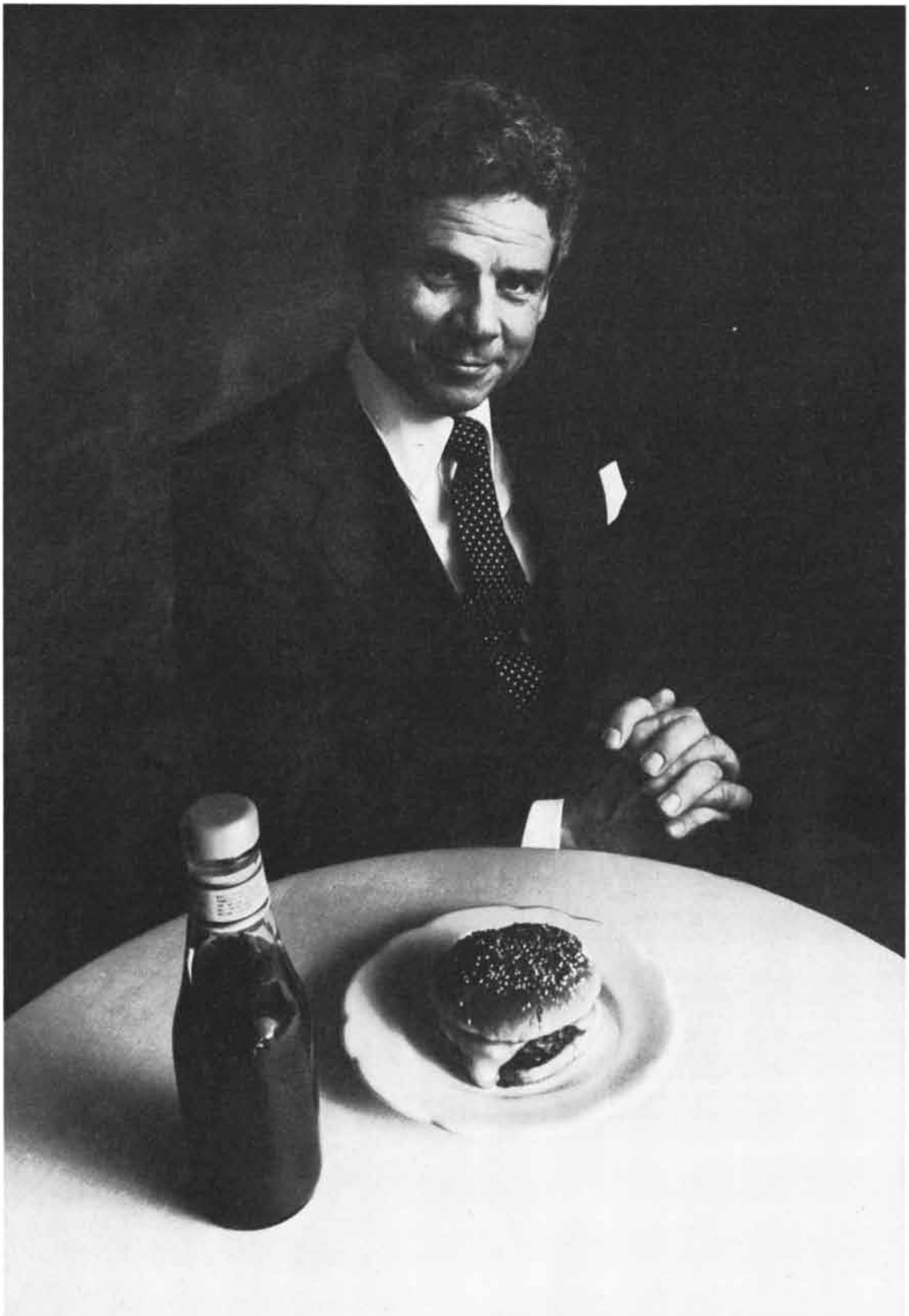
A table showing the number of faces when n *and* k *are known*



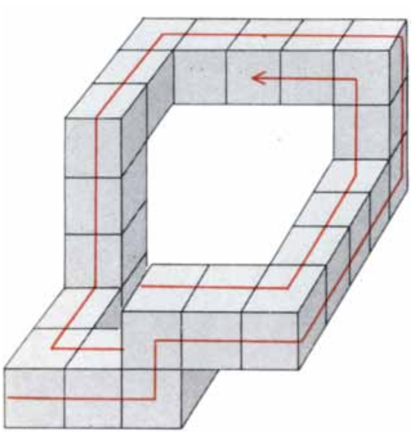
Some ways to cut an n = *4 ring*



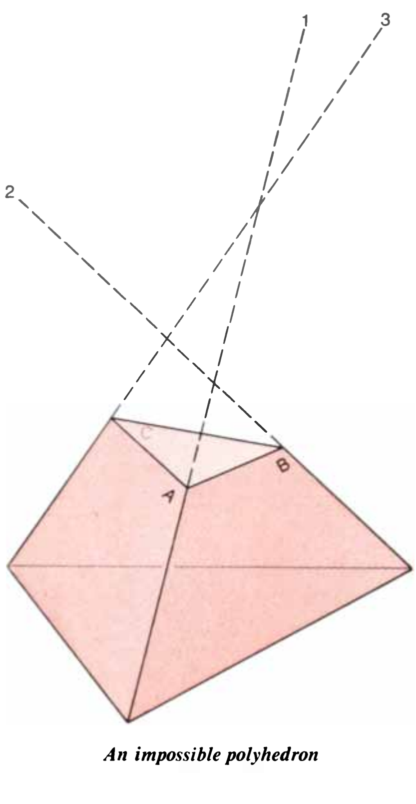
Some ways to cut an n = *3 ring*



iment with some higher values for k. If $k = 6$, for example, and the ring is bisected, the result is a single ring tied in an overhand knot. Also open to exploration are the other ways of cutting the $n = 4$ ring, and the cutting of rings with cross sections that are other than square. Triangular-cross-section rings are easily made with salt ceramic, and the bottom illustration on page 19 shows four ways to cut them. Rings with pentagonal and still more complex cross sections are too difficult to model and cut, but they can be investigated, as Matthews explains, by using appropriate diagrams.



A prismatic one-sided ring made with 22 cubes



An impossible polyhedron

Are there formulas that, given n and k and the method of cutting, will predict the number of twists in the resulting ring or rings? Undoubtedly there are, but I know of nothing published along such lines.

So far we have considered only prismatic rings bent into circular shapes. We can, however, form them in such a way that they are toroidal polyhedrons. What we have called a ring's "face" will then consist of flat four-sided polygons that are joined along their edges in a circular chain. For example, the upper illustration on this page shows a prismatic ring ($n = 4$) of this type. The illustration is based on a photograph in a 1974 article in Spanish on such structures by Gonzalo Vélez Jahn of the architecture department of the Central University of Venezuela. The ring has a single twist; therefore it models a prismatic ring of square cross section that has one face (in our former meaning of "face") and one edge.

Such polyhedral prismatic rings suggest a variety of difficult problems that are only beginning to be explored. Scott Kim, for example, has proved a number of remarkable theorems, not yet published, about polyhedral rings. They are closely related to a class of "impossible objects," such as the Penrose triangle shown at the left in the top illustration on page 24 and its rectangular version at the right.

It is conceivable that figures such as these are not really impossible but are drawings of twisted polyhedral prismatic ($n = 4$) rings. D. A. Huffman, in his paper "Impossible Objects as Nonsense Sentences" (in *Machine Intelligence,* Vol. 6, edited by Bernard Meltzer and Donald Michie, 1971), was the first to devise algorithms for proving whether or not such figures are possible.

There are informal proofs for specific figures. Here, for instance, is Huffman's impossibility proof for the Penrose triangle. We first make some reasonable assumptions:

1. The straight lines in the drawing are straight lines on the actual model.
2. Regions that appear flat in the drawing actually are plane surfaces.
3. Surfaces A and B intersect at line 1, surfaces B and C intersect at line 2 and surfaces C and A intersect at line 3.

Three planes, no two of them parallel, must intersect at a common point P. Therefore each of the three intersection lines must pass through P. Note, however, that lines 1, 2 and 3—the three lines of intersection—cannot meet at one point. Therefore the figure is not possible. It is amazing that this simple proof requires no information about the hidden faces or edges of the picture. A similar argument will show the impossibility of the rectangular "window."

Kim called my attention to the figure in the lower illustration on this page, one of several simple impossible polyhedrons considered by Huffman. Here line 1 is common to surfaces A and B, line 2 is common to surface B and the hidden back surface, and line 3 is common to the hidden surface C and surface A. On the basis of the same argument as before these three lines must meet at a common point. As the dotted extensions indicate, however, they do not. Hence the figure is impossible. It is strange, Kim observes, that this figure looks so clearly possible in contrast to the Penrose triangle, although the reason for the impossibility of the two figures is the same. Huffman develops a general algorithm, based on directed graphs, for testing all such figures.

The problem of distinguishing possible from impossible polyhedral "windows" has also been investigated, along different lines, by Thaddeus M. Cowan, a psychologist. Basing his analysis on braid theory, Cowan has devised a systematic way of generating and classifying such figures and demonstrating various properties. (See "The Theory of Braids and the Analysis of Impossible Figures" in *Journal of Mathematical Psychology,* Vol. 11, No. 3, August, 1974, pages 190–212, and "Organizing the Properties of Impossible Figures" in *Perception,* Vol. 6, 1977, pages 41–56.)

Polyhedral prismatic rings that have square cross sections are easily built from unit cubes. One can rubber-cement together sugar cubes or children's wood blocks, or use the plastic snap-together cubes available from such supply houses for mathematical teaching aids as Creative Publications, 1101 Antonio Road, Mountain View, Calif. 94303. Here is a delightful problem suggested by Kim that makes use of unit cubes. I shall answer it next month.

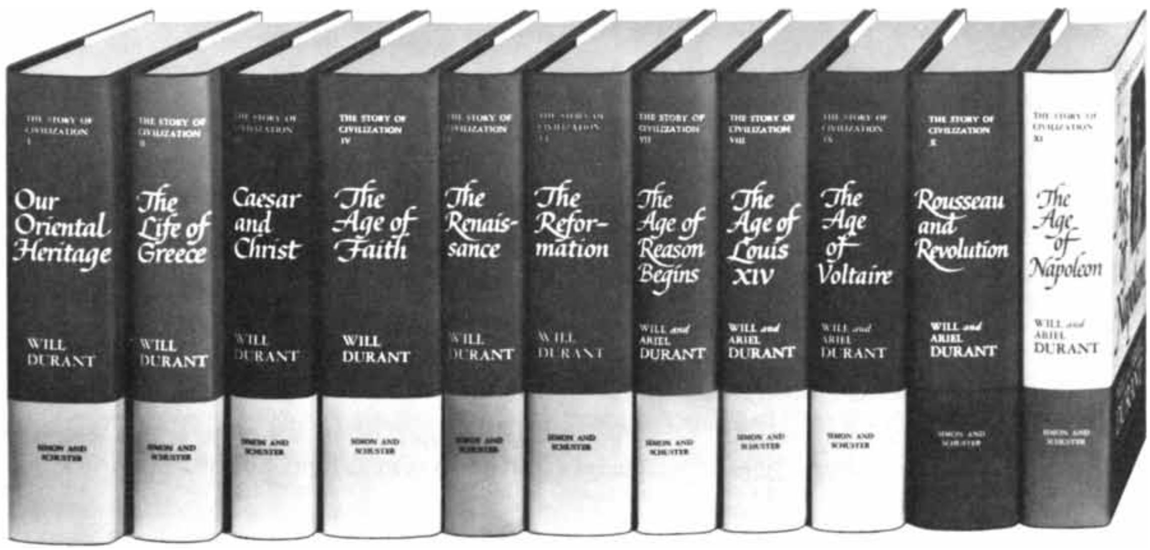
The prismatic ring in the upper illustration on this page clearly can be constructed with 22 unit cubes as indicated. Our problem is to model a figure with the same properties—one "face" and one "edge"—but to do it with the smallest possible number of cubes.

Each cube must have just two faces joined to other cubes to form the prismatic chain. The light lines, which indicate these joins, are not of course part of the model's edge. The single face of the ring circles the polyhedral torus four times, and it is bounded by a single edge that also circles the ring four times. We make one proviso: no point on the edge may touch any other point on the edge. This is to keep the toroid's hole from being partly or entirely closed.

The bottom illustration on page 24 shows a beautiful wood model of a polyhedral ring, of square cross section, in the shape of a triangle. It was made by Ikuo Sakurai of Tokyo. A section of the model can be twisted to four positions.

Rediscover the past in The Story of Civilization.

"Our century has produced no more successful attempt to narrate the whole common story of mankind" —*Clifton Fadiman*



I. Our Oriental Heritage. Ancient Egypt and the Near East. The Far East, early to modern times.

II. The Life of Greece. Explores all facets of Greek life from prehistoric times to the Roman conquest.

III. Caesar and Christ. The rise of Rome and Christianity and the collapse of classic civilization.

IV. The Age of Faith. Christian, Islamic and Judaic civilizations, 325 to 1300, including the Crusades.

V. The Renaissance. Italy's golden age, 1304 to 1576. A turbulent world of intrigue and great art.

VI. The Reformation. Europe's religious conflicts, from two centuries before Luther to Calvin.

VII. The Age of Reason Begins. Europe, 1558-1646. The age of Shakespeare, Rembrandt, Galileo.

VIII. The Age of Louis XIV. The brilliant era of the "Sun King," Milton, Cromwell, Peter the Great.

IX. The Age of Voltaire. Europe from 1715 to 1756. The world of Frederick the Great, Wesley, Bach.

X. Rousseau and Revolution. Europe from the Seven Years' War to the storming of the Bastille.

XI. The Age of Napoleon. France's domination of European history, from the fevers of the French Revolution to Napoleon's defeat at Waterloo. A history of European civilization from 1789 to 1815.

Will and Ariel Durant's masterpiece is now offered as your first benefit of membership in the Book-of-the-Month Club.

The complete 11-volume set: $20

Publisher's price: $248.55

You simply agree to buy four books within a year.

Since the first volume of this remarkable work appeared forty-three years ago, it has been hailed both as a great literary achievement and an epic history of mankind. To read it is to enjoy the past made live with excitement and power. To own it is to have at your fingertips a reference work of undisputed authority. It is hard to imagine a more valuable set of books for your home library.

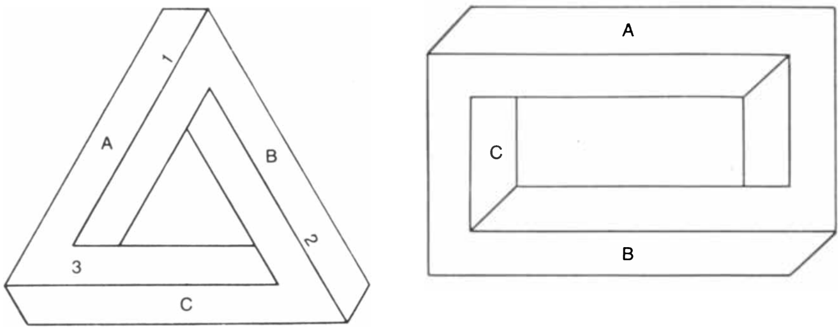
Beginning with this special opportunity to own *The Story of Civilization* at the extraordinary saving of more than 85% (including postage and handling charges), you'll enjoy continuing benefits on the best and most important books published today. As a Club member, you always get bookstore quality at book-club savings. Every book is identical in size, paper, type and binding to those sold in stores. You never settle for the altered or inferior editions that some book clubs send their members. And when you remain a member after the short trial period, each book you buy earns Book-Dividend® credits. These entitle you to enormous savings on current art books, reference sets, literary classics, records, and children's books. Join now, and shop for books the civilized way—in America's Bookstore.

BOOK-OF-THE-MONTH CLUB®

Camp Hill, Pennsylvania 17012

AMERICA'S BOOKSTORE®

Since 1926, 330 million books in 15 million homes.



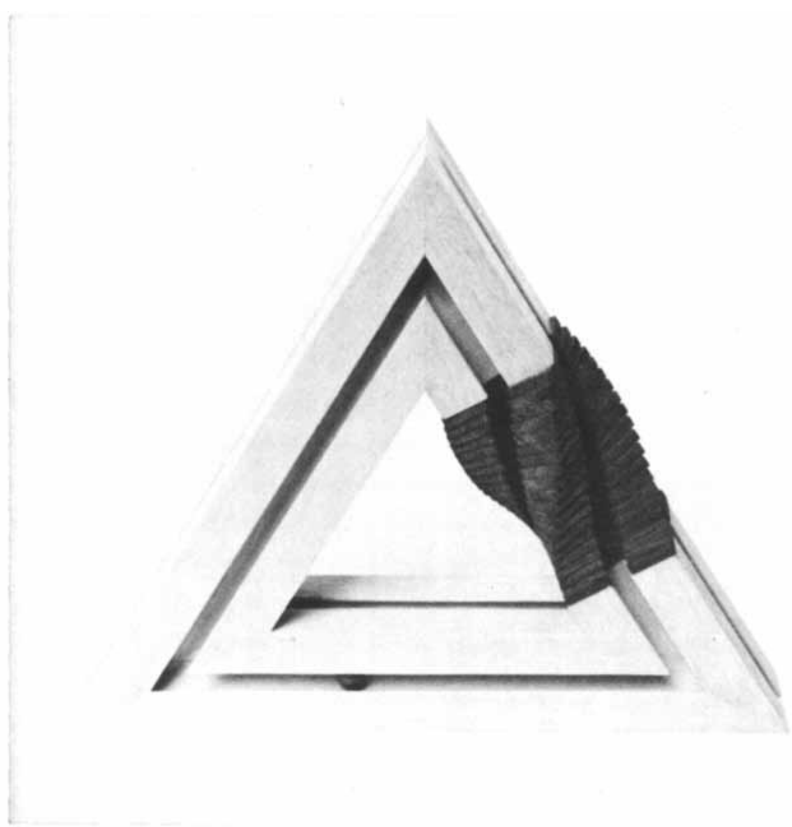
Two impossible polyhedral prismatic rings

Down the middle of each face runs a groove inside which a red marble is trapped. Thus the model can be set for any of its four forms. Then by tipping the triangle you can roll the marble once, twice or four times around the toroid.

Howard P. Lyons once proposed the following prismatic ring. It is hollow with an outside surface that has a square cross section with one twist. The interior hole that circles the ring also has a square cross section with one twist, but the twist goes the other way. In a letter to me Lyons wonders what properties this bizarre ring has. I take my cue from Mark Twain. He once responded to a man who had asked an unusually complicated question about the speed of a cannonball by writing, "I don't know."

The 27-dice problem given in February was to build a cube with the dice so that the sum of the 54 products of the 54 pairs of touching faces is (1) a minimum and (2) a maximum.

Four readers found what almost certainly are the solutions: the minimum is 294 and the maximum is 1,028. The solutions are shown in the illustration on the opposite page. The minimum is unique except for two possible orientations of the central die. The maximum has many variations.



A Japanese model of a polyhedral torus adjusted to display a single side and a single edge

The first reader to send correct solutions was Kenneth Jackson. His computer program was not exhaustive, so that he could not guarantee his answers. A program by Paul Stevens also has this feature. The same solutions were found by hand, also without proof, by Leonard Lupow.

Dave Vanderschel originally tried to solve the problem by computer. After encountering horrendous difficulties in writing a program that would examine all the possibilities in a reasonable time he tackled the problem by hand and eventually found a lengthy, complex proof for 294 and 1,028 that he believes is airtight.

Letters of unusual interest about Bell numbers (my topic for May) are still coming in. I had raised the question of whether the number of Bell primes is infinite or finite. Many readers pointed out that in addition to the first three Bell primes, B_2, B_3 and B_7, the 13th Bell, 27,644,437, is also prime. A conjecture that all Bell primes have prime subscripts was shot down by Vaughan Pratt of the Laboratory for Computer Science at the Massachusetts Institute of Technology. His fast program, which tested Bells through B_{161} (a number of more than 200 digits), found the next two Bell primes, B_{42} and B_{55}. Pratt conjectures that there are infinitely many Bell primes but that he will learn of at most one new one in his lifetime. B_{42} has 38 digits; B_{55} has 54. Incidentally, my statement that B_{100} has 126 digits was incorrect. The number is 116.

Sin Hitotumatu, who translates this department for *Saiensu,* the Japanese edition of *Scientific American,* has provided more details on the Lady Murasaki diagrams. In 1600 Japanese noblemen and ladies played a game called Genji-ko or Monko. An umpire randomly selected sticks of incense from a supply that contained five different kinds. Players sniffed the burning sticks and tried to guess which were the same and which were different. The 52 possible selections were diagrammed by the players as was explained in my May article.

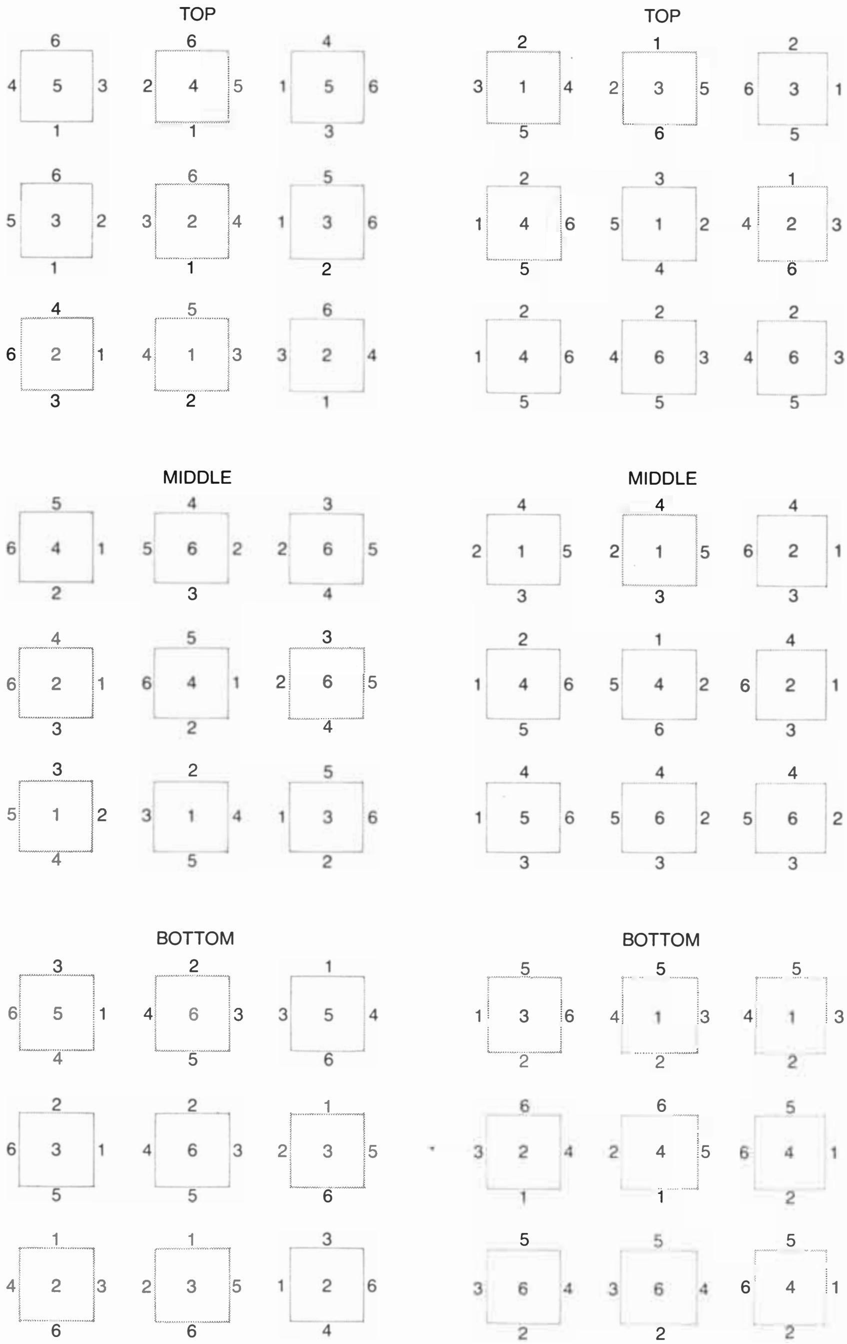
In the early 17th century Japanese mathematicians assigned a mnemonic name to each diagram, using the names of the 52 chapters (between the first and the last) of Lady Murasaki's novel *Tale of Genji.* It is not known whether this assignment of diagrams to chapters was random or was based on some pattern or perhaps a fancied correlation with the characters and events of each chapter. In the late 19th century printed editions of *Tale of Genji* began to carry the Murasaki diagrams as chapter headings.

Andrew Lenard of Indiana University reported success in proving the following curious property of Bell numbers. (The property had been noticed but not established.) The first $2n$ Bells can be arranged in a square matrix for $n = 3$:

$$\begin{matrix} 1 & 1 & 2 & 5 \\ 1 & 2 & 5 & 15 \\ 2 & 5 & 15 & 52 \\ 5 & 15 & 52 & 203 \end{matrix}$$

What is the determinant of such a matrix? It is given by an astonishingly simple formula: $(1!)(2!)(3!)\ldots(n!)$. In this case 203 is the sixth Bell; $n = 3$, and the formula gives $1 \times 2 \times 6 = 12$ as the determinant's value.

Readers who enjoy solving cryptarithms of the SEND + MORE = MONEY type will enjoy Steven Kahan's paperback *Have Some Sums to Solve*. It has just been published by the Baywood Publishing Company of Farmingdale, N.Y., the publishers of the quarterly *Journal of Recreational Mathematics*.



The 27-dice problem. Minimum, 294 (left)*; maximum, 1,028. Bottom numbers are absent*

BOOKS

The art of interpreting artifacts; the state of food and agriculture

by Philip Morrison

EXPLORATIONS IN ETHNOARCHAEOLOGY, edited by Richard A. Gould. University of New Mexico Press ($17.50). Ethnography is where you find it. Informants take shortcuts: ask an amateur handyman what a screwdriver is for and if he has a whimsical turn of mind he may explain how to test for hot circuits in a junction box. Indeed, the little blemishes left on quite a few screwdrivers by arcs will tend to confirm his account. Just this kind of thing happened when an Australian informant, an Aljawara, was asked about a curious class of quartzite blades. They were spoons for eating sweet potatoes, he said. Many of the blades, however, bore marks indicating that their primary function was scraping. This problem was easy, but if you are puzzling out a tool dug from a trench into a deposit going back 10,000 years, that wear mark can deceive.

In this volume (reporting a recent Santa Fe meeting) nine anthropologist-archaeologists offer a nutritious blend of provocative approaches to method and striking specific results. The emphasis on material evidence and on the strong need for controls, in a domain where Occam's razor is often taboo, is healthgiving. Archaeology has been extraordinarily self-conscious in the past decade, full of concern for hypothesis-testing and logical levels. These fieldwise authors, rather more hard-boiled than the computer fans and the logicians of inference, show an engaging lack of cant and a more than frank disdain for "charismatic manifestoes." If their studies are often based on human material discards, these workers are plainly set to defy the old cynicism. Even if garbage is literally what goes in, honest conclusions can come out.

At the University of Arizona they take in modern garbage. The project is an effort to get quantitatively reliable data on household food input and output. Some of the earlier results are striking. Poor people interviewed in Tucson say they never buy beer, but their garbage gives them away. The point is that beer cannot be lawfully acquired with food stamps. (One is glad to see that elaborate precautions were taken to conceal the identity of the households studied.) Student volunteers, dressed and immunized for the task, sort and code the garbage into many categories, from fish to special-interest magazines. The garbage of Tucson, unlike the trash pits dear to archaeology, does not contain many sherds, but 60 percent by weight of the refuse is packaging. What that style of waste has bought the economy of the modern culture is freedom from food spoilage: "It seems fair to say that packaging (along with other technological changes) has cut household food discard from 20 to 25 percent of total input by volume during World War II" to about half that today. To fix the trade-off of oil and iron for renewable crops would require further study.

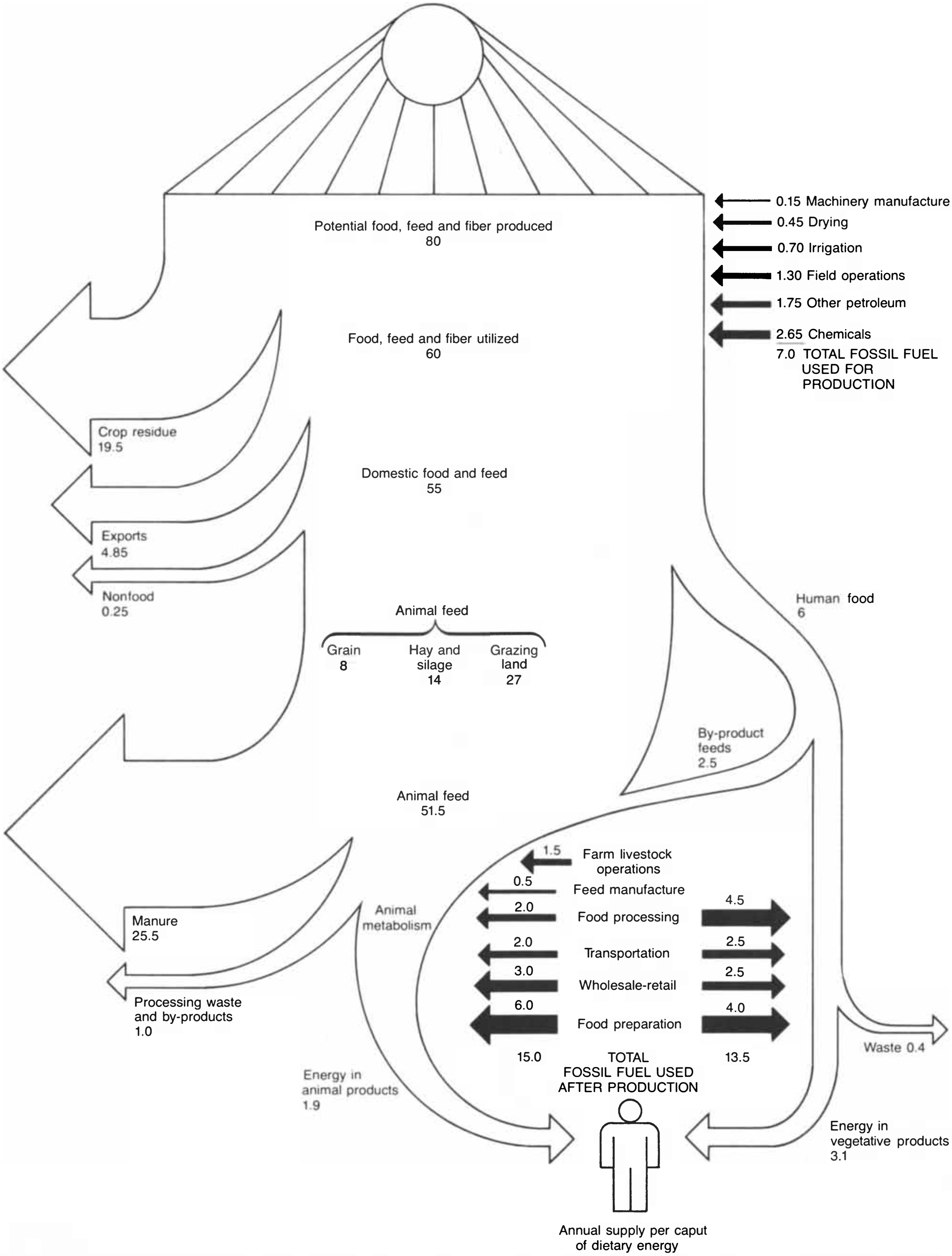
The middens of the coastal Tasmanians are all that remain to tell us of those haunting people. The last Tasmanian died a century ago. About 4,000 of them lived along the shore when first the explorers and then the settlers came. The people had been island-locked and alone since the melting glaciers caused the sea to rise some 12,000 years ago. They lived on seal, shellfish and small game. One ethnographer of goodwill stayed among them for some years. His journal was published only after a century; it confirms the less intimate accounts. These skilled divers and knowing coast dwellers never ate fish during the brief time of direct ethnographic contact. Yet down deep in the middens at many points on the island fishbones abound. Below 4,000 years ago 60 percent of the bones are fishbones. Above 3,500 years ago there is only one fishbone among thousands, and that single vertebra may have been in the gut of a seal or a cormorant. In the old days, Rhys Jones shows, the winter staple was probably fish, a steady supplement to the bonanza of a seal every once in a while. Why this total revolution in habits, the sacrifice of an ecologically valuable resource?

It was not the disappearance of the fish; they are plentiful even today, swimming within reach of the rocks off the midden as they must have when Tasmanians took them easily long ago. It is also clear that the technology was once there. "I feel," Jones writes, "that the Tasmanians made an intellectual decision that had the result of constricting their ecological universe." There are strangely related events; they stopped making bone tools, although less suddenly, at about the time they stopped fishing. They did not use water craft, "nor apparently did they know (or want to know) how to make them," even though boats would plainly have been of use in crossing rivers and bays.

Why did all those shore people for all those years take no fish? They used no wood boomerangs or barbed spears, no hafted stone tools or edge-ground axes, although on the Australian mainland these implements were ancient and widespread. It looks as though it might have been cultural drift: the small, lonely population formed an island of culture whose cultural species was an impoverished, haphazard sample, evolving curiously under little pressure. The dwarf deer and giant tortoises of other islands are the analogues to the fish-avoiding practices of our Tasmanian kin.

THE STATE OF FOOD AND AGRICULTURE 1976, by the Food and Agriculture Organization of the United Nations. FAO Agriculture Series No. 4, Unipub, Box 433, Murray Hill Station, New York, N.Y. 10016 ($11.75). In 1975 the summer monsoon in the Far East was "outstandingly good" and the winter rains were sufficient. Those happy circumstances continued at least marginally through 1976, and in addition there was a real recovery in the U.S.S.R. from the "prolonged and severe drought" of 1975. In other words, we were lucky two years in a row. South American countries had an excellent crop of wheat, and there was a bumper crop in Turkey, Iran and the Mahgreb. Even reckoning per head (there was a good crop of babies too) the world held its own. Meat is in oversupply and fisheries show medium-term prospects for increase, via squid and sardines.

In this large paperbound annual report the United Nations Food and Agriculture Organization thus summarizes the strength of the staff of life worldwide. A world review looks at crops and prices, at fertilizer, pesticides and product, at stocks and nutrition. A regional review adds detail, particularly with respect to agricultural policy in all the main regions. Forty pages of tables country by country complete the volume, a knowing and cool look that lives up to the title. (It is useful to point out that the numbers in the *FAO Production Yearbook*, an austerely statistical collection of computer printout tables giving even more detail, do not always agree with those in the somewhat earlier report reviewed here. The variations are



FLOW OF ENERGY IN THE U.S. FOOD CHAIN, which is diagrammed in *The State of Food and Agriculture 1976,* indicates that the largest energy inputs come after food production (*small horizontal black arrows above the human figure at the bottom*). The numbers are billions of joules. The diagram is based on a study conducted by F. C. Stickler, W. C. Burrows and L. F. Nelson of Deere & Company.

only of a percent or two; they represent the noise behind this urgent signal.)

It is a surprise to read of world reserves policy. Even after the 1976 International Undertaking on World Food Security the ice remains thin. Although we still count on the wayward torrents of the monsoon to feed the poorest and most abundant people, we can manage a world stockpile of only about seven weeks' consumption of grain of all kinds, well under the 10 weeks the FAO has set as a goal. Most of these stocks are held in the exporting countries. And not all countries are above the margin. Most children are still too thin for their height in and out of the camps in the drought-visited Sahel; all East Africa, from the Horn straight south, reports a losing struggle with nutritional minimums. Sri Lanka and Honduras share the problem; behind every rubric in 50 tables there is a long and complex story.

Like many such annuals, this FAO document devotes a section to a monographic review of one important subject. This year it is the role of energy in agriculture. The single topic of most interest here, still awaiting more comprehensive FAO study, is the use of wood and plant and animal wastes as energy resources by the developing countries. The developing world and the developed one make about equal harvests of roundwood (logs, as distinct from pulpwood), but the developing countries burn 86 percent of theirs for fuel and the developed countries burn only an eighth. Fuelwood is "the dominant source of energy for low-income families," giving more than a billion people their daily kitchen fire and much of their home heating, and supplying energy for the drying and curing of foodstuffs, bricks, tea, tobacco and rubber. Most of the wood is collected by the hardworking women, young and old, of the families that burn it, and the daily chore is being made harder by steady depletion. All over Asia and Africa wood is supplemented by the burning of dried cattle dung, amounting to about 150 million tons per year, or 13 percent of the fuelwood energy, but it is a loss to the soil. The biogas generation of methane by the anaerobic digestion of dung saves the plant nutrients and supplies energy, but at a capital cost too high for the poorest villages. The burning of charcoal, now about 5 percent of the fuelwood energy, should grow; its efficiency as a fuel and its light weight allow transport over distance.

A striking discussion (based on a publication by engineers at Deere & Company) analyzes the energy used in the entire food chain of the U.S., say for 1975. It is true that we produce our food, feed and fiber by pouring on the fertilizer, the gasoline and machinery supplied by fossil fuels. By such means we capture some seven times the energy input from sunlight. But farm production is by no means the most important consumer of energy in the chain before it ends at the table. It takes only a fourth of the total. The largest single category of energy consumption, easily double the total use on the farm, is cooking and refrigeration, two-thirds of it in the home and a third commercial. Even allowing an offset for home heating by the leakage of heat from the stove and the refrigerator, this is a big bite, and it suggests that economies in the recipes, pots and appliances of American kitchens would do more for the energy budget than any practical forswearing of fertilizer.

The fertilizer price curves show a big bump in the years 1971 to 1975, peaking at four times the level of 1965. The index was back down to the 120 range by late 1976, after that desperate excursion, which one must hope will not be repeated.

RECONSTRUCTION FOLLOWING DISASTER, edited by J. Eugene Haas, Robert W. Kates and Martyn J. Bowden. The MIT Press ($12.95). The pathology of our cities is much in the public mind. This joint study deals, from the viewpoint of a dozen sociologists and geographers, with the convalescence of cities after life-threatening accidents. Against a brief survey of the long history of cities dead or maimed by war and natural disaster the authors develop in some detail the recovery of three cities hard hit by earthquake and of one small city whose recent trauma was a sudden flood.

The work presents a plausible model of the sequence of the recovery, a careful comparison of the regrowth of jobs and homes, seen both as from the temporary city hall and at a family scale from the ruined living room. Four scenarios of a cautionary kind are offered, one for each case, to show how it might have been if planning had been better. A chapter of recommendations and conclusions ends with a worried prognosis.

Wrecked cities are rarely relocated. After the earthquake of 1773 the capital city of Guatemala moved to Guatemala la Nueva, 30 miles away. The new city in turn suffered heavily from an earthquake only three years ago. Mostly the remaining structures, the underground and surface investments, the natural features and the survivors are patched up and go on. Twentieth-century experience confirms the less precise information from the long past. The time to recovery is a multiple of the time of emergency action, that first period reflecting both the magnitude of the loss and the resources available. The recovery time is a fairly well-defined if steep function of the loss of life and capital stocks.

The authors mark four overlapping periods of recovery. The first is the time of emergency aid, of coping with disaster. The symptoms of the end of this period are the end of search and rescue, and the clearance of the main arteries of traffic. The second stage is restoration, with the patching up of what can be repaired and the beginning of more or less normal function. Its completion is signaled by the return of those refugees who want to return and by the substantial clearance of rubble. The third stage is reconstruction, marked by a restoration of the original number of inhabitants and the working provision of their normal needs. Finally, the stage of commemorative, developmental reconstruction arrives.

Both the second stage and the third last 10 times longer than the preceding one; the final stage lasts perhaps only a couple of times longer than reconstruction proper. The actual times for Anchorage, Alaska (struck in 1964 by an earthquake of magnitude 8.5 on the Richter scale, about as high an energy as has ever been recorded in North America), were in emergency about a week, in restoration 10 weeks, in recovery 100 weeks and in the final urban renewal of a business area totally destroyed by landslides 10 years. Anchorage was lucky: few were killed, plenty of Federal money came, housing stock was in surplus and the city soon took off as a new oil capital.

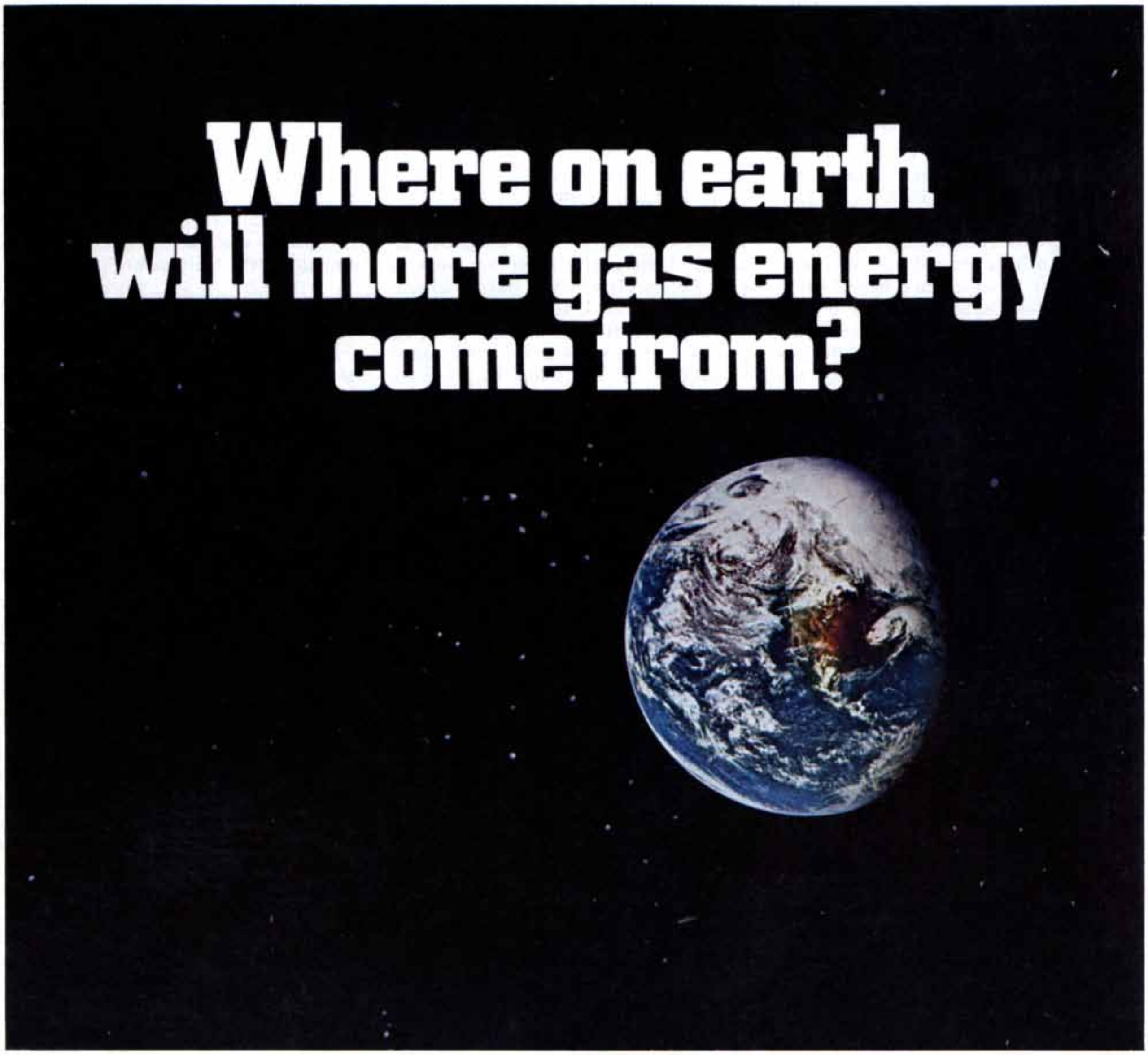
The other cities studied are San Francisco, after the earthquake of 1906; Managua, Nicaragua, earthquake-struck late in 1972, and Rapid City, S.D., devastated by a terrible flash flood in 1972. The simple model of four stages fits these cases rather well, with the main reconstruction period simply scaled to end respectively five, four and two times later than Anchorage. (The empirical model is of course only logarithmically reliable; variations of factors of two or so are typical.) Before 1800 all recoveries in matching cases were slower by a factor of about five. The tempo is faster nowadays. Indeed, the past fixes the recovery. Growing cities recover rapidly; stagnant ones recover slowly or even decline still more. Among modern damaged cities only the completely ash-engulfed St.-Pierre on Martinique has not been rebuilt.

The mechanisms of recovery are examined here in some detail. For example, a table offers the change in mean distance of travel to work by social class after the San Francisco earthquake. The data were obtained by sampling city directories to find address changes of named people. The old low-rent loft buildings suffer heavily, and no one rebuilds them. Their small, often marginal, occupants fail or move away. The low-paid blue-collar workers must go a long way for work; their choices are to leave the region or to remain and scrabble for odd jobs. The pattern is statistically clear, and it is reinforced by the individual families presented in examples. Generalizing, the social structure is shaken no less than the walls; the insecure members fail, the slides follow the lines of easy trend. The planners' hopes for improvements and a new equity fail of accomplishment; the uncertainties and delays they entail come at a high postdisaster cost. After all, there is already complete "a plan for reconstruction, indelibly stamped in the preception of each resident—the plan of the pre-disaster city."

San Francisco seems open to another awful shake along the great fault. The scenario foreseen begins with the death of several thousands, the number depending on the time of day. "The myriad vested interests of the survivors will ensure San Francisco's recovery" even if fire by mischance brings much greater direct damage than the $6 billion estimated. The tallest buildings and the big residential tracts will be little hurt. It is the uniqueness of San Francisco that will suffer most: the low skyline and the human scale of the laissez faire rebuilding after 1906. Then renewal by fire gave the city a new downtown. Now we can expect shifts back to the norms: toward Washington style, wide streets and monumental scale, like the present San Francisco Civic Center; toward Miami style, high-priced and commercialized, "an archipelago of sterilized hotel islands whence tourists take in local atmosphere"; toward a new Manhattan around the financial district, since the tall skyscrapers will stand whereas the fill areas around them and their 12-story buildings will count some failures; finally a drift toward the Los Angeles look. At last the distant shopping centers will thrive off the middle class of the entire Bay Area, the wholesalers and printers now clustered at the borders of downtown will leave for new suburban "parks" and a safer, duller San Francisco will degrade to one more "middle-class city with high-density enclaves of poverty."

Does it have to be that way? Not if modest and open but vigorous plans are laid in advance, for new land-use agreements, for a "special sensitivity to the small and marginal" and for transitional help with benefits, counseling and loans; not if a mobile-home area is prepared in advance and other prepackaged structures are surveyed in wide variety; not if the precious nature of forethought is taken seriously.

These earnest authors nowhere allude to the biggest disaster that lurks in the shadows: simultaneous radioactive earthquakes in 100 cities at once, the man-made ruin of thermonuclear war. Resources to cope with disaster on the



The land.

America still has huge amounts of gas available from conventional underground sources. These include the vast new Alaskan gas fields, and gas recoverable with new deep drilling techniques.

Energy experts say that when the next century arrives, we can still be getting as much gas out of the ground as we do right now.

We not only *can* do it—we must. No other energy is now available in quantities large enough to do all the jobs that gas does.

The sea.

Some of the largest underground deposits lie deep beneath the ocean. Exploratory drilling has already begun on the Atlantic's outer continental shelf, an area that has been virtually untapped till now.

Drilling under the ocean floor is difficult, but the technologies have been developed and proven in the Gulf of Mexico. Twenty percent of our gas supply now comes from offshore, mainly the Gulf.

Drilling platforms are also at work off the Pacific coast and in the Gulf of Alaska.

The sun.

Plant life has always captured energy from the sun—now there may be a way to turn that energy into natural gas. The gas industry and the U.S. Department of Energy are working on test projects with fast-growing seaweed, one of the most efficient solar converters known.

Vast beds of this seaweed would be turned into gas by a process called biomass conversion. It's experimental now but in the 21st century could make a real contribution to energy supply. Wise energy use will help, too.

AGA American Gas Association

scale of Rapid City or Anchorage are improving, but for the grand catastrophe little can be offered except the assurance of prevention, and even that is still far away.

EYE MOVEMENTS AND PSYCHOLOGICAL PROCESSES, edited by Richard A. Monty and John W. Senders. Lawrence Erlbaum Associates, Publishers. Distributed by John Wiley & Sons, Inc. ($35). The eye is no head-held camera, wavering a little around a steady gaze. After all, it "is sitting in a bag of fat in a hole in your head, and there are six big muscles pulling on the sides of it." Those muscles are the speedy effectors of about the most elaborated computer-controlled servo system we know, a stabilized-platform search system, taking signals and stored program out of a parallel-processing computer system we are coming increasingly to admire. Four distinct movements are present and demonstrable. The first remarks of this tight collection of 40 research and review papers by a cheerful and iconoclastic research community are convincing. Try shaking this page through small angles as you read, a few times a second. The text appears blurred. Now shake your head at the same rate by the same amount. The writing on the page stays clear. The same automatic stabilization can be shown to work in all three axes of rotation.

Six semicircular canals in the bony labyrinth are linear-integrating angular accelerometers; they work in zero gravity just as well as they do here below, because they are inertially run, strongly damped by viscous losses and depend on such small motions of the sensing structures that they are hard to saturate. On this syrup-smooth viewing platform, while you walk or ride, the higher programs superimpose three other well-studied movements: one for pursuing moving targets, one for keeping the two separated eyes aimed in the same focal direction and one for generating the swift and mysterious saccades, which scan whatever we look at as surely as, but much more complexly than, the naïve ceaseless scan of the television camera.

We slew our gaze around by saccades often, a couple of times each second during reading, for instance. The motion is the "most rapid somatic movement . . . the body can produce," in angular terms up to 400 degrees per second, three times faster than motion of the forearm, to compare similar large-angle movements. As a rule we do not detect the slewing from within although the nature of that suppression is itself tricky, inconstant and here open to much questioning. Skillful use of the motions enabled one experimenter to read the fine print on a whirling aircraft propeller in flight, he recalls, "in the days when aircraft had propellers."

This year is the centennial of the discovery of the saccades in reading; most of the papers here are devoted to their study one way or another, as clues to the subtle programs of the purposeful mind. "Psychology . . . reinvents its subject matter from generation to generation," writes one historically minded author from the University of Toronto, who demonstrates it by a delightful account of the nearly forgotten work of Guy Thomas Buswell, a "data-oriented" empiricist at the University of Chicago in the 1930's. Buswell studied how people look at pictures and how children and adults read, by the foot-by-foot analysis of continuous motion pictures (not exposed by frames) made by the mile. A silvered bead on spectacles left a baseline record of the head movements; the corneal reflection of a special light beam mapped the absolute direction of gaze and a fan-blade chopper provided timing data. What was hardest in those days was data handling; where now the computer is turned loose on every run, then "heroic analysis" called for plenty of careful choice because the computations took weeks. Buswell adds a charming comment to the account of his work, a generation old but still enlightening.

"Perceiving is in terms of meaning." The spots sampled (scattered but not random?) vary with the viewer, show no easily caught order and require the assembly of picture fragments or word gibberish. Just how the scan goes we do not know, although one deep study suggests that the same internal plans that somehow map what we are about to say are implicated, and not in a simple half-phonetic way. Periphery samples context, meaning guides search; there is need for a buffer store, selective testing and retention, "chunking" of words or scenes and susceptibility to language. This investigator sounds sure, if vague; his peers were less persuaded.

A lot of work is based on the urgent search of expensive scenes, from reading to picking out the shadow of a tumor in a lung radiograph and the faint radar blip of an unexpected aircraft. Block diagrams of rabbit eye-control servos are offered, but the more numerous built-in microprocessors we hominids possess are still elusive. The first 200 milliseconds of searching a chest X ray pretty surely offers most of the success for an expert. Expectation is almost all; photointerpreters who "knew there was an armored division in the area . . . saw tanks all over the road; if they didn't, they saw puddles."

The style and the techniques of this discipline are both delightful. The investigators enjoy being their own subjects, and they frequently cite an exception out of personal experience. The apparatus that measures the direction of gaze to a minute or two of arc for a head quite free to move is one fascinating example. The subject holds between his teeth what amounts to the pickup coil of a magnetic phonograph cartridge. One devoted, emphysemic, middle-aged psychologist, his breath held and his head resting in his hands, showed a steady drift of gaze at the rate of about three degrees a minute. (Rodin's *The Thinker* is figured, without documentation, to emphasize the truly stable.)

QUANTUM PHARMACOLOGY, by W. G. Richards. Butterworth (Publishers) Inc., 19 Cummings Park, Woburn, Mass. 01801 ($24.95). "It is tempting to add a question mark to the title of this book." With this remark an Oxford theoretical chemist disarmingly begins his readable, surprising small tract. By the final pages the author's case has been pretty persuasively made. "The men of genius who developed quantum mechanics" 40 or 50 years back foresaw a chemical science entirely deductive from the Schrödinger equation; that is not what is presented here. Rather we are shown a process of inference in pharmacology whose chief tool is the minicomputer, chewing on apt approximations to molecular orbitals for smallish molecules, real or projected.

The link is this: In the nervous system (other topics in medicine are treated only lightly) we know that a limited number of molecules, each with only a few dozen atoms, specifically bridge a variety of tiny gaps in the circuitry, binding to special receptors throughout the central and peripheral nerve network. Studies of these substances and their antagonists and simulators have defined small classes of molecules that bind to and activate particular sites. Such clever tracing techniques as fluorescent labeling have given primary information on where and how.

Now, a ball-and-stick model of any such molecule is easy to make; the bond lengths and bond angles are today comprehensively and accurately tabulated. But almost no neurotransmitter substance is fully defined in space by such a model. These are flexible molecules, rotating around one bond axis or more to take on a wide range of shapes. Although the model can bend freely at its hinges, the real electron clouds cannot. Instead each conformation of positions of the nuclei of the atoms in the molecule affects the electron energy levels somewhat and gives a softly preferred set of forms to the molecule. The forms bow and bend until they come out right for the hidden sites of their receptors.

The way is clear: compute for the drug in question (it does not matter if it has not yet been synthesized) the molecular energy for each conformation as a function of the few degrees of geometrical freedom in the molecule. It is not hard to weight the energies appropriately for the temperature involved, and so to calculate the relative probabilities of various molecular forms at equilibrium. These patterns are available as the output of carefully developed computer programs relying on one or another of several clever approximations to the full quantum-mechanical description of the molecule. (Almost all the practical methods begin with the assignment of simple molecular orbitals based on hydrogenlike wave functions for the outermost electrons.) Initial forms, sets of comparison molecules and insightful study of the results of different approximations may be more important in this application than big number-crunchers or even professional quantum theorists.

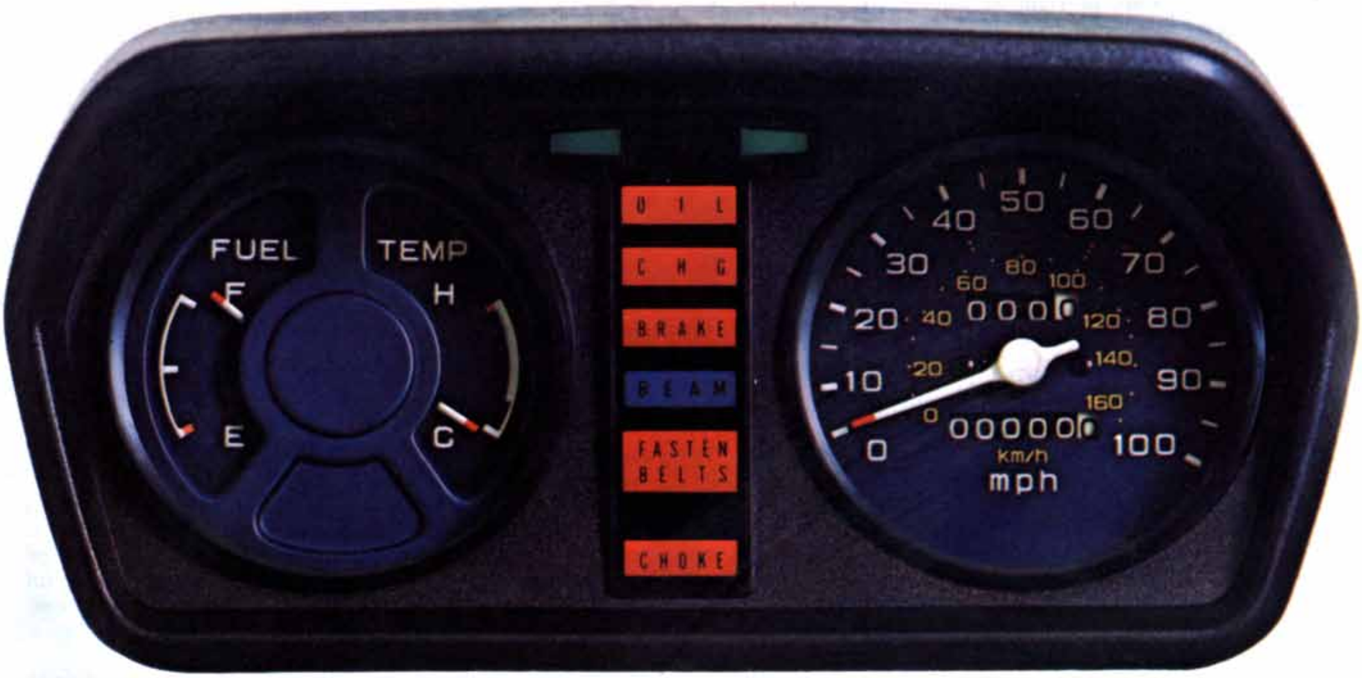
Even more can be tried. Charge distributions and interaction potentials can be found no less than conformation energies. The study of rigid molecules, such as morphine and its analogues, is better done by that approach. The calculated electrostatic properties are deeper than the simple models of localized charges on ions. The effects of ubiquitous water can be tested by a number of tricks, most directly by the study of an enlarged molecule with water hanging on to it at suitable places.

Is the time ripe for applied digital stereochemistry? Richards thinks so, without being able to cite many strong successes in the style, beyond an air of reason and an impressive bibliography of some 300 papers up to 1976, arranged by the compounds studied. He rightly argues that the "best time to enter a research field is just too soon; just before the work becomes possible." That is the aim of his generous and readable missionary work in this book. About half of the book is a graceful summary of neuropharmacology for quantum chemists, a quarter of it is a compact account of molecular wave-function approximations for pharmacologists and the rest is an interdisciplinary examination of how both could apply the new power of the numerical-graphical study of electron distributions in space. Perhaps the work will come to play a role like that of the ball-and-stick and space-filling molecular models. The penetrating, if frankly approximate, quantum descriptions should in no very distant future tell their user more than wood balls snapped together. They should do so in much the same way, as "an invaluable aid to thought and discussion or rationalization but not a field in their own right." The reader wishes it strongly; one comes away with a grasp of what inductive chemistry today is like, a subject very different from any false dream of turning the Schrödinger crank until the truth drops out.

Our lowest priced Honda isn't so simple.

The Honda Civic® 1200 Sedan is our lowest priced Honda.* We hope that statement doesn't put you off.

We know that lots of people tend to be suspicious when they see the words "lowest priced." Especially when it's a car. They immediately think of some stripped-down model calculated to snag the unwary buyer by means of a seductive price tag.



That's why we're running this ad. To let you know that, despite its very reasonable price, the Civic 1200 Sedan gives you such traditional Honda engineering refinements as transverse-mounted engine, front-wheel drive, rack and pinion steering, power-assisted dual-diagonal braking system with front discs, and four-wheel independent MacPherson strut suspension.

And that's not all. The Civic 1200 Sedan abounds with standard features that other manufacturers might charge you extra for.

*Not available in Calif. and high altitude areas. Manufacturer's suggested retail price excluding freight, tax, license, title, and options.

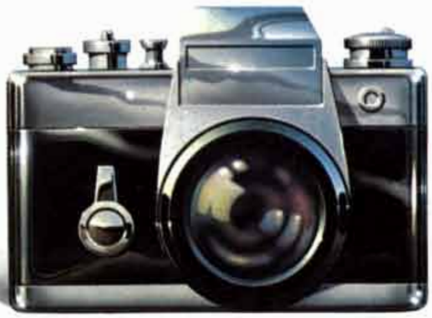
These include reclining bucket seats, adjustable head rests, wall-to-wall carpeting, opening rear-quarter windows, inside hood release, rear-seat ash tray, plus the instrument cluster shown opposite, a simple layout that nonetheless provides the added convenience of a trip odometer.

Like our other two Honda cars – the Civic CVCC® and the Honda Accord® – the Civic 1200 doesn't need a catalytic converter and runs on unleaded or money-saving regular gasoline.

So there you have it. The Honda Civic 1200 Sedan. Because it's a Honda, it's a simple car. But not so simple as its price would lead you to believe.

HONDA
We make it simple.

Add 400 to 35 and subtract the flash.



If you want to get more out of your 35mm camera, put Kodacolor 400 film into it. Kodacolor 400 film is so fast and versatile, you can get the kind of color prints you couldn't get before, without a flash. You can freeze the action in midair, midstride or midgallop, shoot in museums, theaters and arenas, and capture the sun from when it's barely a promise to when it's almost a memory. All in clear, sharp beautiful color prints with a grain that's surprisingly fine.

So if you want to get it right in existing light with a film that's as fast as your camera, just add 400 to 35. And you've got it right.

With Kodacolor 400 film, you can take 35mm pictures indoors, without a flash. More natural-looking pictures, because you're using existing light instead of altering the light.

1/125 @ f/2.0. No flash allowed here.

1/60 @ f/2.0. No flash needed here.

Kodacolor 400 film. You've got what it takes.

Kodak

Established 1845 SCIENTIFIC AMERICAN August 1978 Volume 239 Number 2

The Gas Centrifuge

This machine separates the isotopes of uranium by spinning a gaseous uranium compound. It is expected to play a large role in the Carter Administration's effort to limit the proliferation of nuclear weapons

by Donald R. Olander

A keystone of President Carter's effort to prevent the proliferation of nuclear weapons is his program for blocking the supply of plutonium that could be turned to the making of weapons. To that end the President has sought to shut down the U.S. program to develop a breeder reactor (since the breeder produces plutonium) and to "defer indefinitely" the commercial reprocessing of spent nuclear fuel (since the depleted fuel rods from nuclear power plants contain plutonium that would be separated and recycled). A corollary of this policy, as stated by the President in the spring of last year, is that "we will increase U.S. production capacity for enriched uranium [the fuel for the light-water reactors of most nuclear power plants] to provide adequate and timely supply of nuclear fuels for domestic and foreign needs."

The "enrichment add-on" envisioned by the President would entail the processing of more than 10,000 metric tons of uranium annually, an increase of 30 percent in the country's uranium-enrichment capacity. To achieve this increase by the gaseous-diffusion process that has been the means of enrichment in the U.S. since the beginning of the nuclear age would require two large nuclear reactors dedicated solely to satisfying the voracious appetite of the enrichment plant for electric power. In view of the high cost of electricity the Administration has turned instead to the gas-centrifuge method of enrichment, a system that was tested at the beginning of the nuclear age and put aside because the technology of centrifugation was not up to the task. Now it is. Moreover, enrichment by gas centrifuge requires only 4 percent as much power as enrichment by gaseous diffusion; an increase of 30 percent in enrichment capacity can be handled easily by the existing electric-utility grid. As a result of the new policy the gas centrifuge seems certain to become the nucleus of a multibillion dollar uranium-enrichment industry.

The purpose of any enrichment system (several in addition to gaseous diffusion and the gas centrifuge have been demonstrated, at least on a laboratory scale) is to increase the concentration of the isotope uranium 235, which in natural uranium is .71 percent. Uranium 235 is the only fissionable isotope that exists in nature; the other 99.29 percent of natural uranium consists of the nonfissionable isotope uranium 238. The two isotopes differ slightly in mass, since the uranium-238 nucleus has three more neutrons than the uranium-235 nucleus. The difference in mass provides a basis for separating the isotopes and therefore for increasing the concentration of uranium 235 in a nuclear fuel.

In gaseous diffusion the separation is brought about because a gaseous compound incorporating uranium 235 flows through a porous barrier slightly faster than the same compound incorporating uranium 238. In a centrifuge the separation is achieved by spinning the same gaseous compound incorporating the two isotopes; the compound containing the heavier isotope is thrust against the wall of the centrifuge, leaving the inner core of gas enriched in the compound containing the lighter isotope.

Uranium isotopes were first centrifugally separated in 1940 in the experimental centrifuges of Jesse W. Beams of the University of Virginia. At the time, however, the technology of the high-speed rotating machinery needed to produce enough highly enriched uranium 235 for the weapons program was inadequate to the task, and the centrifuge method was abandoned in 1943 in favor of the gaseous-diffusion process. (Electricity was cheap then, and in any case cost would not have been a major consideration.) Nevertheless, the basic theory of the gas centrifuge was worked out during that period.

Improvements in the size, speed and efficiency of gas centrifuges were achieved by W. E. Groth in Germany during a period of several years after World War II. In the late 1950's Gernot Zippe, who had done gas-centrifuge experiments in his native Germany and later in the U.S.S.R., became engaged in a small experimental program at the University of Virginia. He developed a light but durable centrifuge from which the modern designs are descended.

In 1960, as a result of Zippe's work, the U.S. Atomic Energy Commission authorized a development program (funded at $2 million per year) aimed at achieving large-scale, economically competitive uranium enrichment based on gas-centrifuge technology. At about the same time the secrecy regulations that had been lifted from the program after World War II were reimposed, probably to avoid putting information on how to produce weapons-grade uranium in the public domain. For the past 17 years the U.S. has carried on a substantial classified (that is, largely secret) research and development program centered around the Oak Ridge National Laboratory and the University of Virginia. The present funding of this program is about $80 million per year. Many other industrialized countries are also pursuing the new separation technology. The British, the Dutch and the West Germans have formed a tripartite company (URENCO) to provide enrichment by gas centrifuge on a basis that is economically competitive with enrichment by gaseous diffusion.

The only uranium-bearing compound that is gaseous at ordinary temperatures

ARRAY OF CENTRIFUGES is seen at the component test facility operated for the Department of Energy by the Union Carbide Corporation at the Oak Ridge National Laboratory. Each of the cylindrical casings contains a rotor into which the gas uranium hexafluoride is fed. The compound contains the fissionable isotope uranium 235, which is the minor constituent of natural uranium, and the nonfissionable isotope uranium 238, the major constituent. The gas is spun at high speed in the rotor, and the centrifugal force moves the heavier molecules containing uranium 238 close to the rotor wall, producing a partial separation of the isotopes. Through temperature differences and a stationary bottom scoop in the rotor a vertical circulation is set up, so that the gas enriched in uranium 235 can be withdrawn from the top of the centrifuge and the depleted gas can be withdrawn from the bottom. Only a small amount of gas can be put through one centrifuge, so that the machines are connected in a "cascade." Visible here are connective piping and the upper part of the centrifuges.

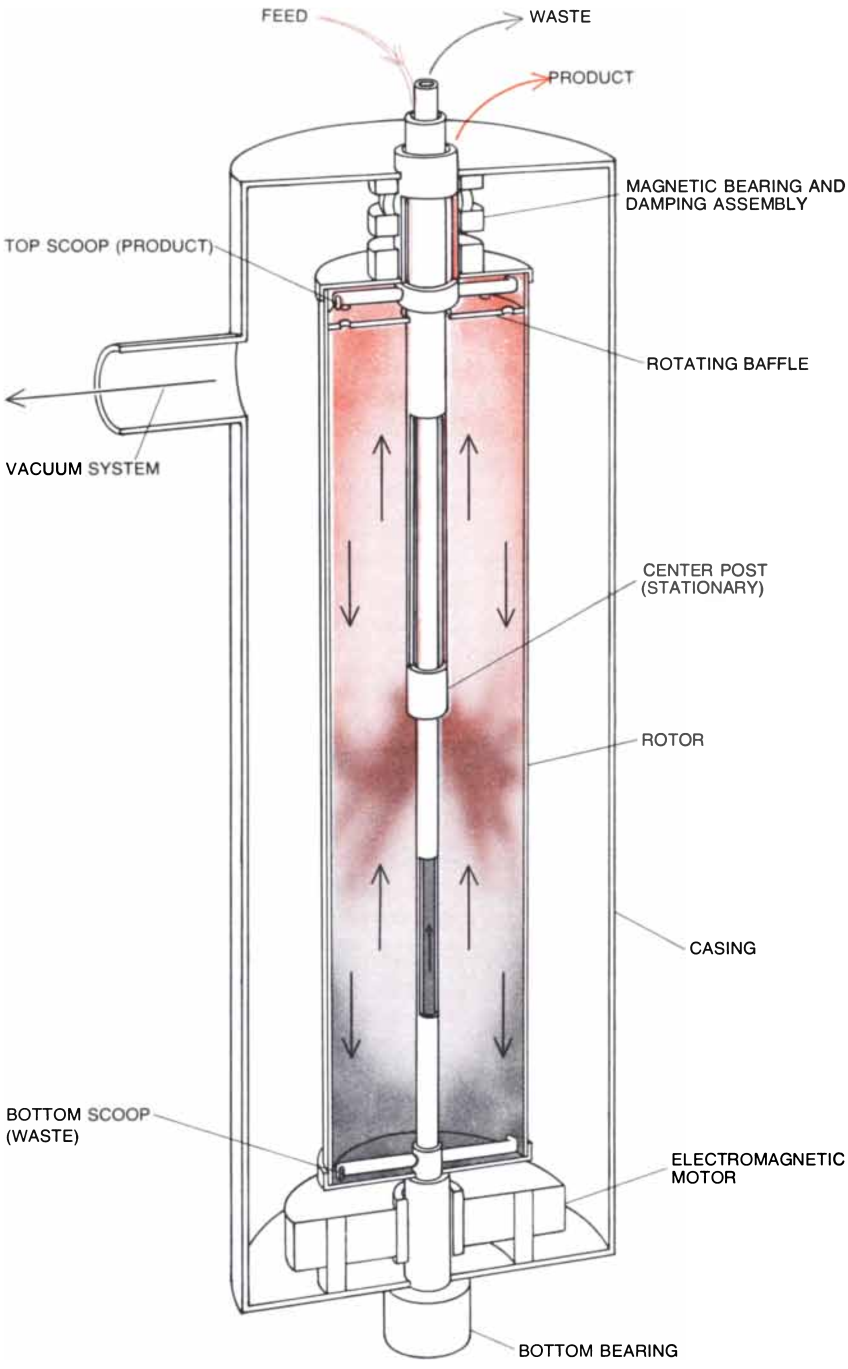
is uranium hexafluoride (UF_6). Thousands of tons of this substance have already been processed by the gaseous-diffusion plants, so that its characteristics (it can be highly corrosive) are well established. In the centrifuge process the feed gas is a mixture of $^{238}UF_6$ and $^{235}UF_6$, and the slight difference in the molecular weights of these species is exploited to separate them. The separation is effected by spinning the gas at very high rotational speeds, so that the heavier $^{238}UF_6$ moves to the outside and leaves the inner core of gas enriched in the lighter $^{235}UF_6$.

The theory of the elementary separation process is simple indeed: it is governed by the same principle that dictates the variation of the density of the atmosphere with altitude. The isotopic separation induced by the centrifugal force established in the vertical spinning rotor is of course radial (horizontal). The success of the gas centrifuge as an economical method of separation is attributable to the trick of changing the direction of enrichment from radial to axial (and considerably augmenting the magnitude of enrichment at the same time) by inducing a very weak axial circulation (or countercurrent flow) in the solid-body rotational motion of the gas. If it were not for this vertical circulation, the removal of the separated isotopes from the centrifuge would be difficult. Because of the vertical circulation the enriched gas can be withdrawn at the top of the centrifuge as the depleted gas goes out the bottom.

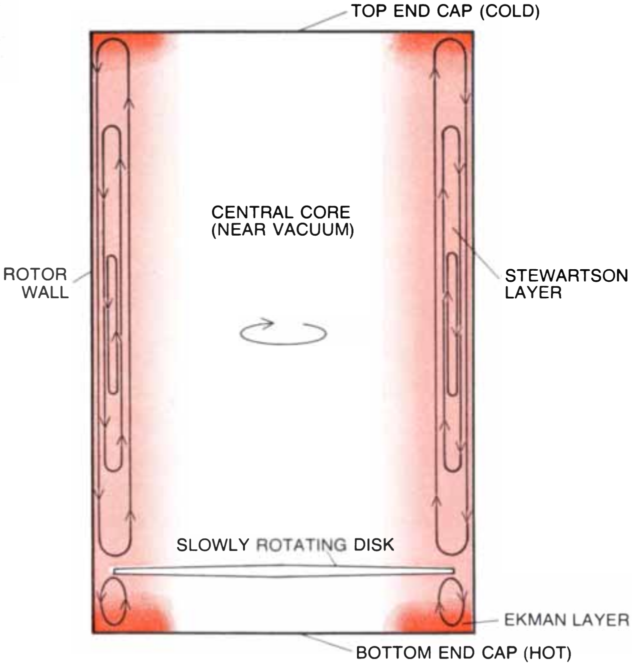
The size and the speed of modern gas centrifuges are still classified, as are the materials employed to build them and the details of their mechanical design. The basic components, however, are well known. A thin-walled cylindrical rotor is spun at high speed inside an outer casing. URENCO states that the peripheral rotor speed of its machines is about 400 meters per second. Feed gas is injected into the rotor near the center, and the product and waste streams are continuously removed at the ends.

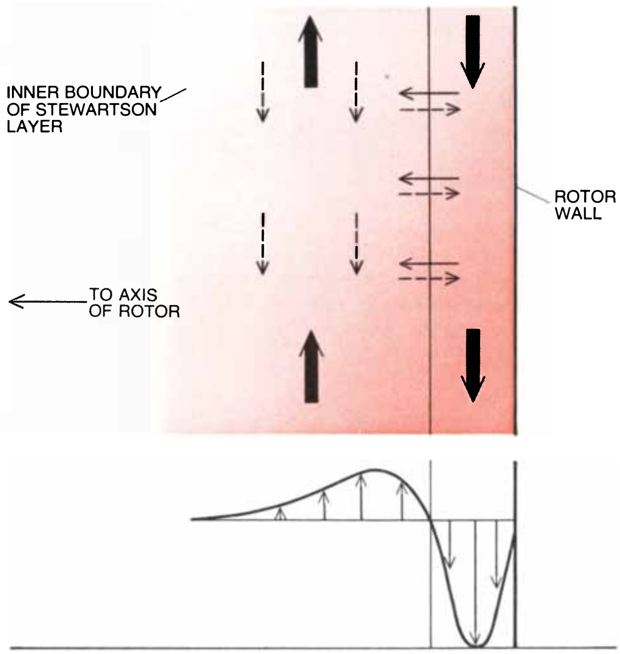
The total quantity of gas in the rotor is limited by the condensation pressure of uranium hexafluoride, that is, the pressure at which the gas becomes a solid; at room temperature the condensation pressure is about one-sixth atmosphere. The gas in the rotor is highly stratified by the rotational motion, so that the pressure in the center is much lower than the pressure at the wall. In addition to the limit on the quantity of gas in the rotor the throughput, or feed rate, is limited to values that will not upset the delicate axial circulation established by a bottom scoop or by other mechanisms, such as a temperature gradient along the rotor wall or a cooled cap at the top. According to information made public by URENCO, the throughput of the European machines is about 100 milligrams of uranium hexafluoride per second.

The outer casing has two functions. First, it provides protection in the event of a failure of the spinning rotor. (Early designs failed so often that they were sometimes humorously referred to as automatic disassembly machines.) Even in modern machines the lifetime of the rotor is crucial; the economic advantage of the gas centrifuge over gaseous diffusion would be lost if the average centrifuge lasted only a few months instead of many years. Much of the develop-



INTERIOR OF A CENTRIFUGE is shown in this cutaway drawing. The thin-walled rotor is driven by a small electromagnetic motor attached to the bottom of the casing. The top end of the rotor is held in a vertical position by a magnetic bearing and does not touch stationary components. Gas is fed into and withdrawn from the rotor through the stationary center post, which holds three concentric tubes for the feed, the product and the waste. The stationary bottom scoop protrudes into the spinning gas and provides a mechanical means for driving the vertical flow of gas. The top scoop, which serves to remove the enriched product, is protected from direct interaction with the rotating gas by the baffle, which has holes allowing the enriched gas to be bled into the area near the scoop. The baffle is needed to keep the top scoop from imposing a vertical flow that would counteract the crucial one generated by the bottom scoop.





GAS MOVEMENT in a centrifuge begins with the effect of rotation (*left*), which drives the gas against the rotor wall in a region known as the Stewartson layer, where the isotope separation takes place. As a result of temperature differences at the top and bottom caps and the effect of the bottom scoop (an effect simulated mathematically by a disk rotating slightly slower than the rotor) a weak vertical circulation is set up (*arrows*). In addition thin gas-containing zones known as Ekman layers, each only a fraction of a millimeter thick, are formed at the extremities of the Stewartson layer, where the circulatory flow changes direction. Countercurrent circulation and molecular transport in the Stewartson layer (*right*) result in the enrichment of the gas. The thin solid arrows indicate the direction of movement of uranium-235 molecules induced by the pressure-diffusion effect arising from the centrifugal force. The broken-line arrows show the directions of transport of uranium 235 by molecular diffusion. Convective motion of the gas as a whole is represented by the thick arrows. Lighter shading indicates areas that have the higher concentrations of uranium 235. The lower curve shows a typical mass-velocity profile (the units of mass velocity being grams per square centimeter per second) that characterizes the countercurrent circulation in the Stewartson layer.

ment work at Oak Ridge is devoted to testing the durability of rotors at operating speeds.

The second function of the casing is to provide a vacuum-tight enclosure for the rotor. Evacuation of the space outside the rotor greatly reduces the consumption of power resulting from gas friction when the rotor is spinning. A vacuum pump is connected to the casing for this purpose.

The isotopic separation achieved by a gas centrifuge is governed by the conservation equations of continuum fluid mechanics, the branch of physics that seeks to describe the motion of a fluid (a gas or a liquid) by employing the principles of conservation of mass and momentum. (Although the uranium hexafluoride in the inner core of the rotor is at such low pressure that rarefied-gas dynamics would be more appropriate in this region, the amount of gas in the core is so small that it can be safely ignored.) The theoretical analysis is divided into two parts: hydrodynamic and separative. The hydrodynamic analysis seeks to determine the flow velocities of the gas in the regions close to the rotor wall and at the end caps. Once the velocities have been found the separative analysis reveals how the $^{235}UF_6$ becomes segregated from the $^{238}UF_6$ as a consequence of the combined actions of the imposed gas circulation and the centrifugal force that is responsible for the primary separative effect.

Historically the development of the theory reversed this logical order. The mathematical description of isotope separation in a pressure gradient is almost the same as that of isotope separation in a temperature gradient. In the 1940's Karl Cohen of the Manhattan Engineer District (the "Manhattan Project") adapted to the gas centrifuge the theory of the thermal-diffusion column developed by Wendell H. Furry, R. C. Jones and Lars Onsager. A separative analysis was thus available but the fluid-dynamical information needed to make use of it was not. In 1960 the Atomic Energy Commission organized a committee (headed by Onsager) that was charged with rectifying this deficiency. The present work in the U.S. on gas-centrifuge hydrodynamics (guided by the Department of Energy) is being conducted by James A. Viecelli of the Lawrence Livermore Laboratory and by Richard A. Gentry of the Los Alamos Scientific Laboratory. European and Japanese work on the theory of the gas centrifuge (much of it published) has proceeded in parallel but in the absence of knowledge about the classified program in the U.S.

If the temperature of the rotor wall were uniform and if the rotor contained no nonrotating solid objects, the gas would simply rotate as a solid body at the angular velocity of the rotor. What actually happens is that small nonuniformities of temperature at the end caps or along the rotor wall (together with the effect of the stationary bottom scoop on the gas swirling around it) perturb the gas and establish a vertical countercurrent circulation pattern that is superposed on the basic wheellike flow. It is this gentle countercurrent that is responsible for the favorable isotope-separation properties of the gas centrifuge. The countercurrent must be described theoretically if the performance of the centrifuge is to be understood and predicted.

Measurement of the pattern of circulatory flow inside the spinning rotor is difficult. Swedish and Japanese workers have been testing laser velocimetry, a technique that utilizes the reflection of light from small particles in the flow to determine the magnitude and direction of the circulatory patterns. The reason for choosing such an esoteric method is that ordinary devices for measuring fluid velocity (such as a Pitot tube attached to the rotor wall) are structurally unstable in the strong centrifugal-force field. Even if they were usable, a system of noncontact telemetry would have to be devised to extract the information they were gathering.

The motion of the gas is analyzed by solving a set of rather complex equations relating to the conservation of mass, energy and momentum. Such a situation is common in fluid mechanics, but in most instances (for example the flow of liquid through a round pipe or the flow of air over a wing) the equations can be drastically simplified. Unfortunately the dynamics of a gas in a centrifuge offers no relief from mathematical rigor. This difficulty is on the verge of being overcome through the use of large computers and numerical techniques in the place of analytical ones. The work of the French and Japanese in this area is particularly advanced.

The crucial axial movement of the gas is activated by the perturbations imposed on the spinning fluid by the nonuniform temperature of the wall and by the presence in the fluid of the stationary bottom scoop. The way nonuniformities of temperature along the wall influence the motion of the fluid can best be understood by analogy with atmospheric physics. When gas comes into contact with a heated end cap, its density is reduced as it absorbs heat; the gas proceeds to rise in the "atmosphere," that is, it moves radially inward. Similarly, a cooled end plate causes the gas to sink toward the rotor wall. This phenomenon taking place at the end caps, which is termed Ekman suction, can drive the vertical countercurrent over the entire length of the machine. A similar circulation is produced by a continuous variation of temperature along the length of the rotor wall; if the bottom of the rotor is hotter than the top, the circulation loops move gas upward near the center and downward near the rotor wall.

The circulatory effect of the scoop is somewhat more complex. As a stationary object in a stream of gas moving at high speed, the scoop tends to warm the gas passing over it by the same skin-friction effect that causes the wings and body of a supersonic aircraft to heat up. The heated gas rises in the atmosphere. In addition the drag of the scoop reduces the tangential velocity of the gas, and the rotating gas contracts radially in order to maintain the same angular velocity. Thus the circulation generated by a scoop at the bottom of the rotor is in the same sense as the circulation generated by an appropriate distribution of temperature along the rotor wall and on the end caps, although the radial shapes of the two types of axial circulation are different.

One therefore sees that either thermal or mechanical stratagems can drive the axial countercurrent flow in the centrifuge. The velocities imparted by the two types of drive can be controlled by the design of the centrifuge and the mode of operation. The thermal drive can be varied by controlling the extent of the cooling of the top of the machine and the heating of the bottom. The mechanical drive is determined by the size and shape of the bottom scoop and by the radial position of the scoop in the atmosphere of gas next to the wall. The object of the theoretical hydrodynamic analysis is to accurately calculate the flow field once the physical nature of the various drives has been specified and to optimize the circulation pattern to achieve the largest possible axial separation of the uranium isotopes in the gas.

Once the distribution of gas velocities has been calculated, the final step is to determine the nature and extent of the separation of uranium isotopes in the gas. The task is accomplished by solving a conservation equation that takes account of all the physical phenomena acting on the two isotopic species in the gas: the centrifugal force (also called the pressure-diffusion effect), which tends to concentrate the heavy $^{238}UF_6$ near the rotor wall; convection, which acts to enhance the separation caused by pressure diffusion by enabling the upflowing inner stream to be continuously enriched as it passes from the bottom to the top, and ordinary molecular diffusion, which is a counterseparative force.

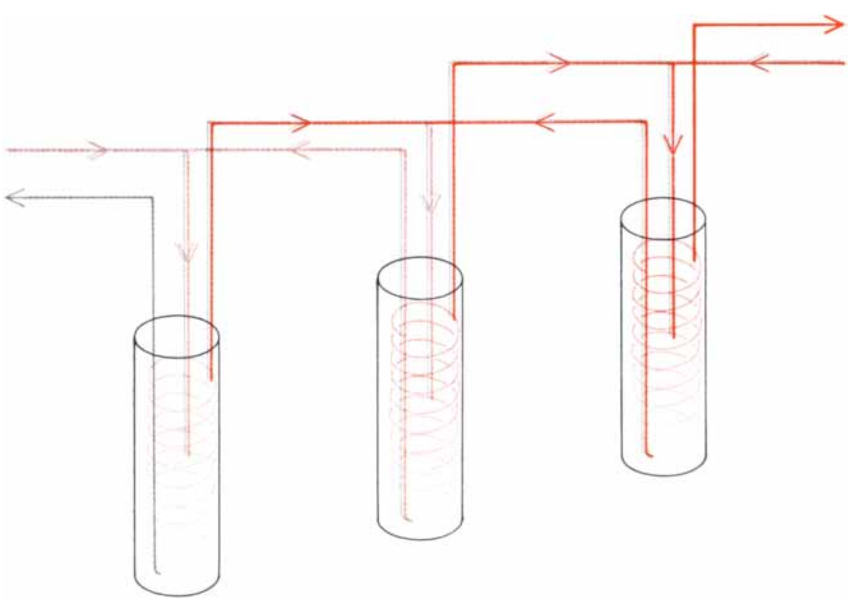
The net result of the interplay between the separative and the convective process is that the light isotope ($^{235}UF_6$) accumulates at one end of the machine and the gas at the opposite end is depleted in that isotope. The axial separation is the reason the product and waste streams are respectively withdrawn from the top and the bottom of the machine, even though the elementary separation phenomenon is radially directed.

One measure of the efficiency of the centrifuge as an isotope separator is the separation factor, which is the ratio of the amount of uranium 235 in the product stream to the amount in the waste stream. A typical separation factor for an early machine was 1.25, which means that if the fraction of uranium 235 in the feed gas is .71 percent, as it is in natural uranium, the product contains .794 percent uranium 235 and the waste contains .635 percent.

At first it might seem that the separation factor is the natural measure of the ability of the centrifuge to separate isotopes, but it is not. An efficient centrifuge combines a high separation factor with high throughput. A centrifuge (or any other device for separating isotopes) is therefore best judged by a single measure of efficiency that takes account of both factors. The measure is termed the separative power, a concept originated in the era of the Manhattan Project by P. A. M. Dirac and R. E. Peierls.

The concept has three essential features. First, it encompasses all the physics of the isotope-separation processes occurring in the separating unit. Second, it is a function of the throughput and the "cut" of the machine (the ratio of the product rate to the feed rate) and also of the internal factors that govern the separation phenomena. (For a centrifuge the factors include the temperature of the rotor wall and the shape and position of the scoop.) Third, it is not a function of the isotopic composition of the gas in the unit.

The separative power is usually reported as kilograms of uranium (not



CASCADE PRINCIPLE is employed to enrich the uranium in stages. In each centrifuge the uranium hexafluoride is partly separated into gas containing the lighter uranium-235 isotope and gas containing the heavier uranium 238. The stream enriched in uranium 235 (*color*) flows to the next stage, which is processing a gas of higher enrichment. The stream depleted in uranium 235 (*gray*) flows to the next-lower stage, which is separating a gas of lower enrichment.

uranium hexafluoride) per year. This unit, called the separative work unit (per year), is the universally accepted measure of the separative power of a single centrifuge device (or of an entire uranium-enrichment plant). The real product of an isotope-separation plant is separative work, not enriched uranium. What this concept reflects is the ability of a plant to produce either a small quantity of highly enriched uranium or a large quantity of slightly enriched uranium.

The separative power, as one might expect, bears some relation to the actual power required to operate an isotope-separation device. It is principally this relation that is at the heart of the President's announcement of the choice of the gas centrifuge for the next generation of U.S. enrichment plants. The specific energy consumption of gaseous-diffusion units is about 2,500 kilowatt-hours per separative work unit, most of which goes to recompressing the gas that has diffused through the porous barrier. (The gas has to be recompressed before it can be sent on to the next diffusion unit.) The centrifuge requires only 100 kilowatt-hours per separative work unit. This power is expended in overcoming the friction of rotating parts, electrical losses due to heating of the motor, pressure losses when the gas is extracted and aerodynamic drag of the process gas against the scoop and the inner wall of the rotor (also the drag of the residual gas in the casing against the outer wall).

If competing processes were to be judged solely on the basis of lowest power consumption, the gas centrifuge would be the clear winner over gaseous diffusion. The cost of fabricating the devices and the frequency of replacement of the separating units must also be considered, however, and here the advantage lies with the gaseous-diffusion method. When the total cost of separative work is calculated by the appropriate weighting of the contributions of power consumption and capital outlay (and a few smaller items), the gas centrifuge emerges as the most economical method.

In sum, it is not enough for an isotope-separation device to provide a large separation factor; it must do so at a throughput of sufficient magnitude to make the separative power acceptable. Although the highly touted system of separating isotopes by laser excitation may provide large separation factors, it has not been shown capable of achieving throughputs of sufficient magnitude to give it a separative power per dollar invested that would represent an improvement over either the gas centrifuge or gaseous diffusion.

Gas-centrifuge technology is mainly a matter of economics. The major question is not whether the method will work but whether it will produce separative work units at a lower cost than competing processes. Even small reductions in the investment per unit of installed separative capacity become magnified into large sums of money because of the great size of the uranium-enrichment industry.

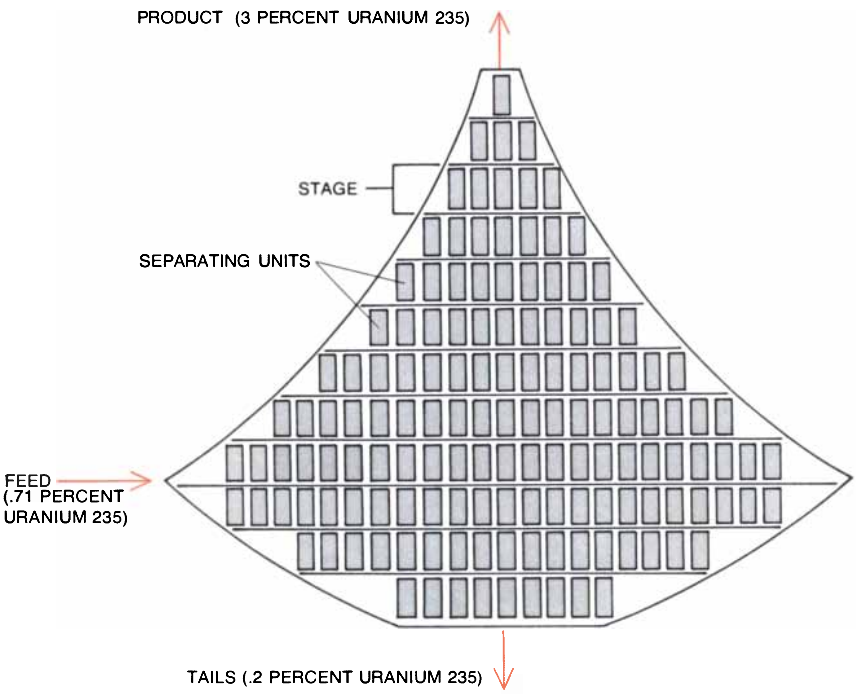
Suppose enriching facilities providing 25 million separative work units need to be constructed in the U.S. between now and the year 2000 in order to sustain the domestic nuclear-power industry and the foreign demand implied by the President's policy. The capital required to build the new enrichment plants will be about $10 billion. Therefore an improvement of 10 percent in the efficiency of each machine would represent a saving of about $1 billion. This amount of money is enough to justify a substantial research and development effort aimed at improving the productivity of the machines.

Performance improvements are to be found in three aspects of gas-centrifuge technology: in machine design, in the mode of operation of the individual machines and in the manner of interconnecting the individual centrifuges. In essence the separative power of a gas centrifuge is determined by its rotational speed and the length of its rotor. The separative power increases rapidly with rotor speed and is proportional to the length.

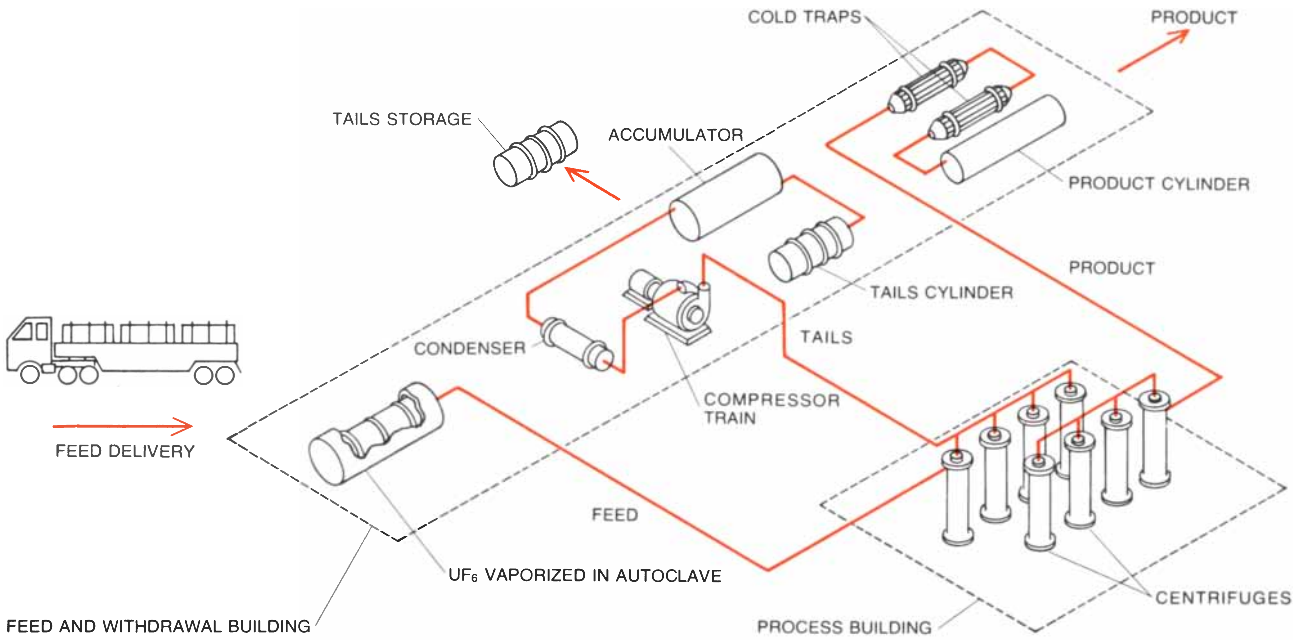
The peripheral speed of the rotor is limited by the ratio of strength to density of the material it is made of. Aluminum alloys and stainless steels are capable of rotor speeds slightly in excess of 400 meters per second. Composite materials can do better than metals. Japanese and European studies suggest using fiberglass-polyester composites of the kind employed in flywheels designed to store energy; with such materials the peripheral speed of the rotor can exceed 500 meters per second. Carbon-fiber composites could achieve even higher speeds, but at present they are much too expensive.

The limit to the length of the machine is determined in part by the difficulty of controlling the straightness of the rotor and the uniformity of the wall in mass-production manufacture and by the durability of the bottom bearing, which must support the weight of the rotor. Another concern with centrifuges that have a large length-to-diameter ratio is the problem of getting them up to operating speed, since they may have to pass through critical speeds at which large-scale vibrations can occur.

The basic objective in designing a centrifuge is to produce the fastest and longest rotor possible at the lowest cost. The next problem is to select the operational conditions once the size and speed have been established. The task requires a judicious combining of the various flow-driving mechanisms that affect the performance of the machine. This aspect of centrifuge technology is by no means trivial, because the separative power depends on at least six adjustable



IDEAL CASCADE is made up of stages, each one consisting of a number of centrifuges operating in parallel to provide the required flow of material. The stages are connected in series to give the desired enrichment of the product and depletion of the waste. The largest stages are near the entry point of the feed gas; stages diminish in size toward the product and waste ends. This configuration ensures that gas streams of unequal isotopic composition are never mixed.



FLOW DIAGRAM indicates how a gas-centrifuge plant might operate. Uranium hexafluoride (UF_6) is received at the plant in 14-ton cylinders. They are heated to vaporize the compound, which is then piped to the centrifuge cascade. The product streams, consisting of uranium hexafluoride enriched in uranium 235, are removed from the cascade for final processing and put into $2\frac{1}{2}$-ton cylinders for shipment to a plant that makes the enriched compound into fuel for nuclear reactors. The stream of depleted uranium hexafluoride is then condensed into 14-ton cylinders, which are put in an outside storage yard. A gas-centrifuge enrichment plant that is now under construction at Portsmouth, Ohio, is scheduled to include eight process buildings (each the size of four football fields) and several other buildings.

factors that influence the internal circulation of the gas in the rotor.

The final step in producing a uranium-enrichment plant is to decide how the individual machines should be connected in the assembly known as a cascade. A cascade is simply the interconnection of individual separating units into a larger network in order to amplify the separation effect of each unit. The design of such a network, however, is far from simple.

Suppose each centrifuge delivers a product that is 12 percent richer in uranium 235 than the feed. One would need to repeat the process 13 times in order to attain the uranium-235 concentration of 3 percent (starting with the .71 percent in natural uranium) that is suitable as fuel for a nuclear reactor. The plant is also stripping uranium 235 from the feed to a concentration of about .2 percent in the waste, and the operation requires 11 more stages.

In this simple example 24 units connected in series are needed in a plant that is to process natural uranium into an enriched product containing 3 percent uranium 235 and a waste containing .2 percent. One should not imagine, however, that an isotope-separation plant consists simply of 24 centrifuges lined up one after the other to provide the necessary multiplication of the separation effect. Such a linear array of individual centrifuges would provide the desired enrichment, but because of the small throughput of each machine the quantity of material processed would be pitifully small. An enrichment plant of large capacity would require hundreds of thousands of the 24-unit linear arrays.

Fortunately this simplistic method of connecting centrifuges is not the best way to construct a cascade. It was discovered early in the Manhattan Project that in order to construct the most efficient plant possible with a specified separative capacity from a large number of individual units of fixed separative power, the connections between units must be arranged in a configuration known as an ideal cascade. The term ideal means that whenever two streams are mixed together in the plant, the isotopic compositions must be the same. If they are not, the entropy, or disorder, of the system is increased. An ideal cascade therefore results in the highest possible cascade efficiency. Many types of ideal cascade for enriching uranium with gas centrifuges can be envisioned; it is the task of the process designer to devise the cascade that most effectively exploits the performance characteristics of the component centrifuges.

No matter what particular method is employed for recycling gas within an ideal cascade, the number of individual centrifuges required is the quotient of the specified amount of separative work per year of the cascade and the separative power of the individual centrifuges that make up the cascade. The URENCO centrifuges are believed to have a separative power of from two to 20 separative work units per year. To construct an enrichment plant with a capacity of nine million separative work units per year would require between 500,000 and 5,000,000 centrifuges.

Dixy Lee Ray, the former chairman of the Atomic Energy Commission who is now governor of Washington, has said that the U.S. models have 10 times the separative power of the URENCO centrifuges. Still, from 50,000 to 500,000 centrifuges will be needed in the next enrichment plant scheduled to be built in the U.S., and it will not be the last plant of its size to be built. It is evident that the manufacture of centrifuges is about to become a substantial industry and will remain one for many years.

The organization of a single plant containing a large number of centrifuges entails many novel engineering challenges. Each centrifuge must be quite inexpensive to avoid inordinately large capital costs. Hence the simplest of vacuum pumps and valves must be utilized, piping runs must be held to a minimum and exotic materials of construction must be avoided. Components must be carefully designed to reduce power losses to the lowest level possible. For the foreseeable future, no matter how advanced the individual centrifuges become or what ultimate configuration is devised for the cascade, a gas-centrifuge plant for enriching uranium will involve large numbers of separating units connected by many kilometers of piping through which large quantities of uranium hexafluoride circulate.

Stellar Winds

Stars of various sizes and types are observed to eject matter into space in a slow, steady stream. A number of mechanisms may be responsible for this surprisingly common phenomenon

by Ray J. Weymann

Anyone who has ever watched a rocket being launched into space cannot fail to be impressed by the enormous force required to free the craft from the grasp of the earth's gravitational field. The energy needed to remove a comparable amount of matter from the surface of a star would be hundreds or thousands of times greater. It may therefore come as a surprise to learn that almost every kind of star, including the sun, is ejecting significant amounts of matter into space, sometimes in sporadic violent outbursts and at other times in a bafflingly slow, steady stream.

Understanding how and at what rate stars shed their matter into interstellar space is important for several reasons. First, the loss of matter at certain stages in the life of a star may have a significant effect on the future evolution of the star. For some stars the loss of mass may indeed be critical in determining whether they end their life peacefully or violently. During most of its life a star shines by drawing on the energy liberated by nuclear reactions in its core in which hydrogen is converted into helium and subsequently into heavier elements. Great as this energy reserve is, it is finite, and sooner or later every star must reach a point in its life at which it essentially exhausts its supply of nuclear energy.

Although astronomers' understanding of stellar evolution is far from complete, it seems fairly certain that there are basically two possibilities for the next stage: the star can become either a supernova or a white dwarf. In the first process the star suffers a catastrophic gravitational collapse accompanied by a violent expulsion of its outer layers. In the second process the star simply shrivels quietly into an object about the size of the earth and slowly radiates away its residual heat. It also appears fairly certain that the latter alternative is not available to stars that are at least two or three times more massive than the sun. Recent work on white dwarfs in open star clusters by J. Roger P. Angel and his colleagues at the University of Arizona has confirmed the earlier suggestion that many stars born with a mass substantially greater than the maximum possible for a star to become a white dwarf have actually ended their lives as white dwarfs. That these stars must have shed a substantial portion of their original mass thus seems inescapable, but at what stage in the course of their evolution they do so, and by what mechanism, is still not clear.

Second, understanding the processes by which various kinds of stars eject matter is crucial to determining whether the gas in our galaxy that is not condensed into stars is being enriched or depleted. Even at present, some 15 billion years after the galaxy was born and the oldest stars in it condensed, new stars are continually condensing out of the interstellar gas. Unmistakable evidence for this comes, for example, from the fact that embedded in huge clouds of gas are many massive, tremendously luminous stars whose rate of energy consumption is so high that they can live for only a few million years. If the interchange of material between stars and gas were a one-way process (from gas to stars only), there would be far less material in the form of uncondensed gas than the 10 percent or so actually found today. Astrophysicists interested in studying the evolution of both the stellar and the gaseous content of galaxies (and the gradual change in the chemical composition of the material) must have an improved comprehension of such mass-loss processes to understand galactic evolution.

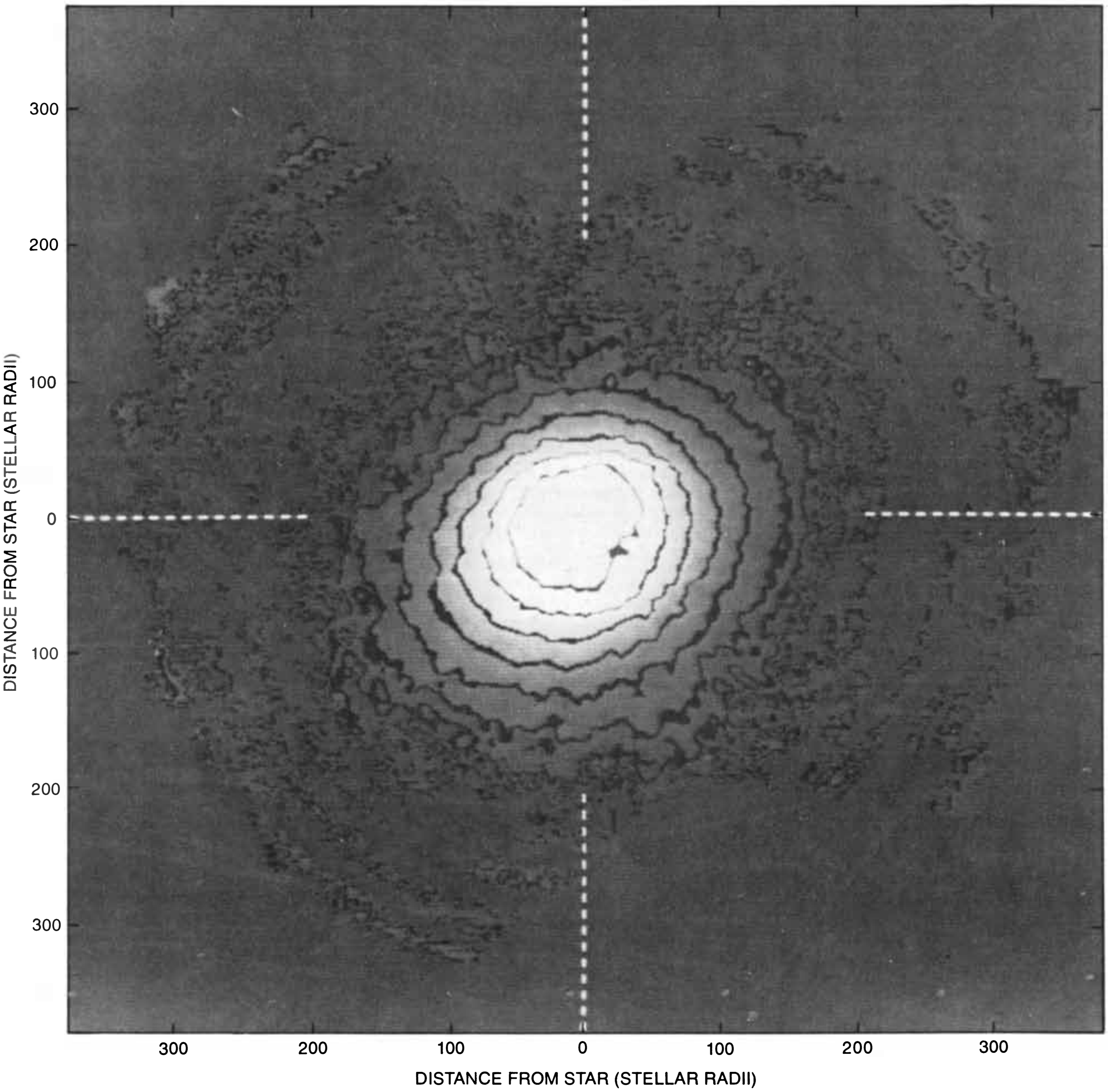
Third, the attempt to understand the mass-ejection process itself is forcing astronomers to examine more carefully the physical processes that take place in the outer layers of stars. These studies involve questions about the circumstances under which the extremely hot envelope of gas called the corona can exist around a star such as the sun, about the circumstances under which solid particles can condense in cool gas surrounding a star and about the correct interpretation of what appear to be very high-speed, large-scale motions of gas in the outer layers of very luminous stars.

The evidence for the loss of matter from stars is based on direct telescopic photographs or (more commonly) on stellar spectrograms. In some instances the matter is ejected in one isolated episode or possibly several. In supernovas and ordinary novas, for example, the material is usually expelled in a single violent explosion. In planetary nebulas (so named because visual inspection through small telescopes gave early astronomers the impression of a planetary disk) a single shell of gas has been ejected but in a much gentler fashion. In other instances the material seems to be streaming away continuously from the stellar surface, sometimes at very high speeds and sometimes at low speeds, suggesting that a variety of quite different mass-loss mechanisms are at work.

The best-known example of such continuous ejection is the sun. It is now almost two decades since it was discovered that a "solar wind" is blowing from the sun and sweeping past the earth. In the remainder of this article I shall concentrate on those mass-loss processes in stars other than the sun that also involve a steady flow of matter, describe the evidence for their existence and then explore the question of whether the process thought to produce the solar wind is also responsible for producing these "stellar winds."

The solar wind has been described in some detail in several previous articles in this magazine, but a brief summary of its properties and origins is appropriate here. The solar corona is a tenuous, roughly spherical halo of extremely hot gas surrounding the sun. The processes that give rise to its high temperatures (between one and two million degrees Kelvin) are not well understood, but they are thought to involve the dissipation of energy propagated upward from the solar surface by waves of various types, with magnetic fields playing an important role in channeling the energy. E. N. Parker of the University of Chicago was the first to draw attention to the fact that the solar corona cannot be static. The temperature of the coronal gas is so high that even at considerable distances from the sun's surface the sun's gravity is not able to reduce the pressure the gas exerts to values low enough for it to be confined by the pressure of the interstellar gas. As a result the solar corona continuously expands into the interstellar gas and is replenished by new material slowly flowing up from lower levels of the solar atmosphere until it enters the zone of intense coronal heating. In this simplified description of the solar wind the push necessary to lift any particle of matter free of the sun's gravitational field is delivered continuously in the form of a gradient in the pressure of the coronal gas and is effective because of the very high temperature of the gas. Such a wind is said to be thermally driven.

How widespread among stars of various types might such thermally driven winds be? By all indications there is nothing exceptional about the sun, and it is reasonable to suppose other stars similar to the sun in terms of surface



HALO OF GAS surrounding the red-supergiant star Alpha Orionis (commonly known as Betelgeuse) can be seen in this computer-generated image made recently by C. Roger Lynds, Leo Goldberg and Jack Harvey, using the spectrograph connected to the four-meter Mayall telescope at the Kitt Peak National Observatory. The image was obtained by aiming the slit of the spectrograph at several positions in the vicinity of the star and recording the intensity of the radiation at or near the characteristic wavelength emitted by neutral potassium atoms. After subtracting the light from the star itself (which is scattered by the earth's atmosphere and by the instrument) and some of the radiation scattered by dust particles in the stellar envelope, the result is a contour map indicating the intensity of the light scattered by potassium atoms in the stellar envelope. As the illustration shows, the gas can be traced out to distances several hundred times the radius of the star before the scattered light becomes too faint to detect. At such distances the star's gravitational field is so weak that the outward velocity of the gas is in excess of that required for it to escape. The radius of the star itself is approximately 800 times that of the sun.

gravity and surface temperature (the two properties that fix most of the observable characteristics of a star's atmosphere) have coronas and attendant thermally driven winds. It is almost impossible to detect such winds in association with other sunlike stars; it is only through direct measurements from spacecraft and indirect observations of the effects of the wind on the tails of comets that one can detect the solar wind, and of course neither of these two techniques is applicable to the stars. Moreover, brisk as the solar wind is, the amount of material in it is quite small, and at the present rate it would take several trillion years for the sun to lose a substantial portion of its mass. This is much too slow to be of interest for studies of either the evolution of the sun or the mass balance of the interstellar medium, notwithstanding the large number of sunlike stars.

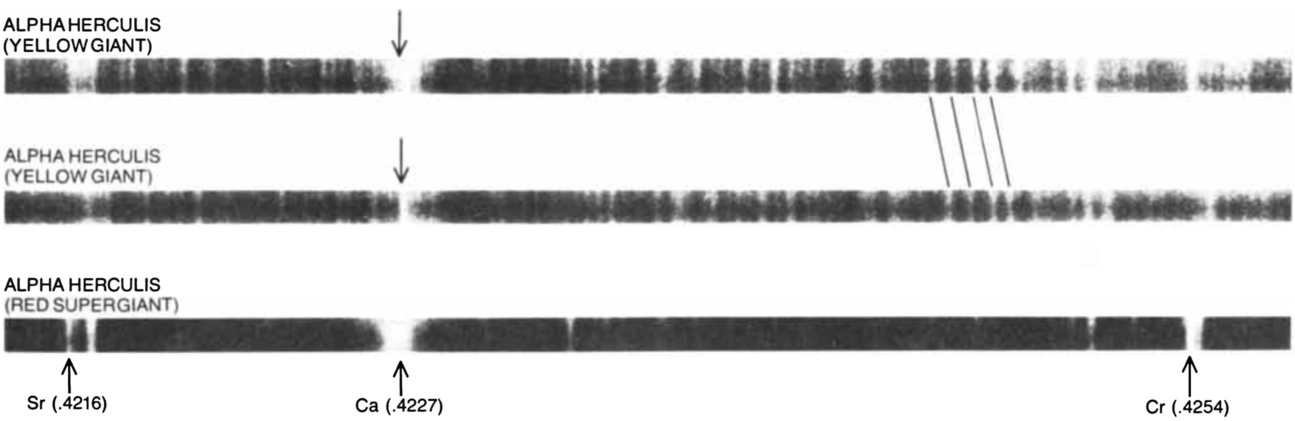
One of the best-established events in the life of a star is a very large increase in the size of its outer layers at just about the time it has exhausted the hydrogen in its core (its principal fuel for the bulk of its life). Stars with such vastly distended outer layers are known as red giants or red supergiants, because of their size and because their comparatively cool surface temperature gives them a reddish hue. One result of their enormous size (the radius of some red supergiants is so large that the earth's entire orbit would fit easily inside it) is that the velocity required for matter to escape from their surface is very much lower than it was earlier in their life. This fact makes the red-giant phase a natural candidate to examine for a stage where one would expect a substantial loss of mass to occur.

In 1956 Armin J. Deutsch, a spectroscopist at the Hale Observatories, made a discovery indicating that the red giants are indeed shedding vast amounts of matter into space. His finding initiated an intensive investigation into the outer envelopes of these stars. To appreciate what Deutsch found, the following basic points of astronomical spectroscopy are useful. Spectrograms of most stars reveal many narrow wavelength regions nearly devoid of light compared with adjacent regions of the spectrum. These absorption lines arise from the selective removal of specific wavelengths from the outward streams of light by atoms in the cooler outer layers of the star, a phenomenon discovered in the sun's spectrum by Joseph von Fraunhofer in 1814. It had been known since the 1930's that in the spectra of many red giants an extra set of absorption lines was present, displaced blueward in wavelength by a small amount from the normal set of lines but confined to those absorption lines that can be produced in a gas only at very low temperature and density. The blueward displacement is caused by the Doppler effect and indicates a net outward flow of gas from the stellar surface.

Earlier spectroscopists had been reluctant to assume that the outflow represented a permanent net loss of matter from the star, since even from the surface of these red giants the energy required for matter to escape is still typically 100 times the energy associated with the measured velocity of the outward-moving gas. What Deutsch discovered was that much of the material is not located just above the stellar surface but is at distances several hundred times the radius of the star, and at such distances the kinetic energy of the moving gas is indeed more than enough to allow it to escape. The discovery was made by analyzing spectrograms of the red-supergiant star Alpha Herculis. When this object is viewed through a telescope, it appears to be actually a double star, the companion star to the supergiant being itself a rather ordinary yellow giant, separated in the sky from the supergiant by an angle sufficiently large to make it possible to obtain a spectrogram of the companion uncontaminated by the light from the red supergiant.

The spectrogram of the yellow companion revealed a set of displaced absorption lines similar to but somewhat weaker than those of the supergiant. The fact that lines such as these are not otherwise observed in association with single yellow stars of the same type was compelling evidence that the absorption lines arise from a gigantic expanding cloud enveloping both the red supergiant and the yellow companion, which is known to orbit the supergiant at a distance that is more than 100 times the radius of the supergiant.

The clinching evidence for this interpretation came from the discovery that the yellow companion itself is a binary star. The fainter star of this pair is too close to the yellow giant to be resolved from it in direct photographs, and it contributes an insignificant amount of light to the spectrum; nevertheless, its presence is detected by its gravitational field, which causes a continuous and periodic Doppler shift of the light from the yellow giant as the yellow star and



SPECTROGRAMS OF TWO STARS in the Alpha Herculis multiple-star system were made in 1956 by Armin J. Deutsch of the Hale Observatories. The spectrograms, which are printed here as negatives, led to the discovery that massive stellar winds flow from the surface of the red supergiants. The spectrogram at the bottom is of the red-supergiant member of the Alpha Herculis system. The upward-pointing arrows mark gaps in the emission spectrum of the star where light has been absorbed by atoms of strontium (*Sr*), calcium (*Ca*) and chromium (*Cr*) in the gas flowing from the stellar surface. (The accompanying numbers give the wavelengths of the absorption lines in micrometers.) The two upper spectrograms are of the yellow-giant member of the system; they were made several days apart. Because the yellow-giant star is itself locked in orbital motion around an unseen fainter companion, the wavelength of the light absorbed in the yellow giant's atmosphere shifts periodically, owing to the Doppler effect; slanted lines indicate typical Doppler shifts. The gas from the red supergiant extends to such vast distances, however, that some of it can be detected in the spectrograms of the yellow giant; for example, the downward-pointing arrows mark the absorption lines produced in the yellow-giant spectrograms by calcium vapor in gas ejected from the red supergiant. The fact that these lines are not Doppler-shifted demonstrates that they are not associated with the yellow giant.

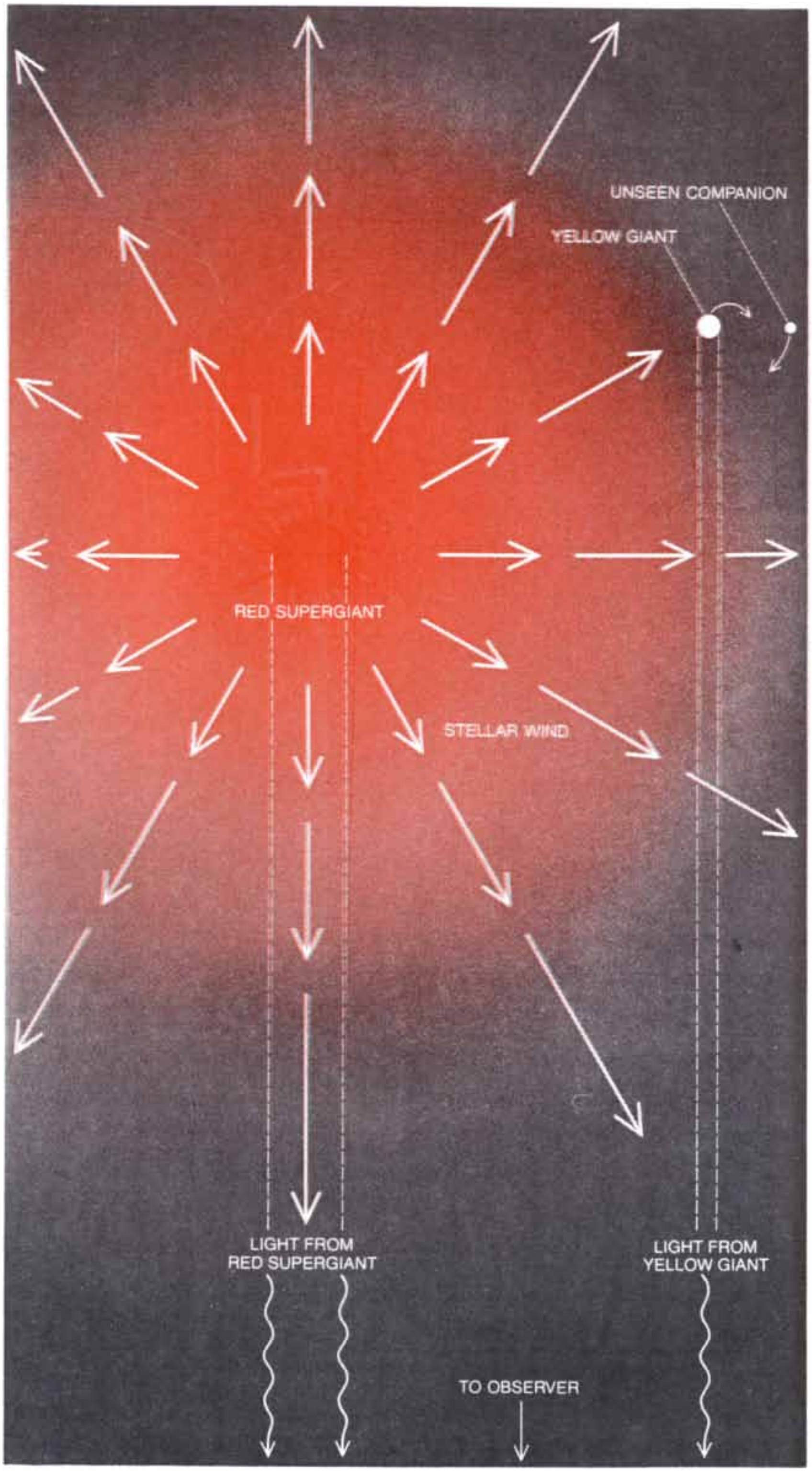
its unseen companion orbit each other. The gas that causes the displaced lines does not participate in the periodic shift, however; it evidently comes from a cloud much larger than the immediate neighborhood of the yellow giant.

With the realization that the displaced absorption lines in the spectra of red giants truly reflect a substantial loss of matter from such stars, Deutsch and several other investigators then surveyed many red giants and established that substantial loss of matter is a common property among them. Recently independent evidence for these large, gently expanding stellar winds has been obtained in several ways. Radio astronomers have found that in the very coolest of the red giants the expanding envelopes are sufficiently cool and extended for the emission of microwaves from hydroxyl (OH), water (H_2O) and silicon oxide (SiO) molecules to be detectable. In many instances a natural maser operates in these clouds to make the molecular emission very intense. The Doppler shifts associated with the microwave spectral emission lines are most readily interpreted in terms of an expanding envelope, and the picture derived from them is in general fairly consistent with the one obtained by the optical spectroscopists.

A second item of striking independent evidence comes from two groups working at the McDonald Observatory and at the Kitt Peak National Observatory. The same atoms in the stellar envelope that selectively remove the outflowing radiation from the underlying star by scattering it, thereby producing the absorption lines in the spectrum, simultaneously produce a faint halo of light around the star. Fog produces a faint halo around a distant streetlight in much the same way, except that in the case of the star only light in narrow wavelength intervals is scattered. Both groups of observers have now succeeded in detecting this glow, thereby reconfirming the enormous extent of the envelopes surrounding such stars.

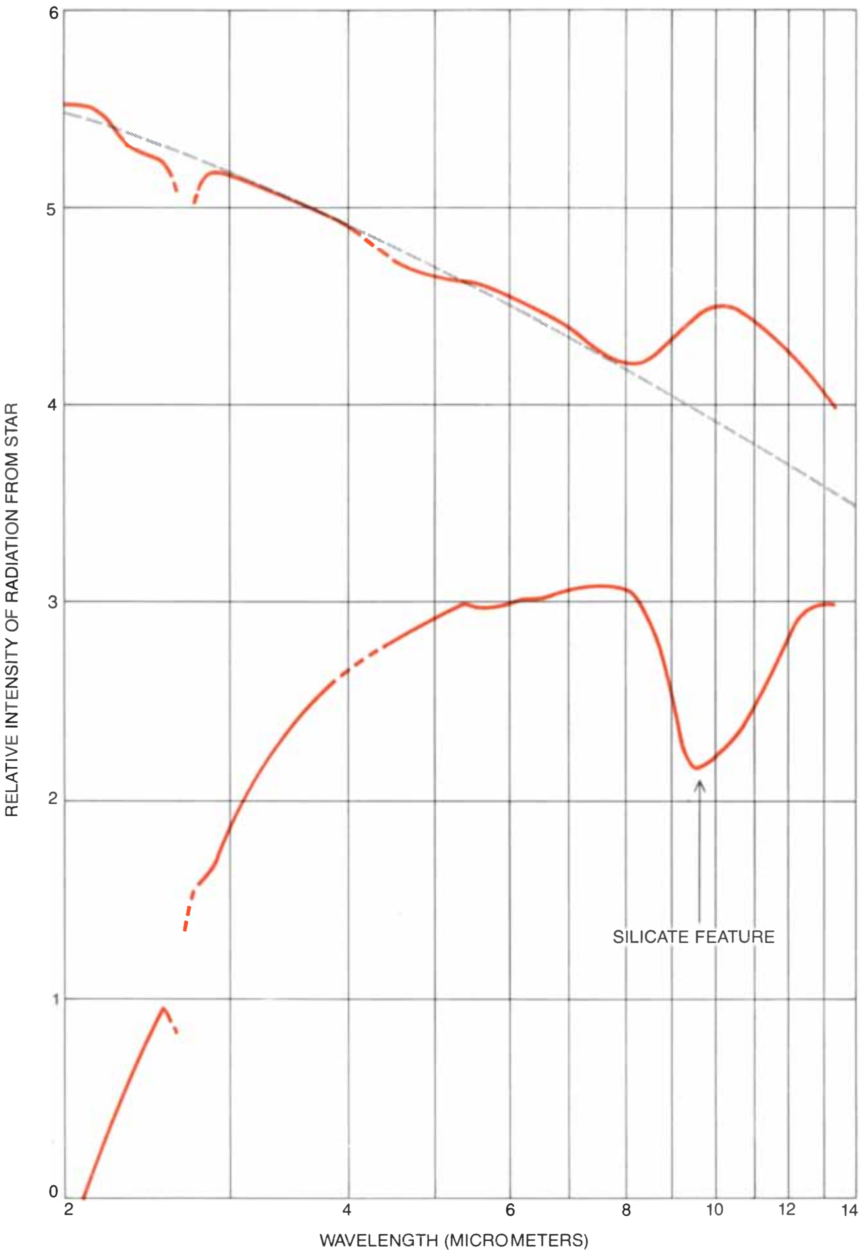
These observations of the red giants established the red-giant stage of evolution as an important one from the point of view of mass loss. They also raised the question of what mechanism is responsible for it. The outward velocities that are actually observed are in nearly all cases much less than those required for a particle of matter to escape from the star's gravitational field at its surface unless there is some extra push. What provides the extra push?

It happened that Deutsch's initial discovery was made at just about the same time Parker's theory of a thermally driven solar wind was meeting with considerable success in explaining solar phenomena. It was therefore widely assumed then (and it still is today) that the



GEOMETRY of the Alpha Herculis multiple-star system is shown schematically in this illustration. The red-supergiant star is surrounded by an envelope of fairly warm, turbulent gas, which is accelerated by some mechanism to form the stellar wind. This outward-flowing gas can be detected in the light from the supergiant star with the aid of a spectrograph. The stellar wind extends to such a large distance from the red supergiant that it encompasses the somewhat smaller yellow-giant companion and can be detected spectroscopically in the light from this star. The yellow giant and an even smaller unseen companion revolve around each other, but the gas from the red supergiant's stellar wind does not share in the motion, indicating that most of this gas is far from the yellow giant. Diagram is not drawn to scale; separation of the yellow giant from the red supergiant is actually more than 100 times the radius of the red supergiant.

same mass-loss mechanism must be operating in the red giants. Nevertheless, there are some serious difficulties with this assumption. If extremely hot coronas with a temperature similar to that of the solar corona surround such stars, it is hard to see why the material should not be streaming away from the stars at a speed comparable to that of the solar wind. In fact, the material in the red-giant winds is moving away at a speed very much less than that of the solar wind. On the other hand, coronas of a more moderate temperature (but still hot enough to drive matter away from the red giants) ought to radiate away far more energy than they are observed to do. Actually the available evidence suggests that the coolest of the red giants do not have a corona at all but only a lukewarm envelope called a chromosphere, which is far too cool to cause any mass ejection.

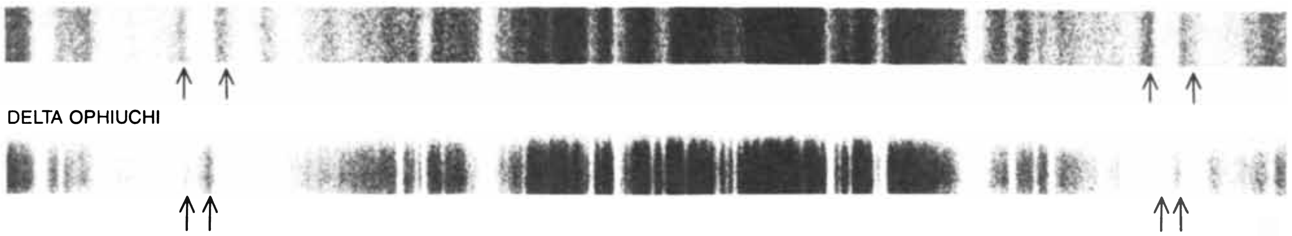


EFFECT OF STELLAR DUST on the spectral energy distribution of two comparatively cool stars is presented in this graph, based on a recent study conducted by K. Michael Merrill of the University of Minnesota. In the case of the star designated Mu Cephei (*upper colored curve*) the amount of dust in the gaseous envelope surrounding the star is rather small, and its only influence on the star's spectral energy distribution is a prominent "emission bump" at a wavelength of about 10 micrometers. In the absence of the dust the spectral energy distribution would approximate that given off by a surface at a temperature of 3,300 degrees Kelvin (*broken black line*). In contrast the star AGL 2205 (*lower colored curve*) is surrounded by a shell of dust so thick that virtually no radiation from the star's surface can penetrate it. The distinctive spectral feature at about 10 micrometers now appears as an "absorption trough," and the radiation that emerges from the outermost layers of dust is characteristic of a surface with a temperature of only about 375 degrees K. The strong interaction of the dust particles with the stellar radiation at wavelengths of approximately 10 micrometers is an indication that silicate particles are an important constituent of the dust. Radiation pressure acting on such small dust particles probably plays an important role in mass-ejection process, particularly in cool stars.

A new element entered the picture in the 1960's with the development of sensitive infrared detectors and telescopes specially designed or modified to operate effectively at infrared wavelengths. When the infrared portions of the spectra of red giants were examined, many of them showed radiation in excess of that expected from the surface of the stars. Moreover, the spectral distribution of the excess radiation frequently had a characteristic "bump" in it. Edward P. Ney and his collaborators at the University of Minnesota pointed out that this broad spectral feature could be understood as infrared emission arising from small solid particles of silicates such as olivine (Mg_2SiO_4) that condensed in the extended outer envelopes surrounding the red giants.

Several other workers then at Minnesota, including N. J. Woolf and R. C. Gilman, realized that the solid particles could efficiently tap the most plentiful source of energy available to drive the mass loss: radiation from the star itself. The essence of the alternative to the thermally driven stellar wind that was proposed by Woolf and Gilman involves the following sequence of events. At even a moderate distance from the surface of the star small "protograins" of silicates are able to radiate away the energy absorbed from the star with enough efficiency for them not to get hot enough to evaporate. In fact, they can grow larger. The radiation pressure exerted on the grains by the stellar radiation is many times greater than the pressure required to overcome the star's gravity, and the grains quickly accelerate. Most of the gas (mainly hydrogen and helium) remains uncondensed and therefore does not respond directly to this radiative force. The outward-moving grains will collide with atoms of the uncondensed gas, however, and will communicate their momentum to the gas atoms; hence energy is quickly distributed to the entire gas. In short, the radiative force on the grains is adequate to drag the uncondensed gas along with it and lift it away from the star.

The key question in this otherwise plausible scheme is what the "moderate distance" above the stellar surface at which solid particles can condense actually is. If the condensation distance is quite large, most of the work against gravity must be done by some other mechanism in order to get a significant amount of gas out to the point where it can condense into grains. If this is the case, the grains present in many of the envelopes of the red giants are simply a



DIFFERENT DEGREES OF TURBULENCE in the atmospheres of stars of different luminosities are manifested in these two stellar spectrograms. The spectrogram at the top is of the extremely luminous red supergiant Alpha Orionis; in this negative print all the bright absorption lines (except those arising from the stellar wind itself) appear diffuse and washed out, owing to large-scale motions in the stellar atmosphere. The corresponding absorption lines in the spectrogram of the smaller, less luminous red giant Delta Ophiuchi (*bottom*) are much sharper, indicating less turbulence in the atmosphere of this star. The effect is particularly apparent in the difference between the widths of the dark emission lines associated with singly ionized calcium atoms (*arrows*). The close correlation between the luminosity of a star and the width of these emission lines was first established by astronomers at the Hale Observatories. The notion of "astronomical turbulence" associated with such spectral line widths may have little in common with the well-known phenomenon of turbulence in fluid dynamics; the true character of these atmospheric motions and their possible role in producing stellar winds are not yet understood.

by-product of some other mechanism causing the ejection rather than the principal cause of the ejection. Many aspects of the process by which solid particles condense out of a gas are obscure. The question of the distance above the star's surface at which solids can condense and survive involves not only the intensity and temperature of the radiation field impinging on the grain but also a knowledge of both the detailed chemical composition and the sizes of the grains.

The last two properties are poorly known. Woolf and Gilman assumed that the grains are quite free of impurities. In that case they will absorb a fairly small amount of the red and near-infrared radiation from the star (although scattering it effectively enough so that a large radiative force is still exerted on the grains), but they will radiate quite efficiently at much longer wavelengths. The result is that the grains will stay cool enough to grow and survive fairly close to the surface of the star for the same reason that houses whose roofs are painted with visually reflecting but infrared-emitting "cool white" paint stay fairly cool even at midday in the summer. Recent work by K. Michael Merrill of the University of Minnesota and Wendy A. Hagen of the University of Hawaii, however, suggests that the grains contain enough impurities for them to absorb so much near-infrared radiation from the star that their formation and survival near the surface of the star is impossible. If that is the case, the Woolf-Gilman mechanism becomes untenable.

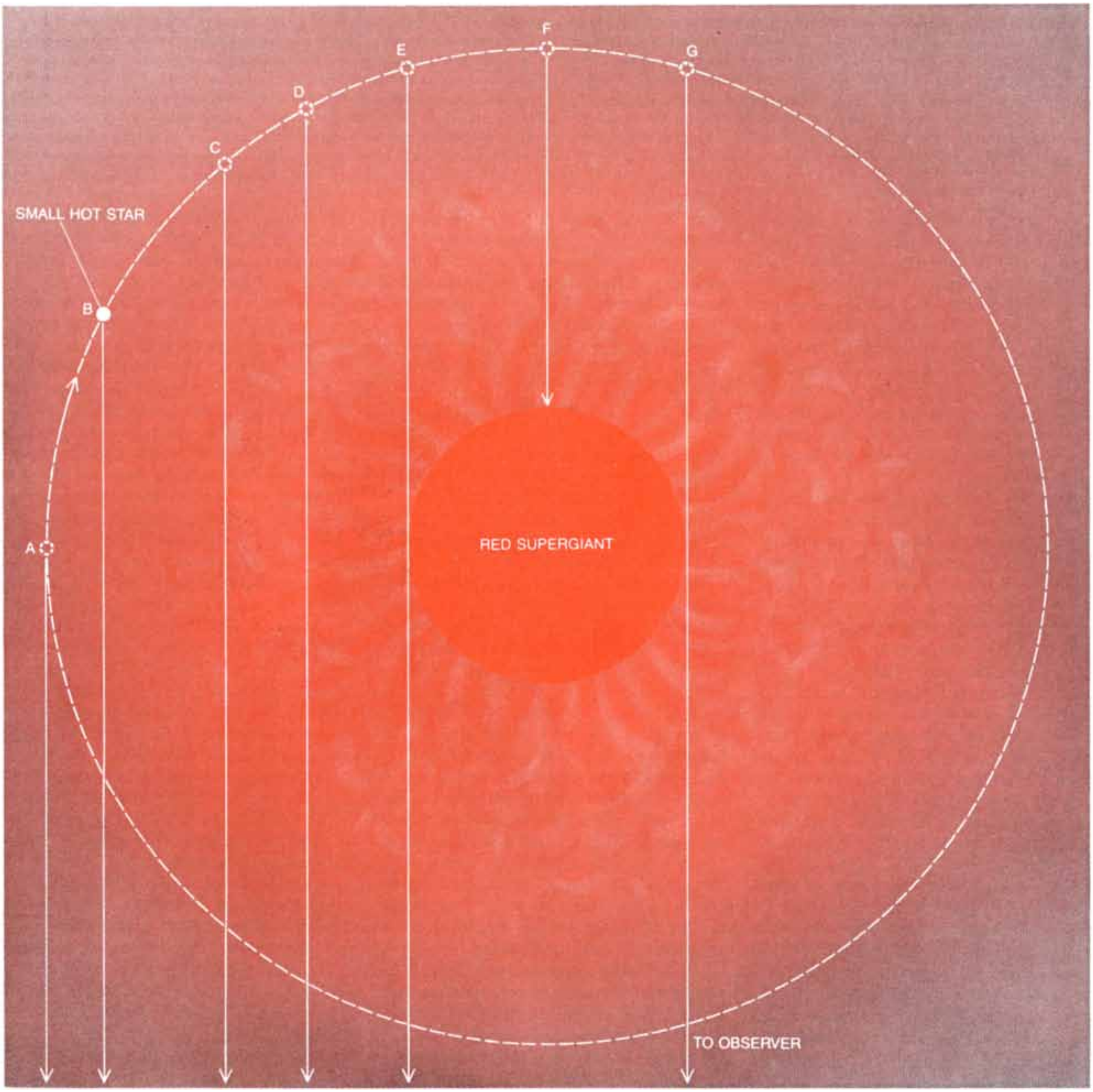
Even in the absence of such impurities an added push to the material is required in the Woolf-Gilman model to supply enough material at the level of the grain-condensation radius so that a non-negligible rate of mass loss will occur. Woolf and Gilman suggest that turbulent motions supply the additional required momentum. The manner in which the word "turbulence" has been used by astronomers has long caused chagrin among aerodynamicists, for whom the word has a fairly well-defined meaning, signifying chaotic, fluctuating flow. Turbulence is suddenly encountered when, for example, a smooth flow of fluid through a pipe exceeds a certain critical speed. Astrophysicists, however, have tended to use the word in discussing stellar atmospheres to explain the widths of either emission or absorption lines whenever the widths cannot be ascribed either to microscopic motions associated with the temperature of the gas or to organized and well-understood large-scale motions such as the rotation of the star as a whole.

Whatever kind of gas motion this "astronomical turbulence" really represents, it is certainly present to a high degree in the luminous red supergiants. At about the same time that Deutsch made his discovery in Alpha Herculis, Olin C. Wilson, another spectroscopist at the Hale Observatories, and V. K. Bappu, a visiting astronomer from India, found a remarkable correlation between the degree of turbulence in the outer atmosphere of a star (as manifested in the widths of ionized calcium lines) and the total energy output of the star. The physical reason for the correlation is still not understood today, nor is the nature or origin of the gas motion responsible for the turbulence.

Nevertheless, suppose one simply assumes that the gas motions contribute to a "turbulent pressure" in the same way that the motions of individual molecules contribute to ordinary gas pressure. One can then show that the turbulent pressure provides enough support of the gas against gravity out to the distance where, when condensation occurs in the most favorable case of impurity-free silicate particles, the radiation pressure acting on the grains will drag along the required amount of gas. The nature of these turbulent motions remains a key unresolved question. Fortunately, nature has provided a handful of objects that make it possible to probe into the structure of the outer layers of the red supergiants and so gain some insight into these motions.

As I have mentioned, a star evolves into the red-giant phase after the prime supply of hydrogen fuel in its core has been consumed. The more massive the star is, the faster the fuel is exhausted. The subsequent evolution into a red giant may involve an expansion of the star's radius by as much as 100 times. Almost as often as not, stars are formed not singly but in pairs, revolving around each other in orbits whose periods range from hours to thousands of years. If the masses of the two stars are unequal, then the less massive of the two will be left behind in its evolution as a small but hot (hence still quite luminous) companion as the more massive one enters the red-giant phase. Finally, if the angle at which astronomers on the earth view the orbit of the pair is almost exactly edge on, then eclipses of the individual stars will be observed. The special nature of such a pair means that the small, hot star, as it moves in its orbit behind the outer layer of its red-giant companion, provides an intense, thin beam of light probing successively deeper layers of the red giant's envelope. Because of the large difference in surface temperature between the red giant and its companion, the small, hot star dominates the energy radiated in the violet part of the spectrum, and its light is not strongly contaminated by that from the red giant. Needless to say, such special circumstances make stellar pairs of this type rather rare. Of all the stars in the sky, scarcely half a dozen are known that satisfy the conditions described above and also are bright enough to be studied in detail.

Nevertheless, study of the eclipses of these objects, primarily by Kenneth O. Wright of the Dominion Astrophysical Observatory and his collaborators, has provided the following important additional items of evidence on the outer layers of the red supergiants and the motions of the gas in them. First, the rate at which the gas density falls off with increasing distance from the red giant's surface ("surface" being defined as the level at which the gas becomes opaque) is quite slow, much slower than would be expected on the basis of ordinary gas pressure but about what one would expect from the large-scale gas motions. Evidently there is something to the naïve concept of turbulent pressure used by the astronomer. Second, there is unambiguous evidence for very-high-speed gas motions in these outer layers, higher even than those measured by Wilson and Bappu. Third, and perhaps most significant, the gas detected in such objects is not distributed uniformly but is clumpy. This phenomenon is not unusual in astrophysics; for example, interstellar gas is often concentrated in clouds. Furthermore, the existence of clouds or filaments in the outermost layers of the red-giant stars would hardly surprise students of the solar corona, who are thoroughly familiar with clouds of cool gas; the solar prominences, which are sometimes quiescent and sometimes moving at high speeds, are



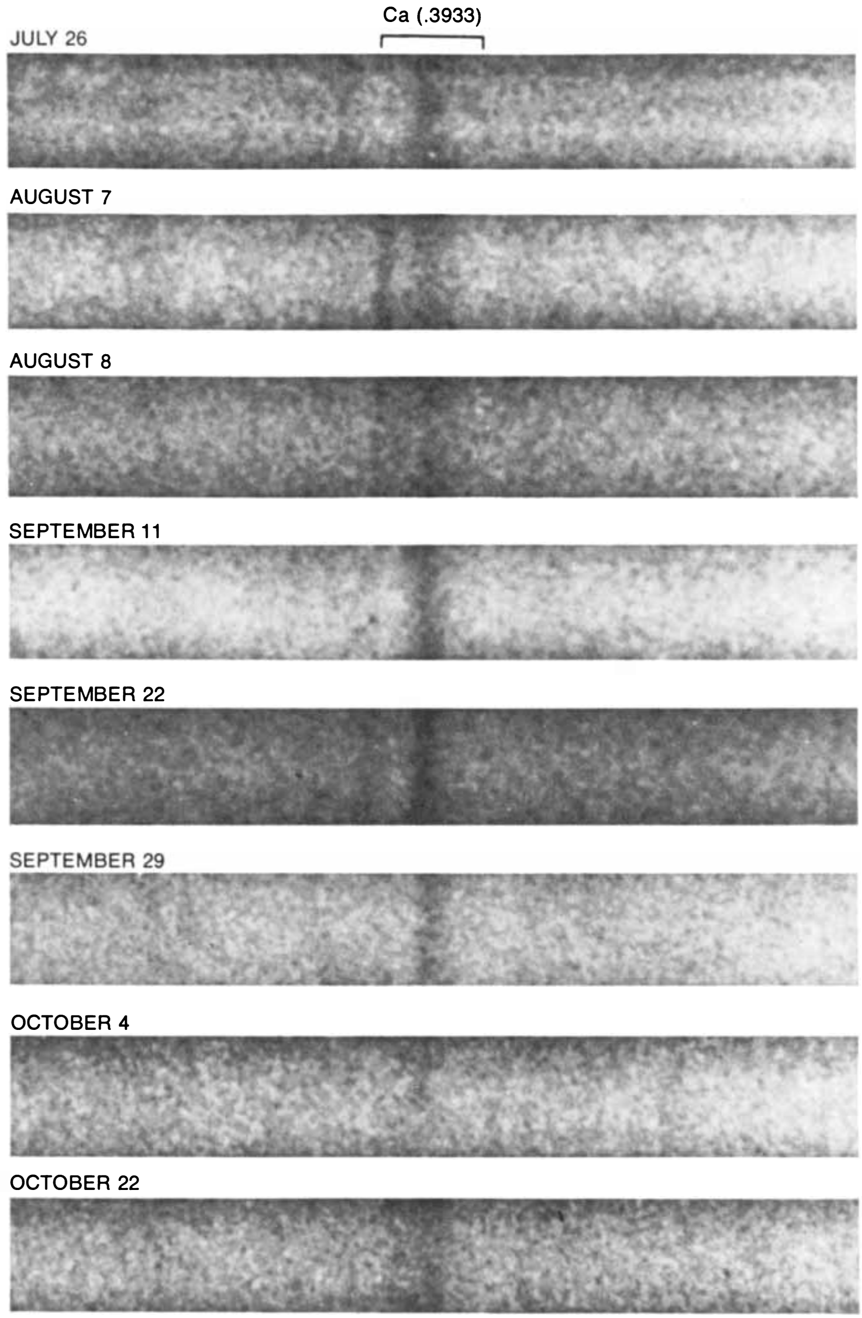
ECLIPSING DOUBLE-STAR SYSTEM, designated 31 Cygni, provides astronomers with an opportunity to probe successively deeper layers of a red supergiant's gaseous envelope. As this schematic diagram shows, the system consists of a small, hot star revolving around a red supergiant. The system is aligned so that the earth lies in the plane of the orbiting star. The period of the system is only about 10 years, making it possible to analyze all phases of the orbit. During the part of the orbit from *A* to *B* the light from the small star misses the red supergiant completely, but from *C* to *E* the light passes through increasingly dense layers of the large star's outer atmosphere. From *E* through *F* to *G* the red supergiant is opaque and the small, hot star is completely eclipsed. Although the distance between the two stars as seen from the earth is so small that the light from them cannot be studied separately, it is known that the small, hot star contributes most of the light at short wavelengths. Spectroscopic analysis of the system during the portion of the orbit from *C* to *E* indicates that outer layers of the red supergiant's atmosphere are composed of a number of discrete gas clouds, some of them moving at very high speeds.

clouds of cool gas embedded in the much hotter solar corona.

The implication of these observations is that the mechanism producing the gentle winds flowing from the red giants is much more complex than the theoreticians, in their zeal to reduce nature to manageable simplicity, have supposed. The picture that is beginning to emerge is that cloud-producing activity akin to the processes at work in solar prominences (perhaps associated with a very hot rarefied gas that confines such clouds) and grain condensation in the cool clouds both play important roles in the mass-loss process at work in the red giants.

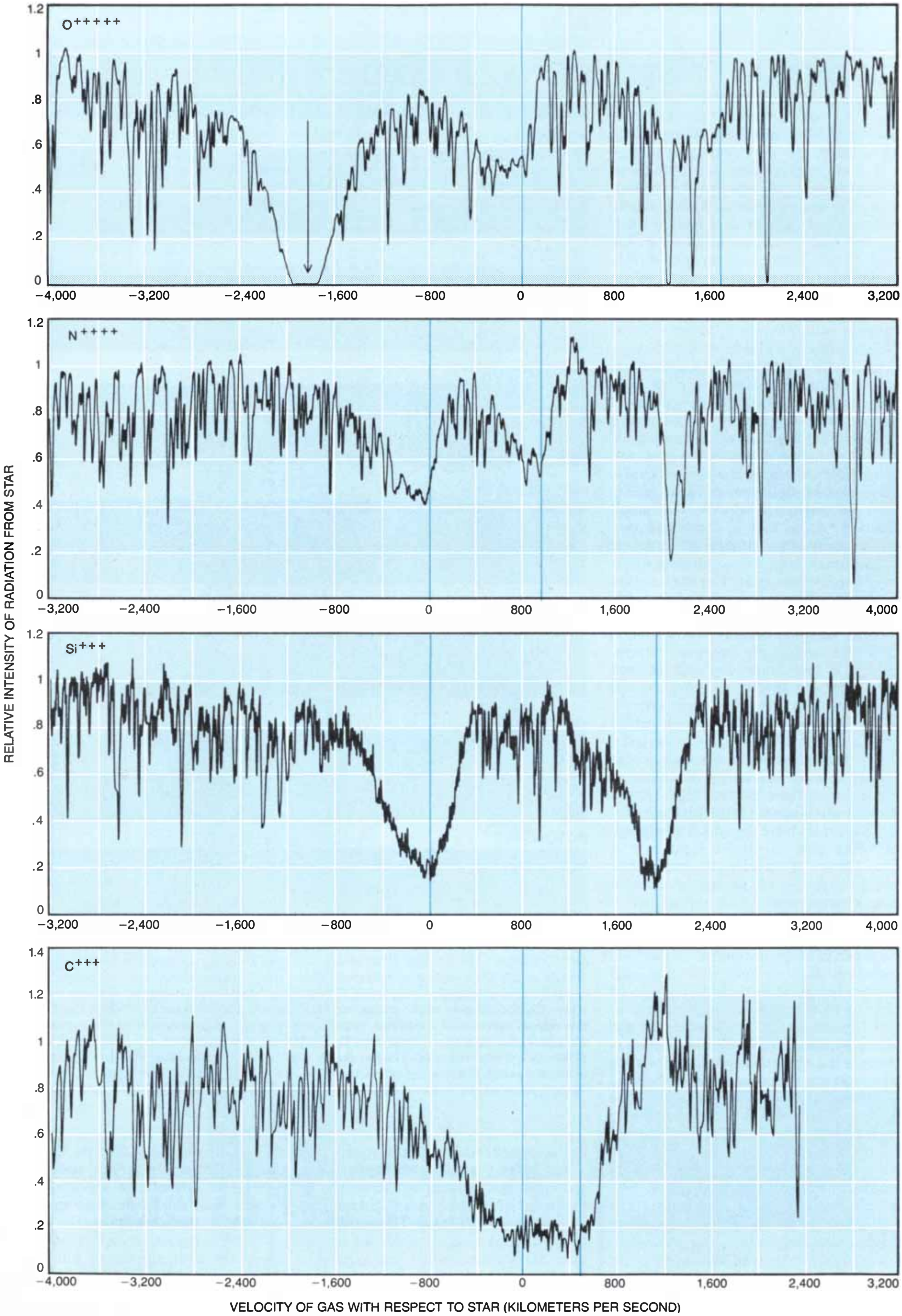
There is an entirely different class of stars that are also known to be continuously shedding mass into interstellar space, although the phenomenon is in some respects quite different. Before I discuss these objects a few words are in order about the effects of radiation pressure. If radiation pressure acting on solid particles is indeed the correct explanation for the stellar winds in red giants, then one might ask why the very small fraction of gas that is condensed into grains—only about a tenth of 1 percent of the uncondensed gas—is more effective than the gas itself. The explanation is that the abundant species of atoms in the atomic gaseous phase (hydrogen, helium, oxygen, carbon, nitrogen and neon) are completely transparent to radiation in the yellow, red and infrared regions of the spectrum where the red giants radiate most of their energy. These atoms are effective absorbers only in the ultraviolet region of the spectrum, where the red giants radiate a negligible amount of energy. In contrast, the grains can intercept red light and infrared radiation with fairly high efficiency. The grains have the added advantage of being able to absorb light over a broad range of wavelengths rather than being limited to the narrow absorption lines characteristic of an atomic gas.

Nevertheless, these considerations raise the question of whether stars that do produce large amounts of ultraviolet radiation (that is, stars with very high surface temperatures) might not also be able to shed matter through the radiative forces exerted on the common gas atoms by ultraviolet radiation. The stars that are the obvious candidates for such a process are the progenitors of the very luminous red supergiants, stars that have not completely stopped burning hydrogen in their core and are therefore on or near the "main sequence." Such stars have radii that are considerably smaller than the red supergiants into which they will evolve. As a result some 100 times more energy is required to remove material from their surface than is required for the red supergiants, but that is compensated for by the intense radiation near the surface of these stars.



SPECTROGRAMS OF ECLIPSE in the 31 Cygni system were made in 1961 by Kenneth O. Wright of the Dominion Astrophysical Observatory. As the small, hot star moves in its orbit around its red-supergiant companion, it acts as a probe of the larger star's tenuous outer layers, where the acceleration of the stellar wind takes place. The presence of discrete clouds of gas in this region, rather than a uniform, smooth flow of gas, is demonstrated by the appearance and disappearance of narrow absorption lines (*dark lines in these positive prints*), which are attributable to calcium vapor (*Ca*) in the clouds. About two months after the last spectrogram was made the small, hot star moved behind the opaque body of the supergiant star. During this "total" phase of the stellar eclipse only light from the red supergiant's surface was recorded.

For many years spectroscopists using ground-based telescopes had suspected that substantial mass loss might be characteristic of such stars. The ultraviolet portion of the spectra, where the important absorption would be observed, however, is blocked by the earth's atmosphere. In 1967 Donald C. Morton of Princeton University succeeded in obtaining ultraviolet spectrograms of several hot stars from instruments lifted above the atmosphere in a rocket. The spectrograms showed displaced absorption lines similar in character to those seen in the red giants, except that



O+++++
N++++
Si+++
C+++
RELATIVE INTENSITY OF RADIATION FROM STAR
VELOCITY OF GAS WITH RESPECT TO STAR (KILOMETERS PER SECOND)

the velocity of the outflow was 100 or 200 times greater than it was in the red giants. These tremendous outflows Morton termed stellar hurricanes. In contrast to the gentle breezes from the red giants, the winds from the hot stars indicate speeds high enough for matter to escape even if it is only just above the star's surface. Of course, the question of what mechanism produces the tremendously high speeds observed, amounting in some cases to about 1 percent of the speed of light, still remains.

As suggested above, these hot stars are prime candidates for examples of mass loss driven by radiation pressure acting on the absorption lines of the abundant elements in the extreme ultraviolet part of the spectrum. The first quantitative theoretical model for this mechanism was developed by Leon B. Lucy of Columbia University and Philip M. Solomon of the State University of New York at Stony Brook. An interesting feature of the mechanism is that the force exerted by the radiation at any point in the flow depends strongly on the velocity of the material below that point. The reason is that the ultraviolet absorption lines of the common elements are so strong that they will quickly attenuate the radiation field at these selective wavelengths if all the material is moving at the same speed. If there is a sufficient gradient in the velocity of the material, however, then, owing to the Doppler shift, atoms in successively higher layers will find "fresh," unattenuated radiation that can continue to accelerate the material. In this way the atoms continue to escape from the "shadow" of the underlying particles. Of course, the faster the flow is, the more momentum there is to be extracted from the star's radiation field and used to drive the flow. By the same token, the faster the flow, the more momentum required for the flow. Because of this compensatory feature Lucy and Solomon showed that the rate of mass loss attributable to this mechanism can be estimated independently of the speed of the wind. It is, in fact, roughly equal to the number of strong ultraviolet resonance lines times the luminosity of the star divided by the square of the speed of light. (A resonance line is one in which the lower of the two energy levels from which the electron jumps when it absorbs light is the lowest energy available to the electron.)

Since the initial work of Morton and of Lucy and Solomon, advances in understanding of the stellar hurricanes have been made in both theory and observation. Elaborate and detailed theoretical models based on the Lucy-Solomon mechanism have been constructed by John I. Castor and his associates at the University of Colorado. Their models can now be compared with detailed data obtained by a group at Princeton using the spectrometer on the *Copernicus* satellite. It seems fair to say that there is certainly qualitative agreement between the predictions of the model and the observations. There is, however, one outstanding anomaly in the observations that has become a rallying point for a group of astronomers, led by Richard N. Thomas of the Institut d'Astrophysique in Paris, who are not convinced that the picture of mass loss driven by radiation pressure acting on the atomic absorption lines is the whole story, or even the most important part of it. The anomaly was first found in the star Tau Scorpio by John B. Rogerson, Jr., and H. Lamers, who used *Copernicus* data.

What Rogerson and Lamers found were absorption lines arising from oxygen in a highly ionized state. This high degree of ionization (in which the outer five orbiting electrons have been stripped off) was indicative either of an anomalously large flux of radiation in the part of the star's spectrum corresponding to soft X rays (a part of the spectrum not accessible to *Copernicus*) or of an anomalously high temperature (on the order of 200,000 degrees) in some regions of the flow. The latter possibility is reminiscent of the solar corona, which shows weak emission from highly ionized material produced by a hot gas, which in turn plays a crucial role in the production of the solar wind. Until this discovery it was supposed these hot stars lacked a corona. (They are considered hot in the sense that temperatures at their visible surface range from 20,000 to 50,000 degrees, compared with the 5,000-to-6,000-degree range of sunlike stars and the 2,000-to-4,000-degree range of the red giants.) This discovery provoked a lively controversy over whether the high ionization was produced by excess soft X rays or by a hot corona. More recent work by Rogerson and Lamers presents a strong case for the view that the high degree of ionization is caused by a moderately hot corona through which the gas flows. The controversy over the question of whether the corona is simply a fairly inconsequential by-product of the radiation-pressure process proposed by Lucy and Solomon or whether it plays a major role in the mass-loss process, however, is still not resolved.

To sum up, a steady outflow of material at comparable rates of mass loss but vastly different speeds is now known to be a ubiquitous phenomenon among both the luminous hot stars and the luminous but cool red giants. The flows are probably massive enough in both cases to give rise to significant effects on stellar evolution and the mass balance between stars and the interstellar medium, but estimates of the rates are still too uncertain to be able to say much more than that. The Lucy-Solomon radiation-pressure mechanism acting on the absorption lines of the gas appears to be the most satisfactory theory proposed so far for explaining mass loss from the hot stars, but it will have to be modified to explain the anomalously high degree of ionization found in the flow. The resolution of this particular puzzle will probably come when the sensitivity of telescopes using X rays, infrared radiation and radio waves is improved just a little. The resulting observations should settle the question of the extent to which hot stars have very hot coronas and clarify the role the coronas play in the process of mass loss.

In the red giants the cause of the winds is still obscure and likely to be complex: radiation pressure acting on grains, prominencelike fountains of gas ejected in atmospheric activity similar to the solar flares, and hot, low-density coronas with cold lumps embedded in them may all play important roles. Here too observations in the new portions of the electromagnetic spectrum opened up to astronomers in the past two decades may be decisive. A satellite developed by a European-U.S. consortium with an ultraviolet spectrometer yielding a significant increase in sensitivity over that on *Copernicus* has just been launched. It will surely provide additional valuable information on the extent of coronas around the red-giant stars. If this new tool is still not adequate to the task, then like so many fundamental questions being asked today by astronomers, the question of the origin of the stellar winds may have to wait until the decisive clue is provided by the very-high-resolution spectrometer aboard the orbiting space telescope to be launched in the early 1980's.

SPECTRAL PROFILES of the outflowing gas from the star Tau Scorpii were obtained by John B. Rogerson, Jr., and H. Lamers, using a spectrometer on the *Copernicus* satellite. The vertical colored lines show the position that each of four pairs of atomic absorption lines would have in the spectrum of the star's atmosphere if the gas were at rest. (The fact that the lines come in pairs is an intrinsic property of these particular kinds of ions and has nothing to do with the stellar mass-loss phenomenon.) As is evident in the traces, the broad absorption troughs extend to large (negative) velocities, which in the case of triply ionized carbon (C^{+++}) reach at least 1,600 kilometers per second, well above the velocity required for matter to escape from the star's surface (about 1,150 kilometers per second). The absorption troughs associated with oxygen ions with five electrons removed (O^{+++++}) appear to indicate the presence of a comparatively warm region near the base of the outflow; the temperature of such a region would be about 200,000 degrees K. The strong absorption line to left of O^{+++++} troughs (*black arrow*) is attributable to hydrogen atoms in interstellar space and is not associated with the star itself.

The Hominids of East Turkana

This region on the shore of Lake Turkana in northeastern Kenya is a treasure trove of fossils of early members of the genus Homo and their close relatives dating back 1.5 million years and more

by Alan Walker and Richard E. F. Leakey

The continent of Africa is rich in the fossilized remains of extinct mammals, and one of the richest repositories of such remains is located in Kenya, near the border with Ethiopia. The first European to explore the area was a 19th-century geographer, Count Samuel Teleki, who reached the forbidding eastern shore of an unmapped brackish lake there early in 1888. Exercising the explorer's prerogative, he named the 2,500-square-mile body of water after the Austro-Hungarian emperor Franz Josef's son and heir, the archduke Rudolf (who within a year had committed suicide in the notorious Mayerling episode). Teleki's party evidently passed a major landmark on the shore of the lake, the Koobi Fora promontory, during the last week in March. They took notes on the local geology and even collected a few fossil shells, but they missed the mammal remains.

Teleki's name has faded into history, and even the name of the lake has now been changed by the government of Kenya from Lake Rudolf to Lake Turkana. Yet today the Koobi Fora area is world-famous among students of early man. Its mammalian fossils include the partial remains of some 150 individual hominids, early relatives of modern man. They represent the most abundant and varied assemblage of early hominid fossils found so far anywhere in the world.

The fossil beds of East Turkana (formerly East Rudolf) might have been found at any time after Teleki's reconnaissance. It was not until 1967, however, that the deposits came to notice. At that time an international group was authorized by the government of Ethiopia to study the geology of a remote southern corner of the country: the valley of the Omo River, a tributary of Lake Turkana. Erosion in the area has exposed sedimentary strata extending backward in time from the Pleistocene to the Pliocene, that is, from about one million to about four million years ago.

Supplies going to the Omo camps were flown over the East Turkana area; on one such trip one of us (Leakey) noticed that part of the terrain consisted of sedimentary beds that had been dissected by streams and that appeared to be potentially fossil-bearing. A brief survey afterward by helicopter revealed that the exposed sediments contained not only mammalian fossils but also stone tools. This reconnaissance was followed up in 1968 by an expedition to the vast, hot and inhospitable area. Out of a total area of several thousand square kilometers the expedition located some 800 square kilometers of fossil-bearing sediments, mainly in the vicinity of Koobi Fora, Ileret and to the south at Allia Bay. The expedition also found the fossil remains of many kinds of mammals, most of them beautifully preserved. Only in the category of hominids were the finds disappointing: the total was only three jaws, all of them badly weathered. Nevertheless, the overall richness of the fossil deposits made it clear that further prospecting would be worthwhile.

The large task of establishing the geological context of both the fossils and the stone tools discovered in East Turkana began in 1969. The work that season was highlighted by the excavation of the first stone tools to be found in stratified sequences there and by the discovery of two skulls of early hominids. It was with these discoveries that the enormous importance of the area for the study of human evolution began to be recognized.

A formal organization was established: the Koobi Fora Research Project. The project operates under the joint leadership of one of us (Leakey) and Glynn Isaac of the University of California at Berkeley. In the years since it was founded the project has brought together workers from many countries who represent many different disciplines: geology, geophysics, paleontology, anatomy, archaeology, ecology and taphonomy. (Taphonomy is a new discipline concerned with the study of the processes that convert living plant and animal communities into collections of fossils.) The interaction of specialists has become a particular strength of the project as the workers have become increasingly aware of the particular outlook (and also the limitations) of fields other than their own.

The project area extends from the Kenya-Ethiopia border on the north to a point south of Allia Bay where the land surface is of volcanic origin. The western boundary is the lakeshore and the eastern is marked by another volcanic outcropping. The promontory of Koobi Fora itself, a spit of land that extends a few hundred meters into the lake, is the site of our base camp. Each of the three principal areas of fossil-bearing sediments has its own natural boundaries; when these areas are seen from the air, they show up as pale patches among the darker volcanic terrain. For reference purposes they have been divided into smaller units that are identified by number on the project maps and that are readily distinguished in the field by their vegetation, dry rivers and the like.

In studies such as these it is of paramount importance to develop a chronological framework that allows the fossil finds to be placed in their correct relative positions. The construction of such a framework is the responsibility of the project geologists and geophysicists. The geology of East Turkana is straightforward in its broad outlines but extremely complex in many of its local details. Among the factors responsible for its complexity are abrupt lateral shifts in the composition of the exposed sedimentary strata, discontinuities, faulting that involves rather small displacements and above all the absence from many of the sediments of volcanic tuff: layers of ash that play a key role in correlating the strata.

The difficult work of geological mapping and stratigraphic correlation has been carried out by many of our colleagues but mainly by Bruce Bowen of Iowa State University, working in collaboration with Ian Findlater of the International Louis Leakey Memorial Institute for African Prehistory in Nairobi, Kay Behrensmeyer of Yale University and Carl F. Vondra of Iowa State.

SURFACE-SCRAPING PARTY of the Koobi Fora Research Project is seen at work in the Ileret area of East Turkana. The discovery of a hominid jawbone on the surface has led to the collection and screening of loose topsoil and rock in the hope of locating additional fossil fragments. Geologists in the party simultaneously record the stratigraphic position of the exposure and set up a site marker.

FOSSIL JAWBONE is photographed where it was found, exposed on the surface by erosion of the sedimentary rock that once surrounded it. This is the fragmented mandible of a robust member of the genus *Australopithecus;* it was preserved through chance burial in sediments almost two million years ago. Mandibles such as this one are so sturdy that a disproportionate number have survived as fossils.

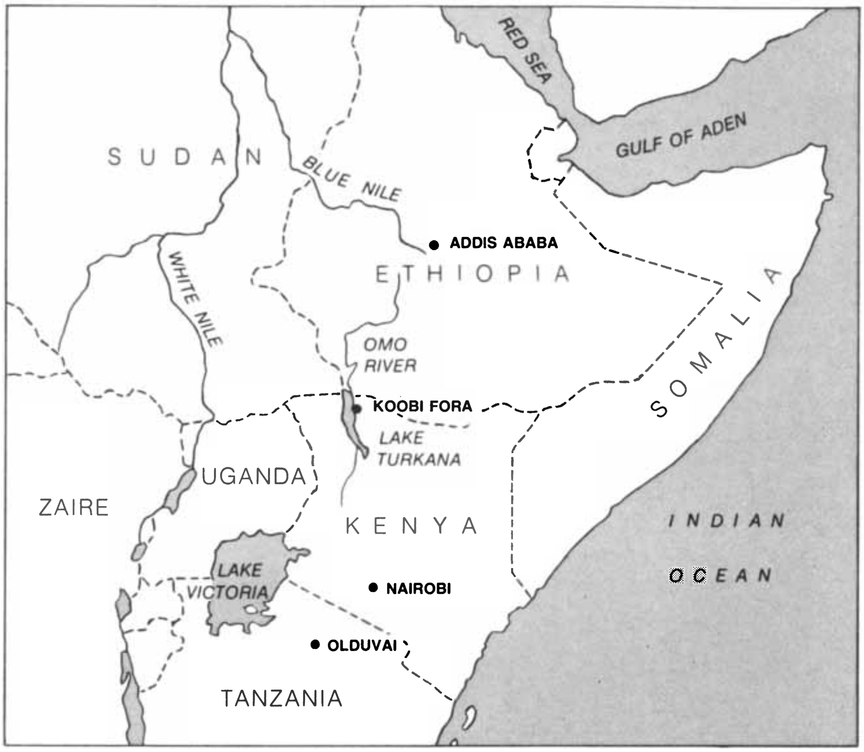
FOSSIL-RICH AREAS in East Turkana appear as pale patches in a satellite image of the region. The deltas of the Omo River, which rises in Ethiopia, appear at the northern end of Lake Turkana (formerly Lake Rudolf). The short, narrow promontory that juts out from the eastern shore of the lake is Koobi Fora, where the base camp for project fieldworkers is situated. Imagery is from the satellite *Landsat 2.*

The complexities have nonetheless prevented the accurate placement of some of the most important early hominid fossils in the stratigraphic framework the geologists established. In such cases we have provisionally assigned the specimens to temporal positions on the basis of criteria other than stratigraphic ones.

The fossil-rich sediments are underlain by older rocks of volcanic origin. The sediments themselves are of various kinds, laid down in such different ancient environments as stream channels and their associated floodplains, lake bottoms, stream deltas and former lakeshores. For the most part the strata dip gently toward the present Lake Turkana. In the past, extensions of various deltas and coastal plains frequently built out westward into the former lake basin; these intrusions alternated with periods when the lake waters intruded eastward. The result is a complex interdigitation of lake sediments and stream sediments.

The major basis for correlating the various strata is the presence of distinctive strata consisting of tuffs; the volcanic material has periodically washed into the lakeshore basin from the terrain to the north and to the east. Some of the tuff beds are widespread and some are not. The uncertainties of correlation between tuff layers in different locations are greatest in the Koobi Fora and Allia Bay areas; these are the areas farthest from the erosional sources of the volcanic ash. At the same time volcanic rocks can be dated by means of isotope measurements, which makes the ash strata particularly important.

Jack Miller of the University of Cambridge and Frank J. Fitch of Birkbeck College in London have conducted most of the dating studies of the tuff (based on the decay of radioactive potassium into argon). Independent chronological data are also available from studies of the magnetic orientation of some particles in the volcanic ash and studies of fission tracks in bits of zircon in the ash. The various measurements are by no means unequivocal, but it can be stated as a generality that the location of a fossil find below one such layer of tuff and above another at least establishes a relative chronological position for the fossil even if its precise age remains in doubt.



KOOBI FORA REGION of East Turkana lies near the border between Kenya and Ethiopia. The fossil material collected there is brought back to Nairobi for preparation and analysis.

There are five principal tuff marker layers. The earliest, which provides a boundary between the Kubi Algi sediments below it and the Koobi Fora sediments above it, is the Surgaei tuff. The next layer of tuff divides the lower member of the Koobi Fora sedimentary formation approximately in half; this is the Tulu Bor tuff, 3.2 million years old. The next layer, the KBS tuff, is named for the exposure where it was first recognized: the Kay Behrensmeyer site. It marks the boundary between the lower and upper members of the Koobi Fora Formation, and its exact age is debated.

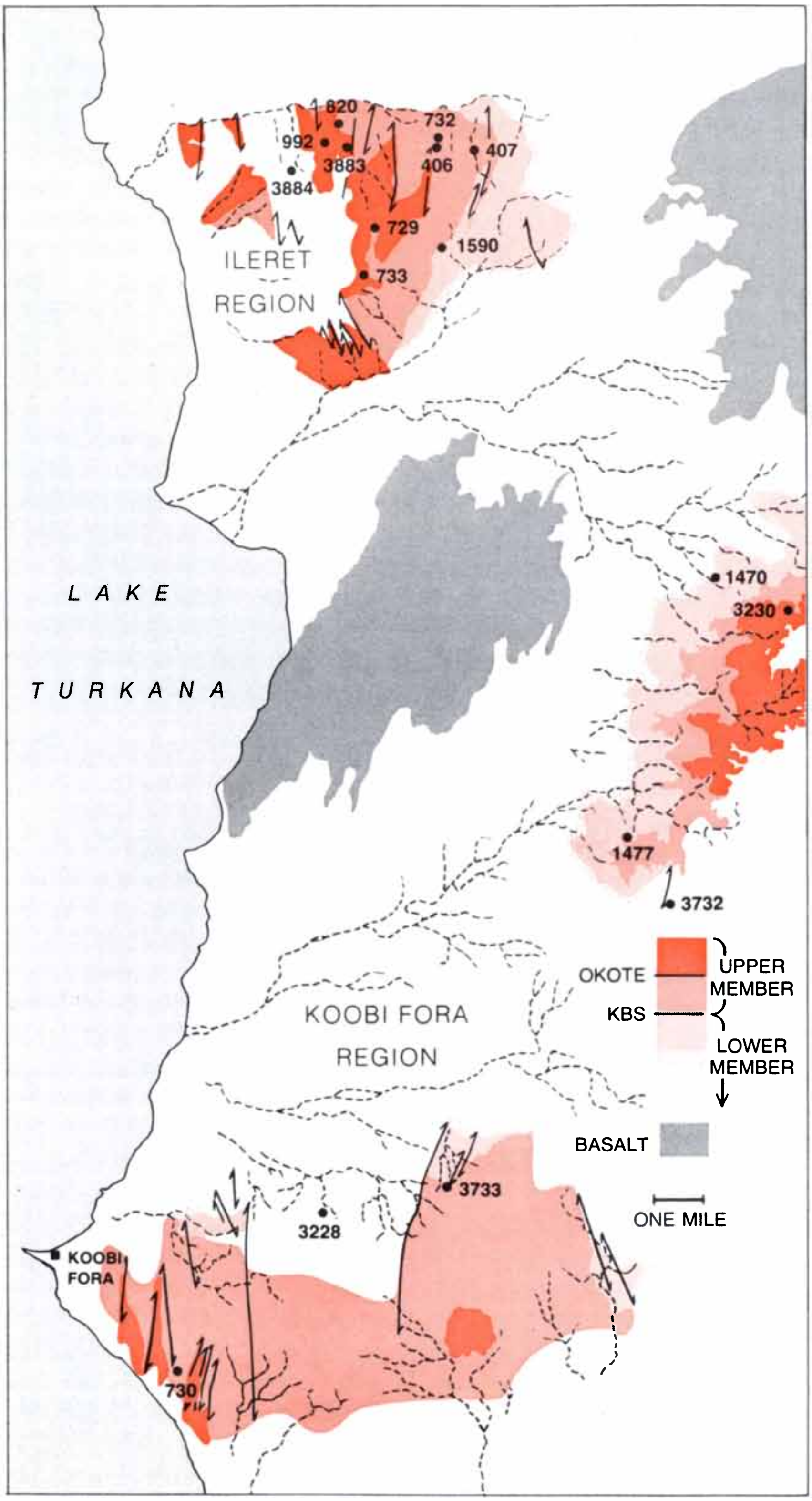
In 1970 Miller and Fitch ran a potassium/argon analysis of the KBS tuff that showed it to be 2.61 (±.26) million years old. They have recently made new calculations indicating that the tuff is 2.42 (±.01) million years old. At the same time an age determination based on other KBS samples, done by Garniss H. Curtis of the University of California at Berkeley, yields two much younger readings: 1.82 (±.04) million and 1.60 (±.05) million years.

Above the anomalous KBS tuff the next marker layer is the Okote tuff, which divides the upper member of the Koobi Fora sediments approximately in half. The Okote tuff is between 1.6 and 1.5 million years old. The fifth and uppermost marker layer, roughly indicating the boundary between the Koobi Fora sediments and the overlying Guomde sediments, is the Chari-Karari tuff; it is between 1.3 and 1.2 million years old.

We present the stratigraphy in such detail because most of the early hominid fossils discovered thus far in East Turkana are sandwiched between the Tulu Bor tuff at the most ancient end of the geological column and the Chari-Karari tuff at the most recent end. Twenty-six specimens, including a remarkable skull unearthed in 1972, designated KNM-ER 1470, come from sediments that lie below the KBS tuff. (The designation is an abbreviation of the formal accession number: Kenya National Museum–East Rudolf No. 1470.) Another 34 specimens, including several skulls, come from sediments that lie above the KBS tuff but below the Okote tuff. Unfortunately for those interested in measuring the rates of evolutionary change in hominid lineages, the difference between the oldest and the youngest proposed KBS-tuff dates is some 1.3 million years.

That span of time far exceeds the one allotted to the whole of human evolution not many years ago. Even by today's standards the KBS-tuff discrepancy is enough to allow uncomfortably different evolutionary rates for various hypothetical hominid lineages. So far evidence of other kinds has not resolved the issue. For example, our colleagues John Harris of the Louis Leakey Memorial Institute and Timothy White of the University of California at Berkeley, who have conducted a detailed study of the evolution of pigs throughout Africa, suggest that the more recent date for the KBS tuff would best suit their fossil data. At the same time the fission-track studies of zircons from the KBS tuff indicate that the older dates are correct. For the time being we must accept the fact that the KBS tuff is either about 2.5 million years old or somewhere between 1.8 and 1.6 million years old.

Relative and absolute chronology apart, other kinds of investigation are increasing our knowledge of the different environments that were inhabited by the East Turkana hominids. For example, most of the hominid specimens can in general be placed either in the genus *Australopithecus* or in the genus *Homo*. Behrensmeyer, Findlater and Bowen



EAST TURKANA FOSSILS are found in a complex interdigitation of sedimentary rocks, some of lakeshore and delta origin and others of streambed origin, that were laid down during alternating periods of lake transgression and land buildup. The major geological marker layers are beds of volcanic tuffs that have been washed from the east and north across parts of the region. Two of the three main fossil-bearing areas are the Ileret region (*top*) and the Koobi Fora region (*bottom*). Allia Bay is not shown. The map is based on the work of Ian Findlater; numbers identify some of the hominid-fossil finds in both regions. Geological faults throughout the region are identified by conventional symbols; also identified (*color*) are sedimentary strata of the upper member and part of the lower member of the Koobi Fora Formation. These strata are separated by two volcanic-tuff marker beds: the Okote complex and the KBS complex.

are engaged in microstratigraphic studies that have enabled Behrensmeyer to associate many of the specimens with a specific environment of sedimentary burial. Preliminary analyses indicate that the specimens identified as *Homo* were fossilized more commonly in lake-margin sediments than in stream sediments whereas the specimens identified as *Australopithecus* are equally common in both sedimentary environments. Facts such as these promise to be of great help in reconstructing the lives of early hominids. In this instance the chance that an organism will be buried near where it spends most of its time is greater than the chance that it will be buried farther away. Thus Behrensmeyer has hypothesized that in this region of Africa early *Homo* exhibited a preference for living on the lakeshore.

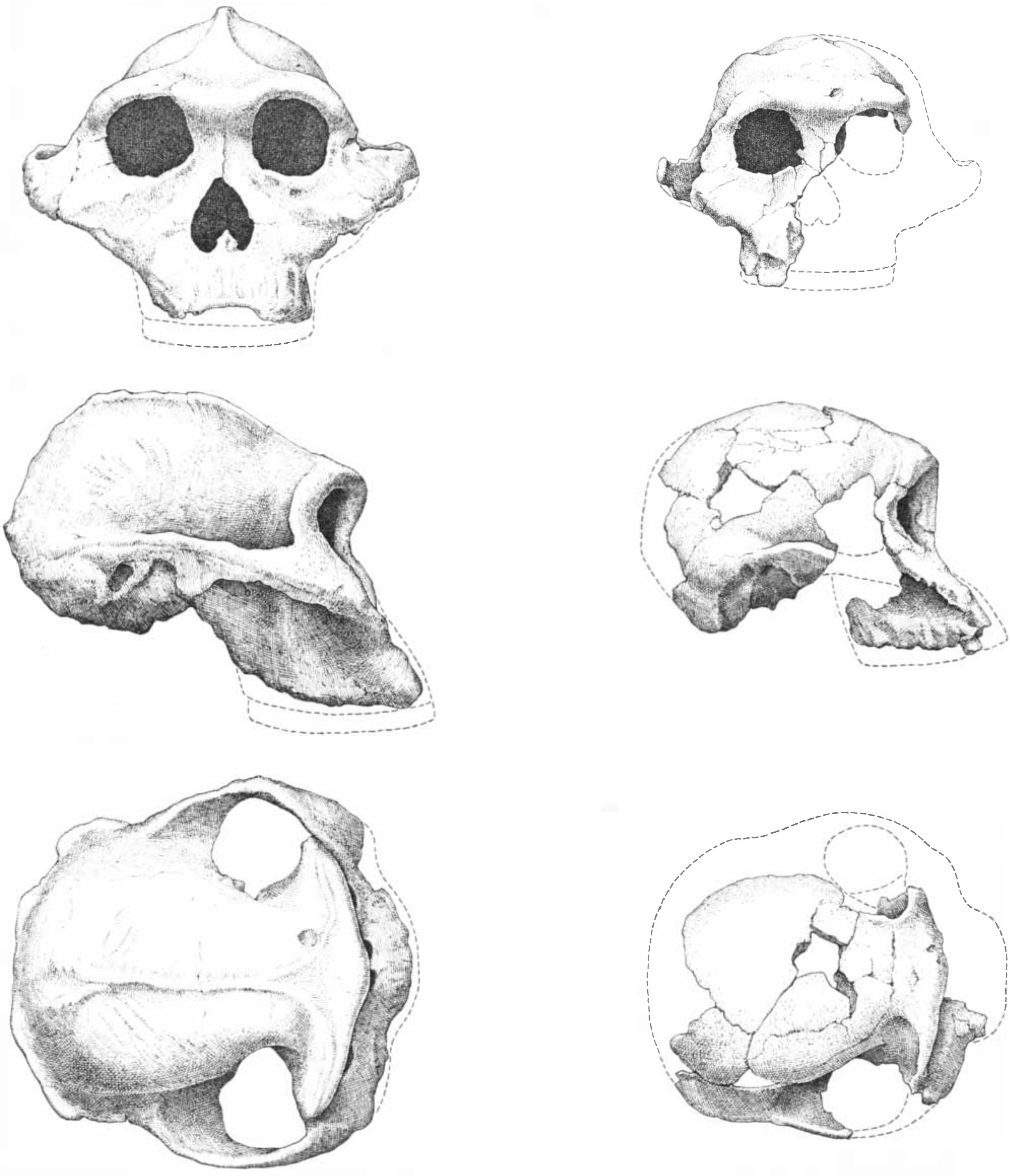
Because of the unusual circumstances in this badlands region it will be useful to describe how the hominid fossils have been collected. The initial process is one of surface prospecting. The Kenyan prospecting team is led by Bwana Kimeu Kimeu; its job is to locate areas where natural erosion has left scatterings of mammalian bones and teeth exposed on the arid surface of the sedimentary beds. Kimeu is highly skilled at recognizing even fragmentary bits of hominid bone in the general bone litter present in such exposures.

Once the presence of a hominid fossil is established by the prospectors one of the project geologists determines its position with respect to the local stratigraphic section and records the location. Thereafter one of two procedures is generally followed. If the bone fragment has been washed completely free from the sedimentary matrix that held it, the practice is to scrape down and sieve the entire surrounding surface area in the hope of recovering additional fragments. As the scraping is done a watchful eye is kept for fragments that might still be in situ, that is, partly or entirely embedded in the rock.

If the initial discovery is a fossil fragment in situ, the procedure is different. Excavation is begun on a near-microscopic scale, the tools being dental picks and brushes. The Turkana hominid fossils are often so little mineralized that a preservative must be applied to the bone as excavation progresses in order to keep it from fragmenting further. Indeed, sometimes the preservative fluid must be applied with painstaking care because the impact of a falling drop can cause breakage.

After excavation each site is marked by a concrete post inscribed with an accession number provided by the Kenya National Museum. The next task, usually undertaken in the project laboratory in Nairobi, is piecing together the specimens. This is rather like doing a three-dimensional jigsaw puzzle with many of the pieces missing and no picture on the box. Any adhering matrix is now removed under the microscope, most often with an air-powered miniature jackhammer. (Cleaning with acid, a common laboratory method, is out of the question because the fossil bone is less resistant to the acid than the matrix.) Finally, the hardened pieces of bone are reconstructed to the extent possible by gluing adjacent fragments together.

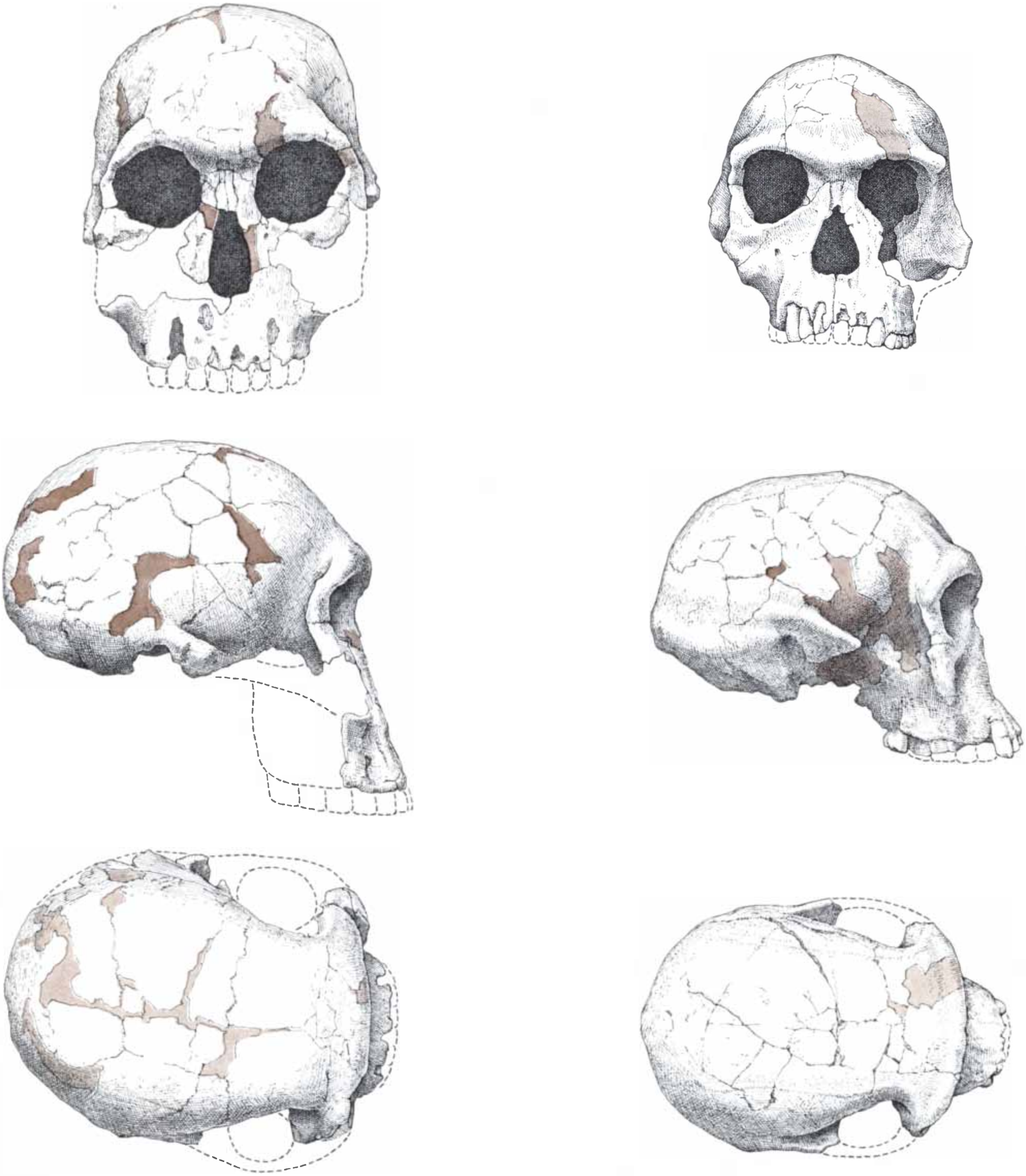
Can the East Turkana collection be considered representative of the hominid populations that occupied the area more than a million years ago? Taphonomy, the relatively new discipline that attempts to define the processes whereby communities of plants and animals do or do not become preserved as fossils, is beginning to provide some helpful answers to the question. The biases that affect fossil samples are many. For example, circumstances may result in the preservation of only some parts of



HOMINID FOSSILS found early in the process of collecting in East Turkana are illustrated. At the left is Kenya National Museum–East Rudolf (KNM-ER) accession No. 406, a robust cranium with well-preserved facial bones. At right is KNM-ER 732, a fragmented cranium that has little of the face preserved and is less robust than KNM-ER 406. Both specimens are placed in the genus *Australopithecus*.

certain individuals. Or a particular specimen may be severely deformed by pressure during its long burial. One bias that is easy to recognize in the East Turkana collection is a disproportionate number of lower jaws of the early hominid *Australopithecus robustus*. This hominid had powerful jaws and unusually large teeth; its lower jawbone is particularly massive. The relative abundance of these jaws and teeth in the East Turkana sediments probably results more from their mechanical strength, and thus their enhanced ability to survive fossilization, than from any preponderance of *A. robustus* individuals in the population.

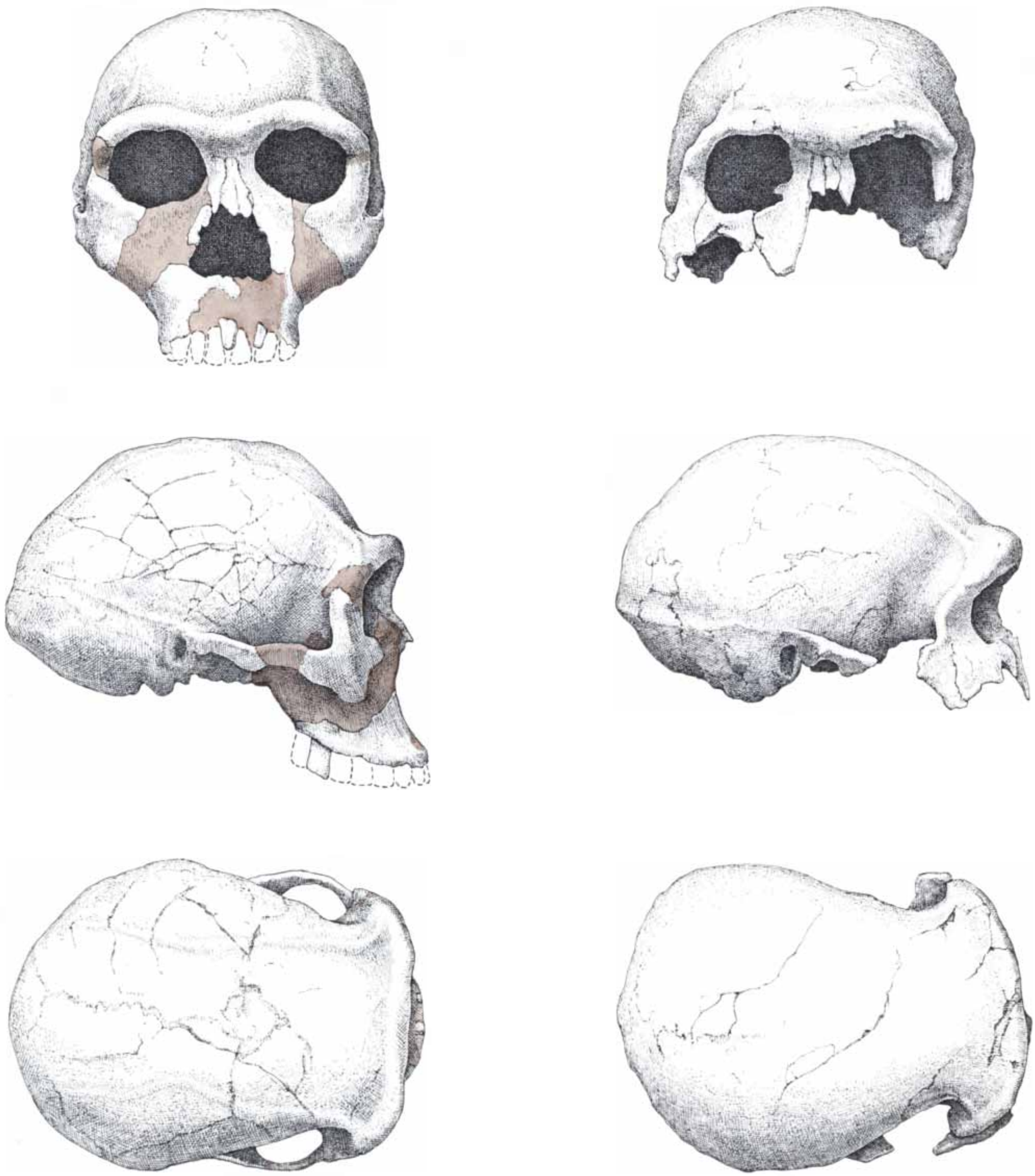
Another example of bias in the hominid-fossil collection is the disproportionate representation of different parts of the skeleton. Teeth are by far the hardest parts, and so it is not surprising to find that teeth account for the largest fraction of the East Turkana sample. In contrast, vertebrae and hand and foot



HOMINID FOSSILS OF DIFFERENT AGE are KNM-ER 1470, at left, and KNM-ER 1813, at right. The first cranium comes from the lower member of the Koobi Fora Formation; it cannot be less than 1.6 million years old and may be more than 2.5 million years old. It has a cranial capacity of about 775 cubic centimeters, compared with the *Australopithecus* average of about 500 c.c. The second cranium is provisionally assigned to the upper member of the Koobi Fora Formation, suggesting that it is no less than 1.2 million and may be more than 1.6 million years old. Its cranial capacity is about 500 c.c. It resembles other African hominid fossils 1.5 to two million years old.

bones are rarely found. Can this bias be attributed to the destructive processes associated with burial and exposure alone? It seems only logical to take into account a third process: carnivore and scavenger feeding on the hominid bodies before the sediments covered them.

What fraction of the hominid population in East Turkana more than a million years ago might our fossil collection represent? The answer, based on modern population studies of wild dogs and baboons, is that the fraction is extremely small. Assuming an appropriate interval between generations, if the hominid population density was low, as it is among living wild dogs, the collection represents two ten-thousandths, or .02 percent, of the original population. If the hominid population density was high, as it is among baboons, the fraction is very much smaller: two ten-millionths, or .00002 percent. If we ask further what fraction of the ancient popu-



TWO SKULLS OF THE GENUS HOMO from the East Turkana fossil beds represent the early human species *Homo erectus*, a fossil hominid first discovered in Java and China. These are KNM-ER 3733 and KNM-ER 3883. Both are more than 1.5 million years old, which makes them a million years older than the specimens from China. Their great age strongly suggests that *H. erectus* first evolved in Africa. KNM-ER 3733 has a cranial capacity of about 850 c.c.; the cranial capacity of KNM-ER 3883 has not yet been measured but is probably about the same. Some specimens of *H. erectus* from Africa and Java have cranial capacities that are greater than 1,000 c.c.

lation is represented by the relatively complete skulls in the collection, it may be smaller still: it is between a hundred-thousandth and a hundred-millionth of the total. The second figure is the equivalent of someone's selecting two individuals at random to represent the entire population of the U.S. today. It is on this small sample that our hypotheses concerning hominid evolution must be based.

In developing such hypotheses we must keep in mind a number of fundamental questions. One question is: How many different species of early hominids were there in East Turkana? Another question is whether those species, whatever their number, existed over long periods of time or were replaced by other species. Again, do any or all of the species show signs of evolution during this interval of perhaps 1.5 million years or perhaps only 700,000 years? If there was any evolutionary change, what was its nature? How did the early hominids feed themselves? Were they relatively low-energy herbivores or relatively high-energy omnivores? If indeed several species were present at the same time, did each occupy a distinct ecological niche? What kept the niches separate?

The questions do not end here. Other questions, more specifically anatomical, also call for answers. Do the hominid fossils possess any morphological attributes that might be correlated with the archaeological record of tool use and scavenger-hunter behavior in the area? For example, can we detect any evidence of significant brain evolution during the period? Are there any morphological changes suggestive of altered patterns of locomotion or hand use that might shed light on the origins of certain unique human attributes? (For the purposes of this discussion we define these attributes as including not only walking upright and making use of tools but also an enlarged brain and the ability to communicate by speaking.)

These questions and many more can be answered only after the first question in the series is disposed of. Basically taxonomic in nature, it asks how many species are represented in the East Turkana fossil-hominid assemblage. We have already said that in general two genera were present: *Australopithecus* and *Homo*. How may the two be subdivided?

The answer to this basic and far from simple question is not an easy one. Conspiring against a clear-cut response are such factors as the smallness of the sample, the fragmentary condition of the individual specimens, the fact that even among individuals of the same species a large degree of morphological variability is far from uncommon and, under this same heading, the fact that a great deal of variation is often found between the two sexes of a single species. Also not to be neglected is the fallibility of the analyst, who is prone to human preconceptions. For example, the very order of discovery of the East Turkana hominids has affected our hypotheses, and we have had to chop and change in order to keep abreast of later discoveries.

It is illuminating in this connection to review the sequence of hominid discoveries in East Turkana. The first specimens to be identified were individuals of the species *Australopithecus robustus*. Fossil representatives of this species were first found at sites in South Africa decades ago. Characteristically they are large of face and massive of jaw; the molar and premolar teeth are very large, although the incisors and canines are small, about the same size as the front teeth of modern man. Although the facial skeleton is large, the brain case is relatively small: the average cranial capacity is about 500 cubic centimeters, compared with the modern human aver-

ILERET

	AGE (MILLION YEARS)	TUFF COMPLEX	1/1A	3	5	6/6A	7A	8	10	11	12	15
GUOMDE FORMATION					● 3884	× 999						
UPPER MEMBER, KOOBI FORA FORMATION	1.22 – 1.32	CHARI										
			● 725 × 739 ● 728 × 741 ● 805 × 993 ● 3883	● 992 ● 1467 × 740		● 731 ● 1466	● 404			● 726 × 1465		
	1.48 – 1.57	MIDDLE/ LOWER	× 1463			● 818 ● 2593		● 729 ● 807 ● 733 ● 808 × 803 ● 806 ● 809		● 1468		
			● 819 ● 820 ● 1817 ● 2595	● 1819		● 727 ● 2592 ● 801 ● 3737 ● 802 × 1464 ● 1170 × 1823 ● 1171 × 1824 ● 1816 × 1825 ● 1818			● 406 ● 407 ● 732 × 815		× 1591 × 1592	
	2.42 (FITCH, MILLER) 1.6 – 1.8 (CURTIS)	(KBS EQUIVALENT)									● 1593	● 2597 ● 2599 ● 2598 × 2596
LOWER MEMBER, KOOBI FORA FORMATION											● 1590	

INVENTORY OF FOSSIL HOMINIDS from East Turkana appears in this chart: specimens, identified by accession numbers, are listed according to the numbered area where they were found. Their positions do not indicate any relative temporal position other than a location between the dated tuff marker layers (or equivalent markers). Such a position means that the specimen is older than the known

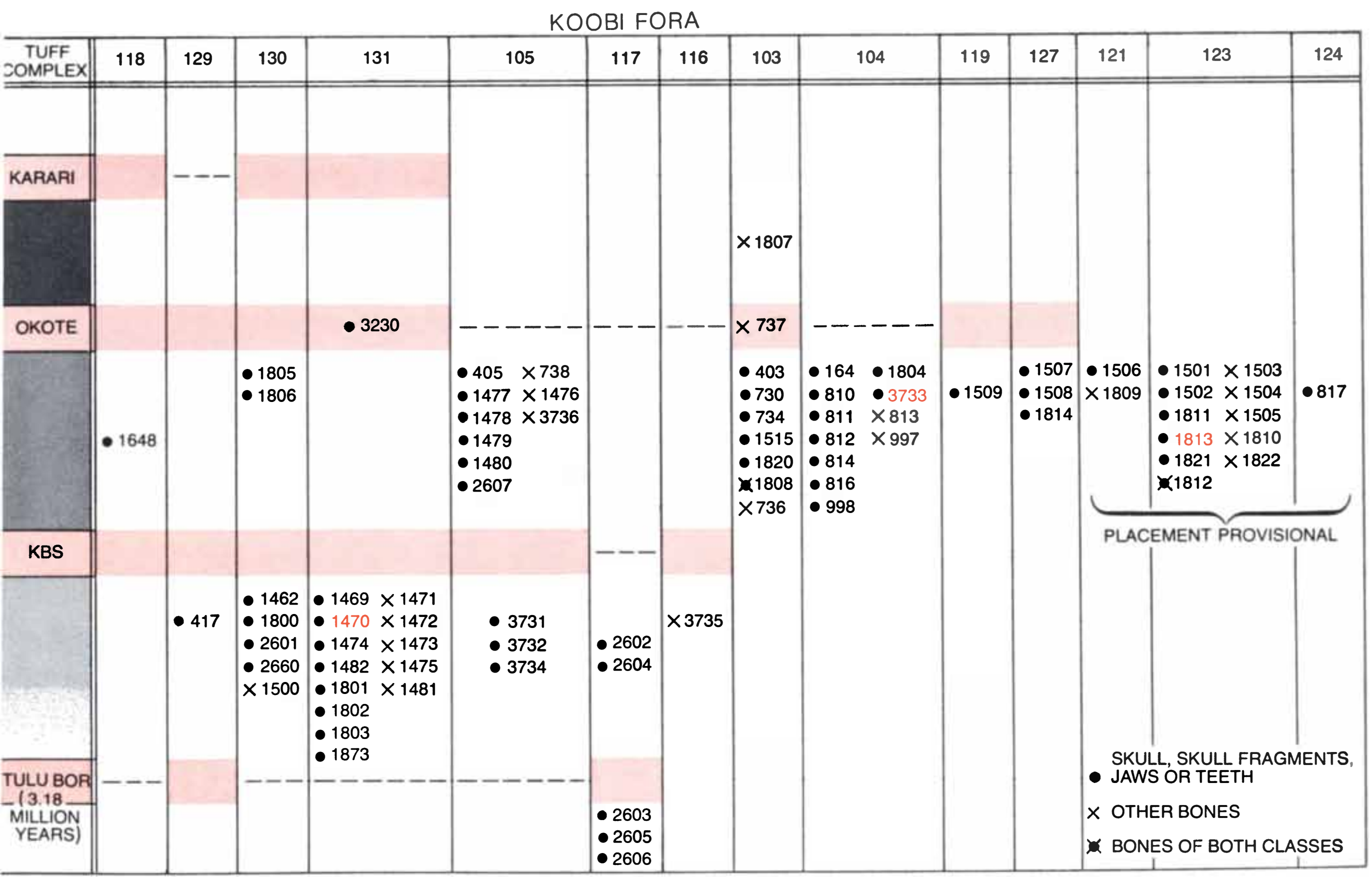
age of 1,360 c.c. Because the chewing muscles were evidently of a size commensurate with the large cheek teeth and massive jaws, many *A. robustus* individuals have not only extremely wide-flaring cheekbones but also a bony crest that runs fore and aft along the top of the brain case to provide a greater area for the attachment of chewing muscle.

Specimens of *A. robustus* have also been found in East Africa, most notably by Louis and Mary Leakey at Olduvai Gorge. The East African examples are on the whole even larger than those from South Africa, and their cheek teeth are more massive. A well-known example is "Zinjanthropus," an Olduvai cranium now accepted by most scholars as being closely related to the *A. robustus* specimens from South Africa. (Some scholars, it should be noted, still assign "Zinjanthropus" to a related species of *Australopithecus, A. boisei*.) Such taxonomic niceties aside, the fact is that the East Turkana deposits have been found to contain a good number of fossils that can be placed in this hominid species.

In the early investigations at East Turkana the skulls of certain smaller and less robust hominids were also discovered. Indeed, one such cranium, deformed by crushing, turned up near (although not in the same stratigraphic horizon as) a robust *Australopithecus* cranium: KNM-ER 406. When this crushed specimen was first discovered, it could not easily be given a taxonomic position. The finding of a second gracile (as opposed to robust) specimen, however, suggested to us that male-female dimorphism might account for both kinds of cranium. In the second specimen most of the right side of the brain case and facial skeleton and part of an upper-jaw premolar tooth and the roots of the molars were preserved. It is evident that although this individual is substantially less robust than KNM-ER 406, its premolar and molar teeth were only a little less massive than those of the robust one. If among the species *A. robustus* the morphological differences between males and females were as great as they are among gorillas, then the robust, crested specimens from East Turkana could be males and the more gracile specimens could be females.

The age of these *Australopithecus* specimens is substantially greater than that of any previously uncovered in East Africa (the age of the South African specimens remains in question), but their discovery presented no taxonomic problems. This happy state of simplicity came to an end in 1972 with the discovery of the cranium KNM-ER 1470. Bwana Bernard Ngeneo came across it on an exposure of older sediments belonging to the lower member of the Koobi Fora Formation. When he found the specimen, all that could be seen was a scattering of bone fragments on the rock surface. The fragments were relatively fragile, which led us to assume that they had been washed out of the matrix quite recently.

The specimen KNM-ER 1470 is a large, lightly built brain case with a considerable amount of the facial skeleton preserved. Our colleague Ralph L. Holloway, Jr., of Columbia University has determined that its cranial capacity is about 775 c.c. The facial skeleton is very large, and the proportions of the front and cheek teeth are indicated by the preserved tooth sockets and by both the sockets and the broken roots of the molars. The proportions are the reverse of those for *A. robustus;* the incisors and canines are very large and the premolars and molars are only moderately large. Even though the tooth size suggests a formidable chewing apparatus, the brain case shows no sign of a crest for the attachment of heavy chewing



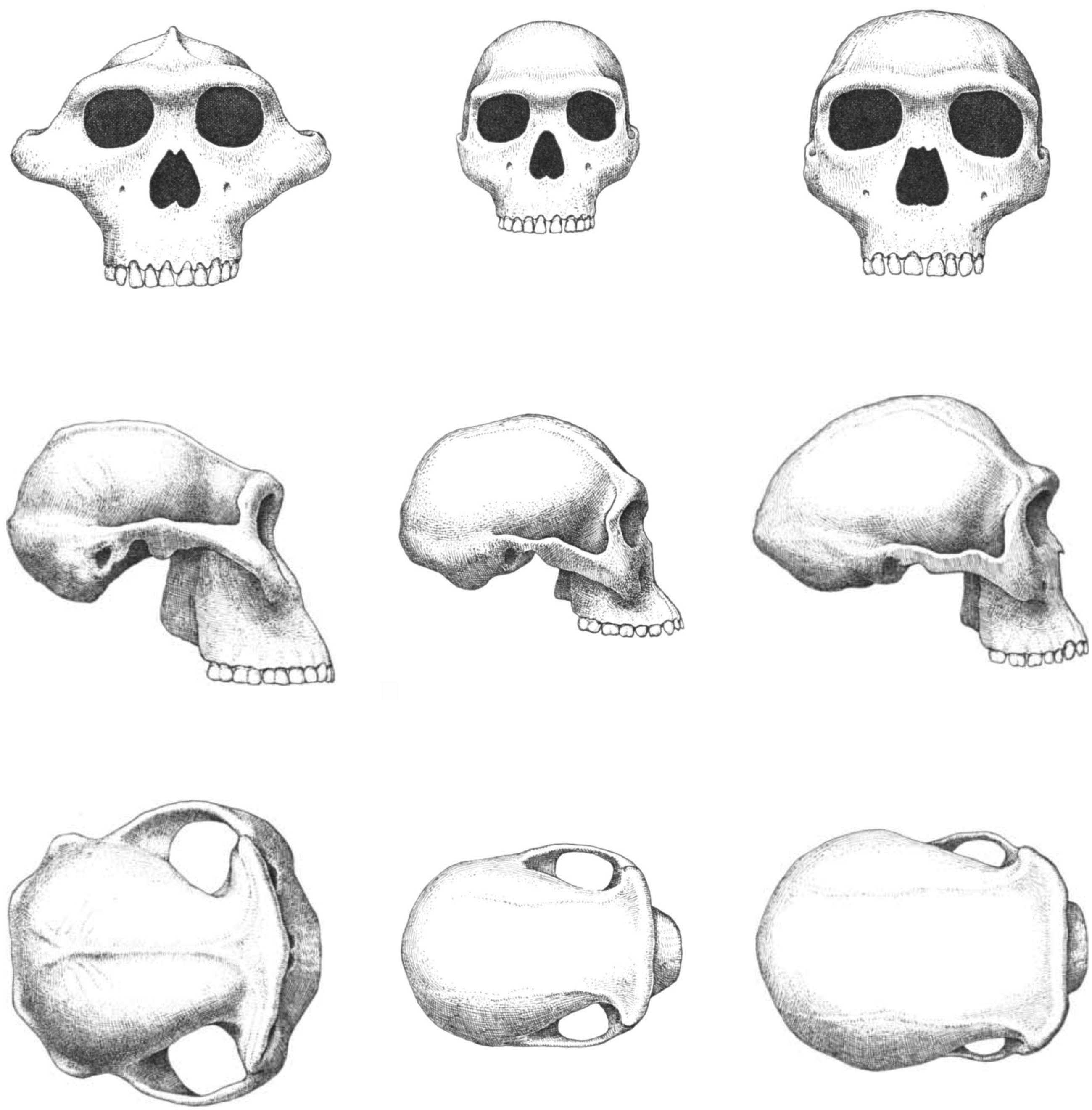
age of the tuff layer above it and younger than the known age of the tuff below. Numbers in color identify fossils illustrated on pages 59 through 61. Two geological columns define sedimentary strata; note ages of the tuff marker layers. Broken lines show correlation between Ileret and Koobi Fora tuffs. Placement of specimens listed at the far right is provisional because marker layers are absent there.

muscles. Similarly, the cheekbones, although incomplete, do not suggest the same great width of face that is characteristic of *A. robustus.*

Much has been written about the significance of KNM-ER 1470. We believe that certain hominid specimens found at Olduvai Gorge in broken and fragmentary condition are examples of the same kind of skull. If it is necessary to decide on a taxonomic term for these hominids, the species name may well turn out to be *habilis.* (*Homo habilis* is the name that was given to an early species of the genus *Homo* by Louis Leakey and his colleagues John Napier of the University of London and Phillip V. Tobias of the University of the Witwatersrand. The name was not accepted unanimously by other students of fossil man and has even caused heated argument.)

We ourselves cannot agree on a generic assignment for KNM-ER 1470. One of us (Leakey) prefers to place the species in the genus *Homo,* the other (Walker) in *Australopithecus.* The disagreement is merely one of nomenclature; we are in firm agreement on the evolutionary significance of what are now multiple finds. Since 1972 two additional partial skulls of this large-brained, thin-vaulted kind have been found in association with strata assigned to the lower member of the Koobi Fora Formation.

It was at about the time of the discovery of KNM-ER 1470 that we and our colleagues began to disagree as to the taxonomic position of certain well-preserved lower jaws from the East Turkana region. The initial source of disagreement was a small cranium: KNM-ER 1813. It is the cranium of a small-brained hominid with the average *Australopithecus* cranial capacity: 500 c.c. It has a relatively large facial skeleton and



THREE FORMS OF HOMINID are represented among the fossils found in the upper member of the Koobi Fora Formation. The illustration shows them as they would appear if they were restored; lower jaws are omitted. The three may be assigned to particular species in five different ways: all three may belong to a single highly variable species, or they may belong to two species in three possible combinations, or each form may be a valid species on its own. The authors suggest that the three-species hypothesis is the most probable of the five.

palate. The upper teeth preserved with the palate are comparatively small, however, and bear a striking resemblance to the teeth of one of the *Homo habilis* specimens from Olduvai, OH-13. It happens that in the initial controversy over the original *H. habilis* specimens, OH-13 represented a species that even skeptics agreed was nearly, if not actually, identical with the species *Homo erectus,* a member in good standing of the genus *Homo* that was first recognized in Java and northern China.

The resemblances between KNM-ER 1813 and OH-13 go further than their teeth. In all the parts that can be compared—the palate as well as the teeth, much of the base of the skull and most of the back of the skull—the two specimens are virtually identical. This leads us to believe the usual reconstructions of OH-13, which have assumed that the specimen had a large cranial capacity and an *erectus*-like skull, are in error. That is not all. The mandible of OH-13 was preserved; its small size and the details of the teeth were major components of the evidence leading to the conclusion that *habilis* was near the *erectus* line. The comparable lower jaws we have found at East Turkana, we can now see, make it clear that the OH-13 mandible could just as well have been hinged to a small-brained, thin-vaulted skull like that of KNM-ER 1813.

To make a final point about these enigmatic East Turkana specimens, we believe there are strong resemblances between them and some of the smaller *Australopithecus* specimens from sites in South Africa, specimens that are usually placed in the gracile species *A. africanus.* Faced with so many possibilities, we argue for caution in the making of taxonomic judgments. Such caution should prevail not only when the evidence in hand is a few isolated teeth but also when the evidence is more generous: lower jaws and upper jaws with the teeth still in place.

In the 1975 season we discovered a remarkably complete cranium: KNM-ER 3733. The find showed unequivocally that a member of our own genus was present in East Turkana when the early strata of the Koobi Fora upper member were formed. The skull bears a striking resemblance to some of the *Homo erectus* skulls found in the 1930's near Peking and is certainly a member of that species. The brain case is large, low and thick-boned. Its principal part is formed by the projecting occipital bone, and its cranial capacity is about 850 c.c. The brow ridges jut out over the eye orbits, and a distinct groove is visible behind them where the frontal bone rises toward the top of the vault.

KNM-ER 3733 has a small face tucked in under the brow ridges. The sockets of some of the upper teeth are preserved. The missing front teeth were relatively large, but the back teeth (some of them still in place) are of only modest proportions. The third molars are among the missing teeth, but the evidence is that they were quite reduced in size. They had been erupted long enough for them to wear grooves in the second molars in front of them, yet the bone forming the sockets indicates that their roots were very small. The point is important because a diminution in the size of this tooth is a common phenomenon in modern human populations.

This fine example of *Homo erectus* from East Turkana predates the example of the same species found at Olduvai Gorge by half a million years and is about a million years older than the examples from northern China. And KNM-ER 3733 is not alone. Another *erectus* specimen was found shortly afterward in Area No. 3 of the Ileret fossil beds. This is KNM-ER 3883. The Ileret specimen is from approximately the same geological horizon as KNM-ER 3733. It has much the same cranial conformation, but its brow ridges, facial skeleton and mastoid processes are somewhat more massive. The cranial capacity of KNM-ER 3883 has not yet been determined, but there is no reason to expect that it will be much different from that of 3733. The similarity of the two East Turkana specimens to specimens from far away that are very much younger strongly suggests that *Homo erectus* was a morphologically stable species of man over a span of at least a million years.

Leaving aside the problems presented by various fragmentary specimens, how can the new hominid finds from East Turkana be assessed taxonomically? One might simply suggest a series of normal taxonomic assignments in the light of what we see at present, acknowledging that as in the past the assignments are likely to be changed. We shall not do so here because we now view the problem of taxonomic assignment in a slightly different way.

Considering only the fossils from the upper member of the Koobi Fora Formation, we think we recognize specimens that might be assigned to three different species. At the same time we may have seriously misunderstood the quantity and quality of variation in any one of the three species. In order to acknowledge this we should like to consider the following possibilities:

1. The three forms are only artifacts of our imagination. Only one hominid species was present, and what we take to be distinct types are merely morphological variants within that species.

This, of course, is the single-species hypothesis. Its strongest proponents have been Loring C. Brace and Milford H. Wolpoff of the University of Michigan. Simply put, the hypothesis states that ever since the human attributes of upright walking, a prolonged childhood, a large brain and small canine teeth became established there has been only one hominid species. What brought this about, those who favor the hypothesis aver, was culture. Once culture became the human domain—and evidence for this can be sought in the fossil bones by looking for the basic human anatomical attributes that developed along with culture—the human ecological niche became so large that the species with culture always had the edge over any other species in competition for the available resources.

2. Two of the three forms represent one species; the third represents a second species. In this hypothesis the separate species is *Homo erectus.* Both of the other forms, then, must be members of a single, highly variable and sexually dimorphic species. The very robust specimens are males and the very gracile ones are females.

To accept this version of the dimorphic hypothesis one must postulate a great deal of variability. The difference in cranial capacity between the robust and the gracile specimens is within the limits seen in living populations of sexually dimorphic apes. The difference in tooth size, however, is outside the observed limits among living apes.

3. Two of the three forms constitute one species, as above, except that the robust forms represent one of the two species and the *Homo erectus* specimens and the gracile forms together represent another highly variable species.

In this dimorphic hypothesis the principal postulated variation is in the size and shape of the brain case. The admissibility of the hypothesis hinges on accepting the fact that among early humans the cranial capacity of the females was roughly half that of the males.

4. Two of the three forms constitute one species, as above, except that the gracile forms represent one of the two species and the highly variable second species consists of the robust forms and the *Homo erectus* specimens.

In this dimorphic hypothesis the principal postulated variations involve the brain case, the jaws and the teeth. The admissibility of the hypothesis hinges on accepting the fact that small-brained but large-jawed forms and large-brained but small-jawed forms can be placed together in the same species.

5. The three forms represent three separate species.

Having listed the five possible hypotheses, we can now assess the probability of each being correct. We shall do so bearing in mind both the fossil evidence and what is known about the variability of living primate populations. We think the probability that the single-species hypothesis is correct is very low. First, the hypothesis involves accepting the fact that there is an enormous amount of intraspecies variability. Second, we think, along with others, that the adaptations apparent in the skulls of both extreme forms (*Homo erectus* and *Australopithecus robustus*) are different. In *H. erectus* the size of the brain case seems to overwhelm the chewing apparatus, as it does in living man. In *A. robustus* the opposite is true.

For the same reasons we think the probability that the fourth hypothesis is correct is very low. In addition difficulties other than anatomical ones stand in the way. One must also ask why no specimens of *A. robustus* are found in Java and China, where *H. erectus* specimens are comparatively abundant. If the answer is that such specimens were present but for some reason were not fossilized or have not yet been unearthed, then how is it that *A. robustus* is the commonest hominid in the East African fossil record?

The third hypothesis is an attractive one, but we think the probability of its being correct is low. It would be difficult enough to account for the enormous dimorphism in brain size without having to supply an answer to an additional question: Why are these gracile forms of *Australopithecus* not found as fossils in Java and China?

The second hypothesis, we think, is more likely to be correct than the third. Although the variability in the dimensions and proportions of the teeth cannot be matched in living dimorphic populations, there is some hint that dental dimorphism might have been greater in extinct hominoids: the superfamily that includes both the hominids and the apes. Louis de Bonis of the University of Paris has collected a series of Miocene hominoid jawbones in Macedonia. His sample is from one small area; if the mandibles all belong to the same species, they represent a degree of dental dimorphism greater than that found in living anthropoid apes.

The hypothesis with the best chance of being correct, we believe, is the last of the five: that three species are present. Sorted out in this way, none of the specimens within each group shows more variability in brain size and chewing apparatus than we see among living anthropoids. Although the fossil record from the lower member of the Koobi Fora Formation is far less rich than that from the upper member, a similar three-species hypothesis could also be advanced with respect to the specimens found in it. In this hypothesis the third species in addition to the robust and gracile ones would be represented by the *habilis* specimens.

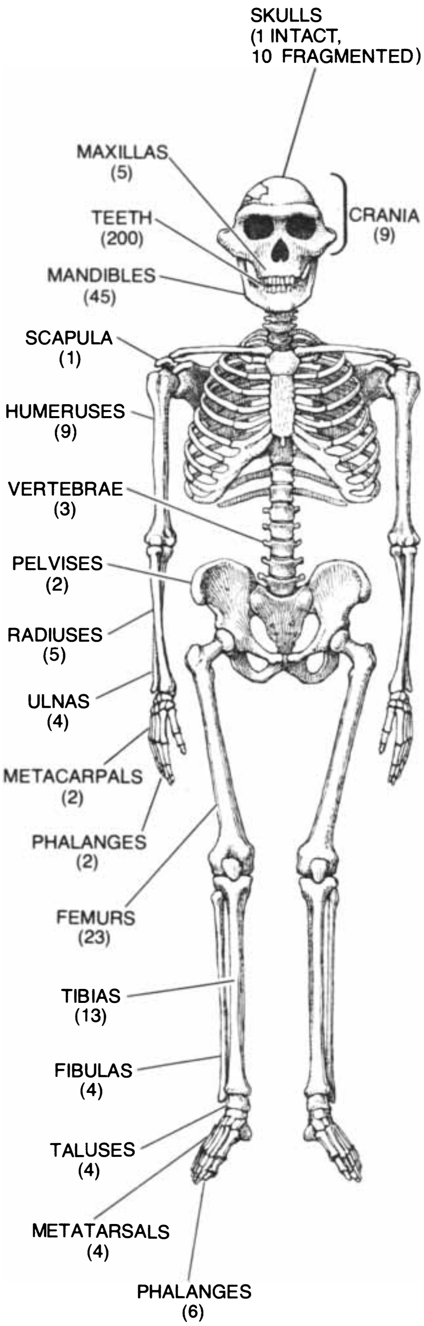
Several consequences follow from our probability assessments. For example, in our view the demonstration at East Turkana that *Homo erectus* was contemporaneous with some of the largest representatives of *Australopithecus robustus* amounts to a disproof of the single-species hypothesis. We believe both *H. erectus* and *A. robustus* had essentially human characteristics. If this is the case, it follows that they occupied separate ecological niches. It would seem either that one of the species did not possess culture and yet still developed human characteristics or that the argument that the advantage of culture would give the cultured hominid dominance within a very wide ecological niche is flawed.

We prefer the first alternative, and we would nominate *Australopithecus* for the role of the hominid without culture. Accepting this alternative requires that we keep searching for the natural-selection pressures that have been responsible for producing the basic human attributes.

Where did *Homo erectus* come from? Some have suggested that the species arose in Asia and migrated to Africa. This seems to us an unnecessarily complicated hypothesis. For one thing, it neglects *habilis*. Worse, it implies that a population of these large-brained hominids, who presumably made the stone tools found in the early East Turkana strata, evolved independently in Africa at the right time to fit into an ancestor-descendant relation with *Homo erectus* and then came to an abrupt halt, without playing any further part in human evolution.

It is our view that the *habilis* populations are directly antecedent to *Homo erectus*. If the earlier range of dates for the strata where the *habilis* specimen KNM-ER 1470 was found proves to be correct, then the transition from *habilis* to *erectus* could have been a gradual one, spanning a period of well over a million years. If the later dates are correct, then the transition must have been very rapid indeed.

We have concentrated here on giving an overview of the hominid record in East Turkana to the neglect of other work in progress that promises to answer some of the questions we have raised. The Koobi Fora Research Project is continuing its field activities, and a number of special studies are also under way outside Africa. Michael H. Day of St. Thomas's Hospital Medical School in London is examining the fossil limb bones and associated parts from East Turkana to see what can be learned about the various species' capacity for upright walking. Bernard A. Wood of the Middlesex Hospital Medical School in London is conducting a full analysis of the skulls and jaws in order to document the extremes and means of variability and dimorphism. Holloway is studying casts of the inside of brain cases in an effort to trace the evolution of the brain. One of us (Leakey) continues as codirector of the research project and the other (Walker) is studying the biomechanics of hominid mastication, examining the fossil teeth from East Turkana with the scanning electron microscope in an attempt to deduce dietary habits from the patterns of tooth wear. These studies and others aim at reconstructing as much as can be reconstructed about the biology of these very early hominids in the hope of determining just what it was that made us human.



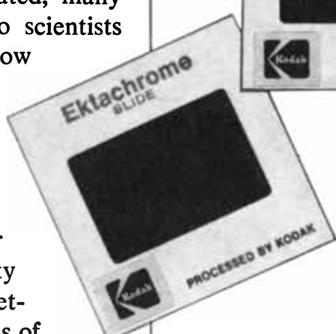
DISPROPORTION in preservation of the remains of fossil hominids unearthed thus far in East Turkana is assessed in this illustration. Teeth, which are the hardest of all body parts, are the most numerous remains. Mandibles, most of them representing the robust species of *Australopithecus*, come next. Rarest of all are pelvic bones, the bones of hands and feet, vertebrae and the bones of the lower arms.

What to do with your slides at a poster presentation

Scientists frequently make amazing discoveries. Now they have discovered, to their own astonishment, that poster sessions at scientific meetings really work.

Ever since the founding of the Royal Society in the 17th century they had assumed that a proper scientist communicates science in a formal lecture. In the 19th century public entertainers introduced lantern slides, and after only a generation or two slides became not merely respectable but *de rigueur* for scientific communication.

Then, a little past the middle of the 20th century, the science fairs hit the schools that the scientists' children attended. Of those children that participated, many have now turned into scientists themselves and have now largely succeeded in convincing their senior colleagues that the trade-fair format for discourse is not only much older than the Royal Society but fosters discourse better pitched to the needs of the parties communicating than with one-way communication. Mutual stimulation is vital to the scientific enterprise, of course.



Trouble is that the passerby at a poster session gets to decide at which presentation to stop and which to pass by. He who is too much passed by may sour on science—a pity. This must be forestalled. He must be warned that at a poster session one does not simply letter on the wall the words that would otherwise have been read to the colleagues from a sheet of paper in the gloom whilst several dozen slides crept up cautiously to the nub of the tale. In a poster presentation the nub must hit the audience first, not last, as the audience is moving.

One selects from all the slides one has accumulated in the course of the work the three or four shots which represent its essence in the most vivid, concentrated, climactic, tangible, curiosity-compelling form. The best of these is selected to occupy a substantial portion of the allotted wall space as a 16″ x 20″ color print, the other two or three a little smaller beneath it, with succinct captioning.

Just take those few key slides into a place where the principal business is photographic. If any cropping or other special attention would help the prints make your point better, tell them. Let their experience determine whether the job should go to us. If they feel it ought to be done by an independent lab on Kodak paper, we'll still be happy if you are.

Don't leave the other slides back home. They may be needed for some deep discussion.

Kodak Publication P-319 will offer more advice not only for presenters at poster sessions but for organizers thereof and even for those merely attending. Drop a line to Kodak, Dept. 412L-118, Rochester, N.Y. 14650 if you'd like to have "Poster Sessions" when ready. No charge for single copy.

Baseless Missile

The future of the U.S. program to develop a mobile land-based intercontinental ballistic missile (ICBM) is in doubt now that the Carter Administration has decided to abandon the Air Force's favorite "basing mode" for the new, improved strategic weapon and to slow development work on the project. The proposed missile, variously referred to as Missile Experimental, Missile X or just plain MX, had been conceived as an eventual replacement for the 1,000-missile, fixed-silo Minuteman force, which some military strategists think will become vulnerable to a massive surprise attack by increasingly accurate Russian ICBM's sometime in the mid-1980's.

As of a year or so ago the Air Force seemed to have settled on a "concealed basing mode" for the MX that called for constructing a vast system of concrete-lined buried trenches, each 12 to 15 miles long, on Government-owned land in the southwestern U.S. The missiles would be moved about at random intervals inside the tunnels, creating a difficult target for an attacking missile force. The earth-covered roof of the tunnels, according to the Air Force, would have segmented "lid" sections that would be lifted by a piston-driven mechanism on the "transporter/launcher vehicle." The cost of such a system, including some 300 of the new missiles and their nuclear warheads, had been estimated to run as high as $36 billion.

In closed-session testimony before the Senate Armed Services Committee, however, Administration spokesmen revealed that they lacked "sufficient technical data... at this point to select a basing mode" for the new missile, and that the continuous-trench approach favored by the Air Force a year ago was "no longer viable." In the words of Lt. Gen. Alton D. Slay, commander of the Air Force Systems Command: "I think that since we are talking of a program that could exceed $25 billion, including operational and support costs for 10 years, we ought to make very, very sure we know what we are doing on the basing mode."

The decision to scrap the MX underground-tunnel basing scheme came after the concept had been subjected to mounting criticism both within and outside of the Government. The Office of Management and Budget, for example, had taken the position that full-scale development of the MX should be delayed until further research could be done on the feasibility of the various alternative basing modes outlined by the Air Force. The Defense Science Board, an advisory group reporting to Secretary of Defense Harold Brown, had urged that the underground-tunnel basing mode be rejected for technical reasons. An independent analysis of the entire MX question by a group of non-Government arms-control specialists under the auspices of the Program in Science and Technology for International Security of the Massachusetts Institute of Technology had in the meantime found that in certain respects the proposed MX tunnel system would be easier for the U.S.S.R. to attack successfully than the present system of fixed Minuteman silos. The M.I.T. group (M. Callaham, B. T. Feld, E. Hadjimichael and K. M. Tsipis) concluded that such "basing modes of limited mobility that rely on concealment for improved invulnerability... do not offer guaranteed long-term security and create insuperable arms-control problems."

An alternative plan has since emerged as the leading candidate for deploying the mobile ICBM. The new scheme, which has actually been under study for some time, calls for the construction of some 4,000 empty concrete launching silos, among which 200 MX's would be moved about in a random fashion, presumably by truck. This approach, called in Pentagon circles the "shell game," is strongly advocated by the Defense Science Board and is said to have the support of Secretary Brown. The projected cost of this MX basing mode is reported to be in the vicinity of $20 billion.

The currently favored MX basing mode was among those considered by the M.I.T. analysts, who in general found it wanting. The key security problem presented by any such system, according to their report, "is that, since the aim points are far apart, an opponent would be able to attack them simultaneously with relatively inaccurate but large-yield warheads, without the impediment of fratricide [in which the detonation of one incoming missile would destroy another incoming missile before it exploded] that forbids the use of numerous warheads against hard-point targets.... Additionally, there is always the question of real-time monitoring of the movements of the missile carriers and the possibility, however remote, that it would allow the Soviet Union to target with more certainty than what random or blanket coverage of the shelters or mobile carriers provides." Besides, the report goes on, "trucks circulating around the country and hiding in unspecified shelters or locations present very serious physical security problems since they could be susceptible to sabotage. In addition, their command and control would present difficulties."

The main trouble with all such "concealed multiple aim-point schemes," however, is that they would "raise very serious arms-control difficulties, since the denumeration of the deployed missiles would not be possible by national technical means of verification. Consequently, any agreement to limit the number of deployed mobile ICBM's would be nonverifiable and the current ceiling on ICBM's that hopefully will be fixed in the context of SALT II would be vitiated.... The entire approach appears to be susceptible to subjective worst-case analyses which will claim that the opponent possesses enough RV's [reentry vehicles] to defeat even the most dispersed multiple aim-point basing scheme. Since the actual numbers could never be verified, such analysis combined with the uncertainty of concealed basing could easily result in an open-ended proliferation of mobile ICBM's."

Instead of the proposed MX shell game, the M.I.T. group suggests that "the most cost-effective means of maintaining the present mutual invulnerability of Russian and American deterrent forces through the next decade and probably somewhat beyond would be through *an agreement* to limit drastically the testing of long-range missiles. Such a measure would impede the transformation of existing ICBM's into effective antisilo weapons, and would slow down the development of any new stragegic missiles. A Comprehensive Nuclear Test-Ban Treaty would be one useful element of such a system of limitations. The achievement of such a missile-test limitation would be the most effective way of pointing the efforts of the U.S. and U.S.S.R. in the desired direction of meaningful restraints and the reduction of the threat of mutual and worldwide nuclear disaster."

Maligned Malpractice System

The medical malpractice system is under widespread attack largely for the wrong reasons because its most significant function is not understood. The purpose of malpractice law is not simply to compensate injured people, according to an article in *The New England Journal of Medicine* by William B. Schwartz of the Tufts University School of Medicine and Neil K. Komesar of the Law School of the University of Wisconsin at Madison. It is also to deter negligent behavior by physicians and other providers of medical care by giving them an economic incentive to avoid careless injuries. Considering the malpractice system from that point of view and in the light of recent experience, Schwartz and Komesar conclude that contrary to received doctrine there are not too many malpractice claims, that damage awards are not too high, that the system is not generally unfair to able doctors, that the customary payment of the lawyer's fee on a contingency basis is appropriate and that no-fault insurance would be an ineffective alternative to the malpractice system.

Schwartz and Komesar base their essentially economic analysis on a definition of negligence they derive from the late Judge Learned Hand: "Negligent behavior is the failure to invest resources up to a level that equals the anticipated saving in damages." Negligence litigation, Schwartz and Komesar write, "signals potentially negligent people that it will cost them more to be careless than to invest in an appropriate level of prevention." Most physicians "guard the welfare of their patients without need for external regulation. The malpractice system exists to discipline the occasional physician who does not (or cannot) protect his patients." A negligence suit or the threat of one might, for example, induce a negligent doctor to take more time for an examination or a procedure, to seek more training or to give up performing a lucrative procedure he is not competent to perform.

"The ideal negligence signal" is transmitted, however, only if noteworthy malpractice always leads to a claim and valid claims always lead to full awards. In the real world, however, "many more instances of malpractice occur, it appears, than result in a claim for damages." Schwartz and Komesar cite several studies indicating that "probably below 20 percent" of all malpractice incidents result in claims. As for damage awards, the most conspicuously high ones usually include compensation for "pain and suffering" almost always from catastrophic injuries such as paralysis or blindness, and they "do not seem excessive." Most payments in the range below $50,000 have been reported to be "substantially below medical expenses and lost earnings." The number of claims and the size of awards thus appear to be insufficient to constitute an "ideal" signal.

Even if every incident led to a claim and every claim led to an award equal to the loss suffered by a patient, the resulting signal "would fail to elicit the appropriate response from physicians." Malpractice insurance "virtually insulates the negligent physician from the damages award," and hence from the signal, because malpractice premiums are usually set for a particular medical specialty in a given region and not on the basis of an individual physician's record of malpractice claims, out-of-court settlements and verdicts. A habitually negligent physician carries no more of the premium burden than his colleagues with good records do. "If a physician were rated by his individual experience, his premium would reflect, to at least some degree, the risk that he poses to the insurer." Schwartz and Komesar concede that experience rating is justifiable only if suits are not brought randomly or capriciously. The common impression that suits are in fact brought randomly is not supported, they write, by a study of 8,000 physicians in the Los Angeles area. "In a four-year period, 46 physicians (.06 percent of the 8,000) accounted for 10 percent of all claims and 30 percent of all payments made" by an insurance plan; an average of five suits was brought against each of the 46 in the four years. "Analysis indicates that doctors against whom multiple suits are brought do, indeed, represent a higher-risk population than their colleagues."

Contingency fees are often attacked as encouraging lawyers to bring malpractice suits that lack merit. In reality they tend to discourage frivolous suits, according to Schwartz and Komesar, because the lawyer loses money on a case unless it results in a settlement or a favorable verdict; he has an interest in filtering out capricious claims, and also an incentive for accepting strong cases in behalf of clients who could not pay a flat fee for service. It is a flat fee, they argue, that would encourage suits regardless of their merit.

No-fault insurance, often urged as an alternative method of compensation, would not be effective as a deterrent if premiums did not reflect the individual doctor's record. Experience rating for no-fault insurance would have "perverse effects," however: it would encourage "defensive medicine" because the doctor would have his premium raised by any "bad outcome," regardless of negligence, and it would discourage risky procedures even when they were appropriate.

Getting More out of Carbon 14

Carbon-14 dating, which has made profound contributions to archaeology and other disciplines, has always had two limitations. First, the samples needed for dating are relatively large (between one gram and 10 grams), which precludes dating smaller samples or precious artifacts that would be ruined by taking a sample of them. Second, samples older than 40,000 years are almost impossible to date because practically all the carbon 14 in them has radioactively decayed. Currently, however, physicists are exploring a new method of carbon-14 dating they believe will do much to overcome these limitations and will also be applicable to dating with other radioactive isotopes.

The new dating method is based on using a particle accelerator as a supersensitive mass spectrometer. In the original method developed by Willard F. Libby and his collaborators in the 1940's the sample to be dated is first burned so that it evolves carbon dioxide, which is converted into pure carbon by being passed over hot magnesium. The carbon is then spread as a thin film on the inside of a proportional counter, which measures its content of carbon 14. This radioactive isotope decays with a half-life of $5,730 \pm 40$ years; if the ratio of its abundance to the abundance of nonradioactive carbon is high, the sample is young, and if the ratio is low, the sample is old.

In the particle-accelerator method a much smaller sample of carbon is ionized and the ions are accelerated as a beam. Each of the three kinds of ion in the sample (the nonradioactive carbon 12, the rarer nonradioactive carbon 13 and the radioactive carbon 14) has the same electric charge but a different mass. By tuning the accelerator so that it accelerates first one kind of ion and then another, the different ions in the sample can be counted and their relative abundance can in principle be determined with high precision.

An accelerator was used in this way as far back as 1939, when Luis W. Alvarez and Robert Cornog of the University of California at Berkeley were looking into the differences between the nucleus of helium 3 and that of hydrogen 3 (tritium). The method was revived by Richard A. Muller of the Lawrence Berkeley Laboratory of the University of California in a 1976 experiment that dated a sample of deuterium (hydrogen 2) by determining the ratio in it of deuterium to tritium. Muller and his co-workers later applied the method to carbon-14 dating. The accelerator he has been using for these experiments is a cyclotron; groups doing similar experiments at the University of Rochester and McMaster University have been working with a Van de Graaff tandem accelerator.

The investigators working with the accelerator method find they need only milligrams of material for dating purposes. The sensitivity of the method should also enable it to detect vanishingly small quantities of carbon 14 in a sample after many carbon-14 half-lives have elapsed. Muller predicts that it will ultimately be possible to establish dates going as far back as 100,000 years.

The early efforts at carbon-14 dating with accelerators have yielded promising results. The Rochester group has dated samples with ages of between 200 and 48,000 years and has achieved close agreement with the ages found for the same samples by the older method. The Berkeley group recently became the first to attempt a "blind date" with an accelerator: to date a sample of carbon dioxide whose age was not known to them. They determined the age to be $6,060 \pm 800$ years, which differed by about 830 years, or one standard deviation, from the age determined by the older method.

Promising as the carbon-14 work is, Muller and others believe the accelerator dating method will be even more useful with other isotopes. The Berkeley group and one at Grenoble in France have worked on dating with the radioactive isotope beryllium 10, which Muller says "has the potential to be the carbon 14 of geology." Beryllium 10 is trapped inside sedimentary rocks as they are formed, and because it has a half-life of 1.5 million years and its radioactivity is therefore low it is almost impossible to detect by counting decays. Work is also being done toward dating with other radioactive isotopes, such as tritium, aluminum 26, silicon 32 and chlorine 36.

Active Optics

Modern astronomy has reaped a rich harvest from new instruments that gather radiation beyond the long- and short-wavelength ends of the visible spectrum, but the great tradition of astronomical instrumentation is the large light-gathering telescope. Since the visible spectrum is centered on a wavelength of about half a micrometer, lens and mirror surfaces must be accurate to about a quarter of that length if they are to produce images of good quality. Moreover, a large telescope mirror must be quite rigid if its surface is to resist distortion. Large telescope mirrors are notoriously difficult to make; the larger the mirror is, the more intractable the problems become. Now a new approach to telescope design is nearing realization at the Mount Hopkins Observatory in Arizona: a telescope made up of several mirrors instead of a single large one. The Multiple Mirror Telescope (MMT) is a joint project of the University of Arizona and the Smithsonian Astrophysical Observatory. When it goes into operation later this year, it will be the third-largest optical telescope in the world.

The Multiple Mirror Telescope consists primarily of six 72-inch mirrors in a hexagonal array. The separate images formed by the individual mirrors are focused at a single point, so that together the six mirrors have a light-gathering capacity equivalent to that of a single 176-inch telescope. The images from the six mirrors are sent up to a six-sided secondary mirror on the central axis of the telescope; that mirror in turn sends the images down to a final common focus. An "active optics" system, or feedback mechanism, keeps the six main mirrors precisely aligned.

This is done by splitting a laser beam into six beams, each of which is directed at one of the mirrors; the relative positions of the reflected beams are sensed by detectors linked to a computer. If there is any discrepancy due to thermal or gravitational stresses, the computer activates a servomechanism that wobbles the secondary mirror or moves it backward or forward to keep the six main mirrors focused on the same point. Although the fluctuations caused by thermal and gravitational stresses are rather slow, the feedback system can correct an error in .05 second.

A more difficult problem is compensating for errors of phase or path length in a light wave front due to atmospheric turbulence, which is the major factor limiting the performance of earth-based telescopes. As an approach to the problem active-optics systems have been developed over the past decade that detect errors in a wave front and compensate for them automatically within milliseconds in "real time," that is, while the fluctuations are in progress. Writing in *Proceedings of the IEEE,* John W. Hardy of the Itek Corporation reviews the state of this new art.

All real-time active-optics systems include a feedback loop consisting of three components: a sensor that detects errors in the wave front, a computer that accepts the measured data and converts it into control signals, and an "active" mirror that can be altered in curvature or tilted to compensate for the detected phase error. Two types of active mirror have been developed: one consists of an array of smaller mirrors, each of which can be moved in two axes of tilt and one axis of focus; the other consists of a thin single mirror that is deformed by means of an array of pistons pushing against its back surface. For high-speed wave-front compensation, such as that required to correct for atmospheric turbulence, light is collected in a large mirror and relayed to a small mirror, which can be deformed on the millisecond time scale required for the phase corrections.

The sensor is a shearing interferometer, which takes the wave front generated by a single intense source (such as a bright star) and splits it into two beams, one of which is displaced with respect to the other. An interference pattern of dark and light "fringes" results, so that errors of phase are converted into differences in intensity. The intensity differences are sensed by photodetectors that are distributed across the aperture of the system and transmit electric signals to a computer. The computer then, for example, commands the array of pistons behind a thin mirror to correctively deform the mirror.

Because the sensor needs a large amount of light in order to analyze turbulence in real time the system can operate only with bright light sources and hence is of limited use in astronomy. There are many other potential applications, however. One is to improve the propagation of a laser beam through the atmosphere or some other gas for communications or other purposes (such as laser weaponry or laser fusion). A laser beam can be focused on a very small area, but the coherence of its wave front may be spoiled by two sources of aberration: thermal fluctuations in the transmitter and turbulence in the gas between

the transmitter and the receiver. These sources of aberration can be corrected for by an active-optics system incorporated into the circuit of the laser. To monitor atmospheric turbulence in real time a telescope is focused on a distant target at which the laser beam has been directed; the distortion of the beam caused by turbulence is then measured and compensated for by distorting the beam before it leaves the transmitter. The concept of laser fusion is to focus an array of laser beams on a pellet of deuterium in order to bring about the fusion reaction; conceivably real-time active-optics systems could improve the efficiency of the process.

The feasibility of large optical arrays in space (such as orbiting solar-energy collectors) depends on the development of highly stable lightweight optical systems. Such a system might have a structure spanning tens of meters that would have to be held true to within micrometers. Since the system would be subjected to disturbing forces, notably in the course of being launched and placed in orbit, it would surely require active-optics control. Hardy concludes: "If the present rate of progress in active optics is maintained, the next decade should bring most of these applications to fruition."

A Locket for Tutankhamen

The ancient Egyptian custom of surrounding buried royalty with a great wealth of funerary furnishings made the occupation of grave robbing a lucrative profession in the latter part of the second millennium B.C. The Valley of the Kings at Thebes was a major robbers' stamping ground, and the kings and queens of the Eighteenth Dynasty, who by then had been buried for centuries, were their chief targets. They pillaged not only the grave goods but also the jewels adorning the mummified rulers themselves.

In a period of particular piety near the end of the second millennium B.C. priests of the Twenty-first Dynasty set about salvaging and rewrapping many of the disturbed royal mummies. The rescued mummies were not returned to their original tombs but were hidden in two principal caches. One was the tomb of an early Eighteenth Dynasty queen, Inhapy, and the other was the tomb of Amenhotep II, one of the immediate predecessors of Tutankhamen, whose treasures escaped the robbers.

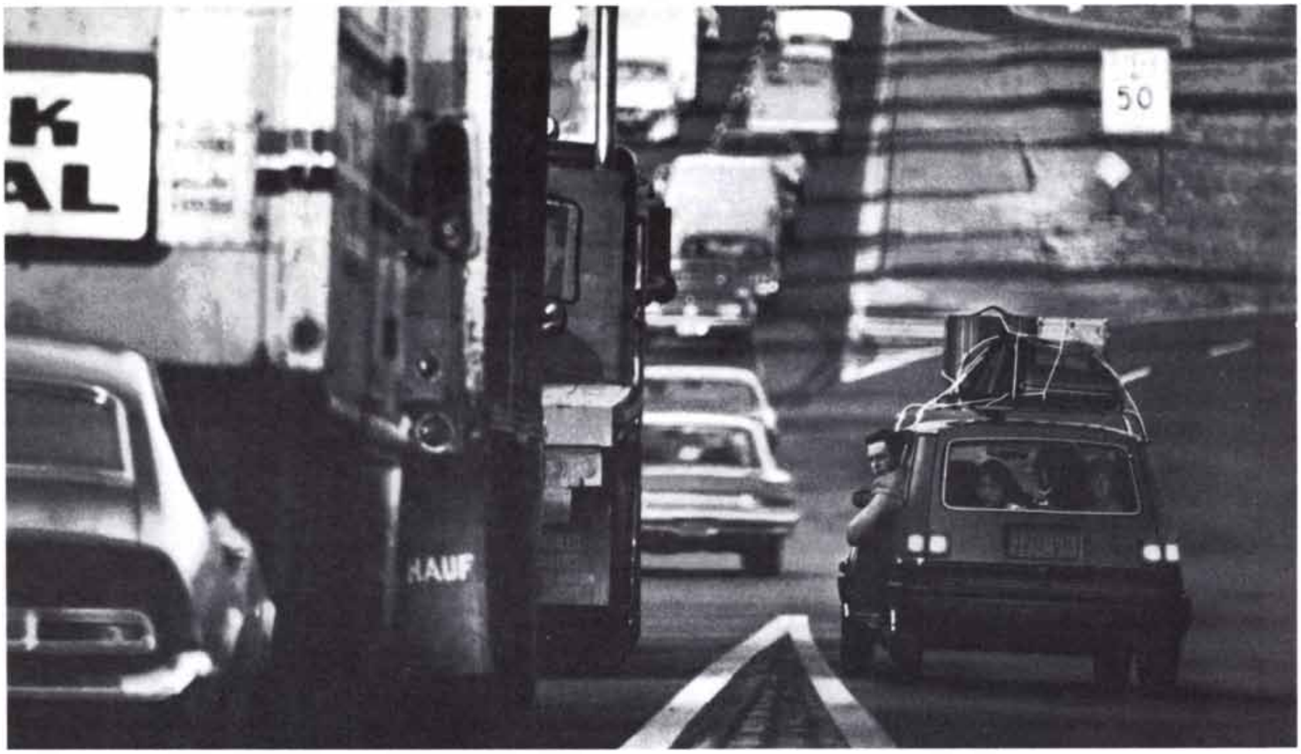
"ELDER LADY," a mummy of Eighteenth Dynasty Egypt, was found in a cache of mummies from Valley of the Kings in 1898. Tests now show her to be Queen Tiye, mother of Akhenaten.

Between 1881 and 1898 Egyptologists at work in the Valley of the Kings discovered both of these ancient caches and undertook to identify the assembled royalty of the Eighteenth Dynasty. It was no easy task: the priests who had salvaged the disturbed mummies some three millenniums earlier had not always used the original wrappings and coffins. Those who wrestled with the problem included the British Egyptologist G. Elliot Smith, whose analysis, *The Royal Mummies*, was published in 1912. Among the mummies whose identity remained ambiguous was one Smith named "The Elder Lady in the tomb of Amenhotep II." Smith described her as "a middle aged woman with long... brown, wavy, lustrous hair.... The left hand was tightly clenched, but with the thumb fully extended... the [left] forearm being sharply flexed."

Since 1967 investigators at the University of Michigan have been conducting an X-ray investigation of the royal mummies in collaboration with the Egyptian Government Department of Antiquities, using Smith's catalogue as a checklist. Most of the collection was in the Egyptian Museum in Cairo, but when the Michigan workers came to look for No. 61070, "the Elder Lady" of Smith's list, they found she was still at the Valley of the Kings. The Michigan workers moved their X-ray equipment to Thebes in 1975, made plates of the Elder Lady and took a small hair sample. Comparison of the lateral skull X ray of the Elder Lady with similar views of other royal females suggested that there was a strong family resemblance between the Elder Lady and Queen Thuya, who was the mother of Queen Tiye, one of the wives of Amenhotep III.

It happens that one treasure from Tutankhamen's tomb is a locket in the form of a miniature coffin, inscribed with Queen Tiye's name and containing a lock of hair presumably donated by the queen. (Tiye was by most reckonings Tutankhamen's grandmother.) The Michigan workers were given a small sample of the hair and submitted both the Tiye locket sample and the sample of hair taken from the Elder Lady to electron-microprobe analysis. Reporting on the result of the investigation in *Science*, James E. Harris of the University of Michigan School of Dentistry and his colleagues conclude that the virtually identical X-ray scatter from the two samples demonstrates that the Elder Lady is Queen Tiye, the wife of Amenhotep III and the mother of the heretical monotheistic pharaoh Akhenaten. The skull X ray further shows that Queen Tiye was only in her forties when she died, indicating that in naming her the Elder Lady, Smith had done her something of a retroactive injustice.

IS IT WORTH RISKING YOUR LIFE FOR 45 MILES PER GALLON?



Tiny little subcompact cars may be great for saving gas. But as accident statistics show, they're not particularly safe.

A Volvo, on the other hand, gets a very respectable 29 m.p.g. highway—19 m.p.g. city.* But ultimately, we put a much higher premium on life than we do on gasoline.

The roomy passenger compartment of a Volvo is surrounded by six steel pillars, each one strong enough to support the weight of the entire car.

Crumple zones, front and rear are designed to absorb the impact of a collision, rather than passing it on to the passengers.

As a matter of fact, the federal government is so impressed with Volvo's crash worthiness they've become one of our biggest customers. They bought more than 60 Volvos, many of which have been crashed into each other at closing speeds of up to 90 m.p.h. in an effort to establish safety standards for cars of the future.

So before you buy your next car, weigh carefully what you have to gain and lose.

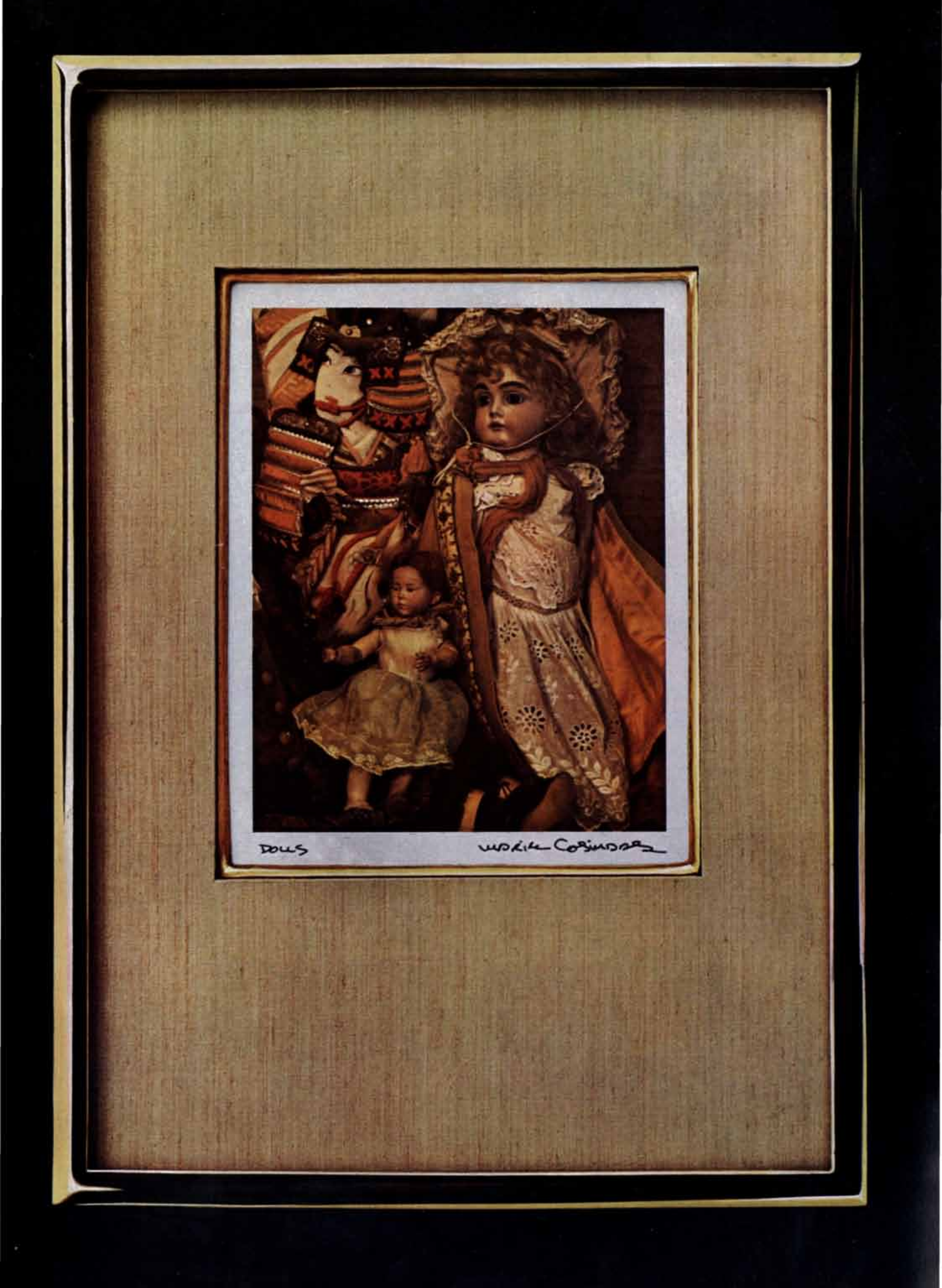
A big substantial Volvo can not only save gas.

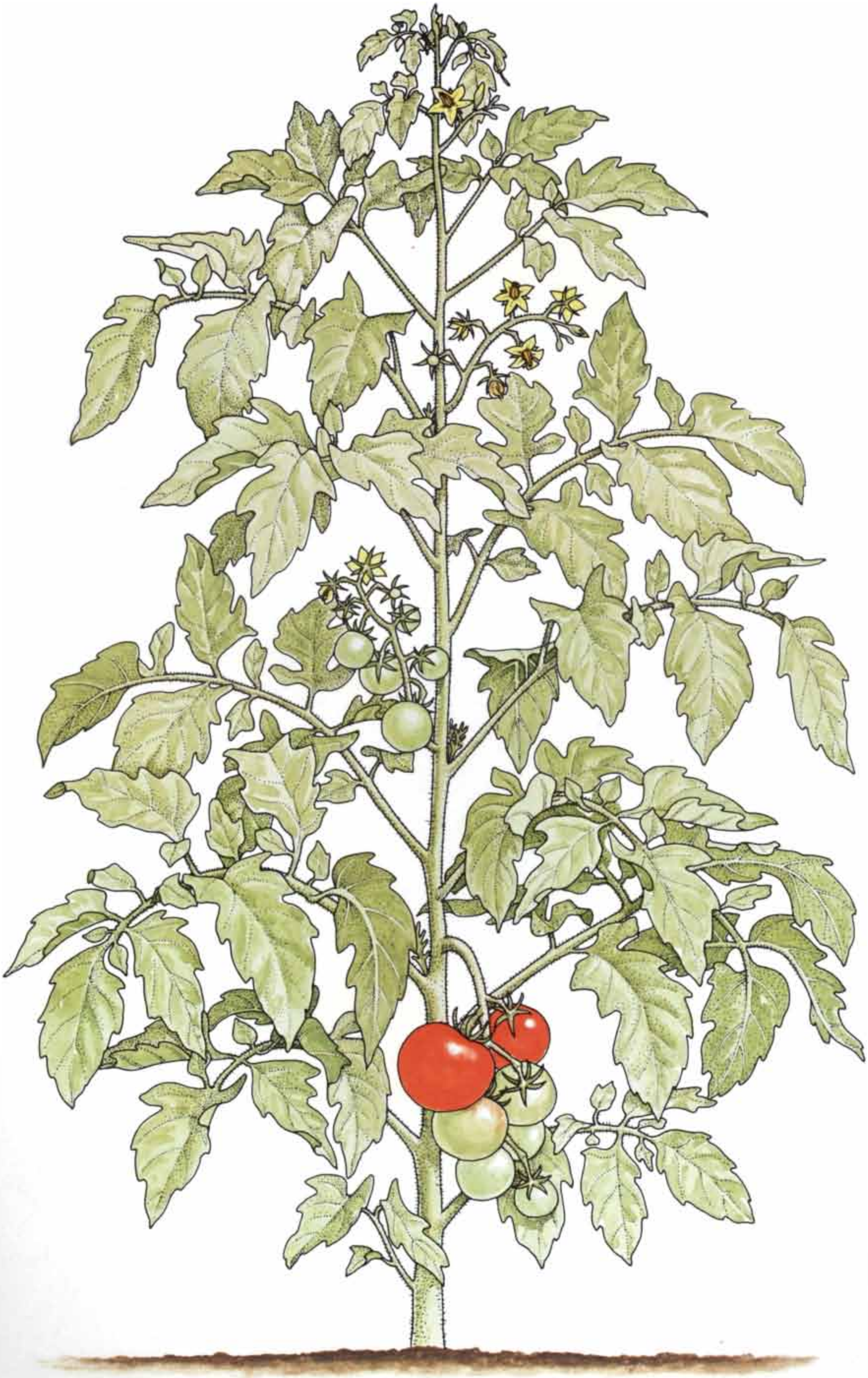
It could end up conserving something much more precious.

*EPA estimates based on 240 Sedans with manual transmission. Mileage may vary on other models or because of your driving habits, cars condition and equipment.

VOLVO ©1978 VOLVO OF AMERICA CORPORATION. LEASING AVAILABLE

VOLVO. A CAR YOU CAN BELIEVE IN.


DOLLS



The Tomato

The esculent Lycopersicon esculentum, long thought to be poisonous, has become a major U.S. food crop and source of vitamins and minerals, thanks largely to genetic modification and new production technology

by Charles M. Rick

The tomato (a fruit that is almost universally treated as a vegetable and a perennial plant that is almost universally cultivated as an annual) is the focus of a large agricultural industry. Among vegetables grown for human consumption in the U.S. in 1977 the commercial production of tomatoes took up 491,080 acres (second only to sweet corn), yielded 8,755,950 short tons of produce (second to none) and had a value of $914.1 million (second to none). The figures do not include the large number of tomatoes grown in home gardens. It is hard to realize now, when the tomato is highly popular not only as a food in its own right but also as an ingredient of soups, sauces (including the ubiquitous catsup) and drinks, that as recently as 1900 it was widely avoided in the belief that it was poisonous because of its known relation to nightshades and other toxic members of the nightshade family (Solanaceae). Even by 1920 the per capita consumption of tomatoes in the U.S. was only 8.2 kilograms per year; now it is 25.5 kilograms (mostly of tomatoes in preserved forms), a rise of 300 percent during a period when the consumption of potatoes declined by 20 percent. The story of the tomato's progress from an exotic fruit to a popular item of diet and a major article of commerce takes many interesting turns.

Essentially all the cultivated forms of the tomato belong to the species *Lycopersicon esculentum*. As with cultivated plants generally, the origins and the early events of domestication are largely obscure. One can be reasonably certain about three aspects. First, the cultivated tomato originated in the New World, since all related wild species of tomato are native to the Andean region now encompassed by parts of Chile, Colombia, Ecuador, Bolivia and Peru. Second, the tomato had reached a fairly advanced stage of domestication before being taken to Europe. Woodcuts in the early herbals reveal that the first types cultivated in Europe bore large fruit. (In all wild species the fruit is small.) According to the descriptions, a good many sizes, shapes and colors were known. Third, the most likely ancestor, the wild cherry tomato (*L. esculentum* variety *cerasiforme*), is spontaneous throughout tropical and subtropical America and has spread throughout the Tropics of the Old World.

The time, place and other aspects of domestication are far less certain. Although definite proof is lacking, the weight of data from several disciplines favors Mexico as the probable region of domestication. A comparison of hereditary enzyme variants reveals much greater similarity between the older European cultivars and the primitive cultivars and cherry tomatoes of Mexico and Central America than between the European cultivars and the primitive plants of the Andean region, which is undoubtedly the center of distribution of the genus *Lycopersicon* and is the other possible area of domestication. Representations of the tomato have not been found in ancient pottery and other artifacts of the Andean region, and parts of the tomato plant have not been encountered in the archaeological remains of the region, whereas parts of most of the native cultivated plants have been. The tomato has no native name in the Andean region, whereas it is known in the Nahua tongue of Mexico as *tomatl*, which is unquestionably the origin of the modern name.

The early chronicles of the New World are disappointingly sparse in their references to tomatoes. The Peruvian chronicler Guamán Poma mentions the sporadic eating of wild-tomato fruit in the Inca empire but does not mention large-fruited types. A good deal of confusion over the region of domestication has arisen from the name *mala peruviana*, which was applied to the tomato after its introduction into Italy. Although the name apparently had no factual basis, it badly misled plant geographers.

The first record of tomatoes in the Old World is credited to the descriptions published in 1554 by the Italian herbalist Pier Andrea Mattioli. The earliness of this date in itself supports a Mexican origin, considering the taking of Mexico City in 1519, the completion of the Peruvian conquest by 1531 and the time probably necessary for the introduction, cultivation and appreciation of the crop in Europe. The 1554 edition of Mattioli's herbal states that the tomato "is eaten in Italy with oil, salt and pepper." The plant was first known as *pomi d'oro* and *mala aurea* (golden apple) and also as *poma amoris* (love apple). These and equivalent names persisted well into the 19th century.

Persistent Superstitions

The references to the eating of tomatoes are quite rare. In most places the plant and its fruit were remarkably slow to gain acceptance except as an ornamental, a medicinal or a curiosity. People already knew it was related to poisonous members of the nightshade family, such as belladonna and mandrake. A typical statement was made in 1581 by the herbalist Matthias de L'Obel: "These apples were eaten by some Italians like melons, but the strong stinking smell gives one sufficient notice how unhealthful and evil they are to eat."

Such unfounded superstitions persisted widely even into the 20th century in many areas, including North America, to which the plant had been taken by colonists. The first mention of it there was made in 1710 by William Salmon in his *Botanologia, the English Herbal, or History of Plants*. The next surviving printed reference was by Thomas Jeffer-

TOMATO PLANT, *Lycopersicon esculentum*, is a member of the Solanaceae, the nightshade family, which includes the potato, tobacco and the petunia as well as the deadly nightshade and other poisonous plants. This is a typical "indeterminate" tomato plant: one that bears flowers and fruits in varying stages of maturity at the same time and grows until killed by frost.

son, who in 1782 wrote of tomato plantings in Virginia. Jefferson also makes frequent reference to the planting and culinary uses of "tomatas" at Monticello in his garden book, which he kept between 1809 and 1814. One Robert Gibbon Johnson, who is not otherwise known to history, achieved a degree of celebrity (and advanced the cause of the tomato as a food) by eating a tomato on the steps of the courthouse in Salem, N.J., in 1820.

The toxicity of certain members of the nightshade family is attributable to alkaloids. The predominant alkaloid in the tomato is tomatine, which is at its highest concentrations in the foliage and the green fruit and is degraded to inert compounds as the fruit ripens. (Even in large doses it is much less hazardous than the alkaloids of nightshade, belladonna and other species.) Possibly on the strength of early observations that plants high in tomatine tend to resist the wilt disease caused by the fungus *Fusarium oxysporum,* strains of *Lycopersicon esculentum* and the closely related *L. pimpinellifolium* have actually been bred to produce high levels of tomatine in ointments for treating fungal diseases of the human skin.

Even after the tomato came to be accepted as a food, debate persisted over whether it is a fruit or a vegetable. (Botanically it is a fruit, since it develops from an ovary.) In 1893 the question came before the United States Supreme Court. The collector of customs for the Port of New York had assessed a duty of 10 percent on tomatoes imported from the West Indies; he classified them as "vegetables in their natural state," which the tariff act of 1883 made dutiable. The same act made "fruits, green, ripe, or dried," duty-free, and the importers contended that their tomatoes should have been admitted to the country without duty.

Associate Justice Horace Gray delivered the unanimous opinion of the court. "Botanically speaking," he wrote, "tomatoes are the fruit of a vine, just as are cucumbers, squashes, beans, and peas. But in the common language of the people, whether sellers or consumers of provisions, all these are vegetables which are grown in kitchen gardens, and which, whether eaten cooked or raw, are, like potatoes, carrots, parsnips, turnips, beets, cauliflower, cabbage, celery, and lettuce, usually served at dinner in, with, or after the soup, fish, or meats which constitute the principal part of the repast, and not, like fruits generally, as dessert."

The present popularity of the tomato in the U.S. is difficult to explain on a nutritional basis. The tomato does not rank high in the concentration of any dietary component. For example, according to a recent survey by M. Allen Stevens of the University of California at Davis, among the main fruit and vegetable crops it ranks 16th as a source of vitamin A and 13th as a source of vitamin C. In terms of consumption, however, tomatoes rank third as the source of both vitamins in the U.S. This apparent anomaly is explained by the relatively large volume of tomatoes consumed. Indeed, the tomato stands at the top in total vitamins and minerals provided in the U.S. diet.

The popularity of the tomato is more readily explained in terms of its attractive color and flavor and its versatility. As an ingredient of a great variety of foods, from pizza to the Bloody Mary, the tomato has become a dietary staple.

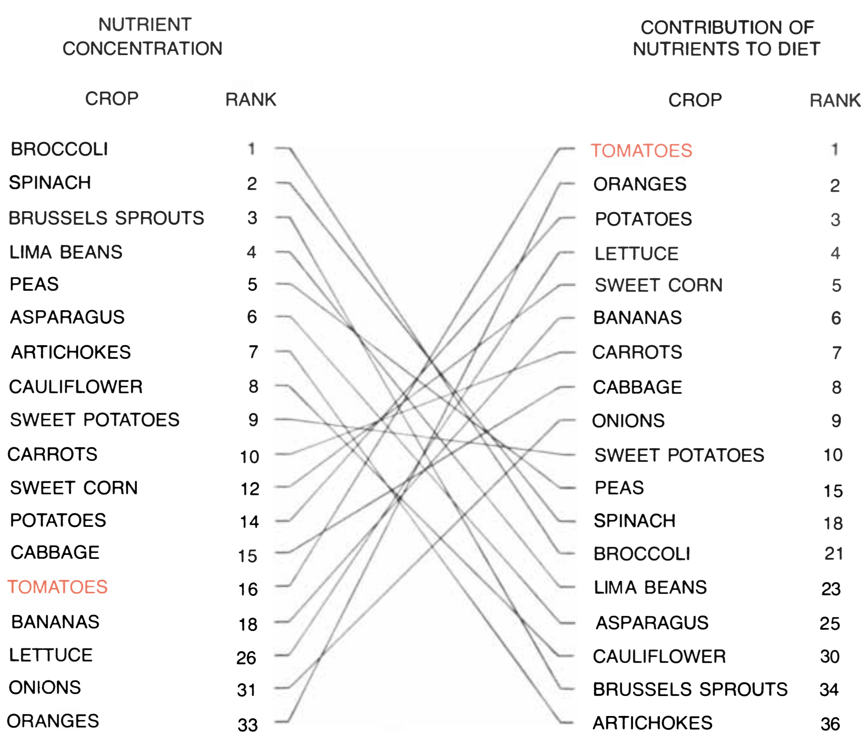
Genetic Variation

The cultivated tomato is one of nine species belonging to the genus *Lycopersicon.* Although consumption is limited largely to cultivars of *L. esculentum,* fruits of the wild variety *cerasiforme* and of the closely related *L. pimpinellifolium* are grown to a limited extent. The fruits of the other species are far less attractive than *L. esculentum* in color and flavor. To most people some of them are downright distasteful.

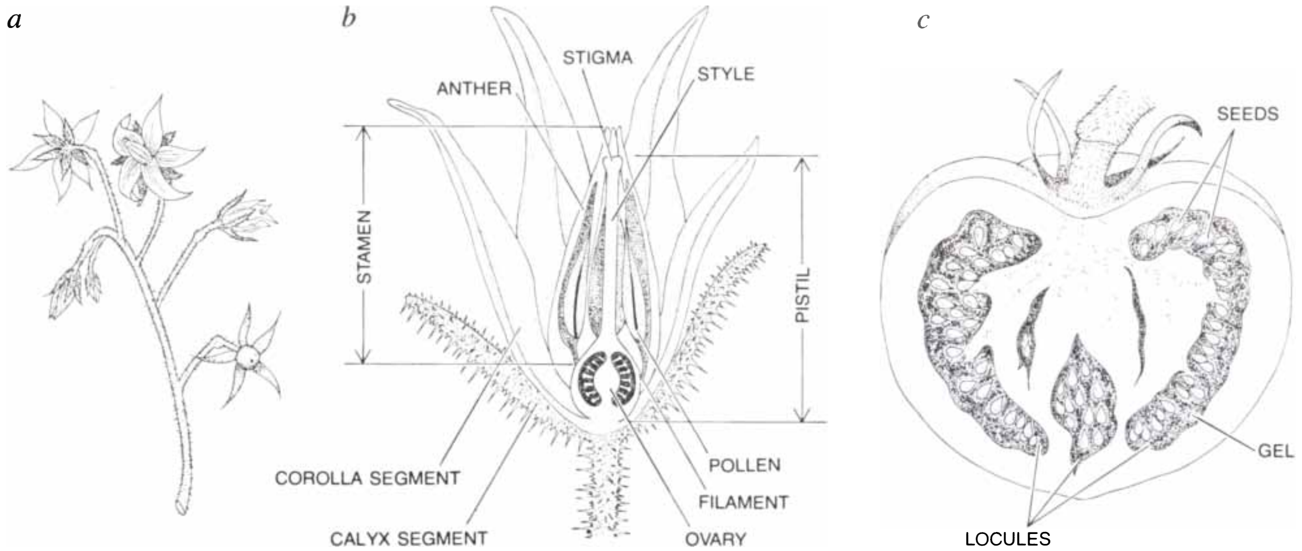
The wild species do have great potential value because of the diversity of their germ plasm. From them genes for resistance to many diseases and for improved color and quality of fruit have been bred into cultivated forms. Many cultivars thus improved are grown on a large scale. In California, for example, plantings are mainly limited to cultivars that resist the wilt caused by fusarium and verticillium fungi. Breeding for resistance to other pests is proceeding apace.

The great diversity among and within tomato species promises many more advances in this area. A case in point is the high tolerance to salt that has been discovered in types of *L. cheesmanii* that thrive on the shores of the Galápagos Islands scarcely two meters above the high-tide line and five meters away from it. Tests by Dale Rush and Emanuel Epstein of the University of California at Davis have demonstrated that these forms will survive in seawater, whereas cultivated varieties die if the concentration of seawater exceeds 50 percent. Rush and Epstein have also found that the tolerance to salt is in the tissues of the plant and does not arise from any barrier to the uptake of salt or to its movement in the plant.

Another species, *Solanum pennellii* (notwithstanding the name, it is related more closely to the tomatoes than to the nightshades), flourishes in exceedingly dry habitats in western Peru, where the main sources of water are fog and mist. Tests by A. T. T. Yu at the University of



NUTRITIONAL VALUE of the tomato is not particularly high. It ranks 16th in relative concentration of a group of 10 vitamins and minerals, according to M. A. Stevens of the University of California at Davis. Because tomatoes are so heavily consumed, however, they stand first in total contribution of those nutrients to the U.S. diet. The chart includes the top 10 foods in nutrient concentration and the top 10 in total contribution. It is notable that although broccoli, spinach and brussels sprouts are "good for you," they are not the country's favorite foods.



TOMATO FLOWERS, borne in clusters (*a*), contain both male and female organs (*b*). Pollen grains from the anther adhere to the stigma, germinate and deliver sperm by way of the style to the ovary, fertilizing an ovule, which develops into a seed. Thus stimulated, the fruit is "set," that is, the ovary becomes the fruit. The mature fruit (*c*) contains a number of locules, or cells, that are filled with seeds.

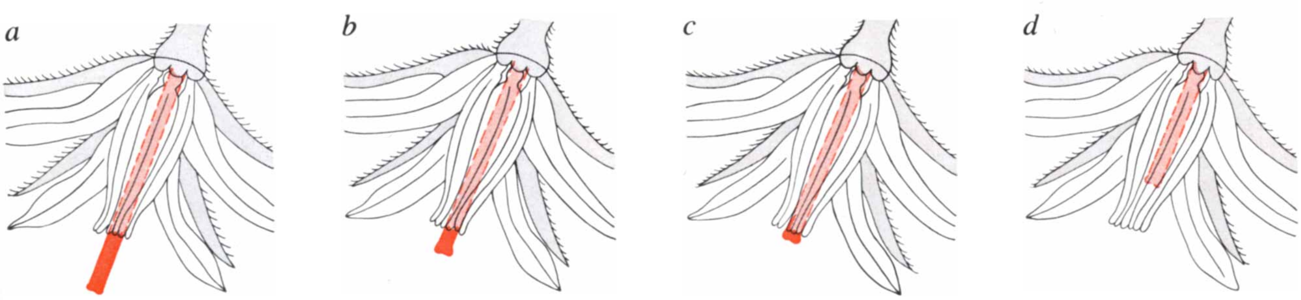
California at Davis have revealed that the plant tolerates the environment largely because its foliage resists water loss. Our own tests at Davis indicate that this trait can be bred into horticultural tomatoes. Other useful characteristics of wild tomatoes are insect resistance in *L. hirsutum* and tolerance of moist, tropical conditions in certain variants of *cerasiforme*.

For a number of reasons the tomato has become a favorite subject for genetic studies. One reason is the great wealth of naturally occurring variability in the species. The supply of variants has been considerably augmented by mutations induced by X rays, ethyl methane sulfonate and other mutagens, notably in work by Hans Stubbe and his colleagues in East Germany. Other features that enhance the usefulness of tomatoes for genetic studies are the plant's high rate of self-pollination, leading to the early expression of recessive mutations; the ease of controlled hybridizations, and the lack of gene duplication in the plant's genetic architecture. Moreover, the plant is easy to grow, has a short life cycle and yields large amounts of seed. Another great asset for these studies is a favorable pachytene stage of meiosis, which makes it possible to identify each of the 12 chromosomes of the plant and their arms. Research to date has produced chromosome maps that rank among the best yet obtained for a flowering plant.

Research on the nature of the genetic variability of *L. esculentum* has facilitated the breeding of improved forms of this species and has thrown additional light on the evolution of cultivated forms. The studies have been based largely on easily detected variations in the form of plants and leaves, in the shape and color of the fruit and on many other morphological characteristics. Recently considerable progress has been made in assaying alleles, or variants, of some 15 genes that code for enzyme synthesis, which is readily detected by starch-gel electrophoresis.

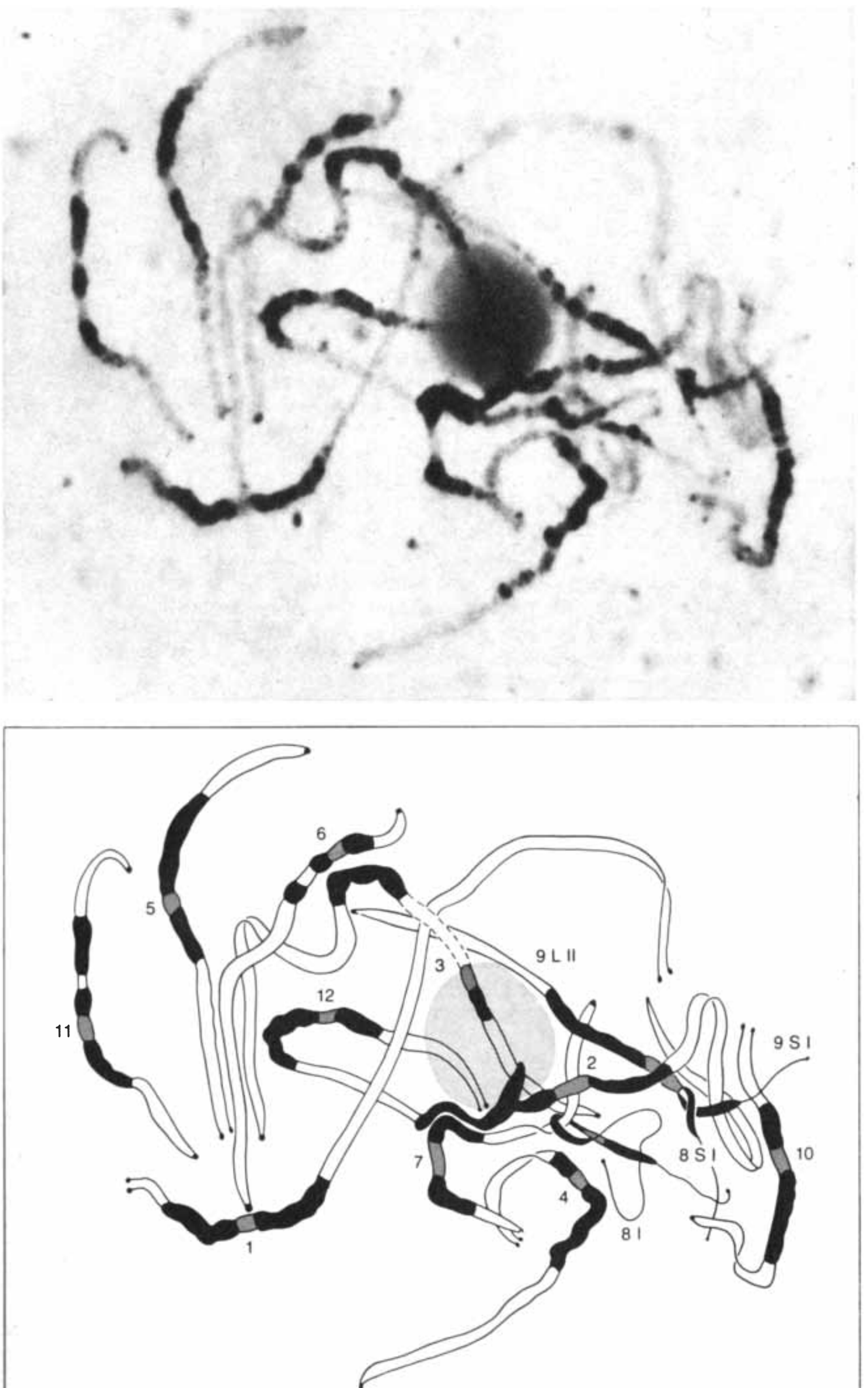
The consensus of such studies is that outside their native region in western South America the cultivated and wild forms of *L. esculentum* are remarkably homogeneous (disregarding the variability resulting from recent hybridization with accessions from the native region). Much more variation is encountered in the native region both within and between accessions. Several reasons might account for these differences in variability: interbreeding with wild species in the native region; more activity by insect pollinators, promoting more outbreeding in the native region; the effects of artificial selection; a longer period of existence in the native region, allowing the accumulation of more variants there, and genetic drift in the long migration from South America through Central America and Mexico, doubtless coupled with further restriction imposed by natural selection for a limited number of successful genotypes.

The genetic drift and selection were accompanied by a sequence of modifi-



LENGTH OF STYLE (*color*) has evolved as is shown here. The normal inclination of the anther tube allows the pollen released within it to drift to the mouth of the tube. In ancestral species (*a*), which could not reproduce by self-pollination, the position of the stigma facilitated cross-pollination by insects. The style is shorter in the *L. esculentum* variety *cerasiforme,* the probable wild predecessor of the cultivated tomato, and in early cultivated varieties (*b*), and is still shorter in most European and North American cultivars (*c*). In new California cultivars (*d*) the retraction of the stigma within the tube increases the opportunity for self-pollination and virtually prohibits "outcrossing."

cations in floral structure [*see bottom illustration on preceding page*]. Like their wild ancestors, the Latin American cultivars generally had well-exposed stigmas that facilitated cross-pollination. Selection for improved fruit-setting presumably resulted in a shortening of the style after tomatoes were introduced into Europe. As the tomato migrated northward, and particularly in places where it was grown in greenhouses, stigmas at the mouth of the anther tube were inadvertently selected to lessen dependency for pollination on insects and wind. The latest step in this sequence was the recent genetic fixation in California cultivars of a much shortened style that places the stigma well within the anther tube, further expediting self-pollination and virtually eliminating the opportunity for outcrossing.



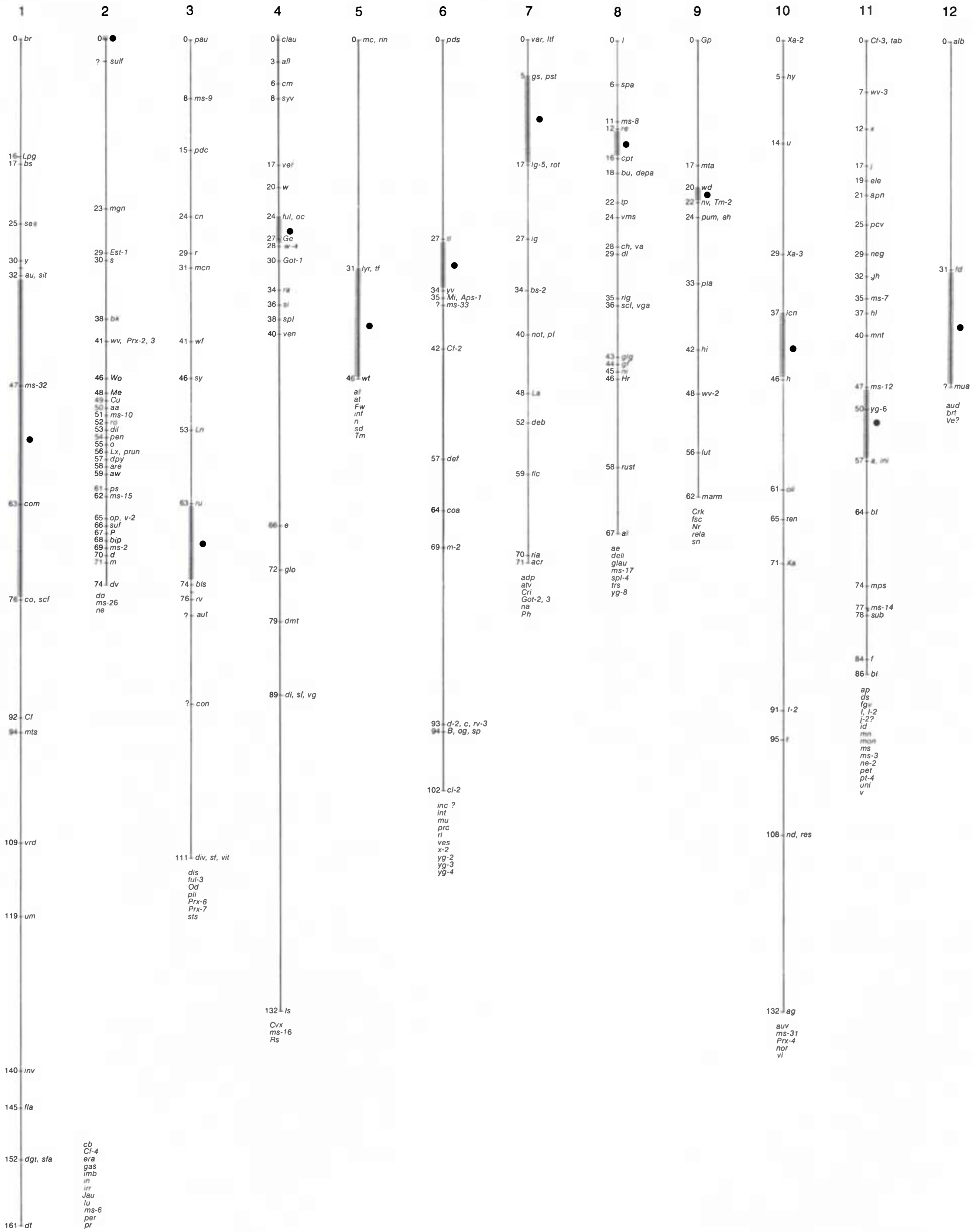
TOMATO CHROMOSOMES are enlarged about 2,000 diameters in a photomicrograph (*top*) and are diagrammed (*bottom*) at the pachytene stage of meiosis, when pollen grains are being formed. Homologous chromosomes are paired, so that the haploid number of 12 bivalent chromosomes is seen. The grayish centromeres are flanked by dark-staining, inactive heterochromatin and lighter-staining active regions (euchromatin). The diffuse body in center is nucleolus. There is a reciprocal translocation between chromosomes No. 8 and No. 9; a long arm of No. 8 and a short arm of No. 9, which were interchanged, are missing and univalent regions (*I*) are not paired. Material was prepared by G. S. Khush of University of California at Davis.

The tomato is a model of spectacular improvement in yield and other characteristics achieved by breeding improved cultivars, coupled with intelligent manipulation of methods of culture. The earliest record of attempts to select better types describes work done in Europe in the middle of the 19th century. Work in the U.S. on stocks imported from Europe followed quickly. The intensity of such efforts increased gradually in the early part of the 20th century and rapidly over the past 40 years. In the latter period the yields per hectare have increased dramatically (by more than four times for cannery tomatoes grown in California, for example). Each year dozens of new cultivars are introduced throughout the world.

Tomato Breeding

The large-scale production of hybrid tomato cultivars is feasible because the flowers are of good size and easily manipulated for hybridization and because the yield of seed per pollinated flower is high. The benefits that have accrued from such crosses include increased vigor, earlier ripening and the rapid development of combinations of desirable traits. F_1 (first filial generation) hybrid cultivars now constitute 95 percent of the plantings for market tomatoes in Japan and 50 percent of those in Israel. They are also a significant factor in eastern Europe and elsewhere. Up to now the hybrid cultivars have not been a major factor in the U.S., partly because the pressure to make maximum use of the available land is less. Hybrids are popular among home gardeners, however, and they dominate the commercial plantings in San Diego County in California. Their popularity in the eastern U.S. and elsewhere is increasing, as is the area devoted to them. The amount of land devoted to them in other regions is increasing.

The substitution of a number of mutant genes in new cultivars has been a feature of the intense breeding activity of the past 40 years. One such gene is *u* (for uniform ripening), which eliminates the dark green "shoulder" of unripe fruit and so prevents the undesired retention of chlorophyll in that part of the ripe fruit. The gene *sp* (for self-pruning) causes the plant to grow in an orderly, compact, determinate fashion, in contrast to the usual sprawling, unlimited, indeterminate growth. The *sp* gene appeared as a spontaneous mutation in Florida in 1914 and has served extensively in tomato-improvement programs. The branches of *sp* plants ter-



MUTANT OR VARIABLE GENES that have been identified with particular chromosomes, and in most cases have been located with respect to one another on an individual chromosome, are displayed on this linkage map of the tomato's 12 chromosomes. The gray region marked with a dot encompasses each chromosome's centromere, the point at which certain chromosome movements are controlled. The numbers are "map units" that show the relative locations of the genes based on frequency of recombination in genetic experiments; genes that have not yet been positioned are listed at the bottom of each chromosome. The genes control a wide variety of traits. At locus 94 on chromosome No. 6, for example, gene *B* codes for orange color; gene *sp* makes a plant "determinate" (*see illustration on next page*).

minate their growth at approximately the same distance from the center of the plant, and the plant flowers more abundantly than the normal (indeterminate) type. As a result the grower has plants in a form that facilitates harvesting, particularly by machine, and the fruiting is concentrated in a shorter season.

Probably the most spectacular modern advance was the development of cultivars adapted for machine harvesting. The work was pioneered in the 1950's and 1960's by Gordie C. Hanna of the University of California at Davis. He succeeded in breeding into a single cultivar (VF 145) smallness and compactness of plant, concentrated ripening and fruits of size and shape such that they withstand machine handling while incorporating the necessary characteristics of disease resistance, yield and quality. VF 145 is still the leading cultivar in California. In terms of tomatoes grown for processing the changes did not degrade the quality.

An indispensable corollary of this achievement was the design of a harvesting machine by Coby Lorenzen at Davis in cooperation with the Blackwelder Company. The imaginative efforts of these workers and of the farmers who rapidly converted to mechanical harvesting were instrumental in preserving and expanding the processing-tomato industry in California, where last year mechanical harvesting was practiced on 113,000 hectares (279,223 acres).

Breeding programs now under way in the industry have the goal of improving yield, color and soluble solids of the fruit, machinability and other characteristics. Cultivars with the "square round" shape that is represented among the tomatoes depicted on the cover of this issue are increasing in popularity because the fruit withstands both mechanical harvesting and decay on the plant before harvesting.

Production Practices

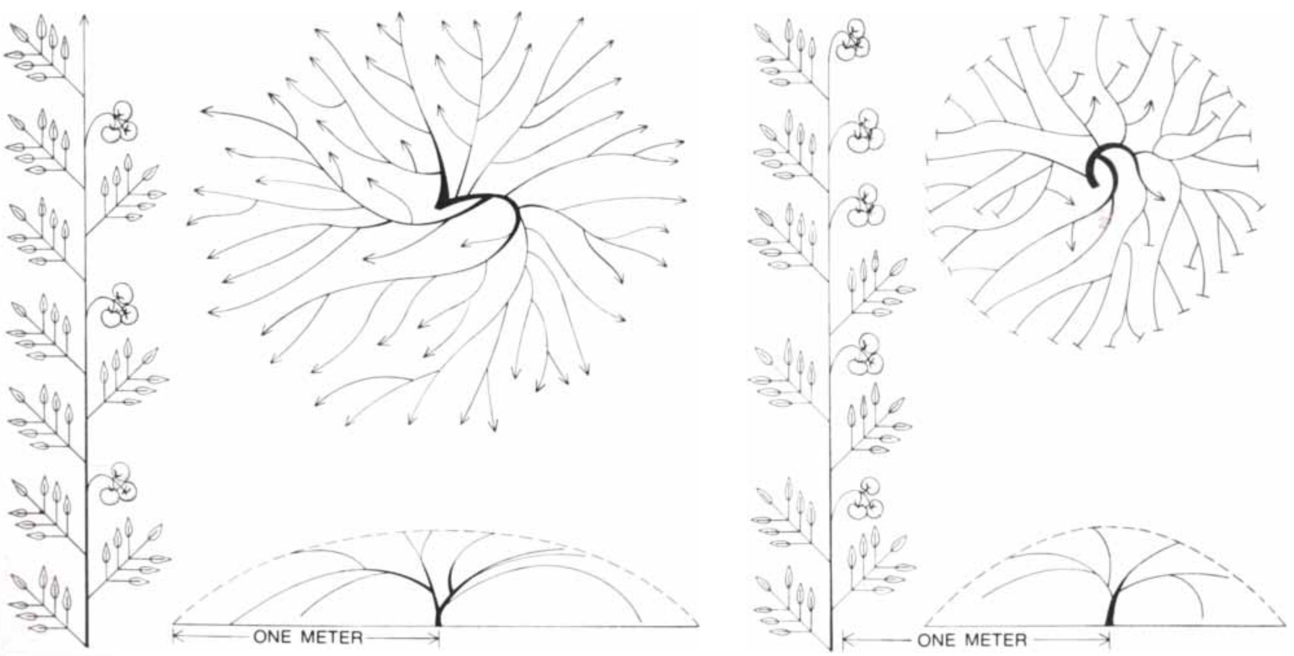
Part of the tomato's great popularity with home gardeners is attributable to the fact that it is so easy to cultivate. It will thrive in an impressively wide range of latitudes, soil types, temperatures and methods of cultivation. Nevertheless, it does have limits beyond which it does not do well. Common problems are lack of sufficient drainage, inadequate light, extremes of temperature and excessive fertilization with nitrogen when conditions otherwise are suboptimal.

The necessity for drainage probably reflects the tomato's adaptation to the desertlike climate of its native region. Good fruit-setting generally requires half a day of sun per day. Poor fruiting can also result from overfertilization under suboptimal conditions. As for temperature, the tomato plant cannot endure either frost or prolonged chilling at temperatures below 10 degrees Celsius (50 degrees Fahrenheit).

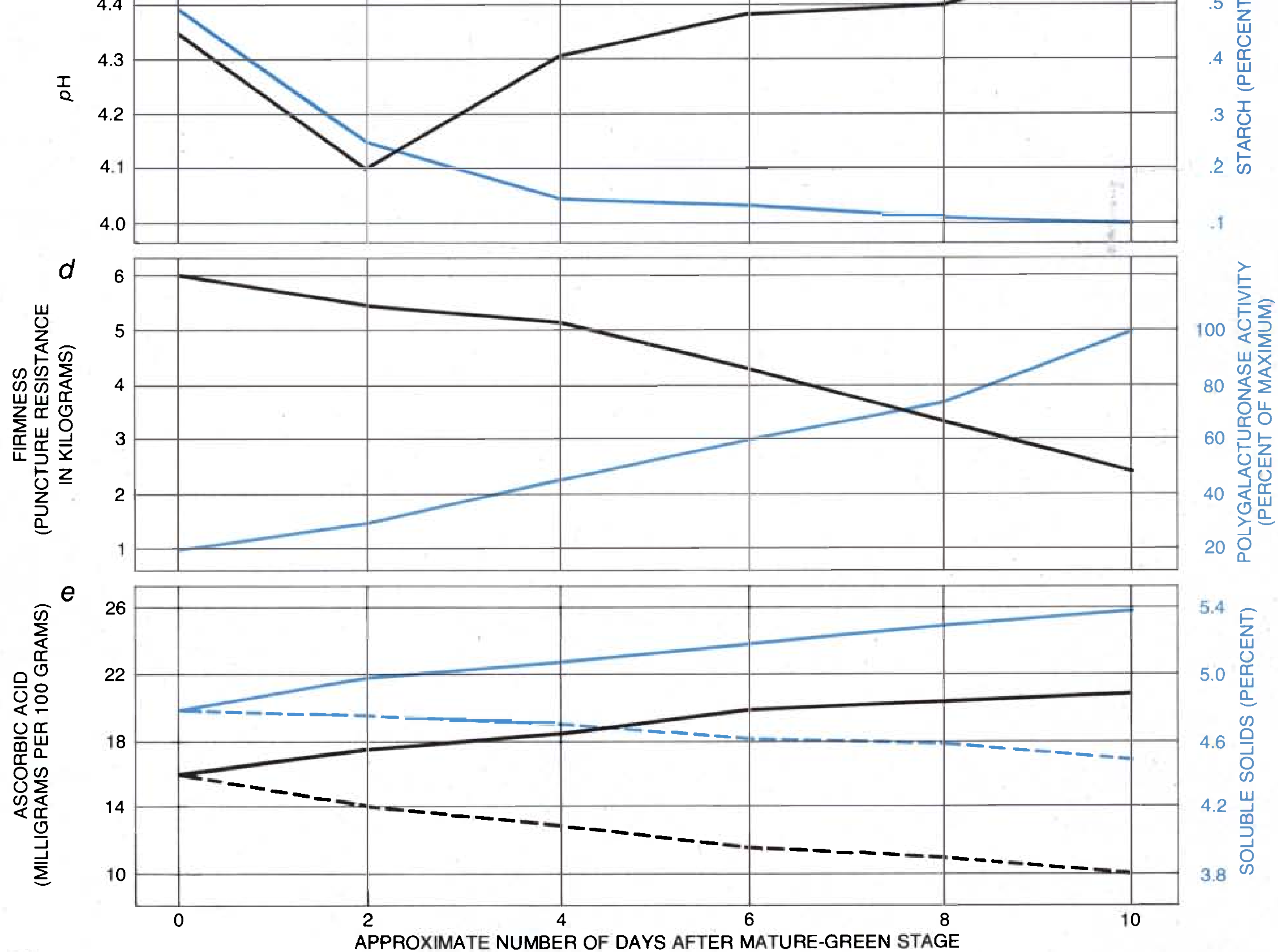
The traditional method of starting tomatoes is to plant seedlings that are one or two months old. The procedure is still widely practiced and is essential where the growing season is short. A recent trend is to drill the seed directly into the soil, as is done with a grain crop. This procedure requires special care in the preparation of the soil and in weed control after the seedlings emerge, but it is more economical and it enables the plant to develop a root system that is not radically disturbed by transplanting. As one might expect, direct seeding has become an integral part of the production of mechanically harvested tomatoes.

The developmental plasticity of the tomato plant makes other methods of propagation possible. For example, tomatoes can be (but seldom are) started from rooted cuttings. Sections of stem with at least one leaf will quickly develop roots when they are cultured in moist sand. Branches will emerge from the stem at the base of the leaf. The method will not work with determinate (*sp*) cultivars because the limitations on further growth are predetermined in each section of stem. Another problem is the difficulty of preventing the spread of certain plant-virus diseases, which can be transmitted by almost anything that comes in contact with the plant.

Since it is easy to graft tomato plants, the method is sometimes employed to combat soilborne diseases that are otherwise difficult to control. Species of nightshade (*Solanum*) are sometimes used as rootstocks because of their resistance to soilborne wilt diseases. An interspecies hybrid (*Lycopersicon esculentum* crossed with *L. hirsutum*) served as a rootstock to resist the corky root disease in Dutch greenhouses. In another situation jimson weed (*Datura stramonium*) was tried as a rootstock to com-



GROWTH CHARACTERISTICS of the normal indeterminate tomato (*left*) and of a determinate plant whose habit is encoded by the recessive gene *sp* (*right*) are compared. The stems differ: instead of the normal three leaf nodes and single inflorescence per segment of main stem, the determinate stem develops progressively fewer leaf nodes per segment and terminates with an inflorescence. The mature plants, shown diagrammatically in top and side views, have different forms: the normal plant spreads indefinitely, whereas the *sp* plant covers less area, being limited to a symmetrical circular shape by the termination of its branches at a fixed distance from the center.


pH
4.4
4.3
4.2
4.1
4.0
STARCH (PERCENT
.5
.4
.3
.2
.1
d
FIRMNESS (PUNCTURE RESISTANCE IN KILOGRAMS)
6
5
4
3
2
1
POLYGALACTURONASE ACTIVITY (PERCENT OF MAXIMUM)
100
80
60
40
20
e
ASCORBIC ACID (MILLIGRAMS PER 100 GRAMS)
26
22
18
14
10
SOLUBLE SOLIDS (PERCENT)
5.4
5.0
4.6
4.2
3.8
0
2
4
6
8
10
APPROXIMATE NUMBER OF DAYS AFTER MATURE-GREEN STAGE

bat infestations of the nematode worms that cause root knot. The procedure was abandoned because toxic alkaloids generated by the root system of the weed were turning up in the tomato shoots. Grafting requires much labor and also has the same limitations as propagation by cuttings.

The tomato plant's trait of sending down deep roots can affect the methods of cultivation. Studies have shown that the roots can go down three meters or more in deep alluvial soils. Hence in the agricultural valleys of California and other areas with deep soil and a desert climate the tomato can be grown by irrigating at intervals of two weeks or longer. All roots absorb water whereas minerals are taken up mainly by the roots in the upper layers of the soil.

A common practice is to support the plant with a stake or a trellis as it gains size. The system is valuable because it prevents the blemishing and the attacks by fruit-decay organisms that can result if the fruit touches moist soil. It is also practiced when the grower wants to attain maximum utilization of land, as in greenhouses. The rising cost of labor and materials has limited trellising in the commercial production of tomatoes and has eliminated commercial production in areas where tomatoes cannot be grown without trellising.

Fruit-setting

The failure of the tomato plant to set fruit is the commonest problem for home gardeners and for commercial growers in areas of marginal production. If the internal or external environment is not favorable for setting, the plant sheds its flowers after they open. A common reason is the failure of pollination, which happens frequently in particular environments: in greenhouses, for example, where conditions are otherwise favorable. This difficulty is usually corrected by devices that vibrate the flowers, stimulating the release of pollen. Another promising approach is to select genotypes with a high capacity for self-pollination under the particular environmental conditions.

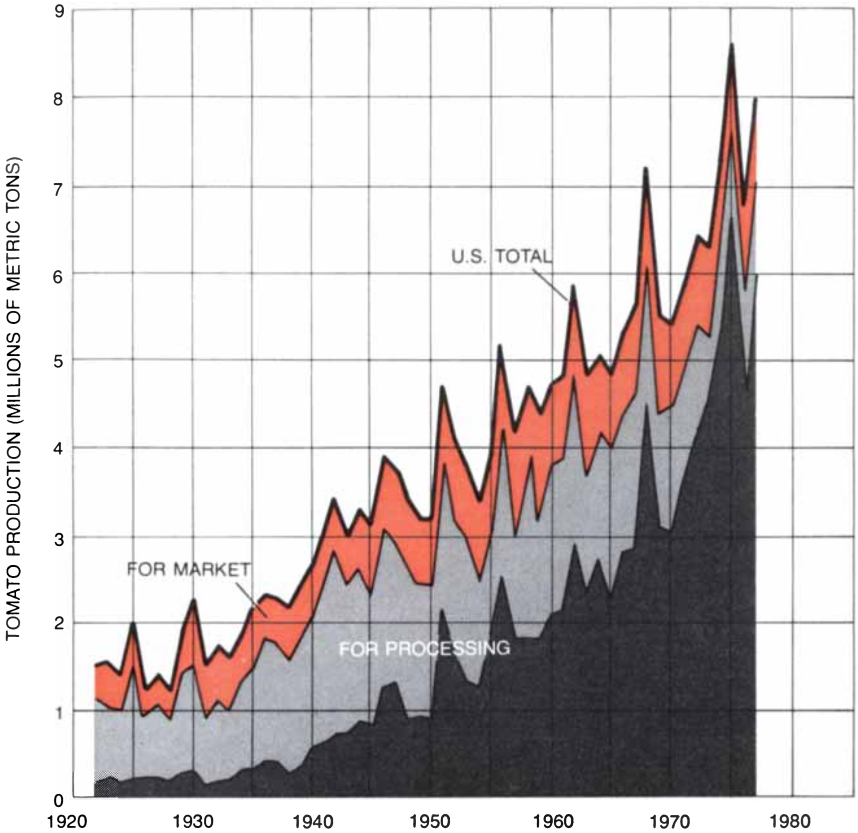
If low temperature is the problem, it can sometimes be dealt with by spraying the flowers with naphthalene acetic acid or other substances that promote growth. Fruits thus set, however, have few seeds or none; as a result they also tend to have inferior shape, firmness and interior color. If high temperature is the problem, the solution sometimes is to plant breeding lines that set well under those conditions. (Several investigations have revealed that plants with an inherited ability to set fruit at high temperatures also set fruit well at low temperatures.)

Other conditions can cause flowers to drop off even when pollination is adequate. They include insufficient light, overfertilization, a previous heavy setting of fruit and excessively high temperature during the day. If the plant is subjected to temperatures that suddenly rise to 42 degrees C. (104 degrees F.) or more, unfruitfulness will ensue and may persist for a week or more. It has been observed that the production of pollen is greatly reduced in extreme heat, but the failure to set fruit cannot be corrected by applying pollen that is known to be viable. Clearly the loss of flowers in high temperatures results from adverse effects on other parts of the plant.

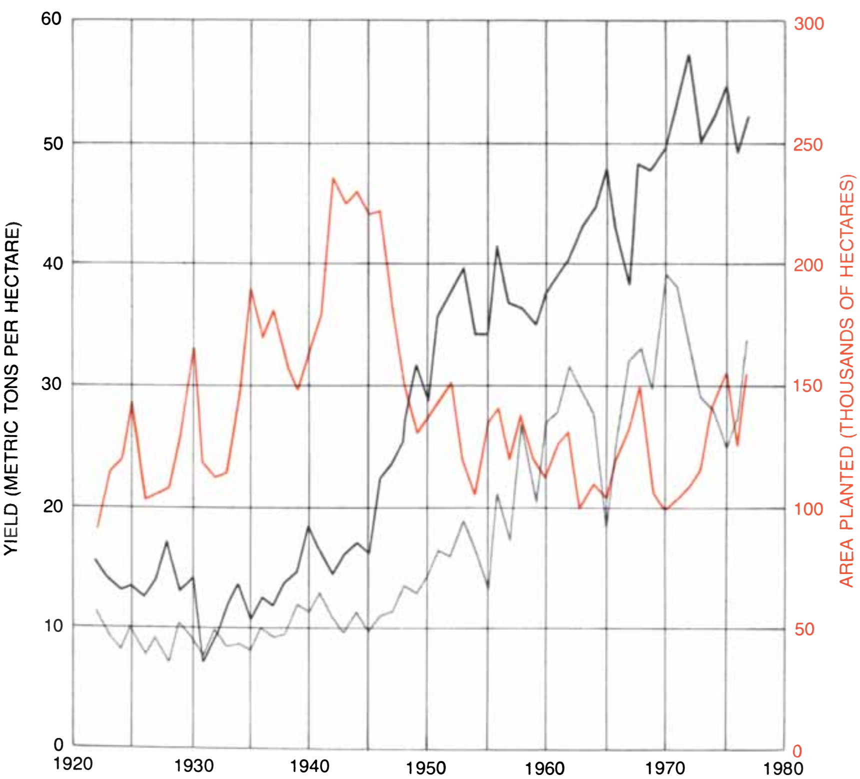
The fruit of the tomato requires from 40 to 60 days from flowering to reach full ripeness. Once set, the fruit develops rapidly to full size in about half of the ripening period. The remaining time is occupied by a series of developmental events that constitute maturation. Starch is accumulated in the fruit from sugars elaborated either in the fruit or in nearby parts. As maturity is approached, chlorophyll is gradually degraded, resulting in a whitening of the fruit, which is a useful harbinger of ripening. During the following week the full red color develops. During this transition the remaining chlorophyll is destroyed, carotenoids (notably the red lycopene) are synthesized, acidity decreases, starch is converted into sugars, essential oils and other components of flavor are elaborated and the tissues become softer. During this time also the rate of respiration suddenly rises and then declines. Ethylene generated by the fruit induces a sudden burst of respiration and is responsible for initiating the entire sequence of events.

The most rapid development of seed takes place during the second half of the maturation period. The embryo reaches full size within its matrix of endosperm, and the seed coat develops, hardens and acquires its characteristic covering of pseudohairs, which are the remains of outer cell walls that had been suberized: converted into a corky condition by the

RIPENING PROCESS is depicted on the opposite page. The six stages are each defined by a range of color standards published for the Department of Agriculture by the John Henry Company; the colors shown here are representative of each range. As the chlorophyll content (*green*) declines lycopene (*red*) and beta-carotene (*orange*) increase (*a*). Ethylene (which triggers the entire process) is generated by the fruit, and respiration (measured by carbon dioxide output) reaches a peak (*b*). Acidity increases briefly, then decreases; starch content decreases (*c*). An increase in the pectin-dissolving enzyme polygalacturonase softens the fruit (*d*). In vine-ripened tomatoes (*solid lines*) the content of ascorbic acid (vitamin C) and soluble solids (primarily sugars) increases; in fruits harvested at mature-green stage and ripened in storage (*broken lines*) vitamin C and sugar content decrease (*e*). Data were supplied by Adel Kader of the University of California at Davis.



U.S. TOMATO PRODUCTION has risen steadily (with short-term fluctuations) since 1922. Most of the increase has been in the crop destined for processing: canning, sauce, juice, catsup and so on (*light and dark gray*). California has accounted for an increasing share of that crop (*dark gray*). The U.S. commercial fresh-market crop has remained about the same (*color*).



INCREASED PRODUCTION of processing tomatoes is the result of improved yields. The amount of land planted to processing tomatoes in the U.S. has decreased since 1940's (*colored curve*). The yield has increased sharply in California (*black curve*), less in other states (*gray*).

infiltration of the substance suberin. Each seed is enveloped in a mucilaginous sheath. Sprouting of the seeds within the ripe fruit, for which the temperature and moisture conditions are ideal, is inhibited by abscisic acid and unidentified substances that are normally present. When the seeds are to be extracted for planting, the inhibitory substances and the gelatinous sheath are eliminated by a process of retting, or fermentation. In commercial extraction the fruits are shredded, the larger fragments are screened out and the remaining fluid mass is fermented in vats. The seeds then sink to the bottom of the retting liquid, and the residue is removed by flotation and screening.

The Tomato Harvest

Ripe tomatoes are perishable, and yet they should be picked at this stage for optimum quality. For long-distance shipment they must be picked at the mature-green stage or the pink stage. They will then ripen unassisted, or the ripening can be speeded up by treating them with ethylene gas, the fruit's natural inducer of maturation. The same result can be achieved by spraying unripe fruit (on the plant or in storage) with ethephon, a compound that generates ethylene. Fruit that is harvested at the mature-green stage tends to lose sugars and ascorbic acid, which is part of the reason it falls short of the vine-ripened tomato in quality.

Around the world most tomatoes are still harvested by hand, but the trend toward mechanized harvesting is clear. The most sophisticated machines (notably in California, where the entire cannery crop is harvested mechanically) handle the complete operation [see "Mechanical Harvesting," by Clarence F. Kelly; SCIENTIFIC AMERICAN, August, 1967]. The machine propels itself, cuts the plant with a blade that travels slightly below ground level, lifts and shakes the plant into an inclined ascending system of eccentrically moving horizontal rods, drops the plants and debris to the ground, diverts the fruit to moving canvas belts on both sides of the machine, where unripe or damaged tomatoes and clods of soil are sorted out, and delivers the sorted fruit to conveyors. The conveyors are fiberglass gondolas (capacity from nine to 11 metric tons) or plywood bins (capacity from 360 to 450 kilograms) that are loaded on trailers.

In temperate climates harvesting continues until cold weather stops it. At the approach of chilling or freezing temperatures the grower can pick tomatoes that are green-white or pink and slowly ripen them at temperatures of from 12 to 20 degrees C. (53.6 to 68 degrees F.), thereby extending the supply beyond the normal harvest period. Sometimes whole plants are cut off and hung in storage; the tomatoes will ripen even without light. Some of them will spoil, however, and they must be culled frequently to prevent the spread of organisms that cause decay.

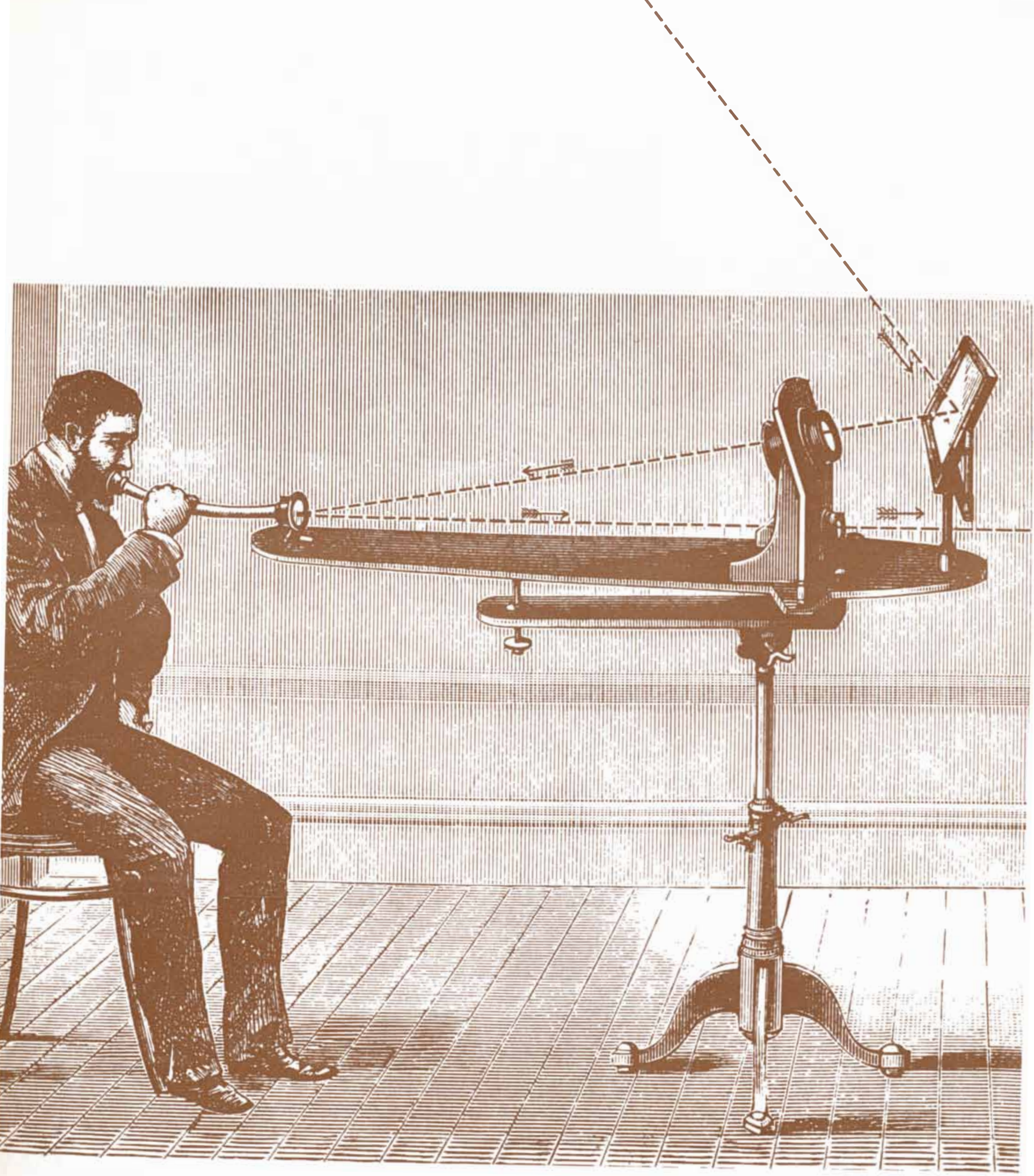
Both the commercial shipper and the home gardener must be careful to avoid storing tomatoes at low temperatures. Prolonged exposure to temperatures between zero and 12 degrees C. (32 and 53.6 degrees F.) causes a slow but progressive breakdown of the cells. The damage appears in the form of sunken, necrotic spots that provide an avenue for various organisms that cause rot. (Flavor deteriorates before rot appears, so that even brief refrigeration can reduce fruit quality.) Damage by freezing or heavy frost can be recognized by the dark, water-soaked appearance of the affected tomatoes, which become soft and deteriorate rapidly.

Tomato plants are often so drained of food reserves by heavy fruit production that little is left of them at the end of the season. The disposal of plant remains is therefore simpler than it is for most other crops. The refuse is minimal and tends to disintegrate rapidly. Its high protein content, however, has stimulated research on the possibility of utilizing the remains as a livestock feed.

During the large rise in the production of tomatoes over the past 50 years California has become dominant. Last year the state accounted for 86 percent of the processing tomatoes and 76 percent of all tomatoes grown commercially in the U.S. Indeed, the output from California represented a substantial portion of the world supply of tomatoes.

As I have indicated, most of the increase has been in tomatoes grown for processing. Tomatoes for sale fresh have had a steady demand, but their gain in total production has been much smaller. An increasing share of this market has been supplied by tomatoes imported from the northwestern coast of Mexico in the off season. One should not overlook the large number of tomatoes produced in home gardens, although reliable statistics on this crop are not available. Recent surveys indicate that the tomato is the most popular vegetable in home plantings.

Perhaps the most remarkable thing about the great rise in the commercial production of tomatoes is that it has been accompanied by a decline in the amount of land planted to tomatoes. The reason is the improvement in yields per hectare. Yields 40 or 50 years ago averaged about 13.5 metric tons per hectare in California and about 10 metric tons in other states; now they are respectively 51.5 and 33.5 tons. This example of improved productivity is equaled in only a few other major crops grown in the U.S. It is a tribute to the research that has led to the introduction of better cultivars and to the investigators and growers whose ingenuity has improved methods of production.



In Bell's "Photophone," sunlight was bounced from a reflector through a lens to a mechanism that vibrated in response to speech. This caused the light beam to vary in intensity. At the receiving end, a selenium detector translated these variations into electrical current to recreate speech through a telephone receiver.

Negative Absolute Temperatures

No substance can be cooled below absolute zero, but some physical systems can have a negative temperature on an absolute scale. Such temperatures are not colder than zero but are hotter than infinity

by Warren G. Proctor

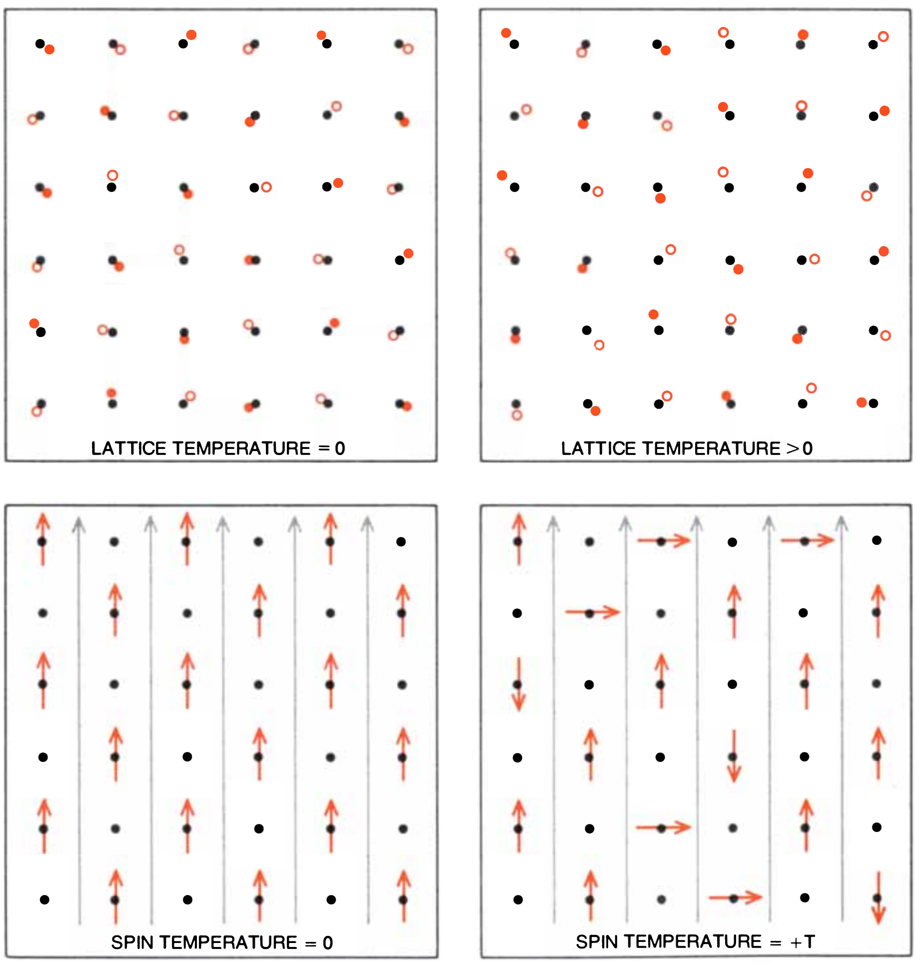
Temperature is a property of matter that seems to have a well-defined range of possible values. If all the heat could be extracted from a body, it would be assigned a temperature of absolute zero, or zero degrees Kelvin, and it could not become any colder. At the other extreme, heat can always be added to a body, at least in principle, and so its temperature can increase without limit. All absolute temperatures therefore seem to fall in the range between zero and positive infinity. (Negative temperatures on the Celsius and Fahrenheit scales arise only because the zero points of those scales do not correspond to absolute zero.)

Temperatures of the kind that are measured with an ordinary thermometer are indeed confined to the range of positive numbers. There are a few physical systems, however, whose temperature when measured on an absolute scale can assume negative values. What may seem even more peculiar, such temperatures are not colder than absolute zero; a system with a negative temperature can give heat to one at an infinite temperature and is therefore hotter. The negative temperatures are not "below zero"; they are "above infinity."

In order to understand how an absolute temperature can be negative, it is first necessary to recognize that a single body can simultaneously include systems of particles that exhibit more than one temperature. In a crystalline solid, for example, the ordinary temperature—the property measured by a mercury thermometer—reflects the average vibratory motion of the atomic nuclei; that temperature cannot be negative. The same nuclei, however, have freedom of orientation as well as freedom of motion, and the set of all the orientations makes up an independent system of particles that can also have a temperature. It is temperatures of this kind that can take on negative values. Because the energy in question is small, a crystal in which the nuclear orientations are at a negative temperature feels neither hot nor cold to the touch. Negative absolute temperatures are nonetheless real; they can readily be measured and can even be manipulated by the methods of calorimetry.

The concept of temperature describes a relation between two quantities, energy and disorder. In fundamental terms what is measured by temperature is the change in the disorder of a system as the energy of the system changes.

In a crystalline solid the major contribution to disorder is the vibratory motion of the atomic nuclei. A crystal of table salt, for example, is often repre-



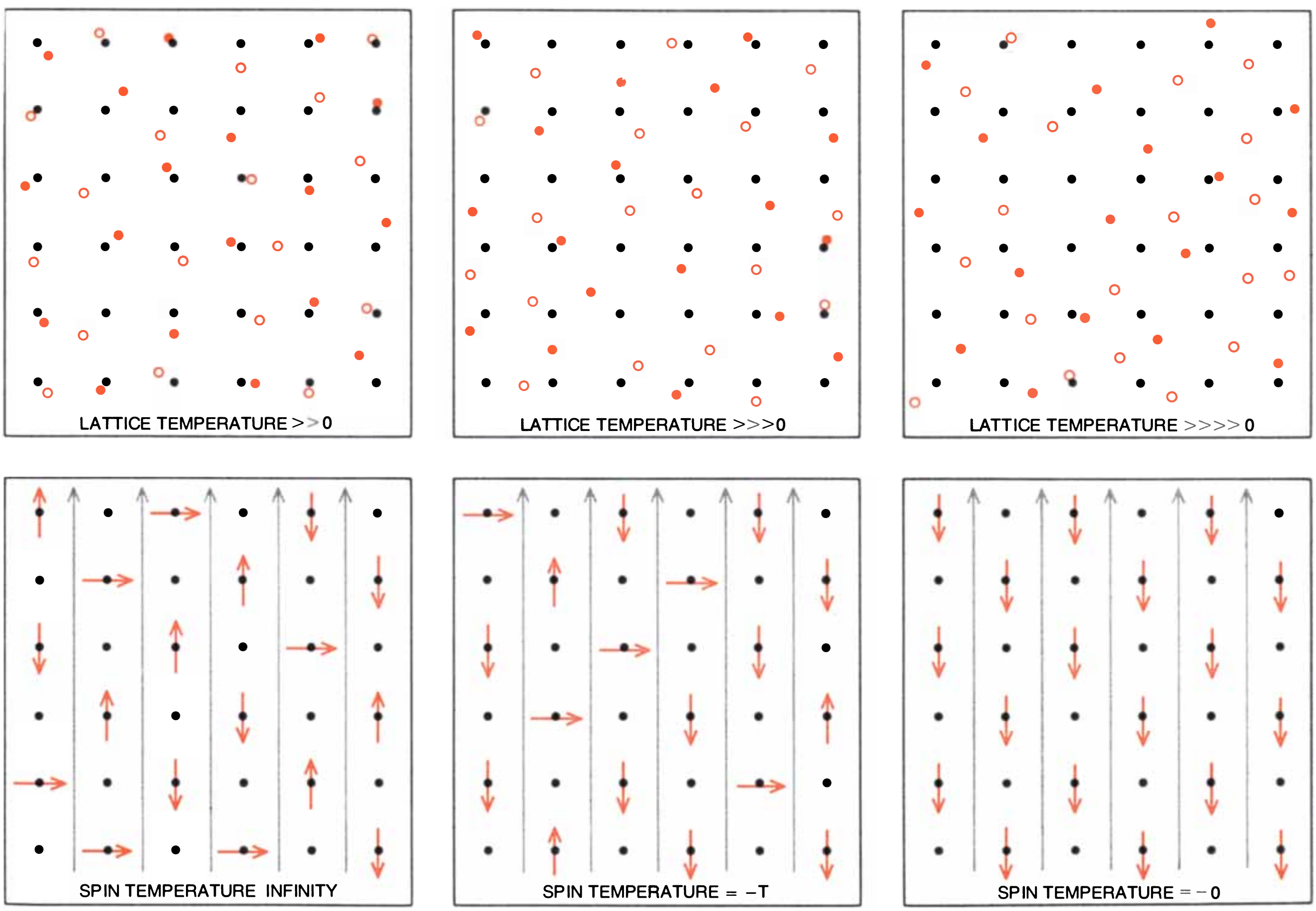
CONCEPT OF TEMPERATURE can be applied to any physical system that has a definite energy and an entropy, or measurable disorder. In a solid the lattice, or ordinary, temperature (*top*) reflects the vibratory motion of the nuclei (*colored dots*) around their average positions (*black dots*). Temperature increases to the right, and so does the average energy of the nuclei. Moreover, every increase in energy is accompanied by an increase in entropy. (The displacements are greatly exaggerated here.) Another system with a defined temperature (*bottom*) is

sented as a perfect three-dimensional lattice in which the nucleus of each sodium and chlorine atom occupies a corner of a cube. Actually the ideal lattice represents only the average or equilibrium positions of the nuclei, around which they vibrate randomly. The amplitude and the frequency of the vibrations vary from moment to moment and from atom to atom; the temperature of the crystal summarizes in a single number the average amplitude and the distribution of frequencies. This number is the ordinary temperature, or what I shall designate the lattice temperature.

As a crystal is cooled the average amplitude and frequency of the nuclear vibrations diminish. Before the development of quantum mechanics it was thought that the vibrations would stop entirely if the crystal were cooled to absolute zero, but in the quantum-mechanical interpretation the vibrations cannot cease: residual vibrations, called zero-point motions, persist even when the temperature reaches absolute zero. Because the zero-point motions cannot be abolished, the residual disorder of the system cannot be eliminated, no more energy can be extracted and the crystal cannot be further cooled, even in principle.

Moving in the other direction on the temperature scale, as heat is added to a solid the average amplitude and frequency of the vibrations increase, but again the motions of the atoms are constrained by the rules of quantum mechanics. Those rules specify that only certain modes of vibration are possible, each with a definite energy, and intermediate energies are forbidden. At absolute zero all the nuclei are in the lowest-energy vibrational state, the state that specifies the zero-point motions. As the crystal absorbs heat some of the atoms jump to higher energy levels, each of which describes a different vibrational mode. As the temperature is raised further, progressively higher levels are populated, but many atoms still remain in the lower levels, so that the population distribution has a pyramidal form. There is no intrinsic limit on the energy that can be acquired by a nucleus, and the number of energy levels available is infinite. It follows that temperature can increase without limit. (Of course any real substance heated indefinitely must melt and then boil, but these complications do not alter the fact that the number of possible energy levels has no upper bound.)

The disorder introduced into a crystal lattice by heating can be described quantitatively in terms of entropy. For the purposes of this article entropy can be defined as a measure of the difficulty of guessing the energy level in which a particular nucleus lies. At absolute zero all the nuclei are in the same energy level (the lowest one possible). Hence there is no uncertainty about their state, and the entropy of the system is at a minimum. As the temperature increases and more levels become available the probability that an atom occupies any given level becomes smaller, and so the entro-



made up of the spins (*colored arrows*) of the nuclei in a crystal. In a magnetic field the spins have only a few possible orientations, in this case three: parallel to the field, antiparallel to it or perpendicular to it. At a spin temperature of absolute zero (*far left*) all the spins line up parallel to the field in a state of minimum energy and entropy. As the spin temperature rises a few spins flip to other configurations, and at infinite temperature (*center*) the spins are equally distributed among the possible orientations; in this state entropy is at a maximum. Adding energy forces more of the spins to assume the antiparallel orientation. As a result entropy is reduced; what is more, the spin temperature becomes negative. When all spins are antiparallel (*far right*), entropy is minimal and spin temperature has reached "minus zero."

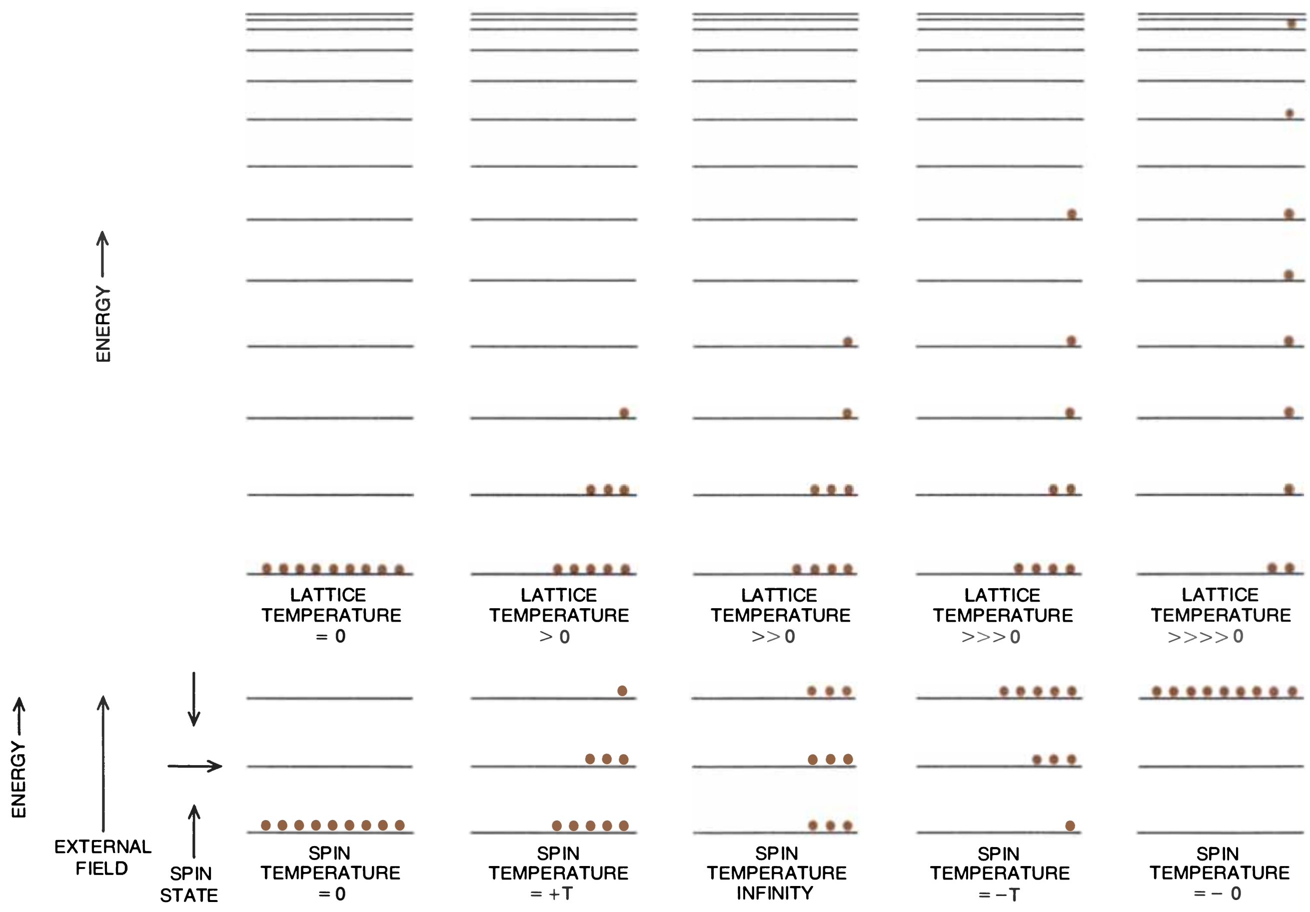
py increases. At infinite temperature any atom would have an equal probability of occupying any one of an infinite spectrum of levels, and so the entropy would also be infinite. It is important to note that as energy is added to the system both the entropy and the temperature increase continuously.

Like the vibrations of the nuclei in a crystal, the orientations of the nuclei are governed by the rules of quantum mechanics. In order for an atomic nucleus to have a defined orientation it must first have some detectable asymmetry: it must "point" in some direction. The orientation is defined by the spinning of the nucleus, which can be imagined as being much like the rotation of the earth around its axis. The presence of the axis designates a direction in space along which the nucleus is aligned. Indeed, each spinning nucleus can be associated with an imaginary arrow (the spin vector) that indicates its orientation. The arrow is drawn parallel to the spin axis, and by convention it points in a direction such that an observer looking along the arrow sees the nucleus spinning clockwise.

Not all nuclei spin, and those that do not spin have no definable orientation. Nuclei that do spin, however, are allowed to take on only certain well-defined quantities of angular momentum. When the angular momentum is measured in fundamental units, it can have values of 0, 1/2, 1, 3/2 and so on; intermediate values are forbidden. For the sake of simplicity I shall assume here that all nuclei have the same spin, equal to 1 unit. This assumption is actually false (sodium and chlorine nuclei, for example, each have a spin of 3/2 units), but it makes the geometry of the nuclear spins much easier to visualize and it introduces no serious errors.

The direction of the nuclear spin vector can be detected experimentally only because it corresponds to the direction of the nuclear magnetic field. All nuclei, of course, bear an electric charge (equal to the number of protons in the nucleus); if the nucleus is spinning, the charge is in motion and it generates a magnetic field parallel to the spin vector.

The magnetic field of a nucleus is detected through the interaction of the nuclear field with an artificially imposed external field. Under the rules of quantum mechanics a spinning nucleus can have only a finite number of possible orientations in an external field, the number being proportional to the spin angular momentum of the nucleus. A nucleus with a spin of 1/2 unit has only



ENERGY LEVELS in a solid correspond to the possible states of vibratory motion (*top*) or to spin orientations (*bottom*). Here the levels are represented as horizontal lines and their energy is proportional to their height above the lowest level; colored dots mark the number of nuclei that occupy each level. For either system absolute zero (*far left*) is a state of minimum energy, since all the nuclei are in the lowest possible energy level; it is also a state of minimum entropy. The entropy can be regarded as a measure of the difficulty of guessing what level a nucleus is in at a given moment; the guessing is obviously easiest when all the nuclei are in the same level. In the vibrational system adding energy forces some nuclei to occupy higher levels, but larger numbers of nuclei remain in the lower levels. As more levels are occupied the entropy increases, since it is more difficult to decide in what level a given nucleus lies. The number of levels available is potentially infinite, and at infinite temperature (*far right*) the nuclei are randomly distributed among them, that is, each level has an equal probability of being occupied. In the spin system there are only as many energy levels as there are possible orientations of the nuclei, in this instance three. As in the vibrational system, adding heat promotes a few nuclei to higher levels, and at infinite temperature (*center*) the nuclei are evenly distributed among the available levels. The entropy is then at its maximum. Because the number of levels in the spin system is finite, however, states of even higher energy are possible; most or all of the nuclei can be driven to the highest, antiparallel energy level. Because the entropy declines as the energy increases in this region (the opposite of the usual relation), the temperature is assigned a negative value. When all the nuclei are in the highest level (*far right*), the entropy is again at a minimum and the temperature is minus zero.

U.S. NAMES RABBIT DIESEL #1 IN ECONOMY.

We always knew we made terrific economy cars.

Now it's official.

Because according to the figures published in the official 1978 EPA Mileage Guide, a Rabbit Diesel gets the highest mileage of any car in America: an incredible 53 MPG on the highway, and 40 MPG in the city.

(Of course, these estimates may vary depending on how and where you drive, optional equipment and your car's condition. That's also official.)

We could go on and on about all the innovations we've built into our newest Rabbit.

Like the fact that it goes on and on (amazingly!) without ever needing a major tune-up. (There are simply no spark plugs, points, condensers, or carburetors to tune.)

Our mechanical wizardry doesn't stop with the engine either.

There's front-wheel drive for better tracking.

There's more room for people than 37 other cars you could buy.

And a Rabbit Diesel goes like a bat out of you-know-where. In fact, it set 31 world records for diesels on a track in Miramas, France.

Come test it for yourself.

For years we've built cars that use very little gas.

Now we've built one that doesn't use any at all.

VOLKSWAGEN DOES IT AGAIN

The cream. The crop.

Olympus introduced the OM-1 and startled the world of photography with the creation of the compact SLR. Today, the OM System is still the cream of the crop.

Because while others have emulated our compact design, OM cameras continue to offer features others can't.

The OM-1 Becomes #1.

Enter the OM-1. Suddenly, the SLR camera is 33% smaller and lighter, yet incredibly rugged to meet the demands of professional wear and tear. Miraculously, the viewfinder is 70% brighter and 30% larger for faster, easier composing and focusing.

And suddenly, the OM-1 became the #1 selling compact SLR. Its metering system is designed to give complete control to professionals and photojournalists. No distractions, blinking lights, or obscured images in the viewfinder.

A Quiet Revolution.

Olympus created a unique shock absorber and air damper system to eliminate noise and vibration, for sharper, unobtrusive photography. Especially vital for long tele shots and macro/micro photos.

The Biggest Smallest System.

More than 280 components, all compact design, include 13 interchangeable screens so you can meet any photographic challenge. Ingeniously designed to change in seconds through the lens mount. And more compact lenses than any other system, each a marvel of optical design and performance.

Olympus "Unlocks" Motor Drive.

OM-1 is still unsurpassed in its continuous-view motor drive capability: 5 frames per second. And a Rapid Winder that fires as fast as 3 shots a second! With no mirror "lock-up," regardless of lens used.

Enter The OM-2. Automatically.

It's the fully automatic OM, with major differences from *all* other automatics! The only SLR with "off-the-film" light measurement for those photographers demanding the ultimate innovation in automatic exposure control. Which means each frame in motor drive or rapid winder sequences is individually exposure-controlled. And it makes possible the unique Olympus 310 Flash whose exposure duration is controlled by the camera's metering system.

And of course, the OM-2 shares every other innovation and system component with the OM-1.

We Wrote The Book On Compact SLR's.

Write for our full color brochure: OLYMPUS, Woodbury, New York 11797. Read it all. Discuss the advantages of an Olympus with your photographer friends.

Visit your camera store. Compare. You'll discover that Olympus is not only the cream of the crop. It's the crème de la crème!

OLYMPUS OM SYSTEM

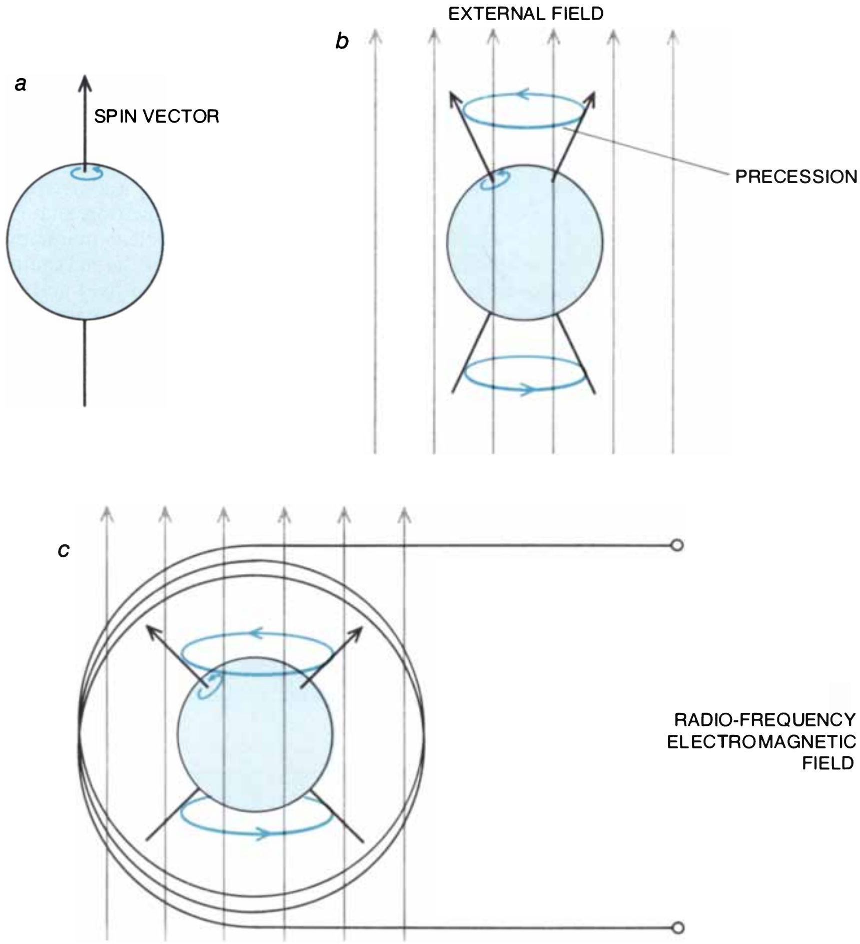
two possible orientations: the spin vector can lie parallel to the field or it can be antiparallel, opposing the field. No other positions are allowed. The nuclei I shall be considering, with one unit of angular momentum, have three possible orientations: the spin vector can be parallel to the field, antiparallel to it or perpendicular to it.

All the nuclei of the same type in a crystal, or even in a solid made up of many crystals, can be considered as a spin system. Thus a crystal of table salt includes two spin systems, one composed of all the sodium nuclei and the other of all the chlorine nuclei. When the crystal is permeated by an external magnetic field, each of the systems has a measurable energy and a definite entropy. As will be seen, each of the spin systems can also be assigned a temperature, which I shall distinguish from the lattice temperature by calling it the spin temperature.

Just as a compass needle tends to align itself with the earth's magnetic field and work must be done to turn it from that orientation, nuclear magnets tend to line up parallel to an imposed field and energy must be supplied if they are to assume any other orientation. Hence the three possible orientations of a spin-1 nucleus represent not only different geometric configurations of the nuclear spins but also different energy levels. The difference in energy between the levels is proportional to the strength of the external field. A nucleus with its spin vector parallel to the applied field is in the lowest-energy state; a nucleus oriented perpendicular to the field has a slightly higher energy, and one antiparallel to the field has a still higher energy.

The distribution of the nuclei in a spin system among the three levels is similar to the distribution of nuclei among vibrational levels. At a spin temperature of absolute zero all the nuclei are in the parallel configuration; that is the state of the spin system in which both the energy and the entropy are at a minimum. As the spin temperature increases, some nuclei change their orientation and jump to the higher perpendicular and antiparallel energy levels. As long as the temperature is finite, however, more nuclei remain in the bottom level than in either of the higher levels. At infinite spin temperature all three levels are uniformly populated, a condition in which the system has maximum entropy.

In the presence of an external field any change in the orientation of a nucleus must be accompanied by a change in energy. In many cases these energy transactions are strictly internal to the spin system. For example, when one nucleus "flips" from perpendicular to parallel, giving up a small quantum of energy, a nearby nucleus may make the opposite transition, from parallel to perpendicular, absorbing the same quantum. The energy of the spin system as a whole therefore remains unchanged, since there are always the same number of nuclei in each state. Such exchange interactions take place constantly and very quickly, and they distribute the energy of the spin system uniformly throughout the volume of the crystal.



SPIN OF A NUCLEUS defines its orientation. An arrow can be drawn along the axis of spin (*a*); by convention this spin vector points in a direction such that an observer looking along the arrow would see the nucleus spinning clockwise. The spinning of the nucleus generates a magnetic field, which is directed parallel to the spin vector. In an external magnetic field (*b*) the interaction of the two fields causes the nucleus to precess. The orientation of the spin vector can be determined by imposing a high-frequency electromagnetic field transverse to the steady external field. When the frequency of the alternating field matches that of the nuclear precession, the nuclei absorb energy strongly. This technique, called nuclear-magnetic-resonance spectroscopy, can be employed to measure spin temperatures and also to alter the spin directions.

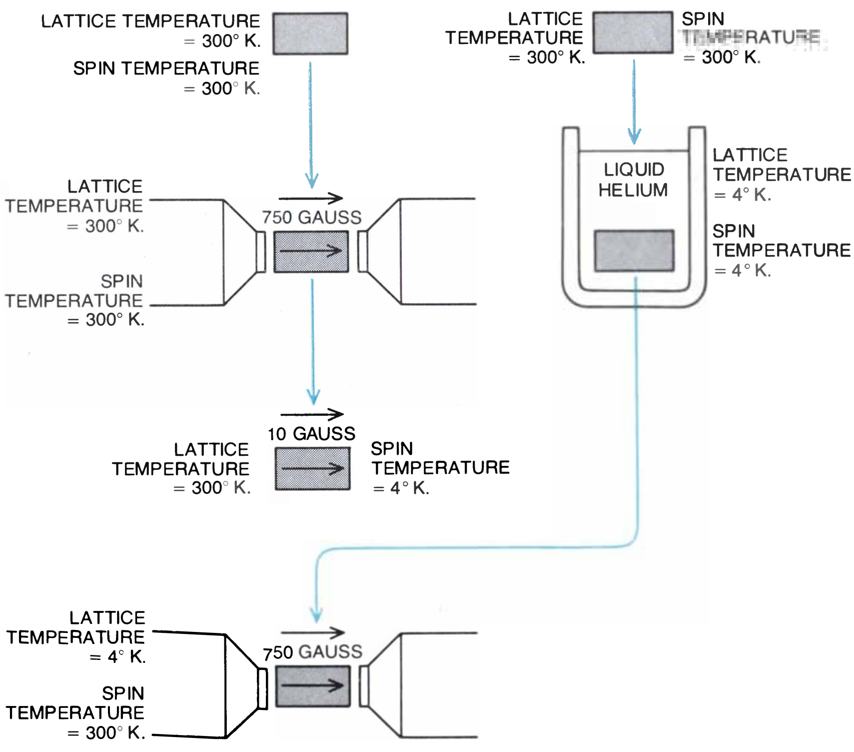
For a nucleus to change its orientation without a compensating transition by a neighboring nucleus, the required quantum of energy must be supplied by or absorbed by some external source. That source is the lattice itself, which can share in the energy available by means of a weak coupling between the energy levels of the spin system and those of the vibrational system. A vibrating nucleus constitutes a moving electric charge, and so it gives rise to a magnetic field that can interact with the spin-generated fields of neighboring nuclei. If the energy of a vibrational mode happens to correspond to the energy absorbed or emitted in a spin transition, then the spin-flip can take place. Such coincidences are rare, but over a period of several minutes they bring the spin system and the lattice vibrations into thermal equilibrium. If the spin temperature is initially the higher one, then most of the spin-flips are to a lower-energy orientation and the spin system gives up energy; it is cooled by contact with the colder lattice. If the spin system is cooler at the outset, then most of the spin transitions represent a gain in energy and the spin system is warmed. In either case the process is called thermal relaxation.

The several minutes required for thermal relaxation in many crystals is a convenience for the experimenter. It is a period brief enough for the spin system to be brought readily into equilibrium with the lattice: the experimenter need only place a crystal between the pole pieces of a magnet and wait for a few minutes; he can then be sure that the spin system and the lattice are at the same temperature. On the other hand, the relaxation time is long enough for measurements to be made on the spin system as if it were isolated from the lattice.

The state of the spin system is measured by the technique of nuclear-magnetic-resonance spectroscopy. In this technique a crystal is exposed to a strong, steady magnetic field, which tends to align the nuclear magnetic fields and also causes the nuclei to precess, or wobble around their spin axes. The precession has a characteristic frequency, which can be determined by imposing a second, oscillating field perpendicular to the first one. This second field is actually the magnetic component of a radio-frequency electromagnetic wave. When the frequency of the applied field is equal to the precession frequency, a "resonance" is observed: the spin system absorbs energy copiously, and the characteristics of the electromagnetic circuit are thereby altered. By measuring the precession frequency in this way several properties of the spin system can be determined. Among the most important of them is the proportion of the nuclei that are aligned parallel to the external field.

At any temperature below infinity there are always a few more nuclei parallel to an external field than there are antiparallel. This observation in effect states that a crystal in a strong magnetic field becomes slightly magnetized; it is this magnetic polarization that is actually measured by nuclear magnetic resonance. The relation of the magnetic polarization to the external field and to the temperature is given approximately by the principle called Curie's law (after Pierre Curie). The law states that the polarization is directly proportional to the strength of the applied field, customarily designated H, and inversely proportional to the temperature, T, of the spin system. Thus the polarization, M, is proportional to H/T. By simply rearranging the terms of this relation it follows that the spin temperature must be proportional to H/M. Both H and M are quantities that can readily be measured.

The relation expressed by Curie's law suggests a simple experiment. Suppose a crystal of salt is placed in a known magnetic field and the spin system is given time to reach thermal equilibrium with the lattice temperature. The magnetic polarization of the crystal is then measured by nuclear-magnetic-resonance spectroscopy. This stable system can subsequently be disturbed by quickly changing the external field ("quickly" meaning in a short time compared with the thermal-relaxation time). Suppose, for example, the field is reduced to half its original strength. Immediately after the external field is reduced the polarization of the spin system remains essentially unchanged; there has been too little time for the nuclear spins to be reoriented. Since the spin temperature measures the nuclear orientations, intuition suggests that the spin temperature too should be unchanged. Curie's law, however, requires that the temperature be proportional to H/M, and since H, the external field, has been reduced by half, the spin temperature must also be half its former value.

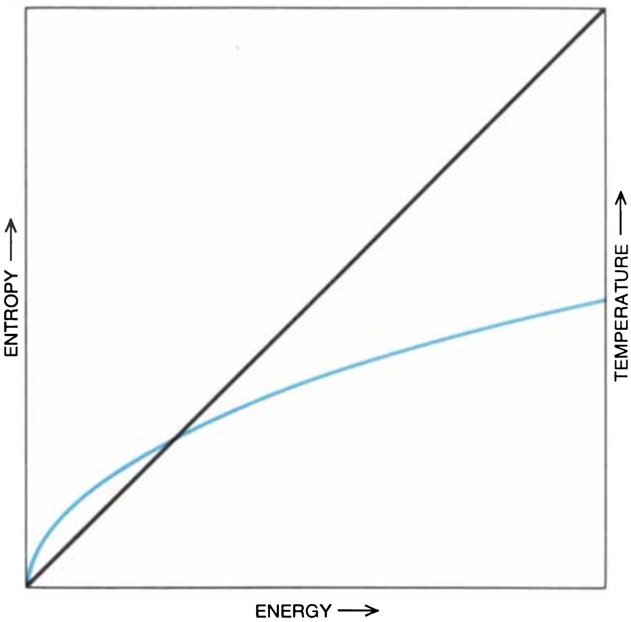


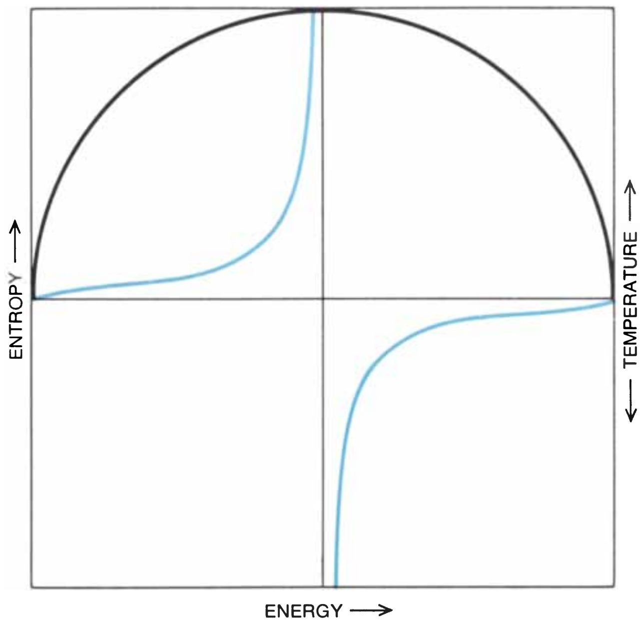
EQUIVALENCE OF SPIN TEMPERATURE and lattice temperature was demonstrated in an experiment carried out by Anatole Abragam and the author. The experiment begins with two identical crystals at room temperature, say 300 degrees Kelvin. One crystal is placed in a magnetic field of 750 gauss, which tends to align the nuclear spins with the external fields. After several minutes the spin temperature reaches equilibrium with the lattice temperature; the magnetic polarization, or the proportion of the spins parallel to the field, is then measured by nuclear magnetic resonance. The second crystal is immersed in liquid helium until the spin and lattice temperatures reach equilibrium; the temperature of equilibrium must be the same as that of the helium, about four degrees K. Removing the first crystal from the magnet reduces the field by a factor of about 75 and reduces the temperature by the same amount, to four degrees K. If either crystal is then placed in the field, the magnetic polarization of the two crystals is found to be identical. Thus there is no difference between the spin temperature of a crystal cooled by removing it from a magnetic field and that of a crystal cooled by contact with a cold lattice.

This behavior is so startlingly unlike the behavior of more familiar thermal systems that it calls into question the very notion of assigning a temperature to a spin system. It could be argued, for example, that the concept of temperature should be employed only where knowledge of a system (such as the distribution of vibrational states in a crystal) is incomplete and cannot be specified any more precisely. In the nuclear-magnetic-resonance experiment, however, the temperature is in fact calculated from a more detailed knowledge of the underlying state of the system. Perhaps it would be better in this case to ignore Curie's law and say that the temperature is undefined.

These issues can be further explored in another experiment. Again a crystal is placed in a magnetic field and the polarization is measured after it has reached its equilibrium value. This time the external field is briefly reduced all the way to zero, and then it is restored to its original strength and the polarization is measured again. A convenient way to produce these abrupt changes in the external field is simply to remove the crystal from between the pole pieces of the magnet, hold it a few seconds and replace it. When such an experiment is undertaken, it turns out that the magnetic polarization is unchanged.

When the external field surrounding a nucleus is abolished, the spin energy levels immediately collapse. There is no longer any energy difference between the three possible orientations; indeed, there is no longer a field to define a set of allowed orientations. The persistence of the magnetic polarization after an interval outside the imposed field might therefore be interpreted as evidence for the validity of the spin temperature. In this argument it is the spin temperature





ENERGY AND ENTROPY are the basic elements in a definition of temperature in which negative values arise naturally: Temperature measures the amount of energy that must be added to a system to yield a given change in entropy. Here entropy (*black curves*) and temperature (*colored curves*) are both graphed as functions of energy. In a vibrational system (*left*) an increase in energy invariably brings an increase in entropy, and so the temperature is always positive. In the spin system, however (*right*), the entropy has a maximum possible value; at that point the change in the entropy is zero, and so the temperature is infinite. With each further increase in energy the entropy is reduced, and so the sign of the relation changes: the temperature becomes negative and at the maximum energy reaches minus zero.

that is "remembered" by the system and that specifies the correct polarization when the field is restored.

In this experiment too, however, there is a counterargument based on what seems to be a more fundamental description of the spin system. The net magnetic field acting on any nucleus in the interior of a crystal is made up not only of the externally imposed field but also of the fields generated by neighboring nuclei. Although the external field is removed, the quantum states of the nuclei may be preserved by interactions with these internal fields. The situation is complex, but it appears that the experimental result could be explained without reference to spin temperature.

The question of whether or not spin systems can have a temperature was resolved through an experiment suggested by Anatole Abragam of the Collège de France and carried out by Abragam and me at the Saclay Nuclear Research Center near Paris. The experiment tested the equivalence of spin temperatures created by magnetic cooling and those achieved by more conventional methods of refrigeration.

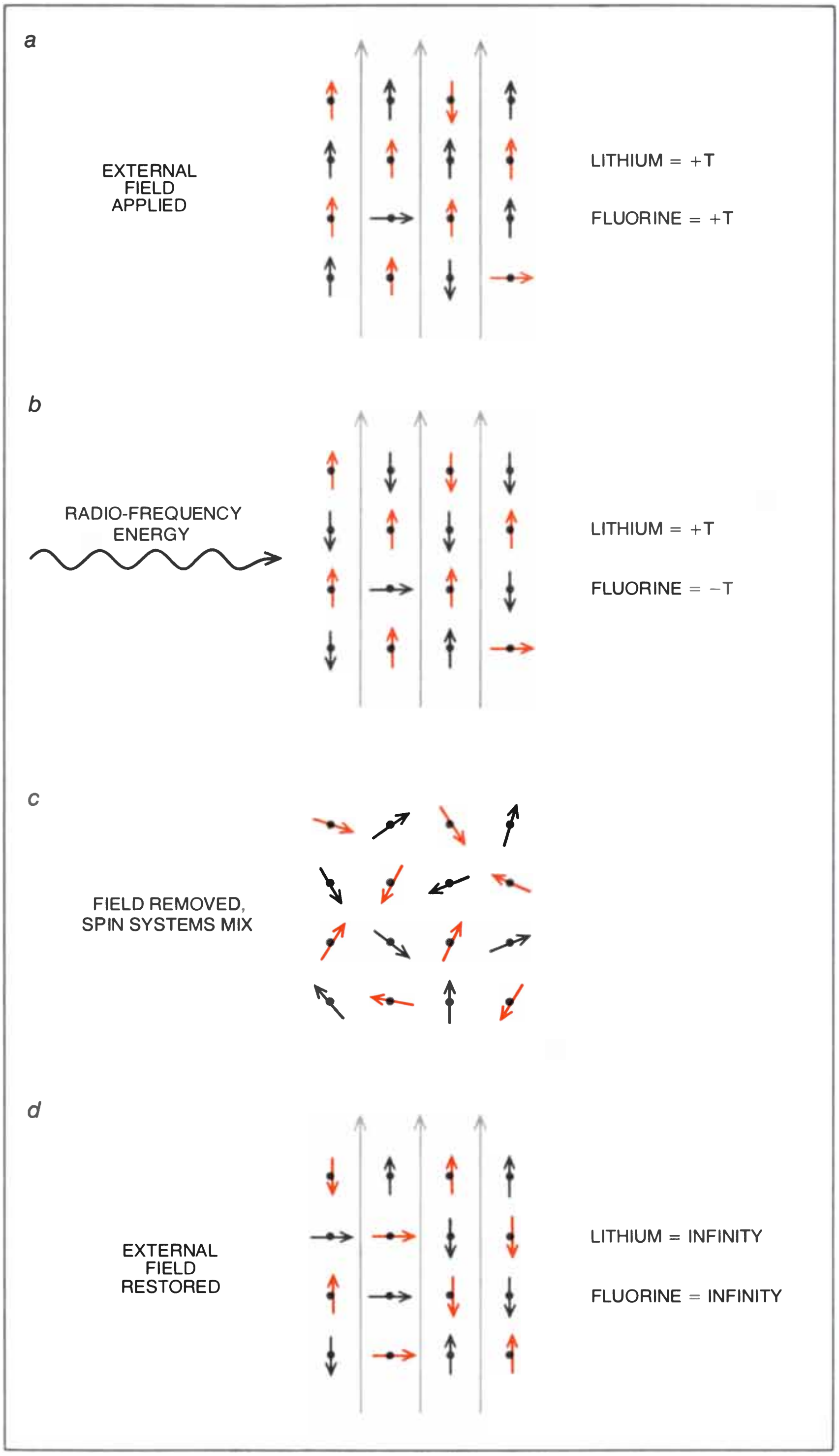
A salt crystal at room temperature (roughly 300 degrees Kelvin, or degrees Celsius above absolute zero) was placed in a magnetic field of 750 gauss, and the spin temperature was allowed to reach equilibrium with the lattice temperature. The crystal was then removed from the magnet, so that the field acting on the nuclei was reduced from 750 gauss to whatever residual field was provided by the local, internal fields; these fields are estimated to be about 10 gauss. The intensity of the field was therefore reduced by a factor of 75, and if Curie's law makes correct predictions under such circumstances, the spin temperature should have been reduced by the same factor, from 300 degrees K. to four degrees K.

It had already been shown that if the same crystal were replaced in the field, the magnetic polarization would be found to be unchanged. Instead of following this procedure another crystal was substituted for the first one; the second crystal was identical with the first in all respects except that it had been kept for some time in a bath of liquid helium at four degrees K. and in the absence of a magnetic field. It could therefore be assumed that the lattice temperature and the spin temperature of the second crystal were in equilibrium at four degrees K. When the second crystal was introduced into the 750-gauss field, the measured polarization was the same as that in the first crystal.

The interpretation of the experiment is straightforward: there is no measurable difference between a spin system cooled to four degrees K. by removing it from a magnetic field and an identical system cooled by thermal contact with a lattice at the same temperature. Although the behavior of the spin temperature in a rapidly changing magnetic field seems bizarre, it is indistinguishable from the behavior of a spin temperature set by conventional means.

Once the validity of spin temperatures has been established, it is a short step to demonstrating the existence of negative spin temperatures. It was pointed out above that as temperature increases, progressively higher energy levels are occupied. At absolute zero all the nuclei are in the lowest level; at temperatures between zero and infinity there is a pyramidal distribution, with more nuclei in the lower levels than in the higher ones; at infinite temperature all energy levels have an equal probability of being occupied. Up to this point the relation between temperature and the distribution of occupied states holds as well for the spin temperature as it does for the lattice temperature. A difference emerges, however, if an attempt is made to add more energy to a system that is already at infinite temperature.

For a lattice system to reach an infinite temperature would require an infinite source of energy, and there is no conceivable state of the system with higher energy. A spin system at infinite temperature, on the other hand, clearly has a finite energy: it is simply the sum of the energies needed to keep a third of the nuclei in each of the three possible spin states. Moreover, it is easy to imagine configurations of the spin system that would have higher energy. All that is necessary is to invert the population



SPIN CALORIMETRY measures the equilibrium temperature reached when two spin systems are "mixed." Each spin system is made up of many identical nuclei; in a crystal of lithium fluoride, for example, there are two spin systems, one consisting of all the lithium nuclei (*color*) and the other of all the fluorine nuclei (*black*). Because the energy levels in these nuclei have different spacings the lithium and fluorine spins do not interact as long as a magnetic field is maintained. When the field is first imposed (*a*), both systems reach thermal equilibrium with the lattice at some positive temperature $+T$, meaning that there is an excess of spins parallel to the field. A burst of radio-frequency energy of the appropriate frequency causes all the fluorine nuclei to reverse their orientation (*b*), but it has no effect on the lithium nuclei. Hence the lithium system retains the temperature $+T$, whereas the temperature of the fluorine system is converted to $-T$. When the magnetic field is removed, all the energy levels collapse (*c*) and the two systems can interact freely. If the field is restored (*d*), each system is found to have a polarization of zero and an infinite spin temperature, the value midway between $+T$ and $-T$.

distribution observed at lower temperatures, so that there are more nuclei in the higher-energy levels than in the lower-energy ones. The state of highest possible energy would have all the nuclei in the antiparallel orientation.

The crucial difference between the lattice temperature and the spin temperature is that the number of energy levels for the vibratory motions of atoms is infinite, but the number of orientational levels is always finite. In the lattice system there is no intrinsic limit to the energy of an individual nucleus, and adding energy merely promotes some nuclei to ever higher levels. In the spin system no nucleus can have an energy higher than that of the antiparallel orientation, and adding energy can only drive more and more of the nuclei into that configuration.

What is the temperature of a spin system in which the majority of the nuclei are in the highest energy level? The measurement can again be derived from Curie's law. At temperatures below infinity the preponderance of nuclei in the lowest-energy, parallel orientation gave rise to a magnetic polarization, M, parallel to the external field, H. The temperature was then found to be proportional to H/M. At temperatures higher than infinity the majority of the nuclei would be lined up antiparallel to the field. It follows that the magnetic polarization would also be opposed to the external field and algebraically would have the value $-M$. Applying Curie's law, the temperature is proportional to $H/-M$, and it must have a negative value.

A better understanding of why negative numbers must be included in the absolute temperature scale can be gained by returning to the fundamental definition of temperature. For a lattice system any increase in energy is accompanied by an increase in entropy. If entropy is related to the difficulty of deciding what state a nucleus is in, then this "monotonic" increase is readily accounted for: with every increase in energy there are more energy levels among which to choose. The temperature can be related to the amount of energy that must be added to the system to yield a given increase in entropy; since both quantities must always have the same sign, the temperature must always be positive.

Up to the point of infinite temperature the relation of energy to entropy is similar in the spin system. Each increase in energy brings a corresponding increase in entropy. At infinite temperature, however, the entropy reaches a maximum: it cannot be more difficult to predict the orientation of a nucleus than it is when there is an equal likelihood of its being in any one of the possible orientations. Further increases in energy now cause a decrease in entropy as the system becomes more highly ordered.

When all the nuclei are opposed to the external field, the system has reached a temperature of "minus zero" and the entropy is at a minimum; every nucleus is certain to be in the highest energy state.

The temperature of the spin system becomes negative as soon as the entropy begins to decrease with increasing energy. Mathematically the temperature curve is said to approach an asymptote at the point of maximum entropy. If the distribution of spin states is nearly uniform but slightly biased toward the parallel configuration, then the temperature is numerically large and positive; with a slight bias in the other direction the temperature abruptly assumes a large negative value. Toward either extreme the temperature approaches zero, but at maximum energy the approach is from the negative side. A more formal analysis of these relations has been carried out by Norman F. Ramsey of Harvard University and others.

Although negative absolute temperatures may be intellectually troubling, they are quite easy to create in the laboratory. In an early demonstration Edward M. Purcell and Robert V. Pound of Harvard converted a positive spin temperature to the numerically equal negative one by suddenly reversing the direction of the applied magnetic field. All the nuclei that were originally parallel to the field then became antiparallel, and vice versa. The energy of the spin system increased (since more of the nuclei were in the high-energy antiparallel state) but the entropy of the system remained constant.

Today the most convenient method for producing negative absolute temperatures makes use of the nuclear-magnetic-resonance apparatus that is also employed to measure the temperatures. The specimen is initially polarized by a magnetic field; then it is irradiated with a short but intense burst of radio-frequency energy. If the appropriate frequency is chosen, the population of occupied states can be inverted almost instantaneously. All the nuclei originally parallel to the field reverse their direction to become antiparallel and the initially antiparallel ones become parallel; the number of nuclei with a perpendicular orientation is not changed.

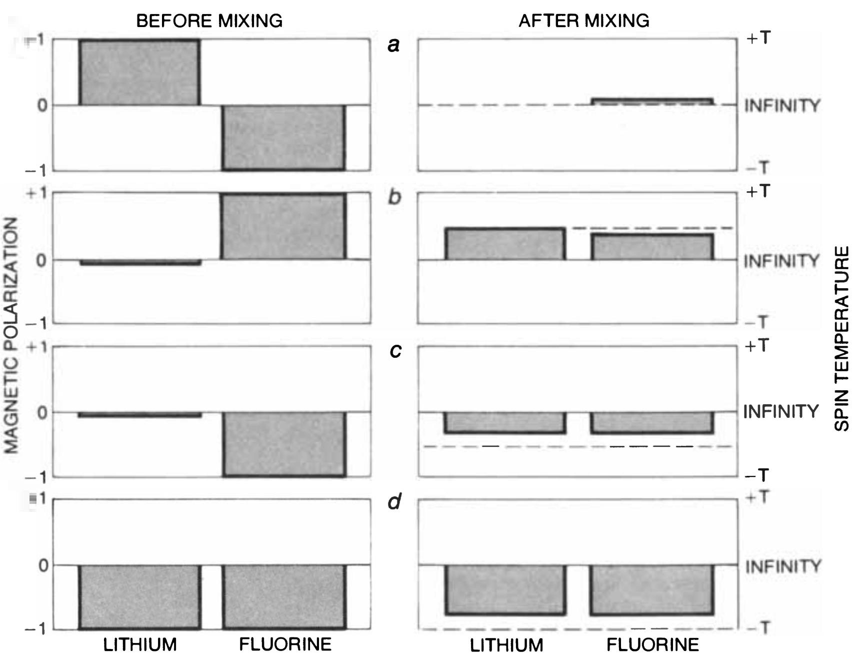
Through nuclear magnetic resonance an infinite spin temperature can be generated with equal ease. The radio-frequency field is merely left on for a few seconds, so that each nucleus flips many times. The result is that the nuclei distribute themselves randomly, and hence for a large number of nuclei uniformly, in all the possible spin states.

If spin temperatures extend from zero to positive infinity and then from negative infinity back to zero, the temperatures should obey a peculiar law of addition. In particular, two temperatures of equal magnitude but opposite sign should add to yield an infinite temperature. That prediction can be tested by an experiment in "spin calorimetry," analogous to the ordinary calorimetry experiments in which two liquids at different temperatures are mixed and the equilibrium temperature of the product is measured.

Abragam and I conducted such experiments in which the two "substances" to be mixed were the separate spin systems of lithium and fluorine in a crystal of lithium fluoride. With nuclear-magnetic-resonance apparatus each of the spin systems could be separately prepared at a positive temperature, $+T$, a negative temperature $-T$ or an infinite temperature. Of course, the two kinds of nuclei are already intimately mixed in the crystal, but as long as the specimen remains in a magnetic field the two spin systems do not interfere with each other. They are effectively isolated because at any given external field the spacing between the energy levels is different in lithium and fluorine. When the crystal is taken out of the field, however, the energy levels of both nuclear species collapse; the two systems can then freely communicate, and they quickly reach a common equilibrium temperature. The crystal can then be replaced in the magnetic field and the polarization of each system can be measured.

A series of such calorimetry measurements showed close agreement with theory. When the lithium system at a temperature of $+T$ was mixed with a fluorine system at $-T$, both systems reached a common temperature near infinity. When one system was at a large negative or positive temperature and the other was at infinite temperature, both reached a common intermediate temperature. Discrepancies from the ideal results could be accounted for by differences in the specific heat of the two spin systems and by relaxation of the spin systems toward thermal equilibrium during the experiment.

The introduction of negative absolute temperatures may seem to accomplish little more than to make obscure and confusing one of the properties of matter that had seemed most simple and most familiar. The experimental results run counter to a strong sentiment that temperature, like volume, is something intrinsically positive. It might even be supposed that the temperature scale was determined too early in the history of science and that some revision of it might encompass all temperatures within a single range of positive numbers. That is not possible; two infinite ranges of temperature are needed. The scale cannot be simplified because the negative temperatures it includes represent real states of any system with a finite number of energy levels.



RESULTS OF SPIN-CALORIMETRY EXPERIMENTS show that spin temperatures behave much like ordinary temperatures when two systems are mixed. If the lithium and fluorine systems are oppositely polarized before mixing (*a*), as in the example on the opposite page, the final polarization is near zero and the final temperature is near infinity. Mixing one strongly polarized system with one at an infinite temperature (*b, c*) yields an intermediate temperature for both systems. If the lithium and fluorine nuclei have the same polarization (*d*), no change is expected as a result of mixing. Discrepancies between theoretical expectations (*broken lines*) and measured values (*gray bars*) can be attributed to differences in the heat capacity of the two spin systems and to the thermal relaxation of both systems toward the lattice temperature.

HP measurement and computer advances

HP-33E Scientific
49 lines of program memory
8 conditionals
3 levels of subroutines
8 addressable storage registers
Go To, Single Step, and Back Step commands for easy editing

HP-38E Advanced Financial
Powerful cash flow analysis
Easy time and money calculations
Up to 99 lines of program memory
2000-year calendar
Advanced statistics capabilities

HP's two new Series E programmables are without equal and within your reach.

If you are an aspiring professional, it makes sense to step up to HP programmability. On the basis of power, capability, and ease of use, the new HP-33E scientific programmable at $100* and the new HP-38E advanced financial programmable at $120* are surprisingly affordable.

For student and educator alike, an HP Series E programmable is the logical choice. Consider: Each has a specialized set of advanced, preprogrammed functions, plus the convenience of programming.

Consider further: Series E calculators make excellence available at a surprisingly low price. Besides the two programmables, Series E includes three preprogrammed models: the HP-31E and HP-32E scientific and advanced scientific calculators, at $60* and $80* respectively, and the HP-37E business management calculator at $75*.

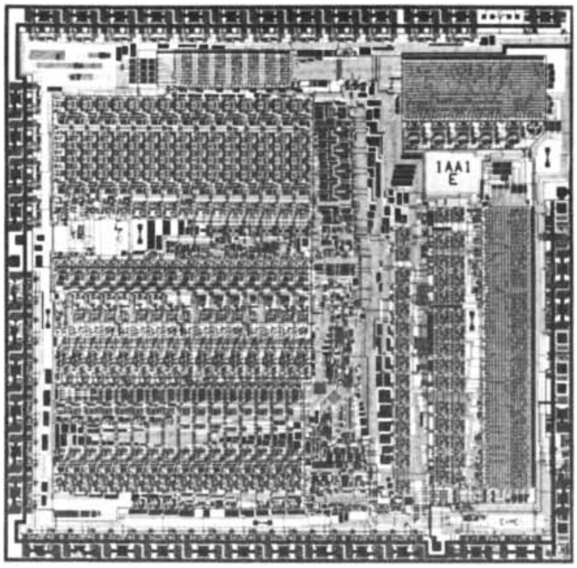
Like their predecessors, they have no "equal." The key to HP's simplified computer logic is [ENTER↑], making possible fast, efficient problem solving with no [=], literally and figuratively. And, like their predecessors, Series E calculators have the feel and reliability born of quality design and construction.

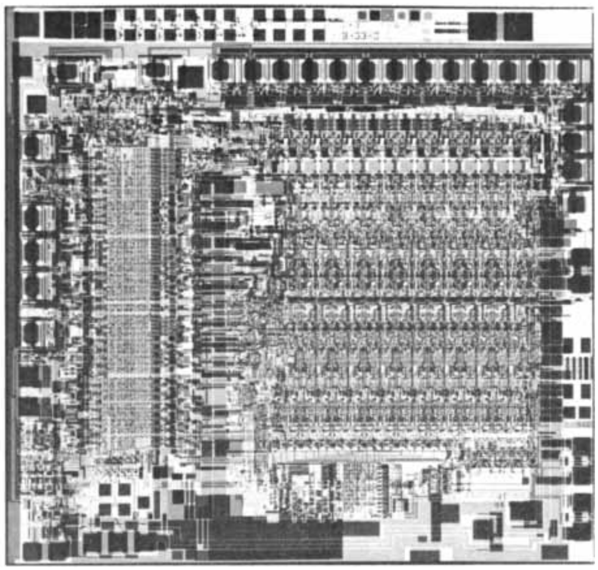
Programming the HP-33E and 37E is easy. Just switch to the Program mode and key in the series of operations you'd normally use to solve a problem. Rather than executing the keystrokes, the calculator stores them in its memory. Switch to the Run mode, key in your data, and press the run/stop key. You can repeat the program with different data as many times as you like. There's no complicated programming language to learn, and no elaborate start-up procedures to memorize.

Series E offers some other helpful advances: Display messages tell you when you've made a procedural error, and what kind. And if you suspect the calculator is at fault, it will perform a diagnostic self-check at the touch of a key. A larger, brighter LED display automatically separates thousands with commas for easy reading. To help you get started, we've prepared modular handbooks for owners, free with the calculator, that let you skip what you already know, and give you what information you *do* need, clearly and simply.

HP has a 16 mm film, *Calculators, Classrooms... Careers*, available on a free-loan basis to teachers. The 14-minute film examines the development of advanced personal calculators and how they help people in a variety of educational and work environments. Narrator is Wally Schirra, astronaut and business executive. To obtain the film for classroom showing, write to Rick Baker, H.P., 1000 N.E. Circle Boulevard, Corvallis, OR 97330.

...extend your possibilities.





These photographic enlargements of two HP microprocessor chips—one with NMOS II large-scale integrated circuits at right and one with CMOS/SOS at left—illustrate a major reason for the rapid growth of digital circuits: increasing functional density at a decreasing cost. These chips contain as many as 10,000 transistors and are about this big ■.

How to operate more confidently in an increasingly digital world.

While analog circuits generally convey data by a continuum of voltages, the digital world of microprocessors and LSI circuits relies on two voltage levels, conveying information by switching between them. Because of this fundamental difference, digital circuits pose a special set of problems for both designers and troubleshooters. And HP has some excellent solutions.

As the solid state revolution increases their functional density at a decreasing cost, digital circuits are infiltrating areas that previously relied on analog technology, and are cropping up in an endless procession of products that include microwave ovens, automobiles, home computers, process control (analog from its inception), telecommunications, and a range of industrial controls.

In short, the technical, industrial, and consumer worlds are turning digital. While this greatly expands possibilities, it also introduces complications—most notably in the design, debugging, software development, and diagnosis of these functional leviathans in minnow-sized packages. While it may not be immediately obvious, this digital data domain requires special measurement tools. And here Hewlett-Packard is making significant strides, continually introducing new measurement approaches and instruments designed to make the digital revolution manageable, and thus enable technical people to operate more confidently.

If you are wrestling with any aspect of digital circuitry, you will be interested in reading a recent issue of the *Hewlett-Packard Journal*, which explores HP's growing family of digital logic analysis test instruments at some length, and offers a few viewpoints of possible interest. Just mail the coupon, and we will be pleased to send you a copy.

HEWLETT hp PACKARD

1503 Page Mill Road, Palo Alto, California 94304

For assistance call: Washington (301) 948-6370, Chicago (312) 255-9800, Atlanta (404) 955-1500, Los Angeles (213) 970-7500

Mail to: Hewlett-Packard, 1503 Page Mill Road, Palo Alto, CA 94304.
Please send me further information on

( ) HP-33E scientific calculator
( ) HP-38E advanced financial calculator
( ) HP digital logic analysis

Name ______
Company ______
Address ______
City ______ State ______ Zip ______

*Domestic U.S. prices only. 00848

The Thermostat of Vertebrate Animals

The hypothalamus, a structure at the base of the brain, monitors body temperature in a wide variety of animals and maintains it at an optimal level by controlling thermoregulatory mechanisms

by H. Craig Heller, Larry I. Crawshaw and Harold T. Hammel

Animals with backbones function most efficiently within a narrow range of body temperatures. The upper limit for survival is about 45 degrees Celsius (113 degrees Fahrenheit), when proteins begin to denature and become inactivated; the lower limit is slightly below zero degrees C. (32 degrees F.), when intracellular water begins to form ice crystals that rupture and kill cells. Even within these extremes slight shifts in body temperature away from an optimal level can have adverse effects because each physiological process involves many integrated biochemical reactions, the rates of which are dependent in a highly specific way on temperature. When the temperature changes, the total integration of the reactions may be disrupted. In spite of these physiological realities vertebrate animals thrive in virtually every habitat on the earth, from the icy waters of polar seas to the fierce heat of equatorial deserts. How vertebrates maintain a favorable internal temperature in such hostile thermal environments is the subject of this article.

Commonly but, as we shall see, inaccurately the vertebrates are separated into two major groups: the "cold-blooded" animals and the "warm-blooded" animals. Biologists refer to these same two groups of vertebrates as ectotherms or endotherms in recognition of their major sources of heat. The ectotherms (fishes, amphibians and reptiles) have poor body insulation and low rates of metabolic heat production. Their body temperature is largely dependent on heat from their environment, and their most important thermoregulatory mechanism is the selection of a suitable environment or "microclimate." In the absence of a strong source of radiant energy such as the sun ectotherms prefer to remain at an ambient temperature almost identical with their optimal body temperature.

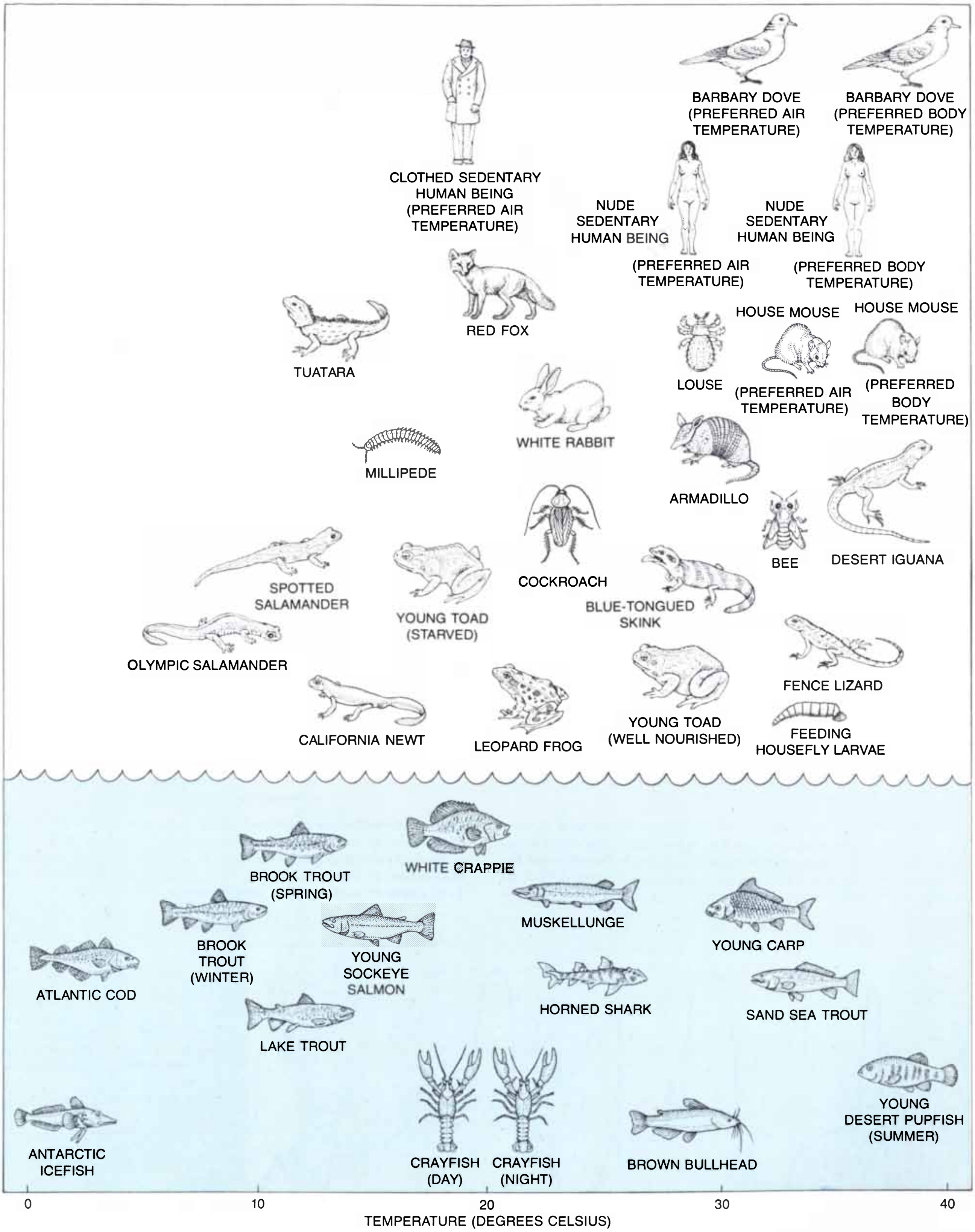
In contrast, the endotherms (birds and mammals) are usually well insulated with fur, feathers or fat and have a resting metabolic rate at least five times higher than that of ectotherms of comparable size at a similar body temperature. When endotherms are exposed to a cold environment, they elevate their rate of metabolic heat production and thereby maintain their body temperature at an optimal level. Because of their high metabolic rate endotherms must lose a significant amount of heat to the environment in order to prevent their body temperature from rising. Therefore endotherms always prefer an ambient temperature somewhat below their optimal body temperature so that heat can be lost passively down a temperature gradient. For example, a clothed human being maintains an internal body temperature of about 37 degrees C. (98.6 degrees F.) but prefers an ambient temperature of about 22 degrees C. (71.6 degrees F.).

Endotherms have evolved physiological mechanisms in addition to variable metabolic rates in order to regulate their body temperature. These additional mechanisms control the rate at which the animal exchanges heat with its environment. The insulative value of fur or feathers can be varied by muscles controlling their attitude. In addition, the insulative value of the skin is variable. Heat is transported rapidly by the blood from the core of the body to the skin, where it is lost to the environment; the rate of heat loss can be controlled by the dilation or constriction of the arterioles of the skin. In many species the loss of body heat from a poorly insulated appendage is controlled by heat exchange between the vessels carrying blood into the appendage and those carrying blood out of it.

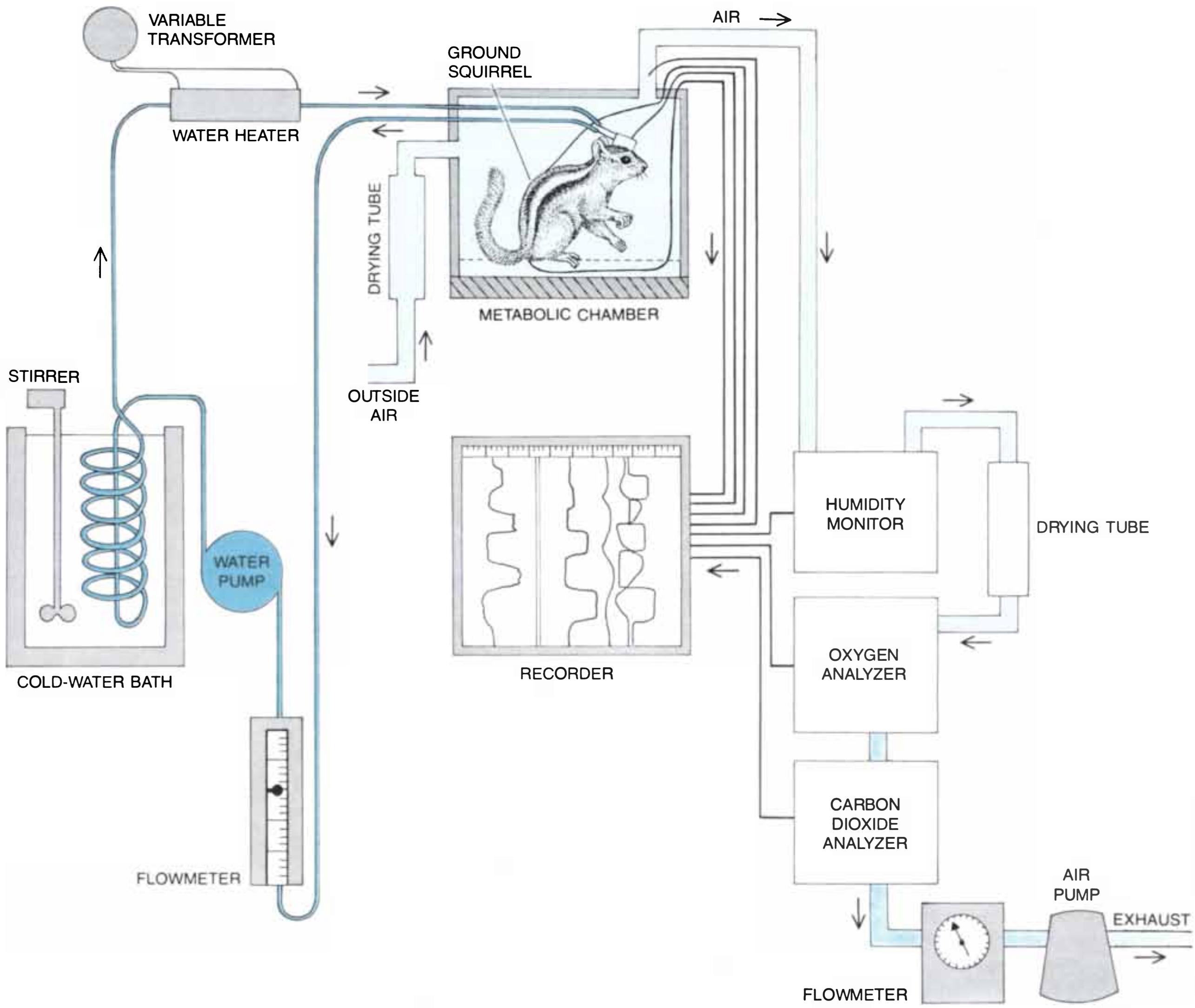
Another way of losing heat to the environment is through the evaporation of water from the skin or the respiratory surfaces. In many species this kind of heat loss is enhanced by the mechanisms of panting and sweating. For evaporation to be effective the animal must direct the appropriate amount of blood to the area being cooled.

Biologists did not fully appreciate the thermoregulatory capabilities of ectotherms until field measurements of body temperature were made on a variety of Temperate Zone reptiles. Surprisingly, regardless of the air temperature the body temperature of active lizards was found to be generally within the range of 34 to 40 degrees C. (93.2 to 104 degrees F.). Since then it has been shown that a wide variety of ectotherms can regulate their body temperature with a great deal of precision through their behavior. For example, John R. Brett, working at the Fisheries Research Board of Canada Biological Station in Nanaimo, B.C., discovered that if young sockeye salmon are given a choice, they will select water at 15 degrees C. (59 degrees F.), the temperature at which their growth, heart output and sustainable swimming speed are all maximal. When a variety of ectotherms (including sharks, bony fishes, amphibians and lizards) are given a choice between two nonoptimal temperatures, they regulate their body temperature at the preferred level with remarkable accuracy by moving back and forth between the two compartments.

Additional thermoregulatory effector mechanisms have recently been documented in ectotherms. Eugene C. Crawford and Billy J. Barber of the University of Kentucky found that the chuckwallah lizard exploits panting as a mechanism to lose heat if its skin temperature or deep-body temperature exceeds a certain level, and Marvin L. Riedesel of the University of New Mexico showed that box turtles can hold their body temperature at 10 degrees C. (50 degrees F.) below the ambient tempera-



PREFERRED AMBIENT AND BODY TEMPERATURES for vertebrate animals vary greatly and depend on the thermal environment to which the animal is adapted. For ectotherms (animals that obtain most of their body heat from their surroundings) the preferred internal and ambient temperatures are the same. For endotherms (animals that obtain most of their body heat from internal metabolic processes) the preferred ambient temperature is considerably below the preferred body temperature because internally generated heat must be continually lost down a temperature gradient. The crayfish demonstrates that preferred temperature may change on a daily basis, the trout demonstrates seasonal change and the toad provides an example of preferred temperature changing with nutritional state.



EXPERIMENTAL APPARATUS diagrammed here is used to measure the temperature sensitivity of the hypothalamus in small mammals. The animal is placed in the sealed metabolic chamber and the temperature of its hypothalamus is manipulated by means of thermodes: implanted stainless-steel tubes through which water is circulated at controlled temperatures. At the same time the animal's rate of metabolism and evaporative water loss are measured by drawing dry outside air through the chamber at a known rate and analyzing the effluent air for its content of water, oxygen and carbon dioxide. Temperatures, measured by thermocouples, are recorded on chart.

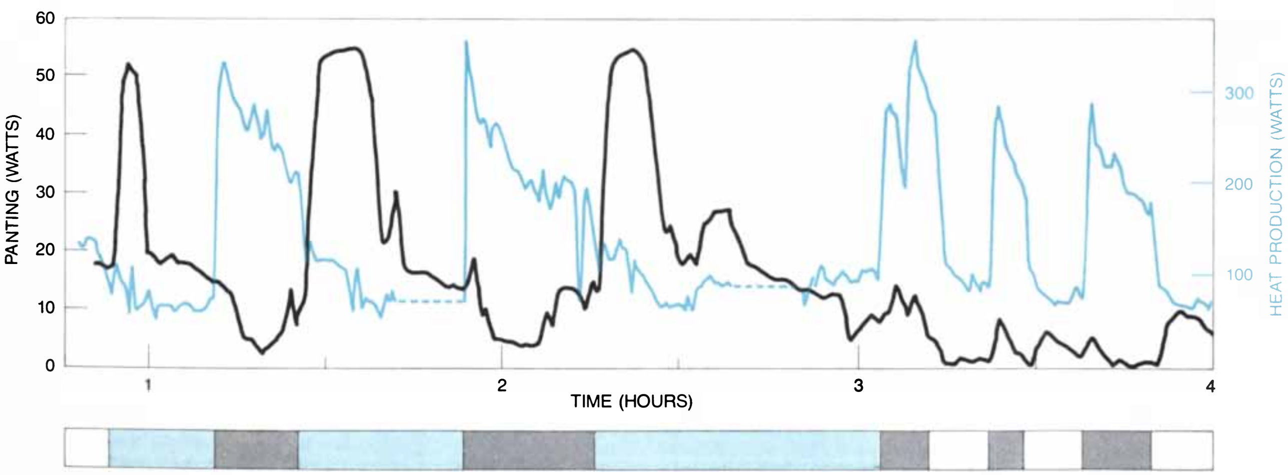


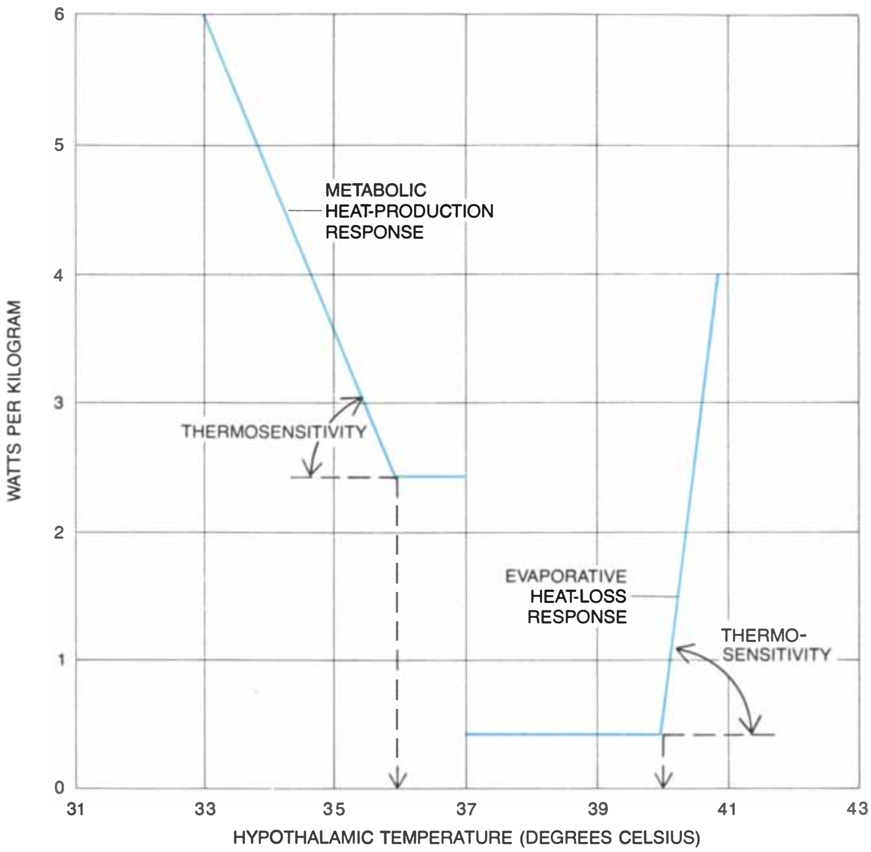
CHART RECORDING shows the results of an experiment in which the hypothalamus of a Labrador retriever was alternately warmed (*color bars*) and cooled (*gray bars*) by means of implanted thermodes as the rates of two thermoregulatory responses were continuously measured. Cooling of the hypothalamus resulted in a rapid loss of body heat by evaporative water loss, primarily from panting (*black curve*), whereas warming of the hypothalamus resulted in shivering and consequently an increase in the metabolic rate (*color curve*).

ture by panting. When ectotherms can no longer find environmental heat sufficient to maintain a high body temperature, they generally seek a cool place safe from predators and become inactive, allowing their body temperature to follow the ambient temperature. This response may also result when the food supply is inadequate, since a lowering of the body temperature means a reduction of the metabolic rate and hence the conservation of vital energy stores. As we shall see, a similar strategy is exploited by some endotherms faced with notably harsh environmental conditions.

In the 1880's Charles Richet in France and Isaac Ott in the U.S. observed that the localized destruction of tissue in the hypothalamus at the base of the brain in dogs gave rise to elevated body temperature. In classic experiments conducted by Henry G. Barbour in 1912 silver thermodes were implanted in the hypothalamus. When the thermodes were cooled, the body temperature rose; when the thermodes were heated, the body temperature fell. Subsequent experiments on a variety of mammals by numerous investigators indicated that the region of the hypothalamus located just over the optic chiasm (the place where the optic nerves from the two eyes converge) is essential for the regulation of body temperature. It is now known that thermoregulatory functions are centered in this part of the brain in all classes of vertebrates.

What is the nature of this internal "thermostat," and how does it regulate body temperature? On the basis of control theory one can postulate that the thermostat must include several features. First, it must have information about the actual temperature of the body by means of at least one feedback circuit. Second, the thermostat must "know" the optimal body temperature in that it must be programmed with some kind of reference point. Third, the thermostat must be able to compare the actual body temperature at any given moment with the optimal setting and, if a difference exists, trigger the appropriate behavioral and physiological thermoregulatory mechanisms.

As we have seen, the hypothalamus is sensitive to temperature. When the hypothalamus of an endotherm is warmed, the animal reacts with heat-dissipating responses such as sweating, panting and dilation of the peripheral blood vessels; when the hypothalamus is cooled, the animal reacts with heat-generating responses such as shivering, erection of the fur and constriction of the peripheral blood vessels. The most basic thermoregulatory mechanism—the seeking of a suitable thermal environment—is shared by all vertebrates. When the hypothalamus of ectotherms and endotherms is warmed, they behave as if they are too hot and select a cool environment that serves to lower their body temperature. Cooling the hypothalamus elicits the opposite response.

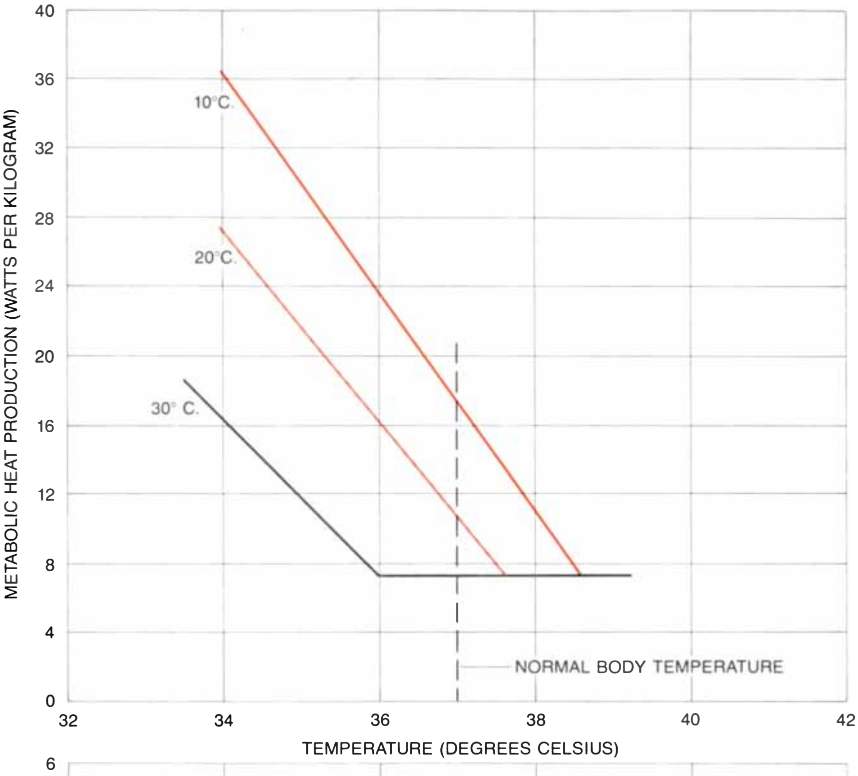


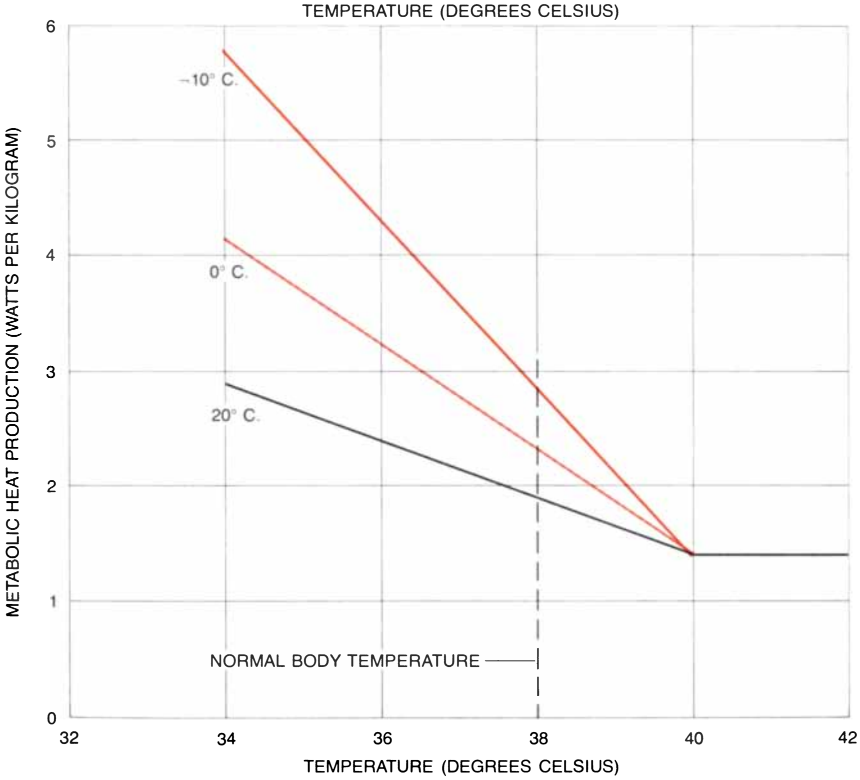
THERMOREGULATORY RESPONSES are triggered when the temperature of the hypothalamus either falls below or rises above specific threshold temperatures. The thresholds may be influenced by a variety of factors including ambient temperature, level of exercise, sleep and microorganisms that cause fever. In the graph shown here (for a dog with an optimal body temperature of 37 degrees Celsius and an ambient temperature of 25 degrees) the hypothalamic threshold for the metabolic heat response is 36 degrees and that for the evaporative heat-loss response is 40 degrees. As is indicated by the slopes of the curves, the intensity of each thermoregulatory response is proportional to the difference between the actual temperature of the hypothalamus and the threshold temperature at which that particular response is initiated.

One of us (Hammel), who was then working at the John B. Pierce Foundation laboratory at Yale University, was able to quantitatively measure the thermoregulatory responses of dogs as a function of hypothalamic temperature. These experiments were made possible by the development of small stainless-steel thermodes that could be permanently implanted around the hypothalamus of intact animals. The thermodes were perfused with water of a specific temperature in order to heat or cool the hypothalamus. When the hypothalamic temperature of dogs was systematically changed, temperature thresholds for the initiation of each thermoregulatory response were observed: shivering was induced when the hypothalamus was cooled below a particular temperature and panting was induced when the hypothalamus was heated above another temperature. Moreover, at temperatures above the threshold level the intensity of the thermoregulatory response was proportional to the difference between the threshold temperature and the actual hypothalamic temperature. These characteristics of the temperature sensitivity of the hypothalamus are qualitatively the same in all the mammalian species tested.

Does the temperature of the hypothalamus provide all the feedback information the system requires to monitor the thermal condition of the body? Two simple observations suggest that the story is more complex. If one enters a cold environment, one begins to shiver almost immediately, before there is any fall in the hypothalamic temperature. Similarly, if one enters a hot environment such as a sauna bath, one begins to sweat before there is any rise in hypothalamic temperature. It is as if the thermostat predicts a change in the internal body temperature and takes corrective action immediately. One explanation for the fact might be that peripheral temperature sensors at the surface of the body are providing the central thermostat with information about rapid changes in the thermal environment.

This hypothesis was tested by measur-





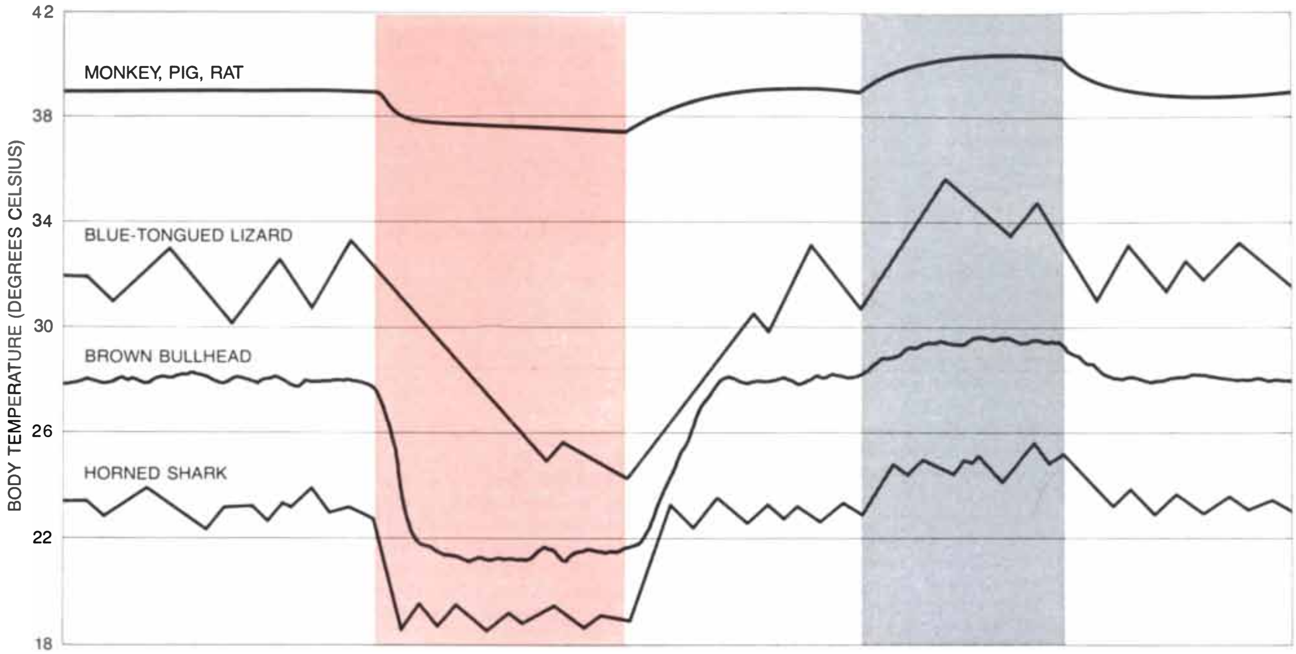
EFFECT OF AMBIENT TEMPERATURE on the rate of metabolic heat production reveals two ways that inputs from peripheral temperature sensors can alter the characteristics of the central thermostat. In the kangaroo rat (*top graph*) the hypothalamic threshold temperature for the metabolic heat response rises as ambient temperature declines. (Response curves at three different ambient temperatures are shown.) This increase of threshold means that a fall in actual core-body temperature is not required at low ambient temperatures to stimulate increased metabolic heat production. Thus the thermostat anticipates future changes in body temperature and makes compensatory responses before those changes actually occur. In the harbor seal (*bottom graph*) a decrease in ambient temperature does not change the hypothalamic threshold temperature for the metabolic heat-production response but instead alters the sensitivity of the thermostat, so that more heat is generated for each unit of temperature below the hypothalamic threshold. This shift in sensitivity is reflected in the changing slope of the response curves.

ing thermoregulatory responses to the heating and cooling of the hypothalamus at different ambient temperatures. Experiments with some species of mammals have shown that as the ambient temperature decreases, the threshold temperature for the metabolic heat-production response increases. In effect at low ambient temperatures the characteristics of the thermostat are modified so that the metabolic heat-production response is maintained at an elevated level without any change in the hypothalamic temperature. In this way the thermoregulatory response occurs before the internal body temperature has declined to a dangerously low level. Changes in the ambient temperature may also affect the thermostat by altering its thermosensitivity, that is, the magnitude of the thermoregulatory response evoked by a unit change in the hypothalamic temperature. This change in sensitivity is reflected by a change in slope of the curve relating the response to hypothalamic temperature.

The effect of ambient temperature on the thermostat differs from species to species. In the harbor seal a decrease in the ambient temperature results in a large increase in the thermosensitivity of the hypothalamus but no change in the threshold temperature for the metabolic heat-production response. The antelope ground squirrel, on the other hand, shows a normal inverse relation between the ambient temperature and the hypothalamic threshold temperature for the metabolic heat-production response but a maximal hypothalamic thermosensitivity at high ambient temperatures. The antelope ground squirrel is a desert animal that is active during the day, so that it is highly adaptive for it to be most sensitive to changes in its deep body temperature at high ambient temperatures.

To sum up, the central thermostat of vertebrate animals is able to obtain information about ambient temperature from peripheral sensors at the body surface; these inputs then modify the characteristics of the thermostat so that a different thermoregulatory output is achieved without a change in the hypothalamic temperature. It is the ability to respond to rapid changes in the ambient temperature that provides the vertebrate thermostat with its predictive capability: the animal is able to anticipate a change in body heat content and take corrective action before that change actually occurs.

There are, however, a few situations where actual changes in the hypothalamic temperature activate thermoregulatory responses in mammals. For example, at the onset of a fever there is an increase in the regulated body temperature and at the onset of sleep there is a decrease in the body temperature.

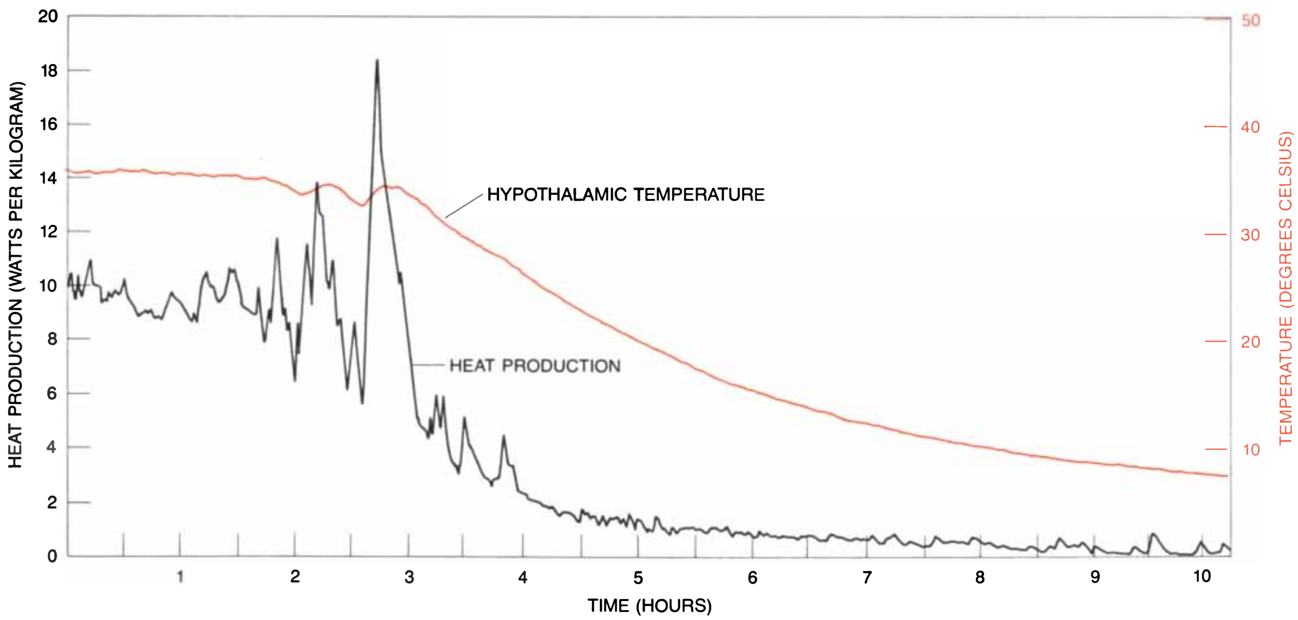


THERMOREGULATION BY BEHAVIOR in vertebrate animals is demonstrated by experiments in which animals are given a choice of two thermal environments, one above and one below their preferred body temperature. Both endotherms and ectotherms maintain their body temperature at the preferred level by moving back and forth between the two compartments. Heating the hypothalamus (*color*) results in a lowering of the preferred body temperature, whereas cooling it (*gray*) results in a raising of preferred body temperature.

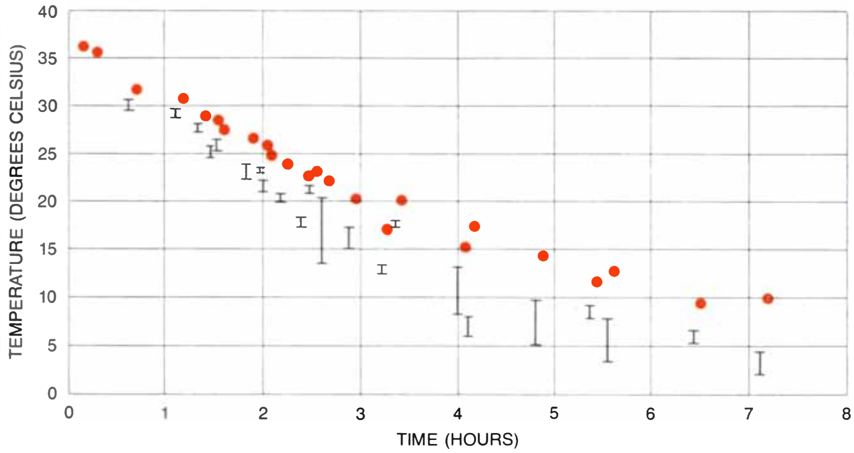
These changes result from shifts in the hypothalamic thresholds for thermoregulatory responses and from changes in hypothalamic thermosensitivity. Until a new equilibrium is reached the thermoregulatory responses are directly influenced by the divergence between the new setting of the thermostat and the actual hypothalamic temperature.

Changes in the hypothalamic temperature may also be the primary activating signal for thermoregulatory responses in small mammals, whose brain and body temperatures are more variable and more closely coupled to the environmental temperature. Small mammals such as ground squirrels, chipmunks, kangaroo rats and wood rats all have hypothalamic thermosensitivities much higher than those of larger mammals such as seals, cats, dogs and rabbits. As the body size of an animal increases, so does the gain or loss of body heat that is required before a temperature change is detected in the deep body core. As a result peripheral temperature sensors become increasingly important with larger body size.

Another situation in which naturally occurring changes in the temperature of



ENTRANCE INTO HIBERNATION of the golden-mantled ground squirrel is monitored in this recording, which shows the rate of metabolic heat production and the hypothalamic temperature of the animal over a period of time. The otherwise passive cooling curve is interrupted by bursts of metabolic heat production whenever the body temperature of the animal temporarily falls below the declining set of the thermostat. This finding and others suggest that the thermostat is not "switched off" during hibernation but rather is "turned down."



RESETTING OF THE THERMOSTAT during the entrance into hibernation of the golden-mantled ground squirrel was demonstrated by a series of experiments in which the hypothalamic temperature of the animal was manipulated. The dots represent actual hypothalamic temperatures at specific times during entrance. Manipulations of hypothalamic temperature showed that the threshold for the metabolic heat-production response was somewhere in the range indicated by the vertical line below each hypothalamic temperature point. It can be seen that entrance into hibernation involves a progressive and continuous resetting of thermostat.

the body core are primarily responsible for inducing thermoregulatory responses is found in physical exertion. In man, as everyone knows, sustained exertion is accompanied by an increase in the rate of sweating. Ethan R. Nadel and his colleagues at the John B. Pierce Foundation laboratory demonstrated, however, that at a given ambient temperature the human sweat rate is poorly correlated with the level of physical exertion but is closely correlated with the temperature of the body core. (It is not known whether the rise in body temperature is having its primary effect on the hypothalamus or on other parts of the nervous system, such as the spinal cord.) In the dog, on the other hand, the hypothalamus apparently receives information directly from sensory receptors in the joints and the muscles, so that panting is initiated soon after the exercise begins and before the temperature of the body core rises.

Hibernation in mammals is an intriguing phenomenon from the point of view of temperature regulation. When an animal is ready to hibernate, it retires to a secluded nest or burrow and becomes inactive. Its temperature then falls to the point where it may be very close to the temperature of the environment. For a long time many biologists assumed that hibernating mammals temporarily abandoned thermoregulation and returned to the more primitive ectothermic state of their reptilian ancestors. Recent studies have revealed that hibernation is not an abandonment of thermoregulation but rather a precisely regulated lowering of the central thermostat for the purpose of conserving energy.

That hibernation is a regulated condition was first suggested by the observation that although ectotherms need an outside source of heat to recover from a cold-induced torpor, hibernating mammals are capable of rousing and returning to normal body temperature even when the ambient temperature is low. For example, the golden-mantled ground squirrel rouses spontaneously many times during the hibernation season; its periods of torpor may last from less than a day to two weeks. Moreover, if the ambient temperature falls to dangerously low levels, the animal will rouse from torpor, a "fail-safe" mechanism that prevents it from freezing to death. The alarm temperature appears to be sensed by the hypothalamus, because if the ambient temperature is held at a level above the alarm temperature and the hypothalamus of the hibernating animal is cooled by means of thermodes, the animal will rouse. Conversely, if the hypothalamus is slightly warmed, the temperature of the rest of the body can fall below the alarm level without inducing arousal. When the warming is stopped, however, arousal occurs.

Charles P. Lyman of the Harvard Medical School has studied sensitivity to low temperatures in three other species of hibernating mammal: the thirteen-lined ground squirrel, the Turkish hamster and the garden dormouse. He found that the ground squirrel and the hamster rouse from hibernation when their head is cooled by an external thermode but that the animals continue to hibernate if the head is warmed and the body is cooled by lowering the ambient temperature. In contrast, the garden dormouse can be aroused from torpor by applying a cold stimulus to its feet; this species appears to depend on peripheral temperature sensors to warn it of dangerously low ambient temperatures. In view of the differences between the alarm systems of the two species it is interesting to note that the garden dormouse hibernates on its back with its feet sticking up, whereas ground squirrels and hamsters hibernate with their feet tucked under them.

The arousal from deep hibernation involves a large increase in metabolic heat production. When the golden-mantled ground squirrel is aroused from hibernation by a cooling of the hypothalamus, the ensuing increase in metabolic rate can be suppressed by a heating of the hypothalamus. The longer the interval following the initiation of arousal is, the higher the hypothalamus must be heated to suppress the thermoregulatory response. These findings suggest that the hibernating animal's thermostat is not simply switched on or off but is capable of a wide range of settings.

In addition to the alarm response to excessive cooling of the hypothalamus, the hibernating ground squirrel may exhibit proportional thermoregulatory heat-production responses without rousing from torpor when its hypothalamus is cooled within a range of temperatures just above the alarm threshold. Whereas the alarm response is characterized by an abrupt increase in the threshold temperature of the thermostat, there is no such change in threshold following the proportional responses. The thermosensitivity of the thermostat, as is indicated by the degree to which a change in body temperature elicits a thermoregulatory response, is much lower when the ground squirrel is hibernating than when it is active, but the temperature-sensitive characteristics of the thermostat are qualitatively the same in the active animal and in the hibernating one.

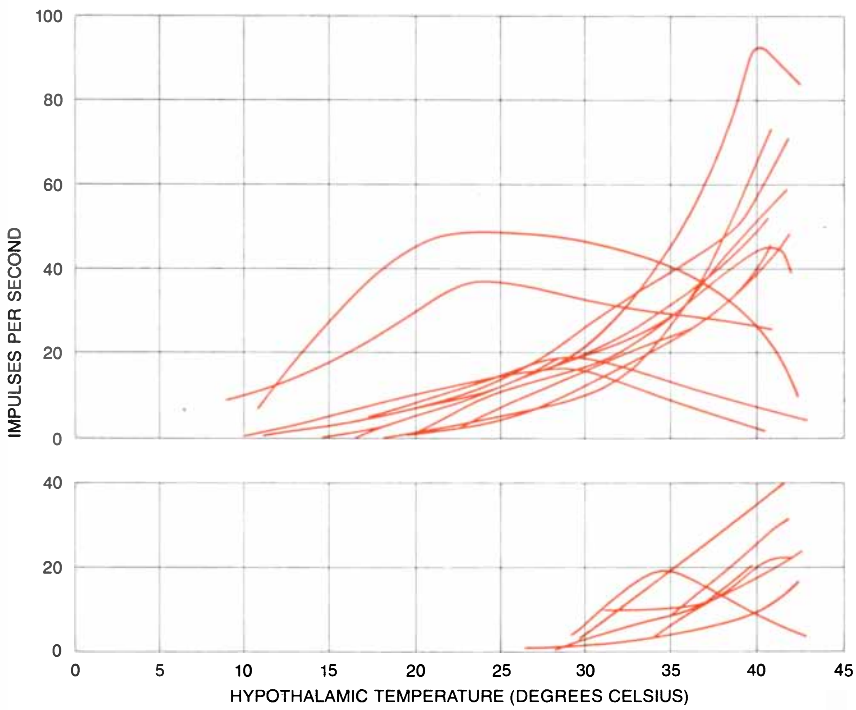
Could the same thermostat be operative over the entire range of body temperatures experienced by the hibernator? We tested this hypothesis by heating and cooling the hypothalamus for short periods while the animal was entering hibernation. At discrete time intervals during entrance we ascertained the highest hypothalamic temperature that induced a metabolic heat-production response and the lowest hypothalamic temperature that did not induce the response. These experiments demonstrated the continuity of the operation of the thermostat over the wide range of body temperatures experienced by the hibernator and showed that entrance into hibernation is associated with a progressive lowering of the thermostat.

When the animal is going into hibernation, the decline in body temperature and metabolic rate is not always smooth. There are sometimes bursts of metabolic heat production and shivering, which causes the decline in body temperature to stop or reverse slightly. Such bursts appear to occur when the temperature of the hypothalamus of the hibernating animal has fallen temporarily below the declining setting of the thermostat.

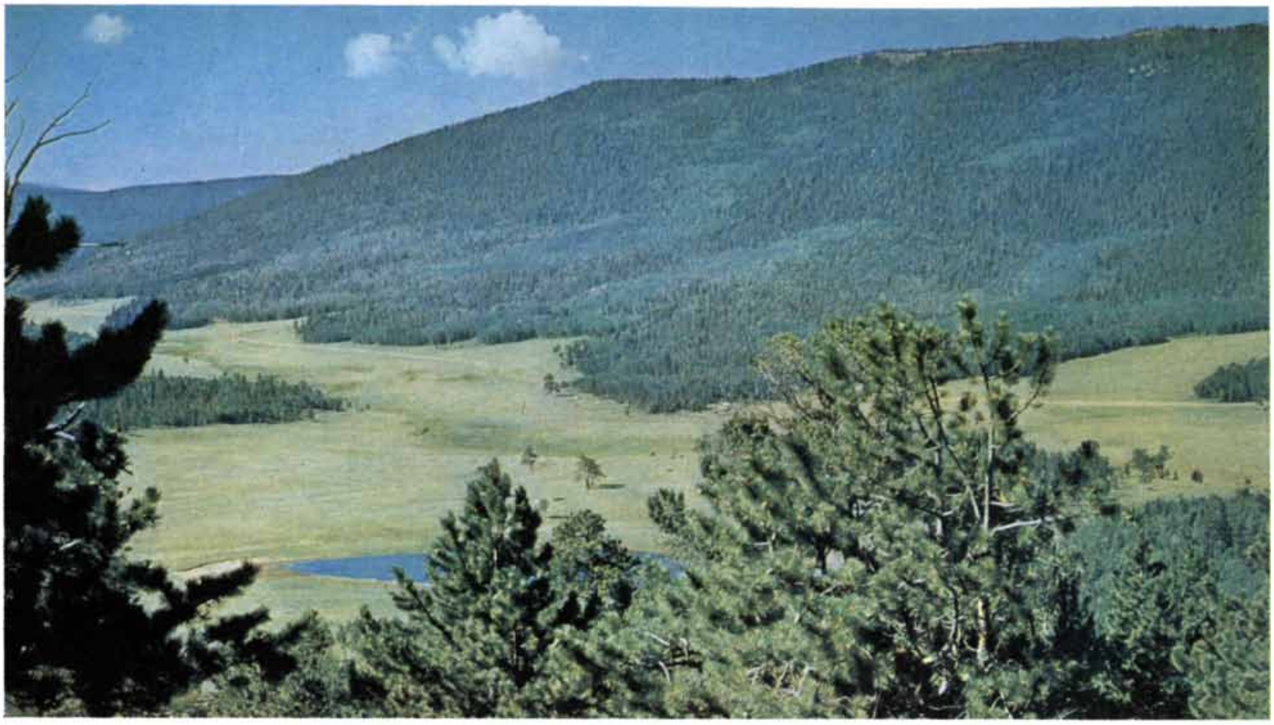
A variety of ectotherms also become torpid under certain conditions: snakes and lizards seek a cool burrow at the end of the day, the brook trout prefers cooler water in winter than it does in the spring and toads seek a cooler environment when they are deprived of food. Whether these animals are also experiencing regulated shifts in the setting of their central thermostat is not yet known.

The birds, which are also descended from reptilian ancestors, evolved endothermy in parallel with the mammals. What little is known about the avian thermostat suggests that it is somewhat different from the mammalian one. For one thing, although the hypothalamus of birds plays an important role in integrating the thermoregulatory responses, it is virtually insensitive to temperature. Instead temperature-sensitive cells in the spinal cord provide the major source of information about the temperature of the body core. Werner Rautenberg and his colleagues at the University of the Ruhr have shown that cooling the spinal cord of pigeons by pumping water through a very thin tube inserted into the spinal canal results in the heat-generating and heat-conserving responses of shivering and constriction of peripheral blood vessels. Warming the spinal cord induces dilation of the peripheral blood vessels and eventually panting. Manipulations of spinal-cord temperature in mammals also elicit thermoregulatory responses, which suggests that the major difference between the thermoregulatory system of birds and that of mammals is that the mammalian hypothalamus directly monitors the internal temperature of the body, whereas the avian hypothalamus receives most of its information about the thermal state of the body from temperature sensors in the spinal cord and perhaps elsewhere in the body.

Fishes too may integrate central and peripheral thermal information in the hypothalamus to control certain physiological responses. Because of the high thermal conductivity of water and the thermal heterogeneity of many bodies of water fishes often experience rapid changes in body temperature. Such changes result in fluctuations of many vital parameters such as metabolic rate. Although the effector mechanisms available to a fish cannot provide it with a constant internal temperature, they may be controlled to anticipate the physiological changes that inevitably accompany thermal changes. For example, a rise in a fish's body temperature means



TEMPERATURE-SENSITIVE NEURONS, or nerve cells, in the mammalian hypothalamus have been detected with the aid of implanted recording electrodes. These two graphs show the firing rate of individual neurons as a function of hypothalamic temperature in a mammal that hibernates, the golden hamster (*top*), and in a mammal that cannot hibernate, the guinea pig (*bottom*). The temperature-sensitive neurons in the hamster respond to a broad range of body temperatures that would be experienced during the entrance into hibernation, whereas the hypothalamic neurons in the guinea pig are silent below a temperature of 30 degrees C. (86 degrees F.). Experiments shown here were performed by Wolf Wünnenberg of the University of Kiel.



Own your own piece of the Rockies!

Five years ago the publishers of Forbes Magazine began offering for sale a portion of their 260-square-mile Trinchera Ranch in five acre minimum sections called Sangre de Cristo Ranches. More than 95% of that land has been purchased to date. Sangre's success exceeds all expectations.

Fifteen miles northeast of that area, deep in the heart of Forbes Trinchera country, is a spectacular alpine valley of unsurpassing beauty. This hidden mountain valley was once the hunting ground of Indians and later conquered by Conquistadores. It remains today almost as it was when man first saw it.

We've named it Forbes Park.

And now Forbes Inc., through its subsidiary Sangre de Cristo Ranches Inc., is inviting people who have a special appreciation for the beauty of nature and who want more out of life, to acquire a homesite in this unspoiled mountain valley.

Obtain the HUD property report from developer and read it before signing anything. HUD neither approves the merits of the offering nor the value, if any, of the property.

Cash prices range from $3,500 to $12,000. Easy credit terms are available.

Forbes Park is a completely private area surrounded by the San Isabel National Forest and the rangelands of the Trinchera Ranch. It is accessible year-round through a main gate restricting entrance to owners and their guests. Hunting is not permitted here, but many sports including fishing, cross-country skiing, boating, swimming, horseback riding, tennis and mountain climbing will be available in the recreational areas of the vast common acreage.

Starting immediately, you may purchase one or more acres for a mountain homesite where you will share with your neighbors the exclusive use of over 5,000 acres of common lands, including ponds, streams, woods, meadows and other recreational areas set aside for the enjoyment of Forbes Park land owners for all time.

For complete information, including pictures, maps, and full details on our liberal money-back and exchange privileges, without obligation, please fill in the coupon or card and mail to: Forbes Park, Box 303, Fort Garland, Colorado 81133.

Not an offer or solicitation in those states where the property is not registered

IL-78-205 AD15618(a) MI-77-361 OAD-77 000337A

FORBES PARK

Development of
Sangre de Cristo Ranches Inc.
Box 303, Fort Garland, Colorado 81133

Obtain the New Jersey Public Report and Broker's Release from the developer or New Jersey broker and read it before signing anything.

6178

Name ______________________

Address ______________________

City __________ **State** __________ **Zip** __________

Telephone ______________________

Preferences: ☐ **$3,500** ☐ **$5,000** ☐ **Higher**

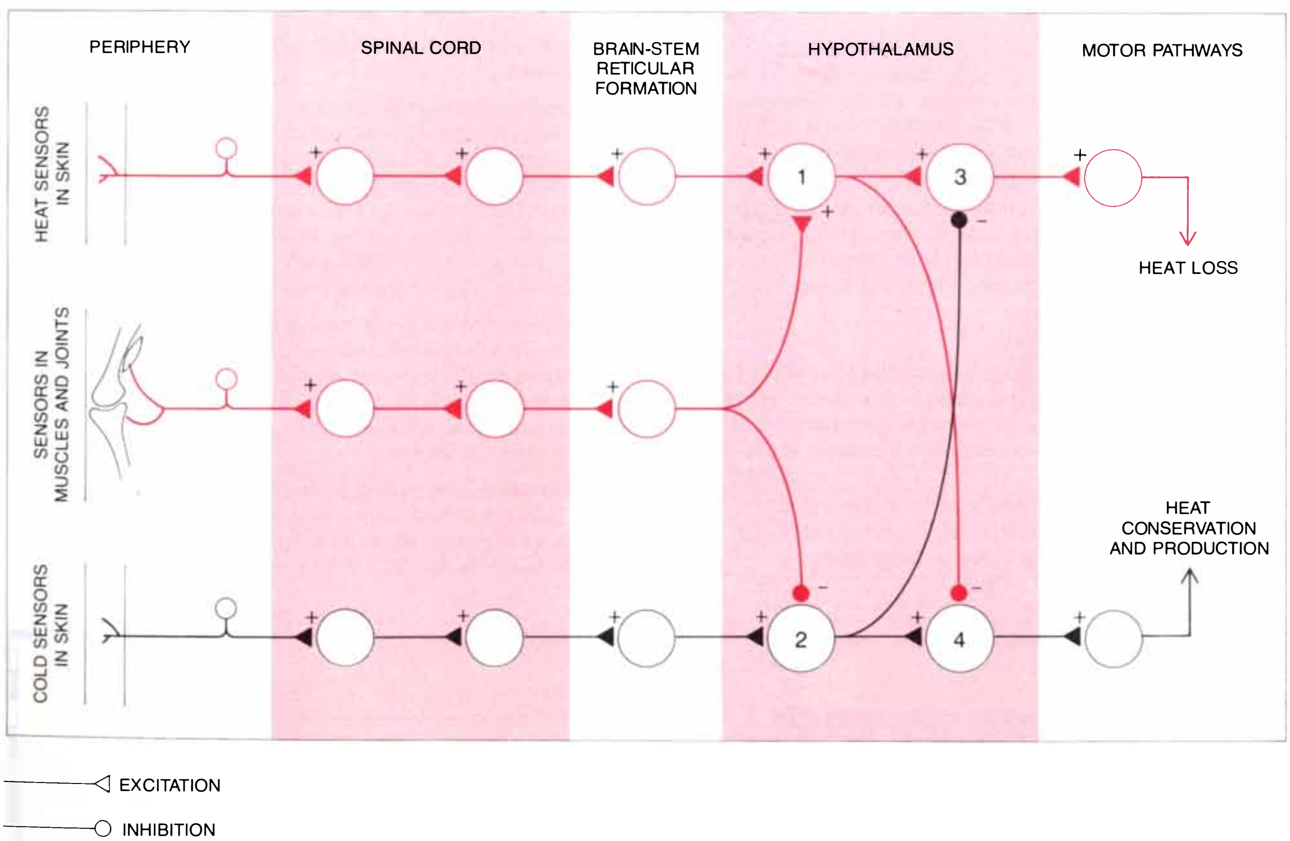
A statement and offering statement has been filed with the Secretary of State of the State of New York. The filing does not constitute approval of the sale or lease or offer for sale or lease by the Secretary of State or that the Secretary of State has in any way passed upon the merits of such offering. A copy of the offering statement is available, upon request, from Sangre de Cristo Ranches Inc. NYA77-432 an increase in its metabolic rate and hence in its oxygen requirement. The fish could either wait until an oxygen deficit occurs and then respond by increasing its gill ventilation, or it could use its temperature-sensing capabilities to relay projected changes in oxygen demand to the respiratory centers. In the latter case alterations in arterial oxygen concentration would be avoided, ensuring a constant supply of oxygen to the tissues.

One of us (Crawshaw) studied the respiratory response in the scorpion fish by manipulating the temperature of its hypothalamus with the aid of implanted thermodes. Heating the hypothalamus caused the fish to ventilate its gills faster, whereas cooling the hypothalamus caused it to ventilate them more slowly. Gill ventilation also appears to be influenced by thermal input from the periphery. On encountering a rapid increase in water temperature carp quickly increase the ventilation of their gills for about 30 seconds; the response then diminishes. Thus fishes appear to utilize both central and peripheral thermal information to project changes in oxygen demand. This regulatory system may well be the evolutionary antecedent of the physiological control system in endotherms that maintains internal body temperature at close to optimal levels.

On the basis of what is now understood about the characteristics of the vertebrate thermostat, it is possible to speculate on its neural design. Neurophysiological investigations have revealed the existence of neurons, or nerve cells, in the hypothalamus that could serve as components of a thermostat. Two populations of thermosensitive neurons have been identified: one that responds to a local warming of the brain tissue and another that responds to a local cooling. The heat-sensitive neurons increase the frequency of their discharge of impulses in direct proportion to the degree that hypothalamic temperature is raised above the normal value for the animal's body. Similarly, the cold-sensitive neurons respond with an increase in the frequency of impulses when the hypothalamic temperature falls below the normal value. It is not yet known whether such hypothalamic neurons are really part of the central thermostat, but this assumption is supported by studies showing that these neurons sometimes also respond to changes in the temperature of the skin and of the spinal cord.

Wolf Wünnenberg and his colleagues at the University of Kiel have shown that the hypothalamic temperature-sensitive neurons in a hibernating mammal and those in a nonhibernating mammal differ in their range of temperature sensitivity, as would be expected if the neurons are actually components of the thermostat. The hypothalamic neurons of the nonhibernator had a narrow range of temperature sensitivity and were mostly silent below a temperature of 30 degrees C. (86 degrees F.), whereas many hypothalamic neurons of the hibernator had continuous temperature-response curves over a much broader range.

A simple model proposed by one of us (Hammel) suggests how neurons in the hypothalamus could be interconnected to achieve the regulation of body temperature. According to the model, when the activity of the heat-sensitive



THEORETICAL MODEL outlined here shows how neurons in the central nervous system might be interconnected to result in a functional thermostat. Four basic neuronal populations thought to exist in the hypothalamus are represented diagrammatically by single neurons. Populations *1* and *2* are respectively sensitive to warming or cooling. These temperature-sensitive populations facilitate and inhibit populations *3* and *4*, which serve to trigger thermoregulatory responses. The relative firing rates of populations *1* and *2* determine the hypothalamic threshold temperatures for particular thermoregulatory responses. These thresholds can be modified by diverse neural inputs: information from peripheral temperature sensors in the skin, from movement and position sensors in the muscles and joints, from temperature-sensitive neurons in the spinal cord and from neurons in the reticular formation of brain stem that control sleep and wakefulness.

neurons in the hypothalamus prevails over the activity of the cold-sensitive ones, heat-loss mechanisms are activated. Conversely, when the activity of the cold-sensitive neurons predominates, heat-generating and heat-conservation measures are activated. The thermoregulatory thresholds of the hypothalamus are determined by the relative activity of the two neuronal populations.

As we have seen, under most conditions the body temperature of an endothermic animal remains constant; thermoregulatory responses are evoked by rapid changes in peripheral temperature. How is the thermal information from the temperature sensors in the skin integrated by the central thermostat? One hypothesis is that heat-sensitive nerve endings in the skin increase their firing rate as the ambient temperature increases and thereby excite the heat-sensitive neurons in the hypothalamus. The cold-sensitive nerve endings could innervate the central cold-sensitive neurons in a similar fashion. Hence if you walk out of a warm room into a cold environment, the peripheral cold sensors could increase the firing rate of the cold-sensitive neurons in the hypothalamus, moving the point of intersection with the population of heat-sensitive neurons to a higher temperature. Because this new intersection is above the actual hypothalamic temperature of 37 degrees C. the peripheral cold stimulus triggers heat-generating responses without requiring a change in hypothalamic temperature. This simple model can therefore account for the predictive capability of the thermostat.

The hypothalamus receives many inputs besides those from peripheral temperature sensors in the skin, and any detailed model of the thermostat will have to account for the integration of thermal signals from the spinal cord and the abdominal viscera, for inputs from tension receptors in the muscles and joints (which signal the level of physical exercise) and for inputs from the reticular formation of the brain stem (which is an important part of the neural system controlling the animal's level of arousal). All these diverse signals are integrated by the hypothalamus and provide information about the temperature of the body core and about the more rapid thermal changes occurring at the body surface.

The value of comparative studies of a wide variety of vertebrate species facing very different thermal problems and having diverse strategies for dealing with them is that they illuminate the general features of the vertebrate thermoregulatory system as well as its specialized features. Continued investigations should enable us to propose, test and refine hypothetical models so that they will better reflect how the thermostat is really designed.

WITH WHAT MINOLTA KNOWS ABOUT CAMERAS AND WHAT YOU KNOW ABOUT YOURSELF, WE CAN MAKE BEAUTIFUL PICTURES TOGETHER.

If you've considered buying a 35mm single lens reflex camera, you may have wondered how to find the right one out of the bewildering array of models and features available.

And with good reason, since the camera you choose will have a lot to do with how creative and rewarding your photography will be.

What you pay for your camera shouldn't be your only consideration, especially since there are some very expensive cameras that won't give you some of the features you really need. So ask yourself how you'll be using the camera and what kind of pictures you'll be taking. Your answers could save a lot of money.

How automatic should your camera be?

Basically, there are two kinds of automatic 35mm SLR's. Both use advanced electronics to give you perfectly exposed pictures with point, focus and shoot simplicity. The difference is in creative control.

For landscapes, still lifes, portraits and the like, you'll want an *aperture-priority* camera. It lets *you* set the lens opening, while *it* sets the shutter speed automatically.

This way, you control depth-of-field. That's the area of sharpness in front of and behind your subject. Many pro photographers believe that depth-of-field is the most important factor in creative photography.

At times you may want to control the motion of your subject. You can do this with an aperture-priority camera by changing the lens opening until the camera sets the shutter speed necessary to freeze or blur a moving subject. Or you can use a *shutter-priority* camera, on which you set the shutter speed first and the camera sets the lens automatically.

Minolta makes both types of automatic cameras. The Minolta XG-7 is moderately priced and offers aperture-priority automation, plus fully manual control. The Minolta XD-11 is somewhat more expensive, but it offers all the creative flexibility of both aperture and shutter-priority automation, plus full manual control. The XD-11 is so advanced that during shutter-priority operation it will actually make exposure corrections you fail to make.

Do you really need an automatic camera?

Automation makes fine photography easier. But if you do some of the work yourself, you can save a lot of money and get pictures every bit as good.

In this case, you might consider a Minolta SR-T. These are semi-automatic cameras. They have built-in, through-the-lens metering systems that tell you exactly how to set the lens and shutter for perfect exposure. You just align two indicators in the viewfinder.

What to expect when you look into the camera's viewfinder.

The finder should give you a clear, bright view of your subject. Not just in the center, but even along the edges and in the corners. Minolta SLR's have bright finders, so that composing and focusing are effortless, even in dim light. And focusing aids in Minolta

Minolta makes all kinds of 35mm SLR's, so our main concern is that you get exactly the right camera for your needs. Whether that means the advanced Minolta XD-11. Or the easy-to-use and moderately priced Minolta XG-7. Or the very economical Minolta SR-T cameras.

Automatic sequence photography is easy when you combine a Minolta XD-11 or XG-7 with optional Auto Winder and Electroflash 200X.

viewfinders make it easy to take critically sharp pictures.

Information is another thing you can expect to find in a well-designed finder. Everything you need to know for a perfect picture is right there in a Minolta finder.

In the Minolta XD-11 and XG-7, red light emitting diodes tell you what lens opening or shutter speed is being set automatically and warn against under or over-exposure. In Minolta SR-T cameras, two pointers come together as you adjust the lens and shutter for correct exposure.

Do you need an auto winder?

You do if you like the idea of sequence photography, or simply want the luxury of power assisted film advancing. Minolta auto winders will advance one picture at a time, or continuously at about two per second. With advantages not found in others, like up to 50% more pictures with a set of batteries and easy attachment to the camera without removing any caps. Optional auto winders are available for both the Minolta XD-11 and XG-7, but not for Minolta SR-T cameras.

How about electronic flash?

An automatic electronic flash can be added to any Minolta SLR for easy, just about foolproof indoor photography without the bother of flashbulbs. For the XD-11 and XG-7, Minolta makes the Auto Electroflash 200X. It sets itself automatically for flash exposure, and it sets the camera automatically for use with flash. An LED in the viewfinder signals when the 200X is ready to fire. Most unusual: the Auto Electroflash 200X can fire continuously in perfect synchronization with Minolta auto winders. Imagine being able to take a sequence of 36 flash pictures without ever taking your finger off the button.

You should be comfortable with your camera.

The way a camera feels in your hands can make a big difference in the way you take pictures.

The Minolta XD-11 and XG-7, for instance, are compact, but not cramped. Lightweight, but with a solid feeling of quality. Oversized controls are positioned so that your fingers fall naturally into place. And their electronically controlled shutters are incredibly smooth and quiet.

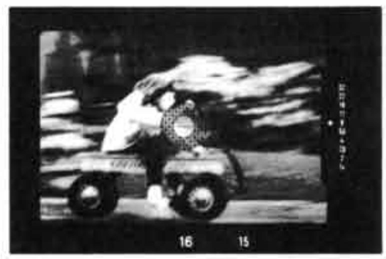
Minolta SR-T's give you the heft and weight of a slightly larger camera, but with no sacrifice in handling convenience. As in all Minolta SLR's, "human engineering" insures smooth, effortless operation.

Are extra features important?

If you use them, there are a lot of extras that can make your photography more creative and convenient. Depending on the Minolta model you choose, you can get: multiple exposures with pushbutton ease (even with an auto winder). A window to show that film is advancing properly. A handy memo holder that holds the end of a film box to remind you of what film you're using. And a self-timer.

What about the lens system?

The SLR you buy should have a system of lenses big enough to satisfy your needs, not only today, but five years from today. The patented Minolta bayonet mount lets you change lenses with less than a quarter turn. There are almost 40 Minolta lenses available, ranging from 7.5mm fisheye to 1600mm super-telephoto, including macro and zoom lenses and the world's smallest 500mm lens.



The electronic viewfinder: LED's tell you what the camera is doing automatically to give you correct exposure.

The match-needle viewfinder: just align two indicators for correct exposure. Because you're doing some of the work, you can save some money.

What's next?

Think about how you'll use your camera and ask your photo dealer to let you try a Minolta. Compare it with other cameras in its price range. You'll soon see why more Americans buy Minolta than any other brand of SLR. For literature, write Minolta Corp., 101 Williams Drive, Ramsey, New Jersey 07446. In Canada: Minolta Camera (Canada) Inc., Ontario.

Specifications subject to change without notice.

MINOLTA

WE WANT YOU TO HAVE THE RIGHT CAMERA.

The Tekite Problem

Did these peculiar bits of glass originate on the earth or on the moon, in volcanic activity or in meteorite impact? The evidence is conflicting and the most reasonable answer is a surprising one

by John A. O'Keefe

In widely scattered locations around the world small glassy pebbles of curious surface configuration and distinctive composition are found, often in vast "strewn fields" that stretch across land masses and, on the evidence of cores drilled from the deep-sea floor, across intervening seas and oceans. These tektites, as they are called from the Greek word *tektos,* meaning molten, are much like the volcanic glass obsidian, but their chemical composition is different from that of any terrestrial lava and they contain far less water and none of obsidian's characteristic microcrystals. And, as the Austrian mineralogist Josef Mayer pointed out in 1787, no one has ever found the *Mutterstein,* or mother lode, of a field of tektites. They cannot, therefore, be the product of terrestrial volcanism.

Many of the most significant findings about tektites are negative findings, and the absence of a mother lode is one of them. Contrary to proposals advanced at one time or another, they are not man-made glass objects; neither are they the remains of some former planet or bits of antimatter or fulgurites fused by lightning from terrestrial sand. Recently acquired knowledge about the moon's surface confirms earlier indications that tektites cannot be bits of lunar soil propelled to the earth by the impact of meteorites on the moon. Painstaking evaluation of the evidence collected over the past century leaves two possibilities for the origin of tektites. One is that tektites are bits of terrestrial sedimentary rock excavated by meteorites striking the earth's surface, melted by the heat of impact and congealed into glass as they travel above the atmosphere to the scattered sites where they are found. The other possibility is that tektites are the remains of gobs of lava fired at the earth by volcanic activity on the moon.

If tektites are terrestrial, it means that some process exists by which soil or common rocks can be converted in an instant into homogeneous, water-free, bubble-free glass and be propelled thousands of miles above the atmosphere. If tektites come from the moon, it seems to follow that there is at least one powerful volcano somewhere on the moon that has erupted at least as recently as 750,000 years ago. Neither possibility is easy to accept. Yet one of them must be accepted, and I believe it is feasible to pick the more reasonable one by rejecting the more unlikely.

In 1975 P. V. Florensky of the Institute of Geology in the U.S.S.R. Academy of Sciences announced the discovery of tektites at the site of the Zhamanshin structure, a shallow crater near the town of Irgiz in Kazakhstan. He called them irghizites, following the custom of naming groups of tektites for the place where they are found. This first major tektite discovery in about 25 years has already proved to be highly instructive. The crater where the irghizites were found is a shallow depression about five kilometers in diameter. It is filled, according to Florensky, with from 100 to 150 meters of lake sediments on top of broken rock. The crater shows strong evidence of impact by an object from space. Some of the local rocks have been shocked, according to Florensky's colleague Yu. P. Dikov, yielding the impact glasses coesite and stishovite: high-pressure forms of silica (silicon dioxide, SiO_2). Such a close association of tektites with an impact crater is not unique; other instances are the Aouelloul tektites in Mauritania and perhaps the "Darwin glass" of Tasmania.

Either the tektites made the Zhamanshin crater or the crater made the tektites. A large block of tektite glass could have come out of space, dug the crater and disintegrated to form the irghizites; alternatively, a meteorite could have landed at the site, forming the crater and at the same time transforming some of the local rock into tektite glass, perhaps by boiling away some of its more volatile constituents. The two hypotheses are being tested by investigators in the U.S.S.R. and the U.S., who are cooperating to compare the tektites (samples of which have been lent by the Institute of Geology to laboratories in the U.S.) with the local rock and with tektites from other sites.

The irghizites are small black objects averaging about half a gram in weight. (Most tektites weigh a few grams.) They are warty, twisted objects that resemble the Aouelloul and Darwin-glass tektites in shape. Their chemical composition, according to Kurt Fredriksson of the Smithsonian Institution, is remarkably uniform from specimen to specimen; it is unlike that of any local rocks and is very similar to that of the tektites found in Java. The resemblance to the javanites extends to the trace elements, it has been reported by Yu. F. Pogrebnyak in the U.S.S.R. and by William D. Ehmann, John A. Philpotts, C. S. Annell, John W. Morgan and other investigators in the U.S. Like all other tektites, the irghizites are homogeneous glasses lacking even small crystals. On the face of it, the evidence is conflicting. The compositional studies seem to support the idea of an extraterrestrial origin, but after examining the samples, Florensky has accepted the hypothesis of meteorite impact.

The ultimate findings concerning this new discovery will have to be fitted into a matrix of facts and calculations accumulated over the years. The javanites are members of a huge family of tektites, perhaps 100 million tons of them, that fell about 750,000 years ago over the Indian Ocean, southern China, southeastern Asia, Indonesia, the Philippines and Australia. The extent of this Australasian strewn field has been defined by the discovery of tektites at one site after another ever since a specimen of one of its most distinctive forms was given to Charles Darwin during the voyage of the *Beagle.* The tektites have been dated to 750,000 years ago by isotope analysis (determining the proportions of potassium 40 and of argon 40, to which potassium 40 decays radioactively), by fission-track analysis (counting the tracks left by the natural fission of uranium) and by the presence of microtektites in sea-floor cores at the level associated with a particular reversal in the earth's magnetic field and in conjunction with particular marine microfossils. (Attempts to date dry-land specimens by stratigraphy have yielded inconsistent results even in the hands of very competent geologists; I think it is because of the inherent difficulty of applying stratigraphy to date fairly recent events.)

In contrast to the contorted tektites found at impact craters, the Australasian tektites are roundish or chunky objects, often shaped like teardrops, dumbbells or disks, that show little evidence of internal strain. Most of them are very similar in composition, and there are subgroups of chemically identical specimens whose discovery sites form long streaks within the overall field. Everywhere the tektites are chemically distinct from the very diverse rocks on which they are found. All of this makes it clear that the Australasian strewn field's tektites originated at a single point, whether on the earth or in space. Some launching mechanism must have fired them from that initial point to their scattered resting places on the earth. The search for the origin of tektites is therefore a search for a launching mechanism that could spread this gigantic array over an area whose longest dimension is some 11,000 kilometers and whose extent is perhaps a tenth of the earth's surface.

Similar considerations apply to two other vast fields of tektites, of which the larger is the North American strewn field. It is the oldest field known, dating from about 35 million years ago, and it

MOLDAVITE, a tektite from Czechoslovakia, glows in a photograph made by Joseph Walters of the Goddard Space Flight Center of the National Aeronautics and Space Administration. The moldavites are found in adjacent "strewn fields" in Bohemia and Moravia; they are named for the Moldau River. This specimen from Koroseky in Bohemia, whose diameter is about 1.5 inches, has the characteristic green color of a Bohemian tektite. Moldavites were once thought to be bits of artificial glass because the region was an early glassmaking center.

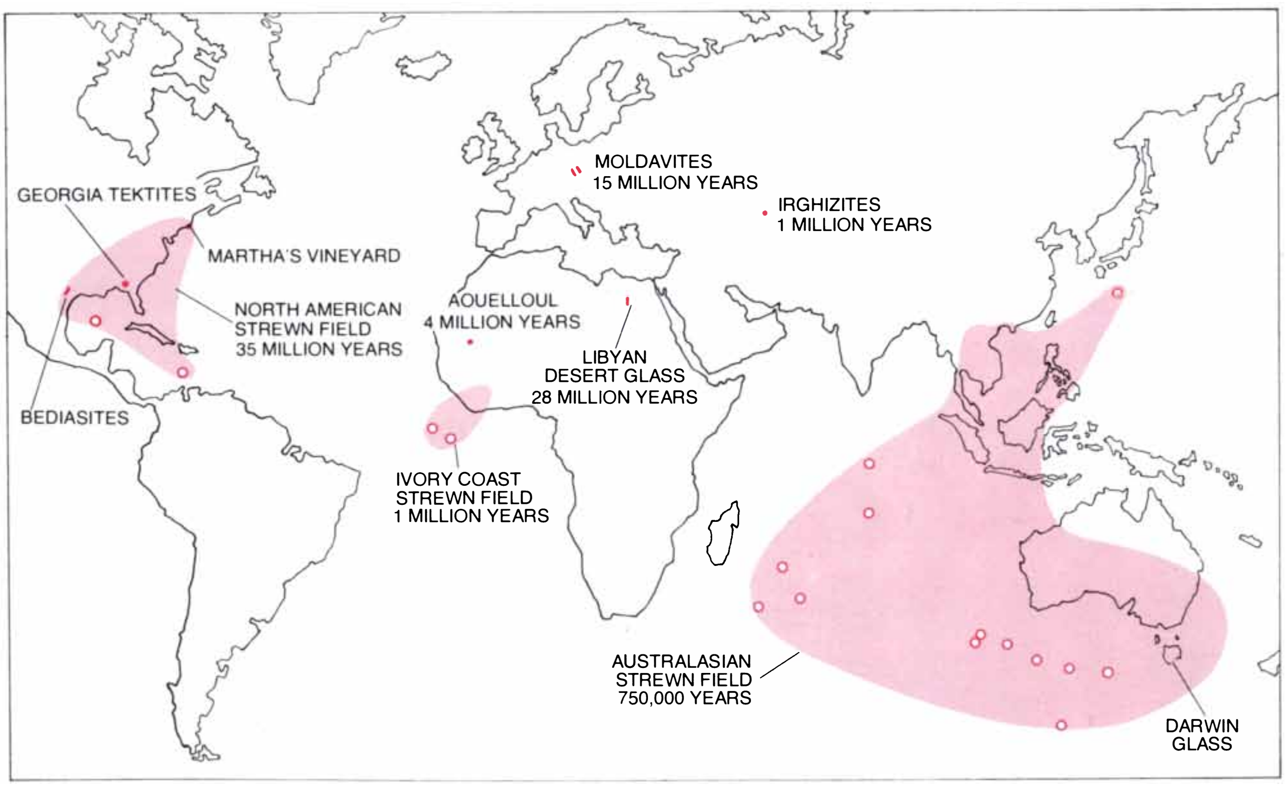
must have had an aggregate mass of about a billion tons, much of which has long since been eroded away. When the field was formed, it extended at least from the southern coast of the U.S. across the Gulf of Mexico and the Caribbean Sea to a point near Caracas in Venezuela. The Ivory Coast strewn field, about a million years old, extends from the Ivory Coast into the South Atlantic and has a mass of perhaps 10 million tons. In each of these large fields the tektites are shaped not like the irghizites but like the Australasian tektites. There are a number of smaller fields and individual finds.

So much for the overall extent of the fields. Let me now summarize the reasoning that ultimately led to the two possibilities I outlined at the beginning of this article. Harold C. Urey of the University of California at San Diego has pointed out that the size of the strewn fields proves tektites do not come from space beyond the moon. A cloud of separate objects approaching the earth from space would (because of the differential effects of solar attraction) become so large and diffuse that it would of necessity cover an entire hemisphere of the earth, as a meteor shower does. On the other hand, the breakup within the earth's atmosphere of a single body would yield too small a distribution—a few tens of kilometers in the longest direction, as in the case of the fall of a meteorite. Further evidence against prolonged travel of a cloud of bodies in space comes from analysis of the tektites. They lack the usual signs of interaction with primary cosmic rays in space: the presence of neon 21, aluminum 26 or beryllium 10 or of the characteristic *V*-shaped cosmic-ray tracks made visible by etching with acid. The source, then, must be on the earth or on the moon.

The only mechanisms anyone considers capable of producing the necessary launching velocities are volcanism and meteorite impact. Winds cannot do it because even millimeter-size microtektites would fall out in a few minutes, which is too short a travel time to account for even a modest strewn field. People do move tektites around, but not hundreds of millions of tons of them, and not into the middle of the Indian Ocean 750,000 years ago. Given two conceivable launching mechanisms and two possible source locations, there are four origins to be considered: terrestrial volcanism, lunar volcanism, terrestrial meteorite impact and lunar meteorite impact.

One of the four origins has long been out of favor: terrestrial volcanism. There are two arguments against it. The first, to which I alluded above, is that the chemical composition of tektites is not the composition of terrestrial volcanic glasses. A typical tektite has a silica content of something over 68 percent; a volcanic glass with such a high silica content is defined as granitic. There are terrestrial granitic glasses, but they have a composition that can be represented as nearly equal amounts of three minerals: quartz (SiO_2 considered as a mineral), soda feldspar ($NaAlSi_3O_8$) and potash feldspar ($KAlSi_3O_8$). In addition no more than 20 percent consists of other minerals.

The reason is clear: for silicates this composition is the ternary (three-way) eutectic, the composition that crystallizes at the lowest temperature. As a somewhat basaltic magma cools, various minerals will crystallize out, each at the appropriate temperature and composition, until the melt reaches this eutectic composition. The liquid phase now has the composition of a granite. If the liquid is somehow separated from the crystals, it may be intruded and cool slowly to form a crystalline rock (granite), or it may erupt and be cooled instantaneously to form a granitic glass (such as obsidian). As an example, the "standard granite," designated G-1, has a mineral composition equivalent to 29



DISTRIBUTION OF TEKTITES is not random. They are concentrated in a few large strewn fields (*light color*) and in smaller groups or as individual finds (*colored streaks or dots*). Microtektites have been found in deep-sea cores (*open circles*), which help to establish the extent of the major fields. The ages of the tektites in each group, which have been established by various techniques, are shown on the map.

percent quartz, 32 percent soda feldspar and 28 percent potash feldspar, with the remaining 11 percent consisting of other minerals.

Not so the tektites. For them the equivalent mineral proportions are typically about 40 percent quartz, 13 percent soda feldspar and 15 percent potash feldspar, with 32 percent consisting of other minerals. Stating it another way, tektites contain more divalent oxides—lime (CaO), magnesia (MgO) and ferrous oxide (FeO)—than terrestrial rocks of similar silica content, and less monovalent oxides: soda (Na_2O) and potash (K_2O).

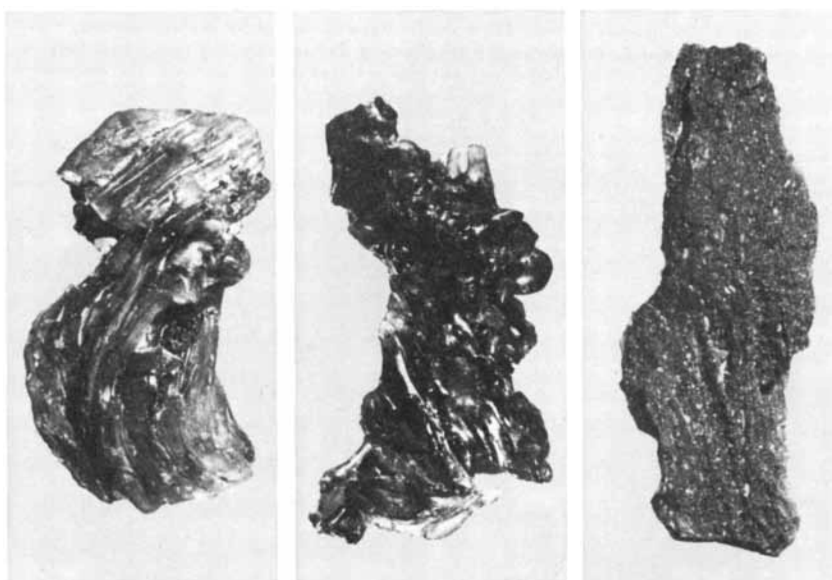
The other argument against terrestrial volcanoes is that they cannot provide enough velocity. They are powered by steam, and at the temperature of a lava (no more than 1,200 degrees Celsius) the velocity of sound in steam is about one kilometer per second. Thermodynamic theory shows that for a gas escaping from an enclosure there is a limiting velocity of about twice the velocity of sound in the gas, or about two kilometers per second for steam. (Higher temperatures would allow higher velocities, but the temperature of a lava is limited by the fact that the lava must be contained in a crucible of rock; if the temperature rises much above the melting point of the rock, the walls will melt and cool the lava.) The maximum "muzzle velocity" that is actually observed for objects thrown from volcanoes is only about 700 meters per second. Yet the velocity required to reach the edges of the Australasian strewn field from a point at the center of the field would be a little more than six kilometers per second! The compositional argument and the deficiency in launching velocity certainly justify eliminating terrestrial volcanoes as a source of tektites.

Of the three remaining possibilities one, meteorite impact on the moon, seems to have been eliminated by the chemical data collected by the Surveyor, Apollo and Luna landings and by remote sensors that have surveyed much of the lunar surface. It seems clear that most of the surface consists of one or another kind of basaltic rock, or rock that is less than 55 percent silica. The lunar maria, or "seas," have iron-rich basalts and the lunar highlands have feldspar-rich basalts. Granitic rock is found only as fragments—including glassy ones—in the soils and in some breccias (rocks composed of fragments cemented together). From this it is deduced that granitic masses do exist, but they must either be small or be in limited areas that have not yet been observed. Almost all the lunar surface soil and rock is basaltic and cannot be the stuff of tektites.

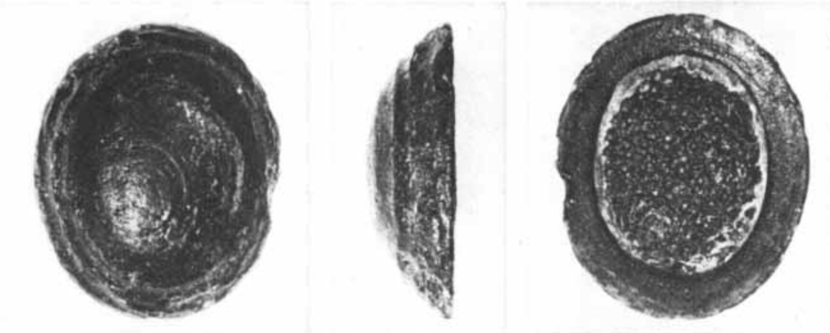
That being the case, impact on the moon would seem to be eliminated as a source of tektites because impact is an undiscriminating process: an object from space is as likely to land in one place as in another. It is true that lunar rock fired at the earth by a meteorite impact would be subjected to some filtering action favoring granitic glasses, which would survive shock, entry into the earth's atmosphere and weathering better than basalts, but the differences do not seem large enough to account for the silicic nature of tektites; indeed, the finding of low-silica microtektites indicates that low-silica material can survive all these perils. (Even those microtektites, however, are not like the lunar basalts.)

It is not only tektites that fail to match the lunar basalts; there are no meteorites that match them either. It used to be assumed that some meteorites striking the earth had been splashed out of the lunar surface by an impact on the moon. Edward Anders of the University of Chicago and his co-workers noted, however, that the principal types of lunar rock that have been studied are unlike anything in the world's meteorite collections. It seems to follow, they pointed out, that a meteorite impact cannot eject matter from an object in the solar system (whether the moon, a planet or an asteroid) at a velocity as great as the lunar escape velocity, or 2.5 kilometers per second. This 2.5-kilometer limit, which I shall call Anders' limit, has become a cornerstone of post-Apollo meteorite theory. It means that meteorites cannot originate on Mars or Mercu-

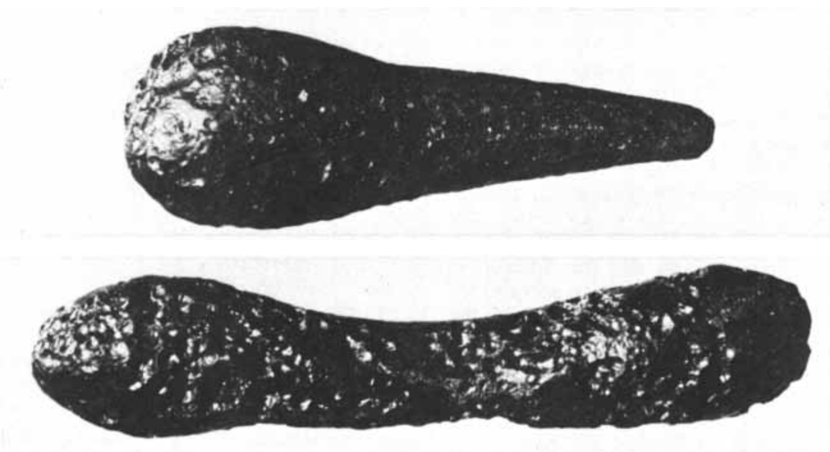
IRGHIZITES, from Kazakhstan in the U.S.S.R., are the most recently discovered tektites. These specimens, which range from about .8 inch to 1.1 inches long (portions were removed from two of them for analysis), were made available to U.S. investigators by Institute of Geology in U.S.S.R. Academy of Sciences through P. V. Florensky, who first reported on them.



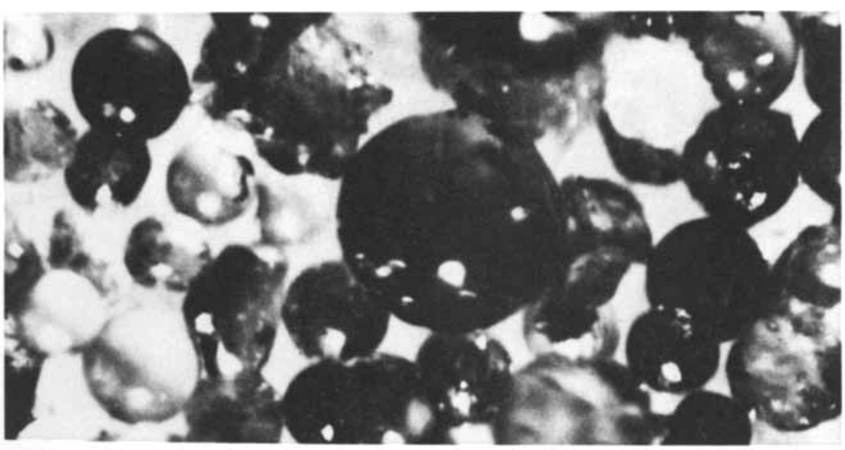
CONTORTED STRUCTURE of many irghizites (such as those in the photograph at the top of the page) is similar to that of some "Darwin glass" tektites from Tasmania (*two at left*) and tektites from Aouelloul in Mauritania (*right*). Like the irghizites, the Aouelloul tektites are associated with a nearby impact crater; the Darwin glass also appears to be. These specimens, two to 2.4 inches long, were collected by Robert F. Fudali of the Smithsonian Institution.



AUSTRALITE, a tektite from Australia collected by Charles Darwin during the voyage of the *Beagle* and now at the Institute of Geological Sciences in London, is seen from the front (*left*), side (*middle*) and rear (*right*); the long dimension is just over an inch. The australites' curious flanged shape is the result of aerodynamic ablation on entry into the earth's atmosphere.



TEARDROP AND DUMBBELL are typical "splash-form" tektites, suggestive of congealed drops of a viscous liquid. These tektites from Thailand are at the Smithsonian. The dumbbell is about 4.5 inches long. Such tektites are thought to have derived their shape from the breakup of a larger molten jet. Most tektites of this type are more massive than irghizites and other tektites associated with impact craters. Their shape suggests control by surface tension, which implies a rather stress-free environment during cooling. Presumably they tumbled as they entered the atmosphere, so that ablation was not concentrated on one face as in the australites.



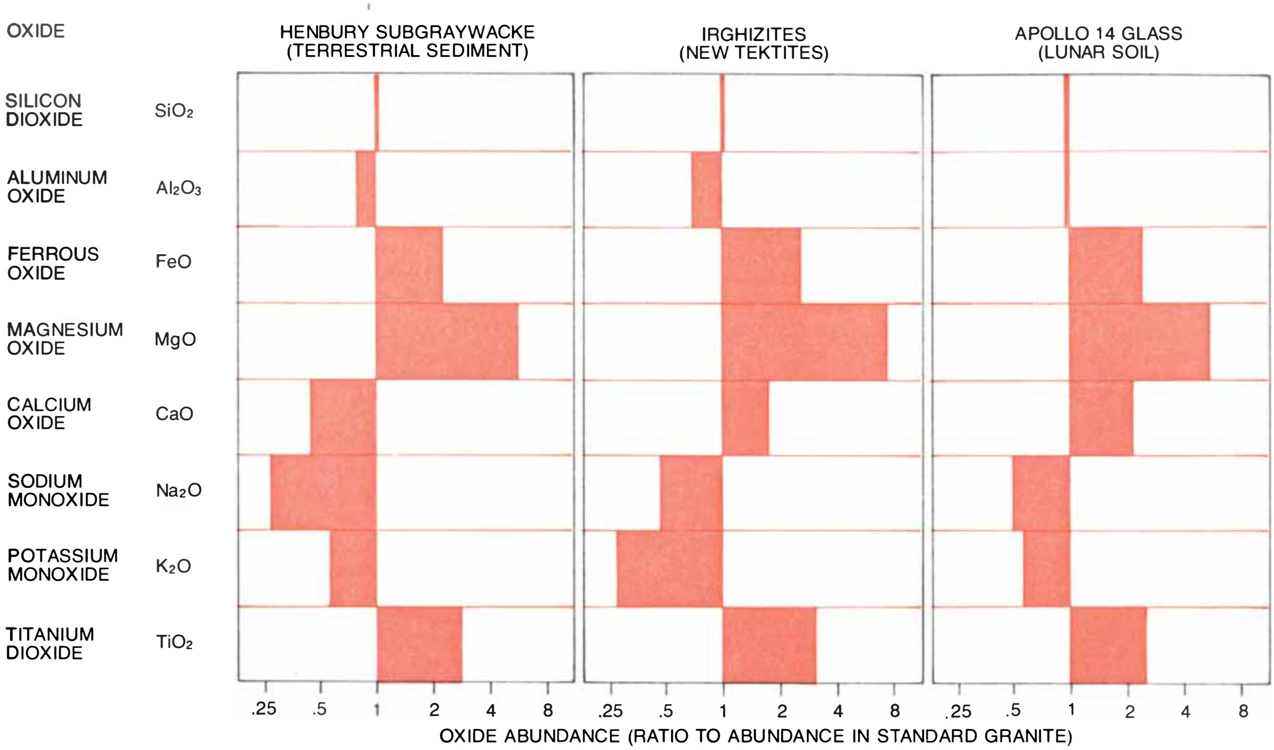
MICROTEKTITES in this photomicrograph were discovered by Billy P. Glass of the University of Delaware and his colleagues in a core from the floor of the Caribbean Sea near the island of Curaçao. Their composition is consistent with that of other tektites in the North American strewn field; stratigraphic and fission-track evidence dates them in the late Eocene epoch, some 35 million years ago. The actual size of the large dark microtektite (*center*) is .25 millimeter.

ry. It means that meteorites cannot be splashed directly from an asteroid in the asteroid belt to the earth: although the initial velocity needed for them to escape from a small asteroid would be insignificant, the velocity change required to put the asteroid material into an orbit intersecting the earth would be about five kilometers per second. And Anders' limit is certainly a further, and final, argument against meteorite impact on the moon as the source of tektites.

The logic does not stop there, however. We must face the fact that Anders' limit also excludes, even more strongly, the possibility that tektites originate from terrestrial meteorite impact! Velocities of from four to six kilometers per second are required by the geographic extent of the large strewn fields, even if one disregards atmospheric resistance.

The elimination of both impact hypotheses is surprising. In the 1960's almost all investigators were convinced—certainly I was—that tektites were generated by impact, whether on the earth or on the moon. What convinced us was the discovery by Edward C.-T. Chao of the U.S. Geological Survey of nickel-iron spherules in certain tektites. Nickel-iron is the characteristic component of meteorites, and such spherules are found in nontektite glasses formed by meteorite impact. The spherules do testify against terrestrial volcanism, since unoxidized iron is rare in lavas. Given the relative frequency of nickel-iron spherules in the Apollo lunar samples, it is now hard to see why, if the spherules are associated with lunar impact, they are so rare. And with the advantage of hindsight it seems strange that they tend to be associated with a particular kind of tektite glass and that their nickel content is low—much lower than that of meteorites. The low nickel content is hard to reconcile with a terrestrial origin for a particular reason. Robin Brett of the Geological Survey has pointed out that the partial loss of iron from a spherule in terrestrial impact glass by oxidation usually results in a spherule with a very high nickel content (and with an iron-rich halo around it, which is not seen in tektites). There is evidence from the Apollo samples that some low-nickel spherules are generated internally by the moon; perhaps the spherules in tektites really point to lunar volcanism.

There are several attractive aspects to the hypothesis that tektites have been fired at the earth by a volcano on the moon. In the first place the inadequate-velocity argument, which knocks out terrestrial volcanism as well as lunar and (I believe) terrestrial meteorite impact, is answered. Lunar rock is so avid for oxygen—so "reducing"—that a lunar volcano seems likely to be powered by



TEKTITES, LUNAR GLASS AND SANDSTONE are remarkably similar in their major-element composition. The bars show how each of the three materials differs from standard granite (G-1) in its content of major oxides. Tektite composition is somewhat closer to that of the lunar glass represented here, certain glassy *Apollo 14* particles, than to that of the Henbury subgraywacke, a terrestrial sandstone.

hydrogen rather than by steam. More specifically, the equilibrium ratio of hydrogen to steam in contact with lunar basalt (or tektite glass) at 1,200 degrees C. should be four to one. The acoustic velocity in that mixture of gases is three kilometers per second. Particle velocities of up to six kilometers per second are therefore possible in principle. The lunar escape velocity is only 2.5 kilometers per second. A hydrogen-powered volcano should be able to attain that velocity, particularly in the absence of an atmosphere.

The lunar-volcano hypothesis also survives the discovery that the moon's surface is chiefly basaltic, the finding that eliminates lunar impact as an origin because tektites are typically silicic in composition. The kind of explosive volcanism required to launch the tektites is associated (on the earth at least) with silicic volcanoes. If tektites are propelled by lunar volcanism, it is reasonable that tektites should be predominantly silicic: they originated not on the lunar surface but in a granitic magma somewhere below the lunar surface, and there is evidence for such magmas even though the moon's surface is almost entirely basaltic.

The hypothesis of tektite origin by meteorite impact on the earth continues to be widely supported in spite of the implications of Anders' limit. Let us suppose my application of the limit to the terrestrial-impact hypothesis is faulty and thus allow that there are two serious possibilities for the origin of tektites: lunar volcanism and terrestrial impact. The next step is to evaluate both possibilities in the light of more detailed data concerning the composition of tektites.

Adherents of terrestrial impact assume that the impact was on sedimentary rock (perhaps metamorphosed) because the chemical objections cited against volcanic glasses apply also to igneous rocks in general. By unfortunate coincidence terrestrial sedimentary rock and a lunar volcanic source are about equally capable of yielding the major-element composition observed in tektites. Among terrestrial sediments the sandstones, and in particular the rocks called subgraywackes, yield the best match because some kind of sorting process has enhanced their quartz content. Quartz is pure silicon dioxide, and because it has no natural cleavage planes it is tougher than other common minerals and tends to form larger grains; any sorting process that respects particle size will tend to separate quartz from the other minerals. The high content of lime, magnesia and ferrous oxide compared with soda and potash, which is observed in tektites and distinguishes them from granitic glasses, is remarkable only in rocks of high silica content; in intermediate or basic (basaltic) rocks it is quite usual. Hence all that is needed to achieve the major-element composition of tektites in a terrestrial rock is to raise the rock's silica content, leaving everything else as it is in more basic rocks. Sedimentary processes can do just that.

It was this point more than anything else that led geochemists to support a terrestrial origin for tektites during the 1960's. They believed that the moon probably would not make granite (and in fact it apparently does not make much). They assumed that if granite were made on the moon, it would be much like terrestrial granite because the laws of magma evolution should be the same on the moon as they are on the earth. And on the moon, of course, there would be no sedimentation to sort the granite into a material having the typical tektite composition.

A number of different granitic compositions have now been found on the moon, and none of them is of the classic terrestrial type. At the *Apollo 14* site about 1 percent of the very fine soil is silica-rich glass. Most of it is also rich in potash and thus quite unlike known tektites; the reason may be that potash-rich tektites do not survive because they are rapidly destroyed by weathering. A significant fraction (some 10 to 20 percent) of the lunar granitic glass has only moderate amounts of potash, however. Billy

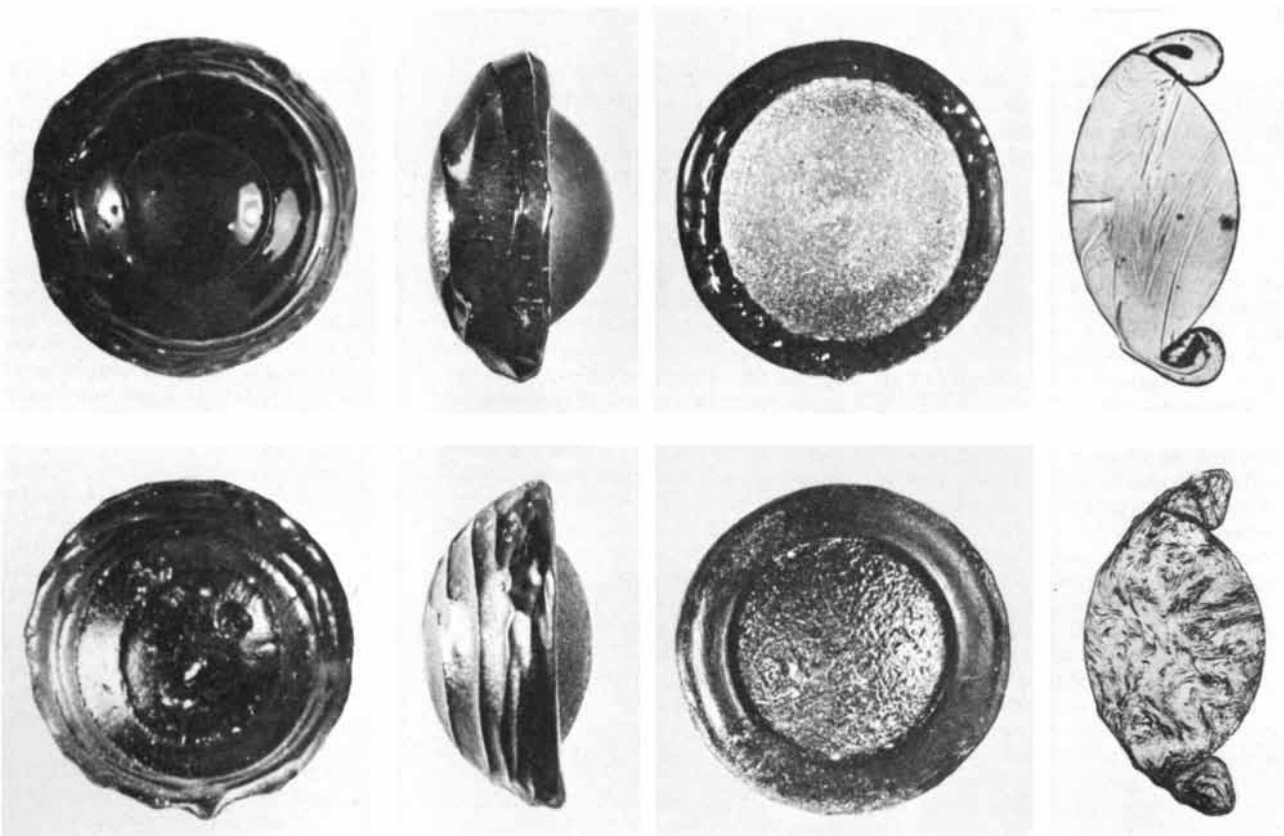
P. Glass of the University of Delaware identified some glassy *Apollo 14* soil particles that are much like known tektites in their major elements. In particular they show the characteristic excess of the divalent oxides lime, magnesia and ferrous oxide and deficiency of the monovalent oxides soda and potash. Similar particles have been noted in the material from *Apollo 12* and *Apollo 15*.

The process by which lunar granites come to be different from terrestrial granites is being studied by Paul C. Hess and Malcolm J. Rutherford of Brown University. They have found in laboratory experiments that the reducing character of a simulated lunar magma makes it tend to become rich in iron as various crystals settle out; the ferrous iron tends to remain in solution, whereas more oxidized ferric iron would be trapped in magnetite (Fe_3O_4) and would settle out. The iron-rich magma then divides into two liquids that, like oil and water, do not mix. One of them is a granitic liquid with the peculiar composition of tektites or lunar granites. (The other is a very iron-rich liquid, evidence for which seems recently to have been found in lunar rocks.)

Going beyond the major-element composition, one can compare the water content, the minor and trace elements and the "oxygen fugacity," or the state of oxidation, of tektites and their possible source materials. Tektites are much drier than terrestrial sediments: they contain about a fiftieth as much water. Compared with terrestrial sediments tektites are deficient in the elements whose compounds are still volatile at 1,000 degrees C., such as lead, thallium, copper and zinc. Finally, there is the strong reducing character of tektites. At about 1,200 degrees C. the oxygen fugacity (or equilibrium oxygen partial pressure) is only about 10^{-14} atmospheres in tektites as against about 10^{-8} in terrestrial basalts.

For a lunar-volcanism origin, on the other hand, those three characteristics present no difficulty. Dryness, low content of volatile elements and low oxygen fugacity are the three hallmarks of lunar-rock composition, as was noted in the first reports on the *Apollo 11* specimens. Brett has recently shown that at magma temperatures the oxygen fugacity of tektites is quite close to the lunar value. (Just before the first lunar landings, incidentally, some of us who had been studying tektites predicted that lunar rock would have a low water content and be deficient in volatile elements.)

It is true that the loss of volatiles, water and oxygen might also be explained by the process of meteorite impact on the earth, at least qualitatively and in principle. If a rock is heated to a sufficiently high temperature and for a sufficiently long time, water, volatile elements and oxygen will escape from it. Serious difficulties arise, however, if one attempts to account for these losses quantitatively. Volatiles can escape at a reasonable rate only by forming bubbles that rise to the surface of a liquid. Within a tektite propelled by a meteorite into a ballistic trajectory the gravitational acceleration is zero; the bubble cannot rise because there is no "up." Even under normal terrestrial gravitational conditions the process of eliminating bubbles from glass, which glassmakers call fining, takes hours or days; it is a major factor in the cost of commercial glass and has been given much attention. Fining is difficult enough with dry, clean glass components and with a glass of low viscosity; in the case of wet, dirty components and a very viscous glass it would be surprising if fining could be

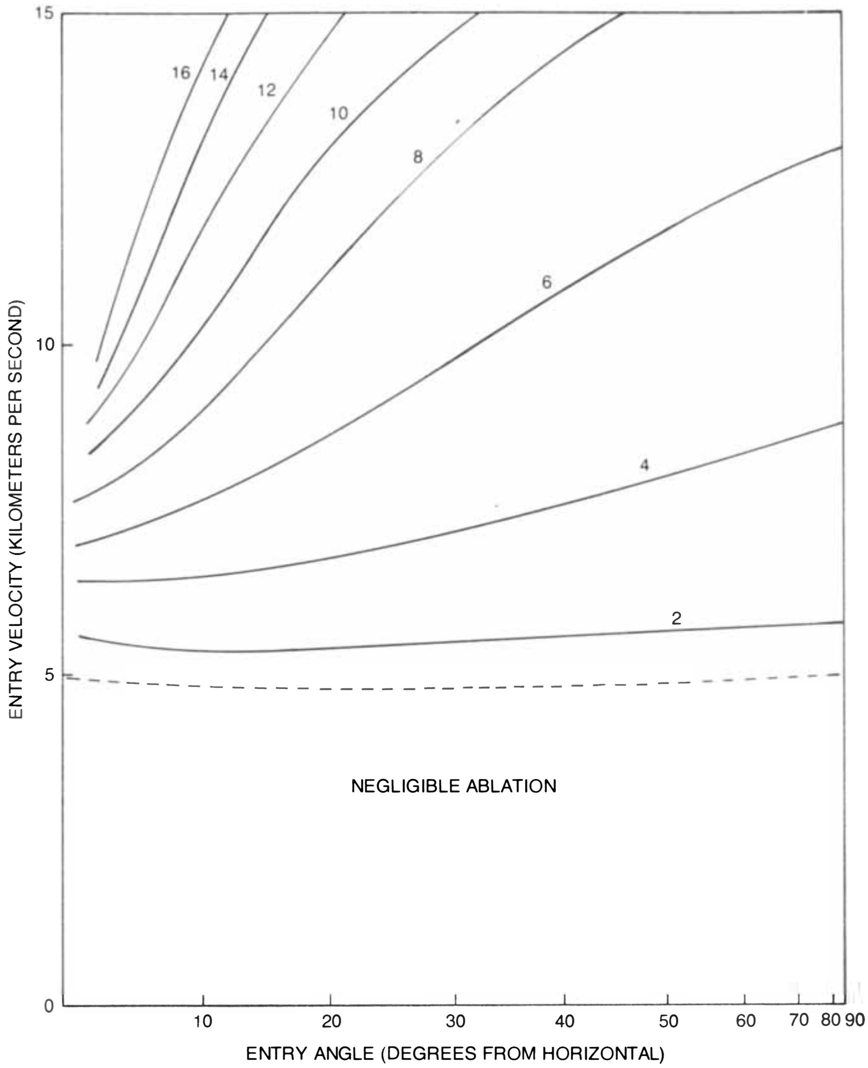


ABLATION OF AUSTRALITES was mimicked in the laboratory by Dean R. Chapman of the Ames Research Center of NASA. He subjected artificial tektite glass to a heated airstream in an arc jet, producing the models shown at the top in (*left to right*) front, side and back views and as a thin section seen from the side. The same views of natural australites are seen at the bottom. The striae in the thin sections meet the back surface at a sharp angle; at the front surface, however, they turn and follow the surface as a result of liquid flow.

accomplished within the few minutes before the tektite material congeals and in the absence of an effective value of gravity.

The large reduction in oxygen fugacity required to transform a terrestrial glass into tektite material leads to a similar difficulty. It means that ferric oxide (Fe_2O_3) must be changed into ferrous oxide (FeO) by the liberation of a volume of oxygen substantially larger (at atmospheric pressure) than the volume of the tektite. Experiments show that at atmospheric pressure the reduction stops at a ratio of ferric to ferrous oxide that is about eight times higher than the ratio in tektites, even at very high temperatures.

There are minor discrepancies between tektite trace elements and the only lunar granite that has been thoroughly analyzed, sample 12013 from *Apollo 12*. The lunar sample is richer in chromium and has a lower potassium-to-uranium ratio, a higher ratio of lead 206 to lead 204, a different set of rare-earth abundances and a lower ratio of oxygen 18 to oxygen 16. On the other hand, some of the more recent analyses have turned up tektites with chromium in the lunar range and lunar material having a potassium-to-uranium ratio in the tektite range, a lead-isotope ratio basically similar to the ratio in tektites and rare-earth distributions much like those in tektites. As for the oxygen-isotope discrepancy, since somewhat similar relations are found on the surface of lunar dust grains it may simply reflect the fact that tektite material has spent some time in the form of very fine particles, a possibility that is supported by some aspects of tektite morphology. On balance, then, the chemical data support a lunar origin rather more strongly than they support a terrestrial one.

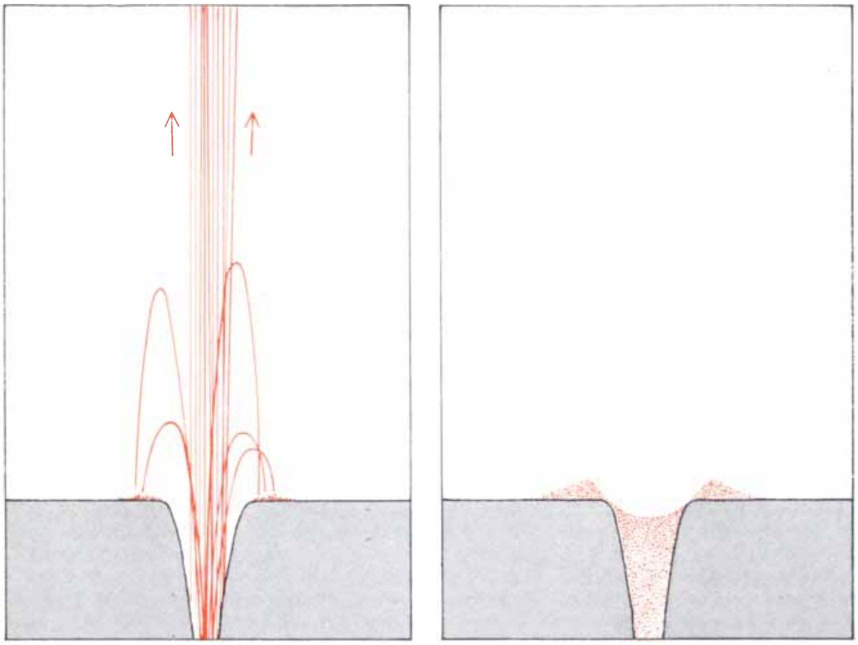


DEPTH OF ABLATION was calculated by Chapman and Howard K. Larson, for various velocities of entry into the atmosphere and for various angles of entry, according to the generally accepted ablation theory. The numbers on the curves give the amount of ablation in millimeters. Since the australites appear to have lost about 10 millimeters to ablation, the calculations indicate an entry velocity of at least 11 kilometers per second at likely angles of entry.

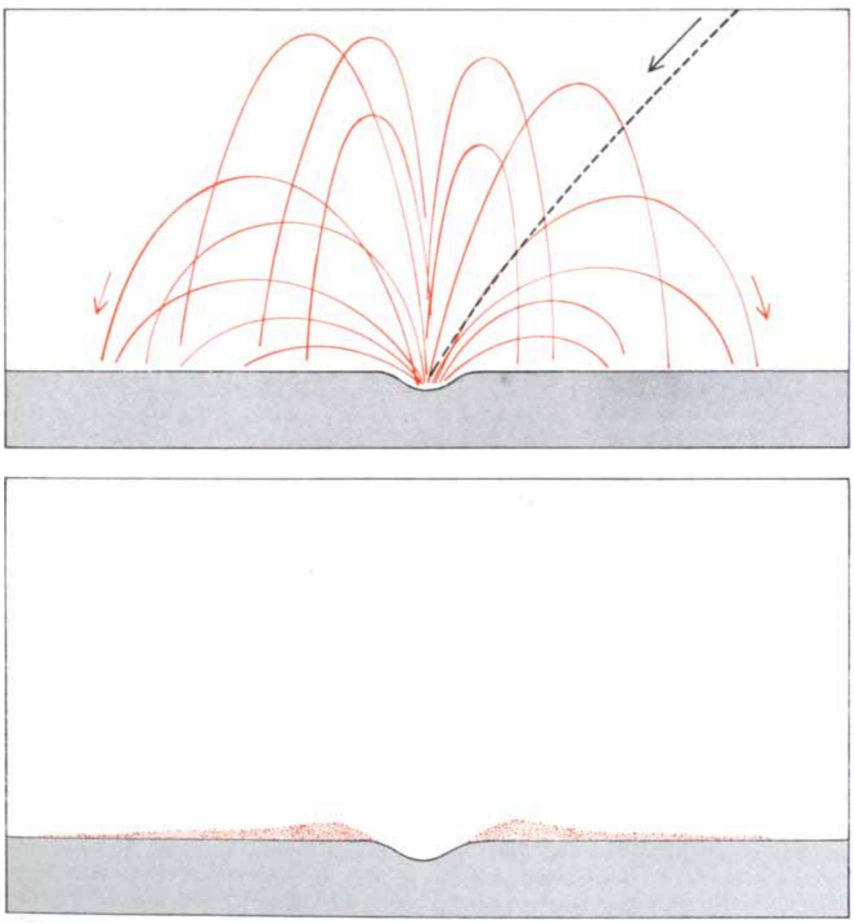
The most telling arguments in favor of the lunar hypothesis, however, are based on striking aerodynamic evidence produced by Dean R. Chapman and his colleagues at the Ames Research Center of the National Aeronautics and Space Administration. Chapman and others had calculated that tektites launched from the ground could travel only a few hundred meters through the undisturbed atmosphere. Shao-Chi Lin of the University of California at San Diego suggested that the atmosphere might not have remained undisturbed: that the air might have been set in motion by a gigantic meteorite impact and so might have conveyed the tektites at ballistic velocity to the top of the atmosphere. Lin calculated that the energy required to launch tektites in this way would have produced a crater some 300 kilometers in diameter and 40 kilometers deep. Chapman and Donald E. Gault confirmed Lin's calculation and suggested that it would be rather difficult to conceal such a crater. Indeed, the existence of three such craters, undiscovered and in the correct positions to account for the Australasian, North American and Ivory Coast strewn fields, is very doubtful indeed.

In a brilliant series of experiments and calculations Chapman demonstrated just how the curious flanged shape of the tektites from Australia called australites, such as the one brought back by Darwin, is generated by aerodynamic ablation. As early as 1893 the German geologist A. W. Stelzner had suggested that the australites had begun as spheres and had been partially melted, with the melt flowing back from the front face of the tektite under air drag. Chapman had glass made up to mimic the chemical composition of the australites and subjected samples to a heated airstream in an arc jet. The glass flowed back from the front surface and coiled into the characteristic flange, identical even to the circular or spiral "ring waves" of the natural tektites.

Chapman showed that the flange, the amount of ablation and the pattern of the ring waves are inconsistent with origin on the earth. The australites could develop a flange only in stable flight. (Most tektites presumably do not begin as smooth spheres; they tumble in flight and are shaped by ablation into irregular forms.) The australites would be stable only during passage through an increasingly dense medium; the amount of ablation indicates a velocity of from 10 to 12 kilometers per second. That is consistent with descent through the atmosphere at the entry velocity (11.2 kil-



LUNAR VOLCANIC ERUPTION, which the author favors as the most likely source of tektites, could eject material at the lunar escape velocity, about 2.5 kilometers per second. The material would be tightly "clustered" in speed and direction (*left*). Not much of it would fall back to the lunar surface and what fell back would not be spread very far (*right*); that would explain why the moon has not been found to be littered with material of tektite composition.



METEORITE IMPACT ON THE MOON, on the other hand, would presumably eject material in many directions and at many different speeds, thus spreading material over lunar surface.

ometers per second) expected for an object coming from the moon. Chapman's calculations, which were confirmed by my colleagues and me, depend on the well-tested theory of ablation that underlies the design of heat shields for NASA spacecraft and military missiles.

In choosing between terrestrial impact and lunar volcanism we are in the position the pioneer French meteoriticist Antoine François de Fourcroy described in 1804: "compelled to choose between ideas that are just as unprecedented the one as the other." Fourcroy wrote that in such a case "it is only by eliminating the absurd or the impossible that one is forced to accept what would at first have seemed almost unbelievable" (that meteorites fall from the sky).

It is absurd to suppose the earth has a number of undiscovered Cenozoic craters larger than Ireland or that the standard heat-shield calculations contain gross undetected errors leading to underestimation of ablation by a factor of five. It is impossible to make substantial chunks of glass of good quality and devoid of water instantly from common rock and soil, or to launch large particles from the earth at six kilometers per second by shock, or to penetrate the atmosphere with gram-size bits of glass at hypersonic velocities. And so one is forced to accept the conclusion that tektites, in spite of their remarkable resemblance to terrestrial rocks, have been fired at the earth by a volcano or volcanoes on the moon.

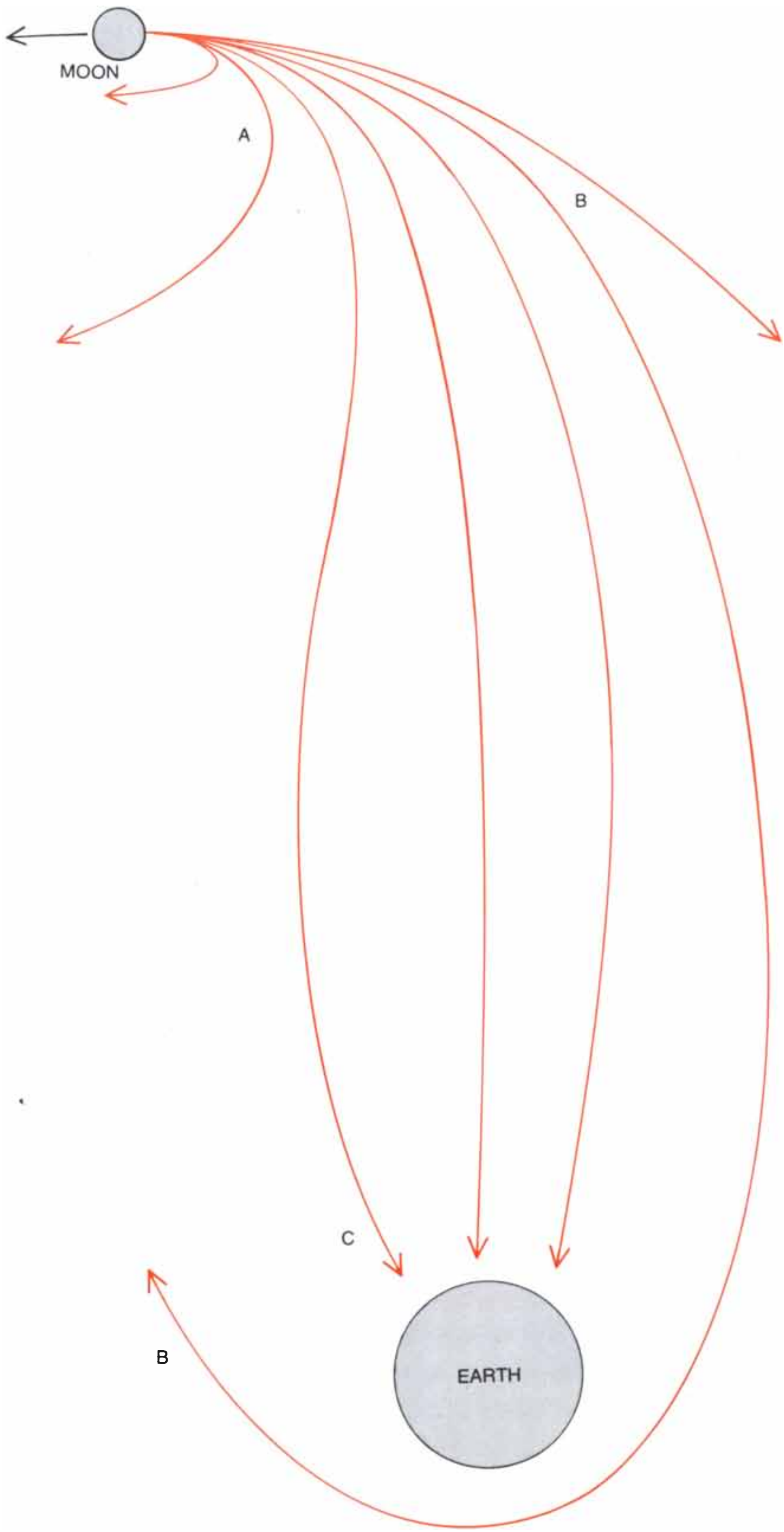
If tektites come from the moon, they are a selected group of volcanic glasses. Volcanism yields nearly vertical launch directions (at least on the earth), which means that material can be propelled to the earth only from a limited set of locations on the moon. That being the case, tektites constitute a limited subsample of lunar granitic materials. It is not reasonable, then, to expect that the average properties of lunar granites would match those of tektites. One can only expect that those properties would not be inconsistent with tektite compositions, and that requirement appears to be satisfied by the lunar granites that have been studied. The compositional argument against a lunar origin for tektites can therefore be met.

Another argument against lunar origin derives from the localized distribution of tektites on the earth. It is held that tektites that miss the earth on a first pass would go into orbit around the sun, would eventually be captured by the earth and would fall out in a uniform distribution over the earth's surface—unless there is some mechanism that destroys them before they are recaptured. Such a mechanism does exist, however. If the albedo, or reflectivity, of a small body moving through the solar system is so distributed that on the average the albedo of the right side is even slightly different from the albedo of the left side, the object will be caused to spin by the differential radiation pressure of sunlight. The spinning will increase to bursting speed in a few tens of thousands of years, thus destroying an orbiting tektite.

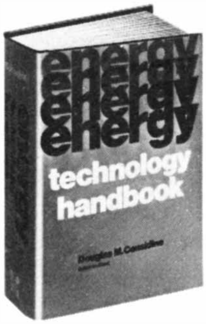
Finally, how is one to account for the few craters that do appear to be associated with certain tektite assemblages: the irghizites, the Aouelloul tektites and perhaps the Darwin-glass tektites? Not all tektites need to have come from the moon as small objects. Large blocks are sometimes ejected from terrestrial volcanoes. It is at least conceivable that a block of tektite material weighing some millions of tons could have been propelled toward the earth by a violent eruption on the moon, could have excavated the Zhamanshin crater and in the process disintegrated, after which its material would have recongealed to form the irghizites.

The key to solving the tektite problem is an insistence on a physically reasonable hypothesis and a resolute refusal to be impressed by mere numerical coincidences such as the similarity of terrestrial sediments to tektite material. I believe that the lunar-volcanism hypothesis is the only one physically possible, and that we have to accept it. If it leads to unexpected but not impossible conclusions, that is precisely its utility.

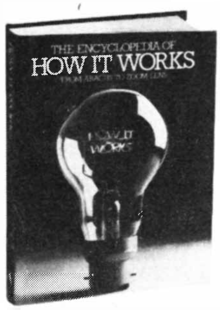
To cite just one example of the utility, the lunar origin of tektites strongly supports the idea that the moon was formed by fission of the earth. Tektites are indeed much more like terrestrial rocks than one would expect of a chance assemblage. If tektites come from a lunar magma, then deep inside the moon there must be material that is very much like the mantle of the earth—more like the mantle than it is like the shallower parts of the moon from which the lunar surface basalts have originated. If the moon was formed by fission of the earth, the object that became the moon would have been heated intensely, and from the outside, and would have lost most of its original mass, and in particular the more volatile elements. The lavas constituting most of the moon's present surface were erupted early in the moon's history, when its heat was concentrated in the shallow depleted zone quite near the surface. During the recent periods represented by tektite falls the sources of lunar volcanism have necessarily been much deeper, so that any volcanoes responsible for tektites have drawn on the lunar material that suffered least during the period of ablation and is therefore most like unaltered terrestrial mantle material. Ironically, that would explain why tektites are in some ways more like terrestrial rocks than they are like the rocks of the lunar surface.



PATHS FROM MOON TO EARTH of tektite material ejected by lunar volcanism are shown. Material ejected at just a little more than escape velocity would (given the velocity of the moon around the earth) go into an orbit like that of the moon (*A*). Material ejected at very high velocity would go into a retrograde orbit (*B*), missing the earth and eventually orbiting the sun. Only material ejected within a narrow velocity range (*C*) would enter an orbit that would intersect, or that could be skewed by the earth's gravitational attraction to intersect, the earth.

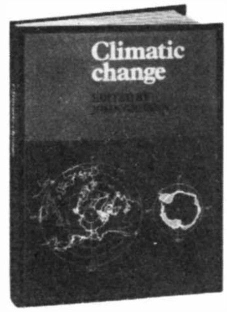


45200-2. ENERGY TECHNOLOGY HANDBOOK. *Douglas M. Considine.* 1850 pages and 1000 illustrations covering a wealth of data on today's energy picture. Reports on major energy sources, transmission and transportation, conversion, and the most promising trends in power technology. *Counts as 2 of your 3 books.* **$49.50**

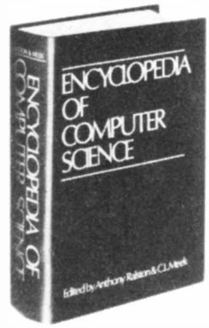


44915. THE ENCYCLOPEDIA OF HOW IT WORKS. From Abacus to Zoom Lens. *Edited by Donald Clarke.* Fascinating explanations of how computers, light bulbs, elevators, and hundreds of other mechanical devices work. Over 300 illustrations and diagrams. **$12.50**

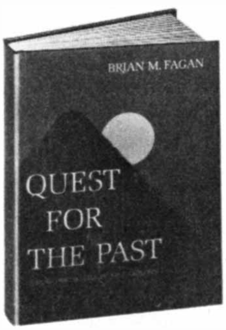
83200-2. THEORETICAL PRINCIPLES IN ASTROPHYSICS AND RELATIVITY. *Edited by Norman Lebovitz et al.* Includes detailed discussions of pulsars, black holes, gravitational collapse, and the mathematical foundations of relativity. *Counts as 2 of your 3 books.* **$23.00**



38362-2. CLIMATIC CHANGE. *Dr. John Gribbin.* What causes climatic changes? How can we learn to predict them? Provides an A to Z report on the basic pieces and patterns of the earth's rich climatological puzzle. *Counts as 2 of your 3 books.* **$32.50**

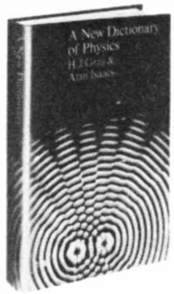


44900-3. THE ENCYCLOPEDIA OF COMPUTER SCIENCE. Monumental 1550-page volume offers accurate information on essential topics ranging from theory to applications, from programming to Boolean algebra. *Counts as 3 of your 3 books.* **$60.00**

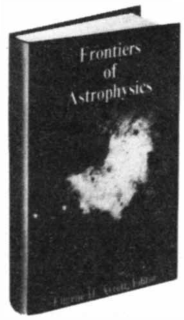


73245. QUEST FOR THE PAST. *Brian Fagan.* Vividly recreates great moments in archaeology from Schliemann's search for Priam's Troy to the discovery of *homo habilis* by the Leakeys in Olduvai Gorge. Beautifully illustrated. **$12.95**

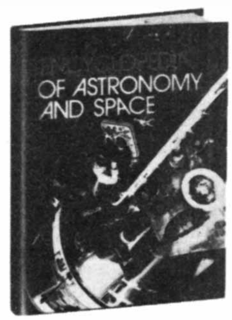
54995. THE ILLUSTRATED ENCYCLOPEDIA OF ARCHAEOLOGY. *Glyn Daniel.* The significant sites, cultures, civilizations, terms, tools, and techniques packed into one all-inclusive volume. 1000 entries. More than 100 illustrations. **$17.95**



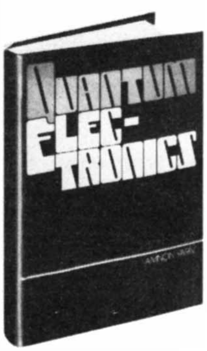
64145-2. A NEW DICTIONARY OF PHYSICS. *H. H. Gray and Alan Isaacs.* More than 9000 entries highlighting the key concepts, formulas, men, and machines. Hundreds of graphs, charts, tables, and illustrations. *Counts as 2 of your 3 books.* **$35.00**



49267. FRONTIERS OF ASTROPHYSICS. *Edited by Eugene Avrett.* A contemporary source of astrophysical topics from recent solar research to neutron stars, supernovae and intergalactic matter. **$20.00**



55000. THE ILLUSTRATED ENCYCLOPEDIA OF ASTRONOMY AND SPACE. *Ian Ridpath.* Covers the full range of astronomical knowledge and space exploration, from ancient times to today. **$16.95**



73165. QUANTUM ELECTRONICS. *Amnon Yariv.* Containing 70% new material, deals with a wide range of important topics, including quantum mechanics, electromagnetism, laser oscillation, nonlinear optics, and much more. **$20.95**

39746-2. COMPANION TO CONCRETE MATHEMATICS. TWO VOLUMES. *Z. A. Melzak.* Goldmine of mathematical techniques and various applications. *The set counts as 2 of your 3 books.* **$46.45**

40910-2. THE CREATIVE ENGINEER. *Winston E. Kock.* Analyzes such 20th century advances as the transistor, radar, the laser, and communications satellites and the innovative thinking behind them. *Counts as 2 of your 3 books.* **$25.00**

34140. APPLIED ANALYTICAL MATHEMATICS FOR PHYSICAL SCIENTISTS. *James T. Cushing.* Deals with topics ranging from spectral analysis of linear operators to complex variables, to group theory. **$19.95**

36395. BLACK HOLES, QUASARS AND THE UNIVERSE. *Harry L. Shipman.* Absorbing exploration of the frontiers of astronomy. **$14.50**

(Publishers' Prices shown)

If the reply card has been removed, please write to
Library of Science
Dept. 2-A9A, Riverside, N.J. 08075
to obtain membership information and an application

Organisms That Capture Currents

A variety of species, from marine sponges to prairie dogs, have harnessed aerodynamic and hydrodynamic forces to increase the flow of air or water through themselves or their abodes

by Steven Vogel

In any ecological system the chemical energy stored in food is a precious commodity. It is not only limited in quantity but also energetically costly for an animal to obtain; hunting, gathering and filter feeding are all energy-consuming activities. As a result a variety of animals have evolved ways of harnessing plentiful forms of mechanical energy in the environment, such as water currents or the wind, to perform tasks that would normally require the expenditure of chemical energy. These animals, ranging from turret spiders to prairie dogs, have achieved a considerable adaptive advantage in the rigorous energy economy of the living world.

The major source of mechanical energy in the environment is the difference in velocity between a fluid and the solid substratum over which the fluid moves. (Ultimately these flows—air currents in the atmosphere and water currents in the ocean—are driven by energy from the sun.) Given the appropriate transducing system, the velocity difference created near the interface between a current and a substratum (such as the ground or the surface of an animal) can be converted into useful forms of energy. Because the energy of a current resides in the movement of a fluid, one simple and direct biological application is to cause part of that same fluid to move through some internal plumbing system. Indeed, most organisms that capture currents do so to pump air or water either through themselves (to facilitate respiration or filter feeding) or through a nest or burrow (for ventilation or humidification).

What are some simple ways in which currents can be harnessed to induce the flow of air or water through an organism or a burrow? One mechanism is based on the principle of the conservation of energy in a steadily moving fluid, first formulated by the Swiss mathematician Daniel Bernoulli in 1738. The principle states that if a fluid moves horizontally so that there is no change in gravitational potential energy, the pressure of the fluid must decrease whenever its velocity increases so that its total energy remains constant. For example, if the fluid is moving through a horizontal pipe that narrows and widens at various points, the fluid must speed up in passing through the constricted areas, and so it exerts the least pressure where the diameter of the pipe is smallest. Bernoulli's principle explains the function of the airplane wing: since air moves faster over the top of the wing than it does over the bottom, it creates a difference in air pressure between the top and the bottom, or lift.

Now consider a small *U*-shaped pipe connecting two points of a larger channel, with both ends of the small pipe perpendicular to the channel. The fluid in the small pipe will move from the end where the flow in the channel is slower (exerts a higher pressure) to the end where flow in the channel is faster (exerts a lower pressure). The velocity of current flow over one of the openings can be increased by elevating the opening or altering its shape so that it is sharper-edged or less sheltered. In this way flow can be induced in the small pipe independent of the direction of current flow in the larger channel. For example, the flow of smoke up a chimney increases when the wind blows, whatever the direction of the wind.

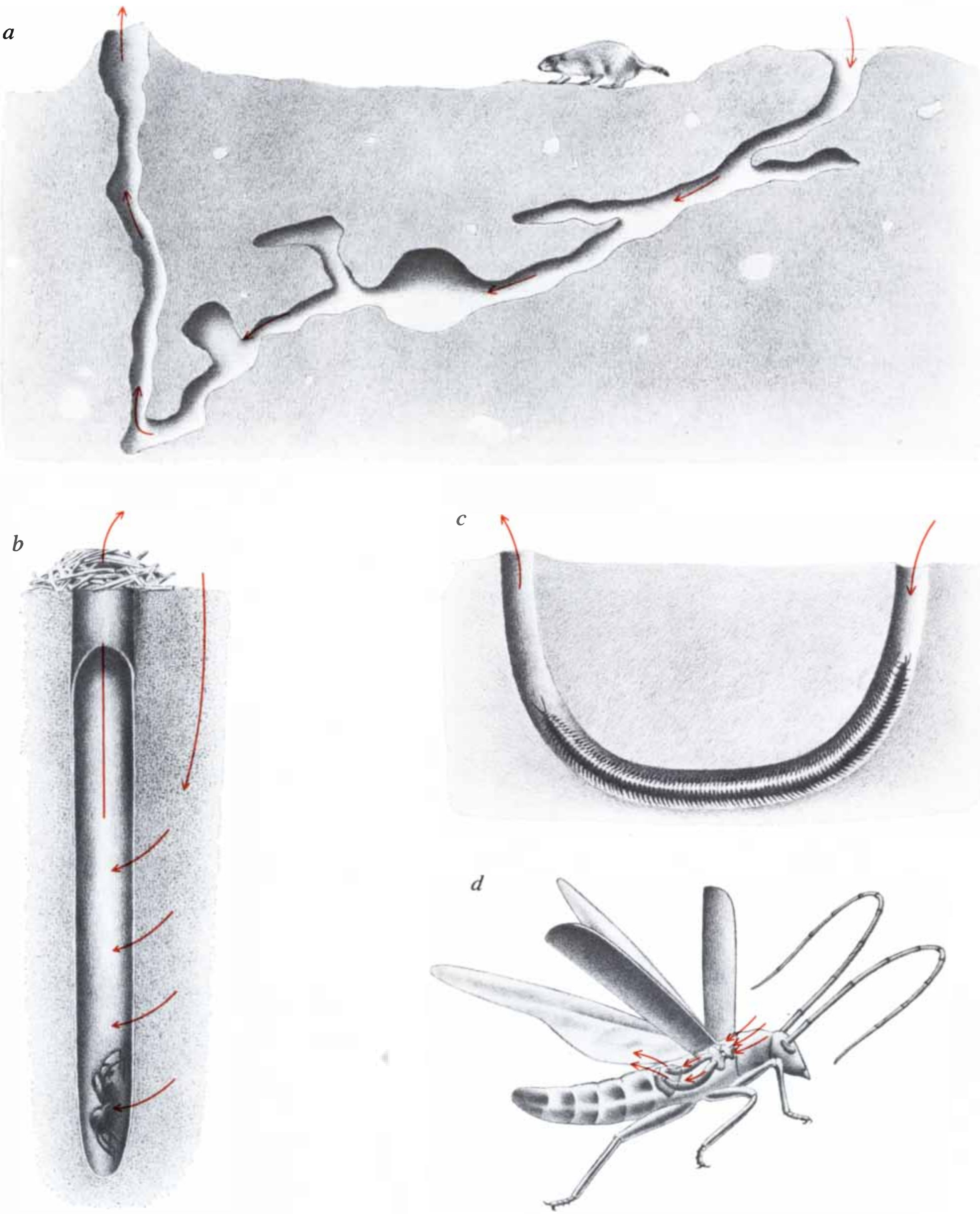
Flow can also be induced by the dynamic pressure of the current as a less direct consequence of Bernoulli's principle. If a small pipe is bent at a 90-degree angle and oriented in a larger channel so that one end is directed upstream and the other opens perpendicularly to the flow of current, the perpendicular opening will be exposed only to the static pressure of the stream whereas the opening facing the current will also be subjected to the dynamic pressure of the oncoming fluid. Flow will therefore be induced in the pipe from the upstream opening to the perpendicular one.

A third mechanism for inducing flow is based on the viscosity (resistance to flow) of real fluids, whether gas or liquid. The movement of a fluid past the opening of a pipe oriented perpendicularly to the current will draw fluid out of the pipe, a phenomenon known as viscous entrainment. The wider the opening of the pipe is or the faster the current is, the greater the entrainment will be.

These three physical effects may work singly or together to induce flow in a biological system. One simple geometrical system often found in nature, which I call Type I, consists of a *U*-shaped pipe through a solid substratum with openings at each end. If air or water is to be persuaded to flow through the pipe, the openings must differ in size, shape, elevation above the substratum or exposure to water currents or the wind. Internal flow will travel from the smaller, blunter, lower or more sheltered opening (where the fluid pressure is higher) to the larger, sharper, higher or less sheltered opening (where the velocity of current flow is greater, the fluid pressure is lower and viscous entrainment is promoted by the larger size of the opening or the greater current).

An excellent example of a Type I system for inducing flow is provided by the burrow of the black-tailed prairie dog, a species indigenous to the Great Plains of North America. These rodents are consummate burrowers, digging tunnels that in well-compacted soil may be as deep as 10 feet and as long as 50 feet. Although a prairie-dog "town" often appears to be a complex network of tunnels, the burrows are usually simple, two-ended passages with only one or two side chambers.

From existing data on the metabolism of the prairie dog, soil properties and diffusion rates it is an easy matter to calculate that the free diffusion of oxygen through either the soil or the tunnel is insufficient to meet the respiratory needs of even a single animal sitting in a nesting chamber at the bottom of the burrow. Moreover, in the warm months cool air is trapped in the chambers, making free convective currents negligible. How then does adequate oxygen reach the prairie dogs inside the burrow? Presumably the animals could force air through the tunnels by erecting their fur and running through the pas-



FLOW OF AIR OR WATER through the burrow of an animal or the animal itself for the purpose of ventilation, respiration or filter feeding can be enhanced by harnessing the energy of external currents. The examples shown here share a common geometry, termed Type I, in which the movement of a fluid over a substratum induces the flow of the same fluid through a *U*-shaped pipe passing through the substratum with openings at each end exposed to the current. Air or water is forced through the pipe because of physical effects resulting from the fact that the two openings differ in size, shape, elevation above the substratum or exposure to water currents or the wind. In the prairie-dog burrow (*a*) air flows from the lower, rounded "dome" mound to the higher, sharper-edged "crater" mound, providing needed ventilation. In the vertical burrow of the turret spider (*b*) air enters through the porous surface of the soil and exits through the elevated burrow opening, bringing up moist air that protects the animal from desiccation. The burrow of a marine worm (*c*) resembles the burrow of the prairie dog in that one opening is elevated above the substratum, inducing the flow of water through the channel. In a large flying beetle (*d*) the wind generated by the forward movement of the insect and the beating of the wings induces the flow of air through large tracheas (internal air pipes) that open directly to the outside, thereby enhancing supply of oxygen to the beetle's flight muscles.

sages like pistons, but as we shall see an energetically cheaper alternative is available.

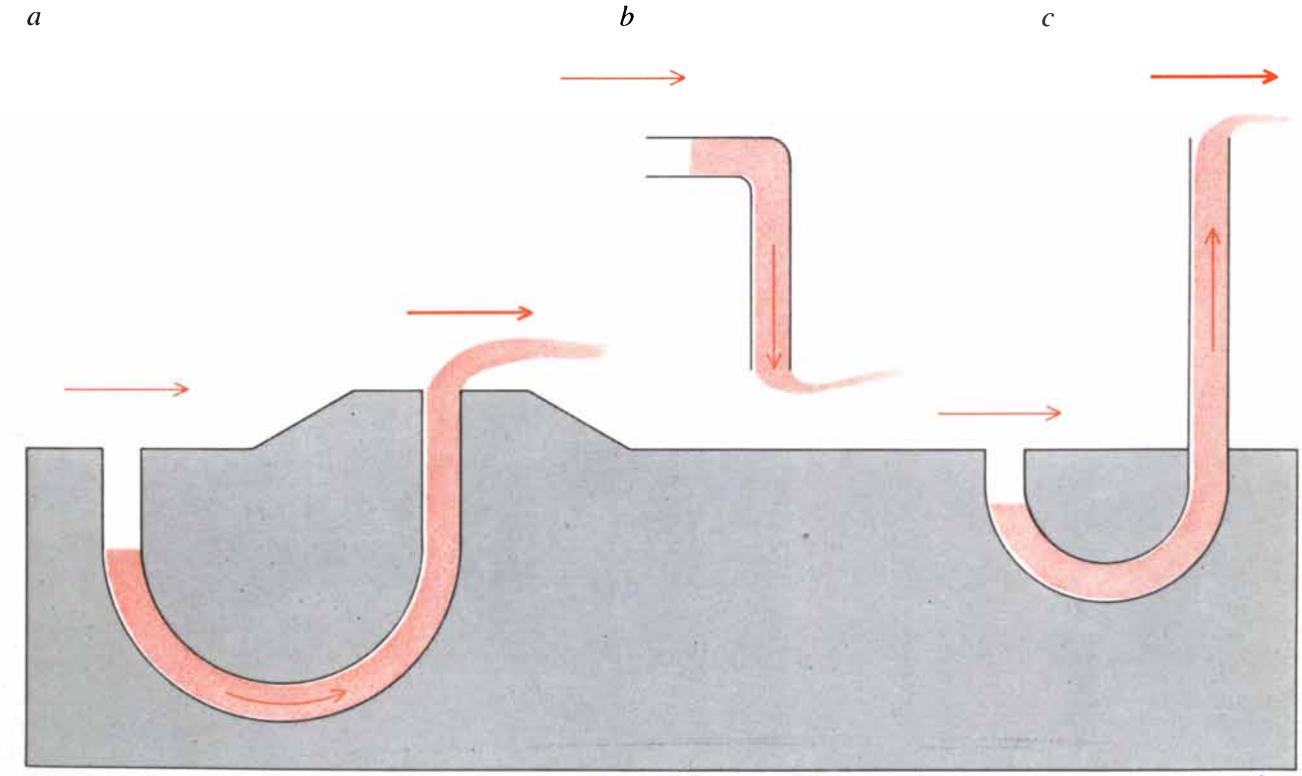
Each end of the prairie dog's burrow emerges in the center of a mound of dirt; such mounds are commonly regarded as lookouts or as protection from flash floods. Two types of mound are scattered through a town: a low, rounded "dome" mound and a higher, sharper-edged "crater" mound. The difference is clearly more than an accident of excavation, since after a rain the animals carefully rebuild the mounds, maintaining the distinction between them. Delbert L. Kilgore, Jr., of the University of Montana identified the openings of a given burrow by the simple expedient of forcing smoke down one hole and observing where it emerged. He found that a typical burrow has a crater mound at one end and a dome mound at the other.

The burrow of the prairie dog therefore meets the geometrical requirements of a Type I flow inducer. Does air actually flow through it? Kilgore dropped small smoke bombs into a few burrows and noted that whenever a breeze sprang up, a plume of smoke appeared above the end of the burrow with the higher crater mound. To test whether the Type I mechanism was involved Charles P. Ellington and I, working in my laboratory at Duke University, built a model burrow in which all the linear dimensions were reduced tenfold and tested the model in a wind tunnel at airspeeds 10 times higher than those expected in nature. (One change compensates for the other.) When scale models of the dome mound and the crater mound were added to the model burrow, it transmitted air in a manner indistinguishable from that of the real burrow. The effect was detectable in the wind tunnel at low wind speeds, indicating that even a barely perceptible breeze of about one mile per hour (.45 meter per second) over a real burrow should suffice to change the air inside about once every 10 minutes.

Measurements made with a series of different mound configurations on the model burrow demonstrated that flow could be induced by either a difference in height between the mounds or a difference in shape; in practice the two differences work in concert. The direction of the wind passing over the mounds did not alter the direction of induced airflow through the burrow. This finding makes "biological" sense: although wind is a dependable feature of the climate in Dodge City, Kans. (the closest Weather Bureau station to Kilgore's burrows), wind direction is notably variable.

Turret spiders (genus *Geolycosa*) construct single-opening vertical burrows in sand or porous soil on roadsides, in open fields and near ocean fronts. Over the entrance of the burrow is the "turret," a craterlike ring of sand, pebbles or bits of vegetation perhaps half an inch high, lined and laced together with silk. During the day the spider stays at the bottom of the burrow, a foot or so below the surface; at night it straddles the turret, poised to pounce on prey. This structure too is no mere accident of excavation: in addition to serving as a handy perch and as some insurance against inedible objects rolling into the burrow, the turret induces a flow of air through the tunnel. If a small flow-measuring device is inserted into the burrow and a second device is placed near the turret, recordings from the two are correlated: whenever a breeze crosses the turret, air moves upward in the burrow.

If the burrow has only one opening, where does the air come from? The an-



PHYSICAL EFFECTS are exploited by living organisms to induce the flow of fluid through some internal plumbing with a minimal expenditure of metabolic energy. According to the principle formulated by the mathematician Daniel Bernoulli, the pressure of a steadily moving fluid must decrease whenever its velocity increases so that its total energy remains constant. Thus in *a* flow will be induced in a *U*-shaped pipe if one end is elevated above the substratum because the velocity of the current there will be greater and the pressure accordingly lower. In *b*, a less direct consequence of Bernoulli's principle, flow is induced through the *L*-shaped pipe because the perpendicular opening is exposed only to the static pressure of the stream whereas the opening facing the current is also subjected to the dynamic pressure of the oncoming fluid. A third mechanism for inducing flow (*c*) is based on the viscosity (resistance to flow) of a fluid: the movement of a current past the opening of a pipe oriented perpendicularly to it will draw fluid out of the pipe, an effect known as viscous entrainment. In cases *a* and *c* the flow of air or water through pipe is induced in only one direction regardless of the direction of external current.

swer is that the surface of the sand or soil around the entrance serves as a second opening, because only very slight pressures easily generated by gentle winds are needed to draw air through the porous substratum in which the spiders live. Thus the turret-spider burrow matches the Type I model.

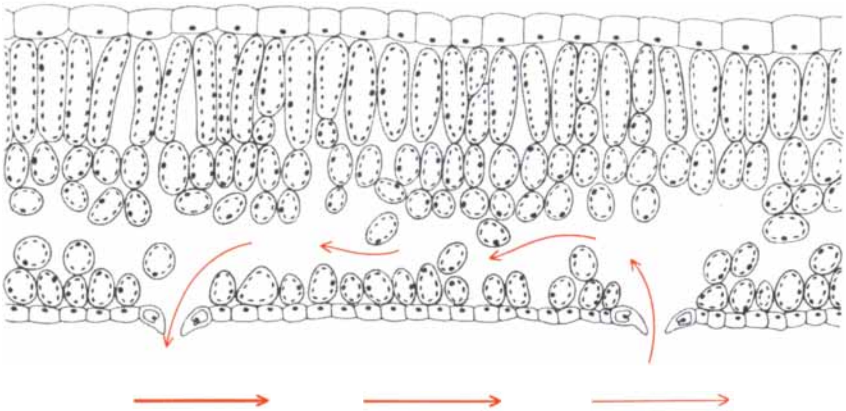
What is the functional significance of induced airflow for the spiders? Is it just a by-product of a turret built for another purpose or does it play a more direct role? Turret spiders live in what amount to local deserts, and since they have no regular access to standing water desiccation can be a serious problem. Even in a desert, however, the air between soil or sand particles is usually saturated with water only a few inches below the surface. A system for slowly drawing moist air into the burrow could enable the spider to spend its days in air of high humidity, thereby minimizing the loss of body fluid. In addition, as will not surprise anyone who has walked barefoot across sand on a relatively calm sunny day, the air in the upper part of the burrow can get very hot. A slow upward flow could prevent the penetration of hot surface air into the deeper part of the burrow.

On the ocean shore many intertidal mud flats are perforated with thousands of openings belonging to the double-ended burrows of marine invertebrates such as lugworms, clam worms and burrowing shrimps. Like the burrows of the prairie dog, these burrows have one opening elevated above the substratum, as can be demonstrated by injecting dye into one opening with a basting syringe and observing where the dye emerges. The worms, at least, will also excavate their burrow in buckets of mud in the laboratory. If they can then be persuaded to leave the burrow, the induction of flow through the burrow can easily be demonstrated; indeed, given the layout of the burrow, some special device would be needed to prevent the induction of flow when the water above the burrow is in motion.

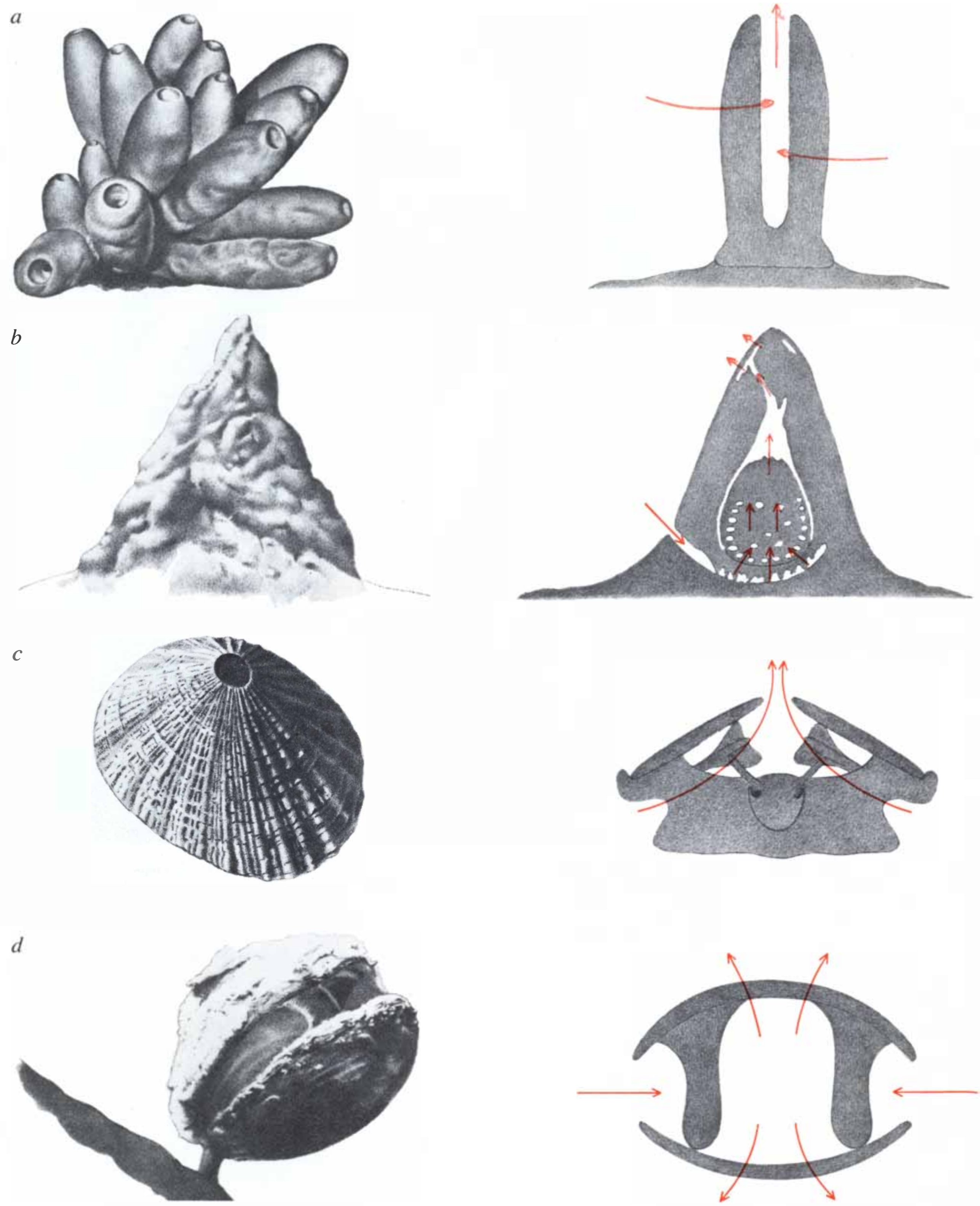
When a marine worm is in its burrow, however, its body plugs or nearly plugs the passage, and the worm displays a wide repertory of rhythmic pumping movements, raising the question of whether induced flow is significant under natural conditions. The worms can clearly sense external currents, but it is difficult to judge whether they are making use of the currents to augment their pumping. On the other hand, some of the burrowing shrimps are reported to leave the main tunnel of their burrow open, residing in a small side chamber after excavation is complete. Under such circumstances flow would certainly be induced and might serve to enhance the feeding and respiration of the animal.

An example of induced flow at the

STOMATE, or leaf pore, of the walnut tree *Pterocarya stenoptera* is magnified 6,700 diameters in this scanning electron micrograph. The stomate is surrounded by craterlike lips that may induce a flow of air within the leaf by a Type I mechanism (much like the crater mound of the prairie-dog burrow), thereby enhancing the diffusion of carbon dioxide into the cells. Perhaps because of the drying effect of the induced flow such craterlike stomates are generally limited to the leaves of hydrophytes (plant species that possess liberal supplies of water).



CROSS SECTION OF THE LEAF of a hydrophyte illustrates the induced flow of air through the continuous air passages that connect the stomates. Since the stomates near the upwind edge of the leaf are subjected to higher wind speeds than those farther from the edge, air should enter the stomates in the center of the leaf and exit through the stomates at the edge of the leaf.



SECOND ARRANGEMENT for the induction of flow, known as Type II, is exploited by a wide variety of living organisms. This configuration involves a conical or cylindrical protuberance of the substratum with small holes around the base or covering the side walls that are connected internally to a large central cavity; air or water flows into the circumferential holes and out through the top of the protuberance. Unlike the Type I system, the entire Type II system is located above the substratum. Marine sponges (*a*) exploit induced flow to increase the rate at which nutrient-rich ocean water passes through their filtration system. The giant mounds of African termites of the genus *Macrotermes* (*b*) induce the flow of air for the purpose of ventilating the brood chamber in the center of the mound. The keyhole limpet (*c*) has an opening, or "keyhole," at the apex of its conical shell so that currents induce a flow of water under the lower rim of the shell and out through the keyhole. The brachiopod (*d*) orients itself perpendicularly to a current so that water is forced into the sides of its gaping shells, passes through the filtration apparatus within and exits through the center of the gape. When the direction of the external current changes, brachiopod will rotate on its stalk in order to continue to benefit from the induced flow. This is its only known behavior.

microscopic level is provided by the stomates, or leaf pores, through which plants exchange gases. In certain hydrophytes (plants that have liberal supplies of water) the stomates are surrounded by craterlike lips and are interconnected within the leaf by continuous air passages. Ellington has proposed that the movement of wind across such a leaf might induce the flow of air into one stomate and out through another, since the stomates near the upwind edge of the leaf would be subjected to a higher wind speed than those farther from the edge. If the flow is induced by a mechanism of Type I, the stomates at the edge of the leaf should function as exit pores and those in the center as entrance pores. Observations of the flow of dye through leaves whose air passages have been filled with water indicate that such bulk flow does occur, although at rather low speeds.

A function for induced airflow within leaves is not hard to imagine. Green plants utilize carbon dioxide as their source of carbon for growth and energy storage, but this gas is only a minor constituent of the atmosphere, about one part in 3,000. In order for carbon dioxide to diffuse into a leaf there must be a still lower concentration of the gas inside it. A forced airflow, however, could circumvent this requirement. The fact that only the leaves of hydrophytes possess the crater-lipped stomates required for the induction of flow suggests that only plants with large supplies of water can tolerate the drying effect of induced airflow within their leaves. Reducing the phenomenon to quantitative terms has proved difficult, because the pores and passages of leaves are less than a tenth of a millimeter in diameter. For this reason the relative contribution of diffusion and induced flow to gas exchange in the leaves of hydrophytes remains to be determined.

What other biological applications of Type I flow induction might there be? Let me suggest a few possibilities. Insects in flight consume prodigious amounts of oxygen, which is supplied to their flight muscles through a set of tracheas, or air pipes, opening directly to the surface of the insect's body. In certain large flying beetles the wind generated by the beating of the wings and the forward motion of the insect has been shown to induce airflow through the thoracic tracheas. Flying birds also require rapid gas exchange and have evolved lungs through which air passes in only one direction. In the most advanced birds air capillaries loop off the fine passages of the lungs in such a way that induced flow enhances the diffusion of oxygen into the blood.

One species of kangaroo rat (*Dipodomys spectabilis*) of the American Southwest builds mounds several feet in diameter in the middle of cleared areas. Up to a dozen relatively large holes perforate the face of the mound; in passages radiating outward from the holes the rat stores supplies of seeds and other plant materials. The mound appears to function as a flow-inducing device to draw moist soil air into the passages and thereby increase the water content of the stored food, one of the kangaroo rat's few sources of water. Still other instances of induced flow may well be found in the diverse burrows of ants, wasps, rodents and moles and even (for water) in the nostrils of certain fishes. There is certainly no dearth of possibilities awaiting investigation.

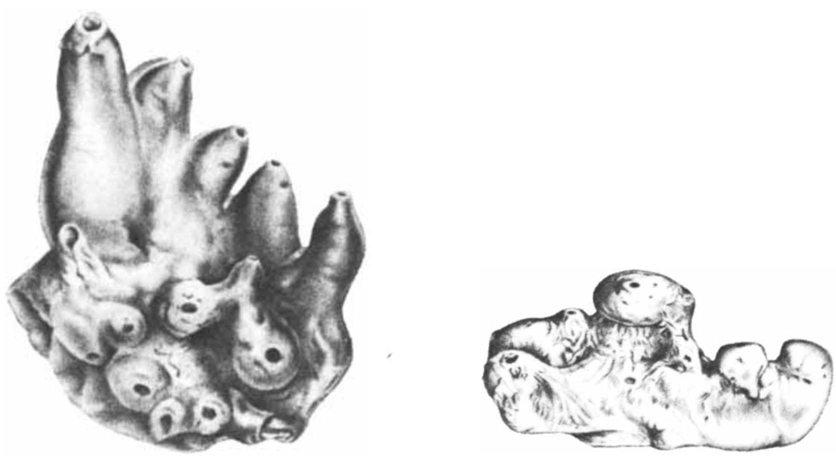
Another geometrical arrangement for inducing internal flow is called Type II. This configuration involves a conical or cylindrical protuberance of the substratum with small holes around the base or covering the side walls; the holes are connected internally to a large central cavity that opens to the outside at the top of the protuberance. Air or water flows into the circumferential holes and flows out at the top of the central cavity. Unlike the Type I system, the entire Type II system is above the substratum.

Marine sponges provide the best examples of the induction of flow by a Type II mechanism. Sponges make their living by passing enormous quantities of water through themselves (10,000 to 20,000 times their body volume per day) and filtering out microorganisms and other nutrient particles. Water enters through tiny holes in the animal's surface and is forcibly expelled through one or more large openings at the top of the animal. To the surprise of early investigators the flow is generated not by muscular action but by the uncoordinated beating of a very large number of tiny flagella.

According to many accounts, marine sponges prefer habitats in which the water is normally moving. Do they under certain circumstances derive benefit from induced flow while actively pumping with their flagella? To answer this question William L. Bretz of the Duke University Marine Laboratory and I built a cylindrical plastic model of a sponge with one large hole near the top and a ring of small holes near the base. When the water around the model was moving, water flowed into the small holes and out of the large hole just as if the "animal" were actively pumping.

At the Bermuda Biological Station I was able to record flow within and adjacent to eight species of sponges without removing them from their site of attachment or even touching them. In all cases any increase in the velocity of the current around the sponge was closely mirrored by an increase in the rate at which water passed through the sponge. Even local currents well below the pumping velocity of 10 to 20 centimeters per second were effective, and higher currents nearly doubled the flow rate. The advantage to a sponge of such augmentation of flow is obvious: filter feeding in most places is a marginal business, with the energy cost of processing water not far below the energy yield of the filtrate. Any device that increases the filtering rate without direct metabolic cost should therefore prove profitable.

Further investigation revealed that sponges, supposedly among the simplest of macroscopic animals, are exquisitely designed to take maximal advantage of induced flow. In most species



SHAPE OF A SPONGE is adjusted to the magnitude of the prevailing local currents to allow the maximal utilization of induced flow. Two colonies of the same species (*Halichondria panicea*) are shown. The sponge at the left grows in calm water and has chimneylike extensions on its output openings, which are exposed to the more substantial currents located well away from the substratum. The sponge at the right, on the other hand, grows on surfaces swept by fast-moving water and is therefore lower, more rounded and has smaller and less elaborate exit holes. The extreme plasticity of this sponge species increases its ecological versatility.

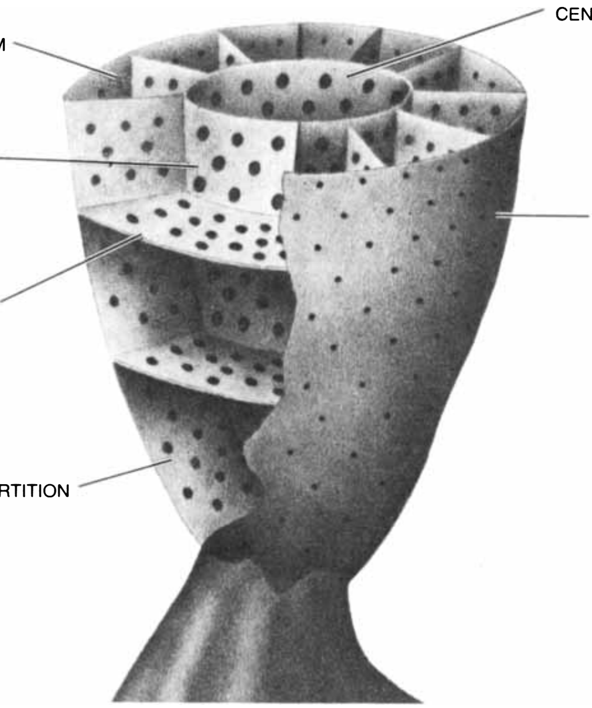
there is an open space just under the outer skin that allows water entering the upstream holes under the pressure of the current to enter the parts of the filtration system on the downstream side of the animal as well; a set of valves in the skin appears to prevent backflow where the pressure inside the animal is greater than that outside. As a result no matter what the direction of the current around a sponge is, the animal can take advantage of both the positive water pressure on its upstream entrance holes and the negative pressure at its exit holes.

Moreover, the very shape of sponges is adjusted to the magnitude of prevailing local currents in just the manner appropriate for maximal utilization of induced flow. Even within a single species individuals growing in relatively calm water are commonly taller and may have chimneylike extensions on their output openings; these arrangements expose the exit holes to the more substantial currents located well away from the substratum. In contrast, sponges of the same species that grow on surfaces swept by faster-moving water are lower, more rounded and have smaller and less elaborate exit holes. The extreme plasticity of the shape of sponges is therefore a remarkable strategy for increasing ecological versatility.

The archaeocyathids ("ancient cups") were a group of sessile animals that occupied a major portion of reefs during the Cambrian period, some 500 million years ago. Although nothing is known of their behavior or of the soft parts of their anatomy, the typical archaeocyathid was about four inches high, was shaped like a vase and had a central cavity that opened wide at the top and was surrounded by two concentric walls of calcified material. The outer wall was perforated by a large number of holes; the inner wall was also perforated, but the holes were larger. Archaeocyathids were probably filter feeders, and since their remains are found in coarse sediments they probably lived in moving water.

William L. Balsam of Brown University first pointed out that archaeocyathids were quite reasonably shaped to take advantage of induced flow, somewhat in the manner of their possible relatives the sponges. To test this hypothesis Balsam and I made an approximation of an archaeocyathid out of aluminum: a double-walled inverted cone with small holes in the outer wall and larger ones in the inner wall. When we tested the model in a flow tank, the model induced flow much more strongly than even the best sponge models. In effect it had a two-stage induction system: the external current not only drew water up out of the central cavity but also drove a large vortex in the cavity, which in turn drew more water through the holes in the inner wall.

The contrast between sponges and archaeocyathids raises some interesting questions. Modern sponges that are the size of fossil archaeocyathids have a more constricted exit opening, a feature that reduces the efficiency of flow induction but acts as a nozzle to increase output speed and lessen the chance that when outside currents are slow, the sponge will simply be reingesting water it has already filtered. If the archaeocyathid lacked a metabolically powered pump, then it could filter feed only in the presence of an external current and thus would never have confronted the problem of reingesting already filtered water. Is the wide opening of the central cavity of the archaeocyathid evidence that it lacked an active pump? If it is, then perhaps these primitive animals were more completely dependent on currents than sponges are, a lack of behavioral flexibility that may ultimately have contributed to their extinction. A corollary of this hypothesis is that flow induction in macroscopic filter feeders may be a device more ancient than metabolically driven pumping.

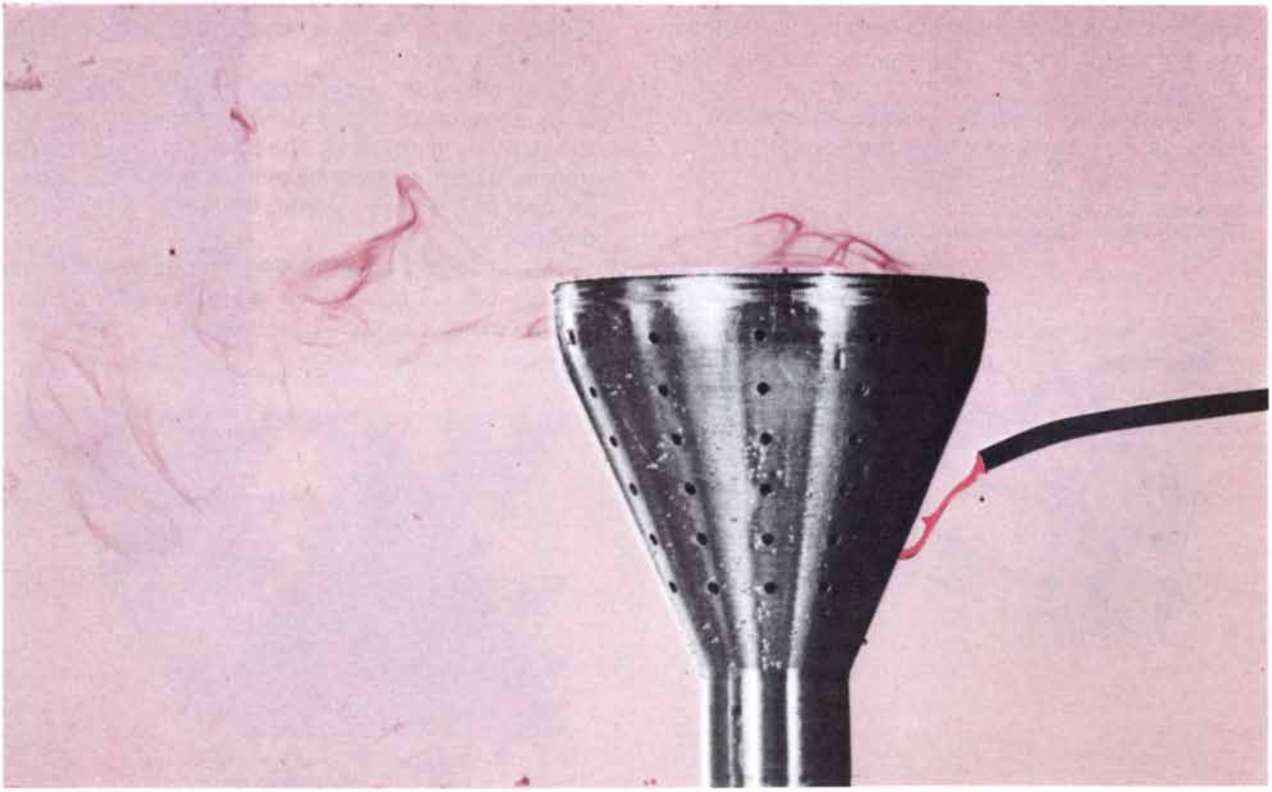


ARCHAEOCYATHIDS were sessile animals, somewhat akin to modern sponges, that thrived in the Cambrian period some 500 million years ago. A typical specimen was about four inches high, was shaped like a vase, had a central cavity that opened wide at the top and was surrounded by concentric walls of calcified material held together by horizontal and vertical partitions. The walls were perforated by a large number of holes whose diameter increased from the outer wall to the inner one. Archaeocyathids were almost certainly filter feeders, and because their fossil remains are found in coarse sediments they appear to have lived in moving water. It seems likely that they used the Type II mechanism to enhance the flow of water through themselves.

The giant mounds of certain African termites (genus *Macrotermes*) combine the ventilation problem of the prairie-dog burrow with the geometry of the marine sponges. Millions of termites together with their associated fungi may live in a mound as much as 16 feet high; their substantial demand for oxygen makes some kind of ventilation system essential. In some mounds found in forests the heat generated in the central brood chamber drives a convective air current up through the middle of the mound and down through peripheral passages located just under the outer walls [see "Air-conditioned Termite Nests," by Martin Lüscher; SCIENTIFIC AMERICAN, July, 1961]. In other mounds found in more open country, however, the peripheral passages are replaced by porous-walled tunnels around the base of the mound and in a turret at the top. According to R. Loos of the National University of Zaire, in such mounds the convective flow of air is augmented by a wind-driven component, apparently another instance of the induction of flow by a Type II system.

A different species of African termite builds mounds equipped with large, funnel-shaped holes that are connected with the network of passages within and under the mound. John S. Weir of the University of New England in Australia found that air entered such mounds through holes without rims around the periphery of the mound and exited through holes with rims near the top. Some holes on the sides of the mound could act as either entrances or exits, depending on the direction of the wind, but on the whole the airflow inside the

MODEL OF AN ARCHAEOCYATHID was made out of aluminum by William L. Balsam and the author at Duke University to enable them to test whether the animal was the right shape to take advantage of induced flow. In the assembled model the cone at the left fits snugly into cone at the right when they are fastened with a screw through their apexes. Two small holes in the outer wall communicate with one larger hole in the inner wall through the grooves in the surface of smaller cone. Model is seven centimeters high and five wide.



INDUCED FLOW through the model archaeocyathid is demonstrated in a flow tank by injecting dye near the outside pores. The external current draws the dye into the upstream pores and out through the central cavity, as is shown in this photograph. A vortex in the central cavity is also created that draws more water through the holes in the inner wall. The efficiency of this two-stage flow induction suggests that the archaeocyathid exploited external currents for filter feeding. Indeed, the animal may have been totally dependent on induced flow.

Save up to 15% on these 4 outstanding titles

Animal Physiology

Roger O. Eckert, University of California, Los Angeles
with chapters by
David Randal, University of British Columbia

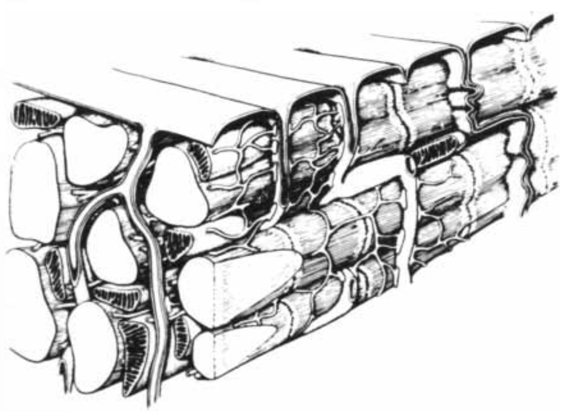
Animal Physiology clearly illustrates the major physiological concepts that will provide students with an in-depth understanding of living systems. The text avoids encyclopedic detail, emphasizing instead the physical and chemical principles and mechanisms that living systems use to meet their functional needs. *Animal Physiology* is presented in a broadly based biological context, hence the examples were selected from a wide spectrum of animal life.

Each chapter concludes with a summary, a set of exercises, suggested readings, and references. The text is beautifully illustrated in two colors, contains an extensive glossary, and incorporates related material in boxes throughout the text. It is designed for use in general, animal, and comparative physiology courses.

"Outstanding in-depth coverage, illustrations, and up-to-date information. I liked the book . . ."

—S. B. Chaplin, University of Missouri

558 pages, 588 illustrations.
Regularly $19.95. **Special Price: $17.50**



Fractals

Form, Chance, and Dimension

Benoit B. Mandelbrot, IBM Fellow,
Thomas J. Watson Research Center

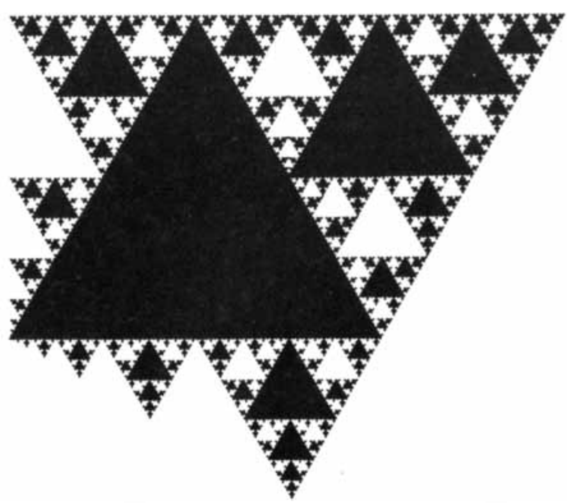
In nature fractals are represented by highly irregular shapes, like the coastline of Britain or the pattern of Brownian motion. Mandelbrot invented the term fractals and is a leader in applying the mathematics used to describe Brownian motion to other problems in science. Irregular structures which can be described by this approach include hierarchial stellar clustering, river networks, turbulence, and soap films. Each has the property that when suitably magnified, the resulting pattern is similar to the original. Mandelbrot discusses various ways of calculating the measure and dimension of these objects. He draws on classical works in real analysis and probability, as well as on more speculative insights of the philologist George Zipf.

Because of its wide range of applications, this book is exciting not only to mathematicians, but also to scientists in any field. The book's striking illustrations and unique ideas make it delightful reading for anyone who enjoys viewing the world with a fresh perspective.

"Mandelbrot's small volume is a good bet to become a classic of its deep yet elementary genre."

—Phillip Morrison, SCIENTIFIC AMERICAN

365 pages, 68 illustrations.
Regularly $14.95. **Special Price: $12.75.**



USGS

Second Edition

Principles of Paleontology

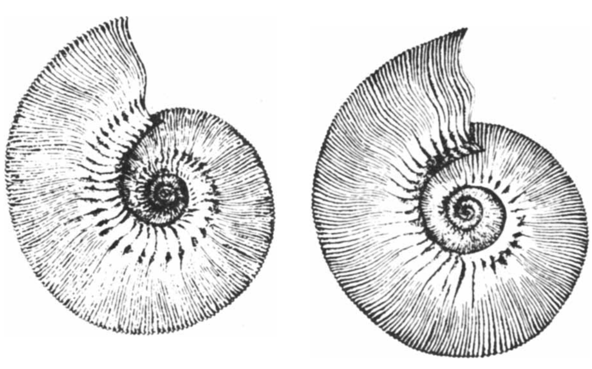
David M. Raup, University of Rochester
Steven M. Stanley, The Johns Hopkins University

During the last seven years the field of paleontology has undergone some dramatic changes. This second edition of *Principles of Paleontology* presents a synthesis of these changes, and at the same time retains the widely acclaimed organization and information of the first edition. This volume emphasizes theory and concepts, rather than details of morphology and fossil records. It gives students a comprehensive view of the field of paleontology, beginning with measurement, identification, and classification of specimens and populations, and includes discussions of the most recent conclusions paleontologists have drawn about evolution.

Changes have occurred in a number of chapters. Part I now includes an introduction to the basic concepts of multivariate analysis and other techniques, some that are designed for use with a computer. Part II discusses the latest advances in the area of large scale evolution. The chapter on functional morphology has been revised to include several new studies. And a thorough discussion of plate tectonics was added to the final chapter giving students an appropriate background for studying biogeography.

This new edition is profusely illustrated with photographs, charts, graphs, and tables. It is essential reading for both students and professors interested in paleontology.

481 pages. 158 illustrations.
Regularly $19.95. **Special Price: $17.50**



Callomon, 1963

Second Edition

Earth

Frank Press, Massachusetts Institute of Technology, and
Raymond Siever, Harvard University

Frank Press and Raymond Siever have produced a second edition of their highly-acclaimed geology book. It maintains the comprehensive treatment and high standards of the first edition, while adding the newest discoveries of plate tectonics, oceanography, geochemistry, geophysics and planetary geology.

"A book as good as this, and with such a wide appeal, can arise only from the most fortunate match of authors, publishing techniques, and subject matter . . . The basic approach has been to describe geological processes within the conceptual framework of plate tectonics. It is a happy fact that this framework has recently stimulated methods so elegant, so useful, and so simple . . ."

—Geophysics, about the first edition.

649 pages.
Regularly $16.95. **Special Price: $14.50**

W. H. FREEMAN AND COMPANY
660 Market Street, San Francisco, California 94104

MAIL THIS COUPON TODAY

Please rush my copy of the books I've checked below at your special 15% discount. I enclose payment in full. *Publisher will pay postage.* (California residents add appropriate tax.)

☐ Fractals: Form, Chance and Dimension, $12.75
☐ Animal Physiology, $17.50
☐ Principles of Paleontology, $17.50
☐ Earth, $14.50

Name ____________________

Address ____________________

City ____________________

State ____________ Zip ____________ 878SA

Rare-earth magnets: Making their strong attraction more attractive.

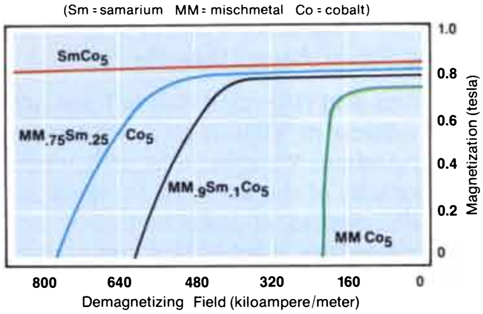
Better magnets make smaller electric motors. Smaller motors can mean lighter cars. (Some cars have over 20 motors.) And lighter cars can mean fuel savings.

Rare-earth magnets offer the hope of achieving such a weight reduction. And that's why, years ago, magnet R&D at the General Motors Research Laboratories turned in this direction.

The program posed many challenges . . . to our physicists, our electrical engineers, our metallurgists. For example:

☐ The most promising rare-earth magnet (samarium + cobalt) is quite expensive. However, our research has determined the trade-offs between material cost, availability, processing, and magnetics. This led to a formulation using lower-cost mischmetal (a natural mixture of rare earths) for most of the samarium, yielding magnets with three times the energy product of the best ferrite magnet used in motors.

RARE-EARTH MAGNET PROPERTIES
(Sm = samarium MM = mischmetal Co = cobalt)

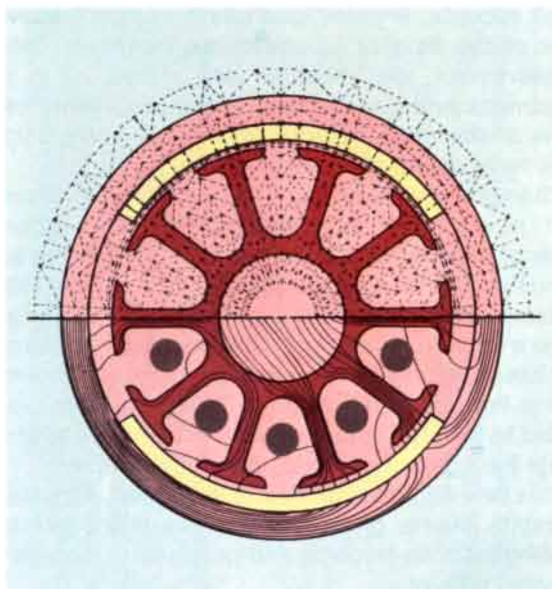


☐ Techniques for designing conventional motors fall short when applied to motors using rare-earth magnets. But finite element methods enabled us to attain effective motor and pole-magnet designs by accurately defining the flux distribution (main illustration) and effects of iron saturation and armature reaction.

☐ A practical way of making pole magnets in thin, curved shapes did not exist. Thus evolved an original powder metallurgy process consisting of special pressing in a radially aligned magnetic field followed by constrained sintering to prevent distortion. It forms precise, wafer-thin magnet arcs with good handling strength.

The overall result? A strong technological base for supporting GM's efforts to shrink motors . . . to help lighten cars . . . to promote energy conservation.

We currently have openings for Ph.D.s in engineering or the physical, mathematical, or biomedical sciences. If interested, please send your resume to: GMR Personnel, Dept. 810. An Equal Opportunity Employer.



General Motors Research Laboratories
Warren, Michigan 48090

mound was unidirectional, as it is in other Type II systems.

Keyhole limpets are a family of marine mollusks in the same class as snails, but instead of being coiled-up cones they are squat and uncoiled: they are two or three times as broad as they are high and an inch or two in overall length. They make their living by scraping algae from the rocks of the intertidal or subtidal zone. Keyhole limpets differ from other species of limpet in that they have an opening, or "keyhole," at the apex of their cone. John Markham of the Bermuda Biological Station first suggested that keyhole limpets might make use of flow induction, and Gordon R. Murdock of Duke later found that the flow of water through the mollusk, from the lower rim of the shell to the hole at the apex, was faster when the surrounding water was in motion. Moreover, some of the limpets oriented themselves to face an oncoming current and partly withdrew their gills to allow an even faster flow of water through themselves. The function of such induced flow in the keyhole limpet is unclear; the animals are not filter feeders and live in well-oxygenated water, so that little metabolic energy is required to satisfy their modest respiratory needs. Still, gills do take up space, and induced flow might confer a real selective advantage if, for example, it made it possible for the animal to have smaller gills and larger ovaries or testes.

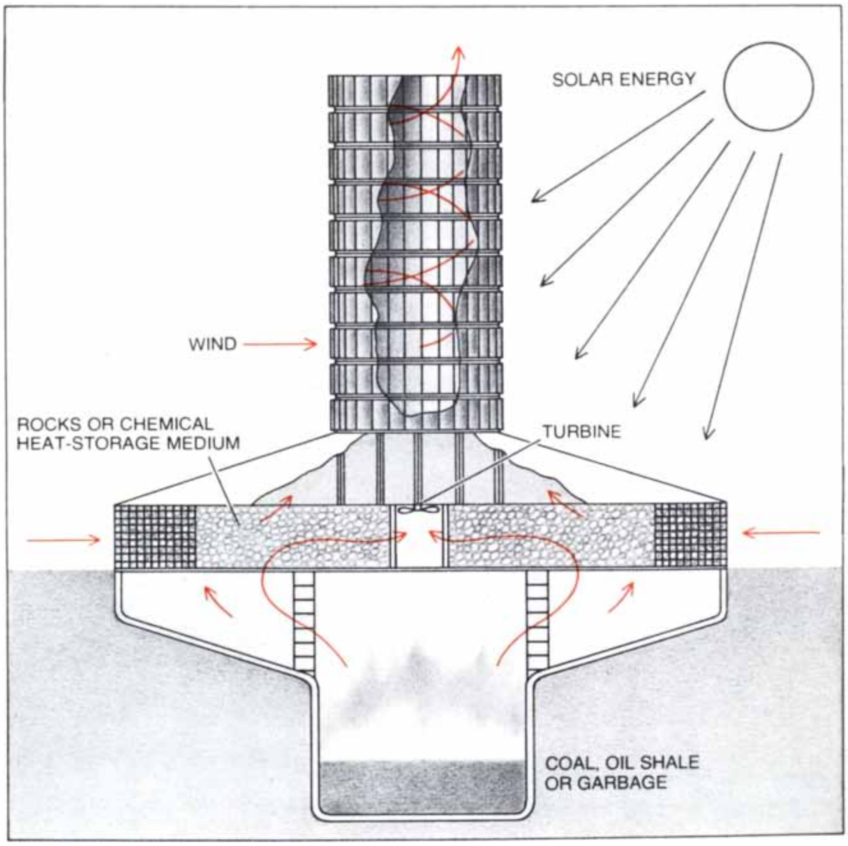
The brachiopods are another group of marine invertebrates. Although they were enormously abundant in Paleozoic seas, some 300 to 500 million years ago, only a few hundred species now remain. At first glance brachiopods resemble clams, but their hinged shells form the anatomical top and bottom of the animal rather than the left and right halves. The shells gape at one edge, revealing an elaborate filtration apparatus within. Water leaves the filter at the center of the gape and enters on each side.

Michael C. LaBarbera of the University of Chicago investigated several species of brachiopod that are attached by a stalk to subtidal rocks. He observed that the animals rotated on their stalk so that the widest central portion of the gape was oriented perpendicularly to the fastest flow of current, with the sides of the gape exposed to lower speeds. In this orientation the circulation of water through the filtration system was enhanced by flow induction. A change in the direction of the current resulted in a shift in the animal's orientation, so that it continued to benefit from the induced flow. These findings contradicted the widespread assumption that brachiopods possessed little in the way of sensory equipment or overt behavior, since current flow as an environmental stimulus had been overlooked by earlier investigators. Organisms care about flow, and biologists should too.

Undoubtedly the most elegant and sophisticated biological exploitation of an environmental velocity difference does not fit into either Type I or Type II but should not go unmentioned. Certain birds practice what is called dynamic soaring, in which they are able to stay aloft without either beating their wings or having an upward current of the air around them. By alternately diving and climbing they are exposed to different wind speeds and from the difference can extract the energy to remain aloft.

Human beings have also learned to extract energy from velocity differences. The operation of windmills, carburetors, chimneys and certain types of ventilators depends at least in part on flow induction. James T. Yen of the Grumman Corporation recently devised a wind-energy extractor that avoids the long blades of conventional windmills; his device bears some functional resemblance to a sponge. The traditional architecture of many cultures also exploits the induced flow of air for heating or cooling purposes. Sydney A. Baggs of the University of New South Wales has pointed out that opal miners in the Australian outback often ventilate their dugout dwellings with a wind-driven system resembling that of the prairie-dog burrow. The tepees of the American Indians of the Great Plains are conical structures with a porous lower edge and an opening near the top that induce flow by the Type II mechanism; such structures are naturally well ventilated and are both comfortable in the heat of summer and capable of containing a fire in winter without asphyxiating the inhabitants. The traditional architecture of the Middle East provides some even more sophisticated examples of cooling by flow induction [see "Passive Cooling Systems in Iranian Architecture," by Mehdi N. Bahadori; SCIENTIFIC AMERICAN, February].

What conclusions can be drawn from this excursion into what might be called experimental natural history? One important point is that careful and imaginative consideration of the physical world is imperative when investigating the adaptive strategies of living organisms. The particular physical opportunities as well as the constraints under which an organism lives must be appreciated. Moreover, the biologist should hesitate to invoke explanations of phenomena that require the expenditure of chemical energy until simpler physical mechanisms have been ruled out. In matters of energy nature seems to love a bargain.



WIND-ENERGY EXTRACTOR designed by James T. Yen of the Grumman Corporation avoids the long blades of conventional windmills; instead it exploits Type II flow induction much as a sponge does. Convective hot-air currents from solar heating or the burning of fuel, together with the currents induced by the wind, drive a central turbine to generate electricity.

THE AMATEUR SCIENTIST

Observations on grinding glass by hand and on making the most of a fireplace

by Jearl Walker

Although the optics industry has sophisticated machines for grinding glass into lenses and mirrors, many amateur makers of optical instruments prefer grinding by hand. The term for the task is free abrasive grinding. This hand technique involves rubbing a hard object in a random motion over a fine abrasive grit on the object being ground. One usually begins with a grit of fairly large size and then shifts to progressively smaller grit sizes until the ground object reaches the desired size, shape and smoothness.

Two seeming paradoxes are associated with grinding. One of them has been understood in essence since the time when the wave nature of light was accepted. The question is: How can grinding the surface result in a smoother surface, one that is optically good enough to be useful in an instrument such as a telescope? The answer lies in the transition to increasingly fine grit. The deviations from smoothness on the surface correspond roughly with the size of the grit. Therefore they are gradually reduced during the grinding and polishing until they are the same size as or smaller than the wavelength of light. In that condition the surface then appears smooth and optically true.

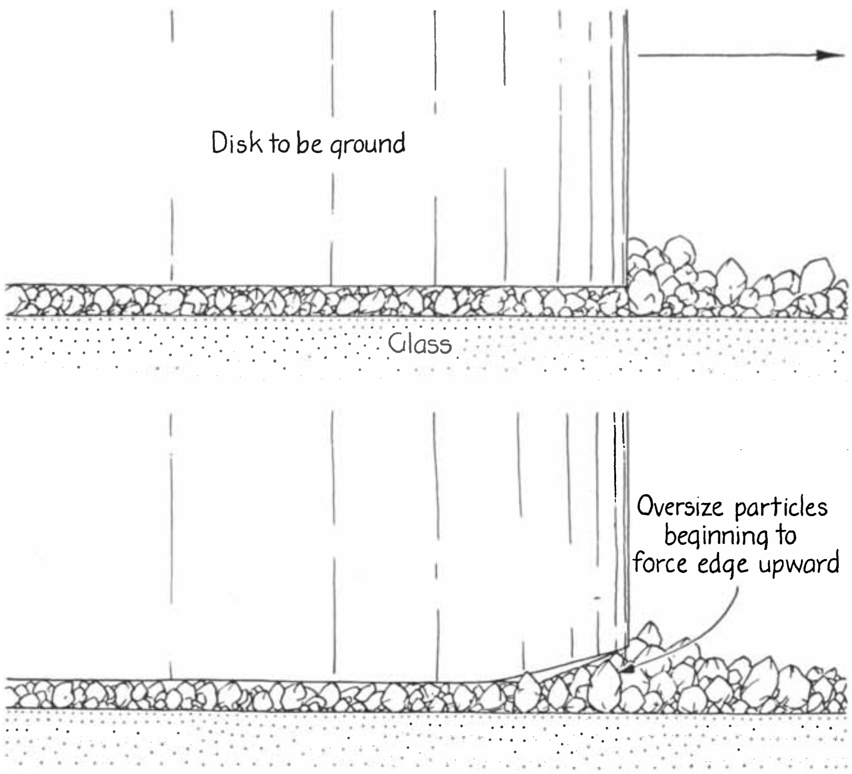
The second paradox has, I believe, only recently been understood. Indeed, more work could be done to firm up the understanding. This second question is: How are scratches on the ground surface avoided in spite of the fact that in any abrasive compound the distribution of grit sizes ranges over about two orders of magnitude, so that the compound surely contains some relatively large particles that should scratch the surface? Free abrasive grinding appears to have become so much of an art, depending a great deal on the grinder's experience and sense of touch, that few people seem to have investigated why more scratches are not made.

Recently Edward J. Saccocio of Columbus, Ohio, developed a model that seems to resolve this second paradox, at least in part. The success of free abrasive grinding appears to depend on the shape of the edge of the object moved over the grit. Normally the lapping tool on top is a metal or a glass that is as hard as or harder than the material being ground. Between the two lies a slurry of the grit. Saccocio found that the leading edge of the tool determines the size of the particles that can get between the tool and the ground object.

If the edge is sharp, oversize particles will enter the space between the tool and the ground object less easily than if the edge is rounded. The height of the space is largely determined by the average size of the grit. Scratching appears to result when a particle that is much larger than average enters the space, becomes stuck on one of the surfaces and then gouges a scratch in the other surface. Hence a tool with a sharp leading edge should produce fewer scratches on the average than one with a dulled and rounded leading edge, which will more readily allow oversize particles to get under the tool.

To test these ideas about the causes and prevention of scratches Saccocio performed several simple experiments. I have repeated them, as you can; moreover, you can use them as a base for more experiments to track down the causes of scratches. The material chosen to be ground was a molybdenum disk (brass would be more convenient for you), the harder material was glass and the abrasive grit was aluminum oxide in three nominal grit sizes: 30, 15 and three microns. The specimen of molybdenum was a short cylinder 15 millimeters in diameter; it was worked over a large plate of glass in an inversion of the normal grinding arrangement. The grinding compounds are available from the Edmund Scientific Company (7778 Edscorp Building, Barrington, N.J. 08007), from Buehler Ltd. (2120 Greenwood Street, Evanston, Ill. 60204) and from certain hobby shops.

A grit was spread on a portion of the glass plate to provide a grinding medium for the molybdenum disk. Saccocio mixed the aluminum oxide with water to form a slurry, which is the usual prac-



Effects of a sharp leading edge and a rounded one in grinding tice in free abrasive grinding, but I left my grit dry. The molybdenum surface was first made smooth and scratch-free by grinding it with each of the three grit sizes in descending order. I moved the cylinder around in the grit randomly, applying pressure as evenly as I could. One of the difficulties in the experiments, something that requires experience, is determining what pressure is appropriate. Too light a pressure prolongs the grinding process; a pressure that is too heavy results in scratches.

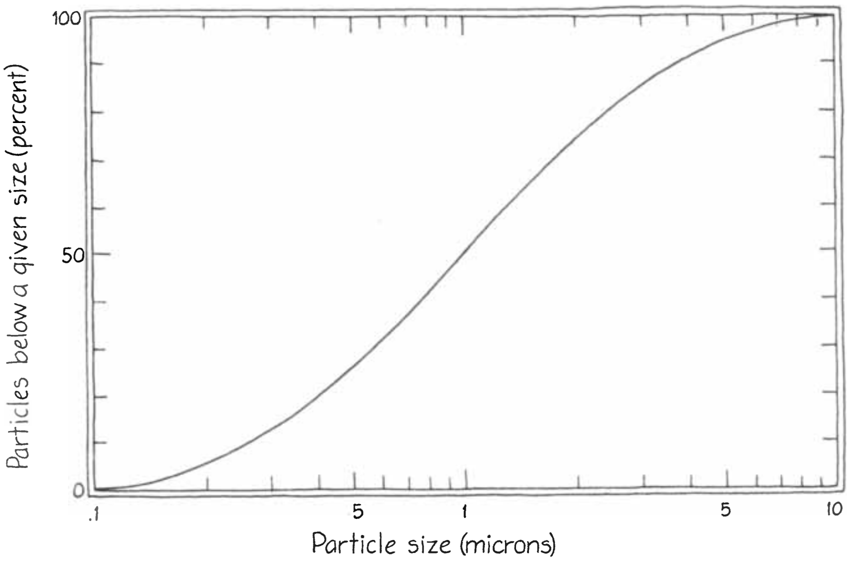
I had to replenish the grit frequently to avoid scratches. Perhaps this was partly because pieces of metal that had been abraded from the disk contaminated it. In addition the amount of grit under the disk seemed to diminish.

Why do scratches occur? Presumably with too much pressure the abrasive particles can no longer roll or tumble smoothly between the moving surfaces while removing small amounts of material, and the larger particles in the grit catch in one surface and gouge the other. Saccocio believes this catching occurs when the two surfaces are at an angle to each other, which happens when the pressure is applied unevenly or when the grinding has already worn the surfaces into concave and convex shapes. I could sense when a scratch was being made because the disk gave a slight jerk and the grinding sound changed noticeably. With an appropriate amount of pressure the abrasive particles grind away the molybdenum surface either molecule by molecule or at least by craterlike sections that are too small to be seen; moreover, the abrasion is confined to the surface.

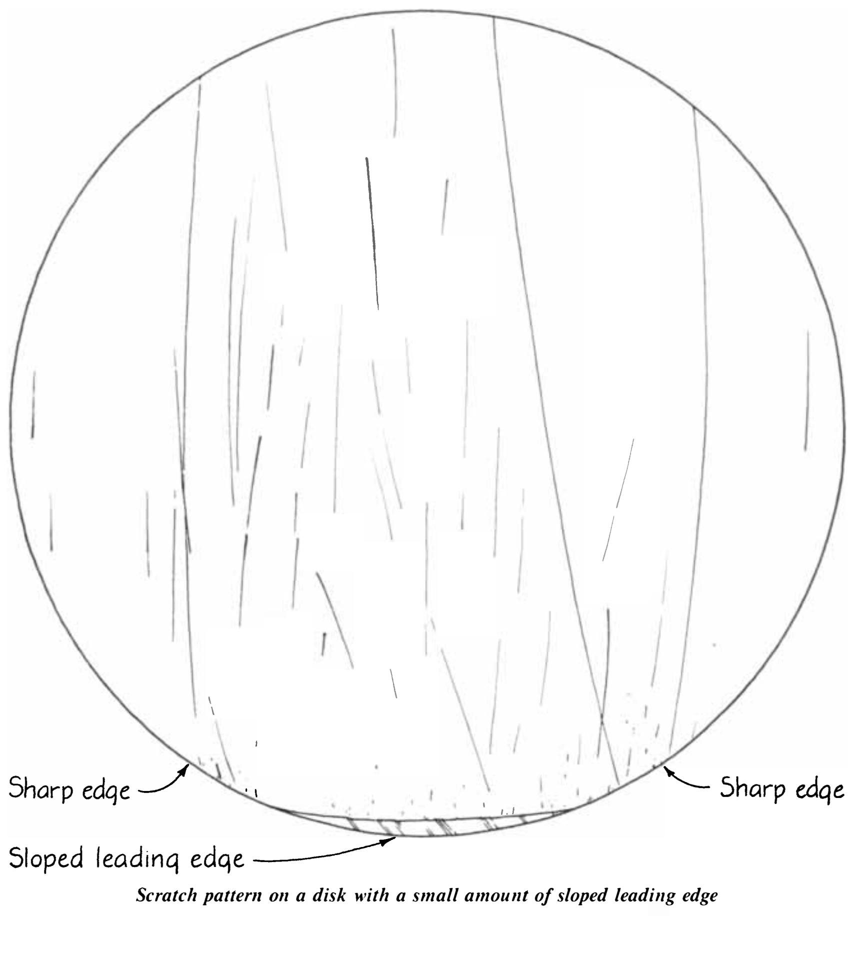
Following Saccocio's lead, I checked the idea that oversize particles cause scratching. I put some 30-micron grit into grits of the two smaller sizes. I first ground the disk over a region of the uncontaminated smaller grit and then, without lifting the disk, moved it into the contaminated region. With about equal amounts of the two grit sizes in the contamination, mixed as uniformly as I could mix them, scratching seemed no more frequent than normal. As I repeated the experiment with smaller amounts of the 30-micron grit in the same amount of smaller grit, however, scratching seemed to be more likely.

In each of the trials I had to maintain about the same amount of pressure on the disk. Some variation in scratch frequency surely resulted from my inability to do this precisely. Still, it appears that if larger particles are present in small amounts, they are more likely to cause scratching. Apparently in such cases the spacing between the disk and the glass is largely determined by the smaller grit. If the larger particles enter the space, they must force the disk up. Since the disk would then be supported in fewer places, the contact points of those larger particles must sustain more pressure. With more pressure on a few spots the larger particles are more likely to catch and gouge the surface.

I next tried to rub the disk in one direction only and then only once or twice before blowing off the grit to examine the molybdenum surface. In that way I could tell where the scratches began and stopped. Some of the scratches started at the leading edge of the disk; others began at random places inside the perimeter. Some of them appeared to be-



Distribution of particle sizes in a "one-micron grit"



Scratch pattern on a disk with a small amount of sloped leading edge gin in a rather wide track that later narrowed. This narrowing could be due to the fracturing of the gouging particle. Other tracks seemed, at least when I inspected them with a simple magnifying lens, to remain about the same width. All the tracks were significantly wider than the particles. One can imagine that the point of a gouging particle catches in a moving surface and that in the course of its subsequent motion it pulls and distorts a fairly wide area of the surface before the distortion is severe enough for the surface to rip. It is possible that some tracks could be narrow at the beginning and could grow wider if the gouging particle digs into the moving surface progressively more, although I could not find any such tracks.

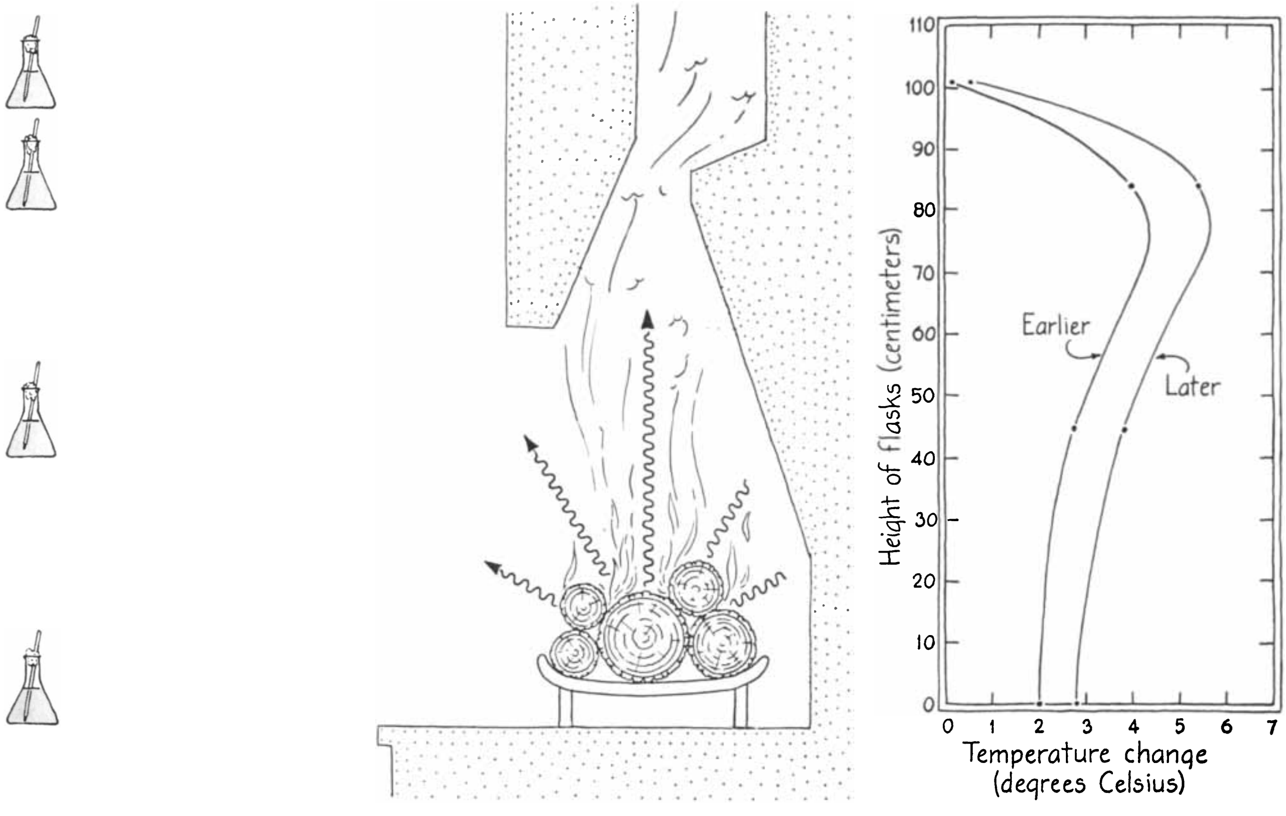
The fact that some scratches begin at the edge and others begin on the inside is probably due to chance. The gouging particles are not round (they would not be likely to gouge if they were) and they have sharp points. It is a matter of chance when one of these particles happens to tumble into an orientation where one side catches and the other side gouges. A degree of chance must also have been introduced by the random motion and pressure I applied to the disk. Occasionally I must have borne down more on one side of the disk and increased the likelihood of scratches on that side.

To test the effect of the leading edge on the frequency of scratching Saccocio performed several experiments with a molybdenum disk encased (except for the surface to be ground) in a hard plastic or epoxy material. The leading edge was the plastic ring surrounding the perimeter of the molybdenum surface. The ring had a slope of from three to five degrees, arranged so that the outside of the ring was farther from the glass than the inside, which was adjacent to the molybdenum surface. Saccocio did his grinding with uncontaminated slurries, and he found that with the sloped leading edge the molybdenum surface was scratched more often than when another (unencased) molybdenum disk having a sharp leading edge was rubbed in the same slurry. Once again, applying the same pressure on both specimens is difficult, and failure to do so can lead to spurious results. Nevertheless, Saccocio believed the difference in the scratch frequencies was due somewhat to the shape of the leading edge. The sloped leading edge gave more opportunity for the oversize particles responsible for scratching to get into the space between the disk and the glass plate.

I repeated this experiment with some 30-micron particles contaminating one and then the other of the smaller grits and found the same general results. I also tried the experiments with an unencased molybdenum disk, purposely sloping one edge, keeping another edge sharp and rubbing the disk through the grit in a single direction, first with the sharp edge leading and then, after inspection, with the sloped edge leading. Although the results varied, they were generally consistent with Saccocio's.

I also noticed a curious scratching pattern that appeared occasionally when the sloped edge was leading. Scratches were mostly to the sides of the sloped edge rather than directly behind it. If this pattern is indeed characteristic (more trials will be required to show whether it is), it might be because the disk pivots over a central ridge of oversize particles running from the sloped edge and across a diameter to the rear. More of the oversize particles will enter by way of the sloped edge than by way of the sharper sides, and this selection tends to raise the center of the disk. Thereafter I shall have to push inadvertently and randomly more on one side than on the other, causing the disk to pivot around the central ridge. Because the sides have fewer oversize particles than the central ridge does, the increase in pressure on one side will result in more gouging there than in the central ridge.

Much more could be done on these experiments. You could get more data with the molybdenum-and-glass system, and you could try other abrasive compounds (both wet and dry) and other grinding materials such as glass (on glass), aluminum, copper, brass and



Flask arrangement for measuring the radiation pattern from a conventional grate lead. The data would be easier to interpret if you could standardize the pressure applied to the specimen. You might try placing a known weight on the disk and then rubbing the disk through the slurry by pushing horizontally on the disk and the weight. For a given grit size, how does the scratch frequency depend on the applied weight? Since the scratching is random, you will need lots of data properly interpreted statistically to answer the question.

When I first put my molybdenum disk on the grit, the full range of particle sizes in the grit was under the disk. Does the full range of sizes remain after you have ground for a while or does the grinding somehow eliminate the larger particles while a sharp leading edge prevents new ones from entering the space under the specimen? Why did I have to replenish the grit in order to avoid scratches? When I ground with the 15-micron grit for a while, the glass plate became opaque, and it was covered with a fine dust that seemed to decrease the frequency of scratching. Does this result arise because the disk is then so close to the glass that oversize particles have even more trouble getting under it?

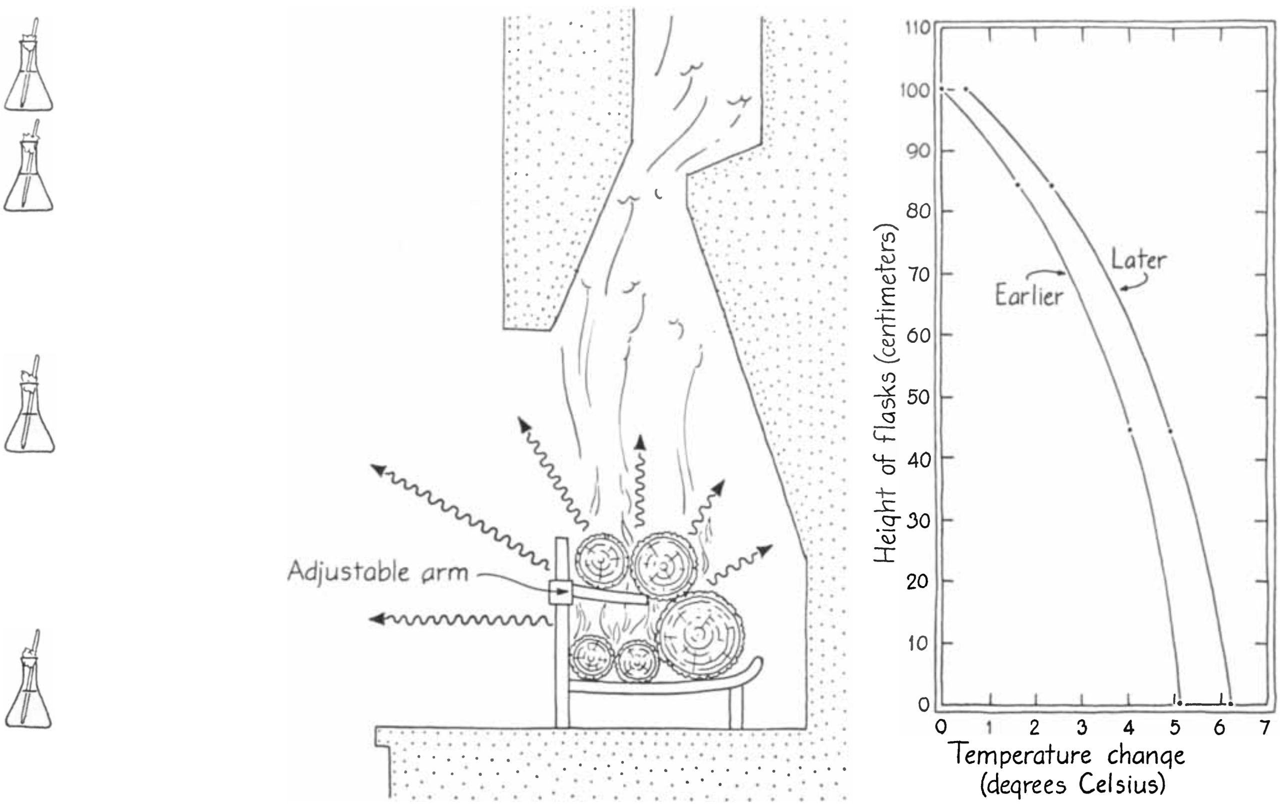
The fireplace is an ancient and common source of heat for the home. The output of the fireplace is mainly limited to one of the three possible types of heat transfer: convection, conduction and radiation. Little of the heat comes into the room by convection because the heated air is lost up the chimney. Hardly any heat enters by conduction for want of solid objects between the fire and the room. Most of the heat enters by radiation, largely in the infrared region. Most of this radiated heat comes from the hot coals in the fire rather than from the visible flames, which is why the fire is hot only after enough wood has burned to create the coals.

The fireplace is considered to be an inefficient source of heat because only a relatively small fraction of all the energy released by the fire ends up in the room. Even some of that energy is often wasted because it has become fashionable to place a glass screen in front of the fireplace to prevent sparks from flying into the room. The glass is rarely designed to transmit infrared radiation.

Until the conservation of energy became a matter of widespread interest little was done about improving the efficiency of the fireplace. Lawrence Cranberg, a physicist of Austin, Tex., began investigating the stacking of logs in the fireplace. He concluded that the conventional stacking (three or so in a triangular cross section) was a poor design because much of the heat radiated by the fire appeared to be sent upward and therefore lost to the brickwork above the fire or to the chimney. Cranberg also noticed that the coals, which are responsible for most of the radiated heat, are not exposed well to the room but are instead hidden under the unburned wood. Thus in order to heat the room enough one is tempted to burn the logs at such a rate that flames leap upward into the chimney and the logs are consumed fairly quickly.

With this poor radiation pattern in mind Cranberg designed a new fireplace grate that holds the logs in a slotlike arrangement. His thought was that if the fire were contained in a slot of logs, the exposed area of coals would be greatly increased and the radiated heat could be beamed into the room (instead of the chimney) by the slot. Since the radiated heat would be produced more efficiently, the fuel would burn slower, with less open flame and better exposure of the coals to the room. The slot would even be partly self-controlling to provide a uniform horizontal distribution of radiation, because if the slot widened at some point, more heat would be radiated out of the slot from that point and the local temperature of the area would drop somewhat. Eventually the rest of the slot would widen to the same extent. Cranberg's grate, which is now patented, is available from the Texas Fireframe Company (P.O. Box 3435, Austin, Tex. 78764) in five sizes.

I arranged an experiment to measure the radiation patterns from the conventional grate and stacking pattern and



The radiation pattern measured from a "slot design" grate from Cranberg's slot design (Texas Fireframe Model U-25, priced at $39.95 plus 10 percent for shipping in the U.S.). Lacking an infrared detector, I decided to measure the radiated heat with a simple arrangement of flasks containing water and thermometers. Directly in front of a fireplace I placed four flasks at various heights from the floor and in a vertical line. They were 250-milliliter Erlenmeyer flasks containing 200 milliliters of water. I put a Celsius thermometer in each flask, and then (to avoid losses from evaporation) I closed the top of each flask by stuffing the opening with plastic food wrap. Some of the infrared radiation emitted by a fire would be absorbed by the water in the flasks and would warm it. I could monitor the warming with the thermometers and thus have a rough idea of the vertical radiation pattern from a particular fire.

One of the problems with this method is that the warming of the water lags behind any rise in the intensity of the radiation from the fire. I was less concerned with how the radiation pattern varied with time, however, than with how it varied vertically. I therefore made the assumption that the general shape of the vertical distribution of radiation into the room could be roughly determined by monitoring the changing temperature of the water in the flasks.

I tried to choose logs of the same type and sizes for both fires. The sizes were determined by the recommendations in a publication Cranberg sent me: there should be one large log about six inches in diameter in the back of the slot, one or two logs of medium size on the top and two smaller logs on the bottom. Although my wood was all the same type, I did not know or find out what kind it was. Moreover, my effort to keep the sizes of the logs the same for both fires was at best an approximation, since logs are far from being standardized objects like machine parts. Still, I think my results were not critically dependent on the type or sizes of the logs. The amount of heat radiated into the room by a fire laid to Cranberg's design, however, can be greatly increased by using a bigger rear log so that the slot is wider.

Because I wanted to measure the radiation pattern I was determined not to monitor the air temperature in the room. The air temperature would vary vertically because warmer air rises. The air temperature would also depend on the size of the room and on its contents and so would be of little general value. Hence I determined the change in the temperature of each flask about every five or 10 minutes after I had lighted a fire. A slight cool breeze was blowing through the room from a door partly open to the outside throughout the experiments, which helped to keep the room temperature constant. I lighted both fires on the same evening to avoid possible variations in the experiment because of changes in the outside air temperature and hence the draft rate through the chimney. Some variation probably occurred anyway because the outside air became cooler.

I began with the conventional arrangement, monitoring the water temperatures for about an hour. Typical temperature changes are plotted in the illustration on page 142. Note that the radiation pattern peaked in intensity in the flask that was 84 centimeters from the floor, was almost zero in the top flask at a height of 101 centimeters and was below the peak value in the lower two flasks. Apparently the heat was radiated strongly upward but not much horizontally. Evidently the top flask received little radiation, partly because the overhang from the fireplace shielded it from the fire and partly because it presented a smaller cross-sectional area to the fireplace than the other flasks did. The overall radiation pattern from the conventional stacking of logs would seem to be directed upward. As Cranberg suspected, much of the radiation was being lost to the brick in the overhang and to the chimney.

I then replaced the conventional grate with the slot design, installed fresh logs in the slot arrangement and lighted the fire. I tried to keep the logs lying tightly against one another to avoid slits or holes in the slot arrangement. The patented grating is helpful in this respect because the arms holding the upper logs are adjustable in height, so that the upper logs can be positioned snugly against the rear log. The adjustable arms are also useful because the rate of combustion can be controlled by varying the height of the slot. The burning was slower with this arrangement, and flames appeared only on the small lower logs but were uniform across the length of the slot and required no rotation or stirring of the logs. The upper logs were partly charred on the surfaces facing the slot, and the one pressed against the rear log was glowing with red coals. The rear log was largely red coals inside the slot. Presumably this was a good design because little of the wood was being consumed in the relatively useless open flames. The coals necessary for good heat radiation were there instead.

The radiation pattern was again monitored for about an hour. This time the vertical profile appeared to peak at about floor level. As Cranberg had calculated, the pattern was largely determined by the slot directing the radiation into the room. Less radiation was directed upward to the higher flasks. Thus the overall pattern implies that little of the radiated heat was lost upward to the overhang or the chimney and that nearly all of it must have been coming out into the room.

The radiation pattern from the slot

design is easily determined by hand. I moved my hand upward in front of the fire at floor level and then held it directly over the top logs. The heat was very noticeable in the first position and much less so in the second.

Cranberg's conception of the radiation pattern from his log holder is correct. You might want to experiment with the adjustment of the arms, the height of the slot and the length of the logs. You could also measure the radiation pattern and intensity against time for the two types of log stacking. Does the conventional pattern always produce radiation slower than the new design? How does the rate of production from the slot fire depend on the size of the rear log and thus on the height of the slot? (To do this experiment with reliable results you will have to standardize the logs and whatever technique you use to start the fire. Cranberg recommends starting with newspaper if the wood is seasoned and with kindling if it is green.)

You could replace the wood logs with the pressed-paper logs that are available commercially. (Similar logs of waxed paper are designed to burn by themselves.) Can you find an advantageous method of stacking the logs other than in Cranberg's arrangement? While you are at it you could also test the transmission properties of a glass or metal screen placed in front of the fire. The radiation pattern should not change much (although the screen will distort it to a certain extent), but you should find your water flasks warming somewhat more slowly.

Last December I described Haidinger's brushes (the hourglass figure you can see when you look into linearly polarized light) and various experiments you can do with the figure and a homemade quarter-wave plate. Part of the explanation for that series of experiments depended on the birefringence of your cornea, which I attributed to a preferred orientation of the collagen fibrils in the cornea. I have been corrected by Frederick A. Bettelheim of Adelphi University. According to research by him and by others, only 25 percent of the cornea's birefringence is due to such a preferred orientation of the fibrils; the other 75 percent results from the fact that these long, thin fibrils do not have the same index of refraction as the material in which they are embedded. The fibrils do lie parallel in a layer, but their orientation from layer to layer through the thickness of the cornea is almost random. Although the fibrils are birefringent and would contribute a much larger net birefringence to your vision if they were parallel throughout the thickness of the cornea, the random orientation from layer to layer diminishes that contribution.

BIBLIOGRAPHY

Readers interested in further explanation of the subjects covered by the articles in this issue may find the following lists of publications helpful.

MATHEMATICAL GAMES

IMPOSSIBLE OBJECTS AS NONSENSE SENTENCES. D. A. Huffman in *Machine Intelligence: Vol. 6,* edited by Bernard Meltzer and Donald Michie. American Elsevier Publishing Company, Inc., 1971.

SOME NOVEL MÖBIUS STRIPS. Charles Joseph Matthews in *Mathematics Teacher,* Vol. 65, No. 2, pages 123–126; February, 1972.

THE THEORY OF BRAIDS AND THE ANALYSIS OF IMPOSSIBLE FIGURES. Thaddeus M. Cowan in *Journal of Mathematical Psychology,* Vol. 11, No. 3, pages 190–212; August, 1974.

ORGANIZING THE PROPERTIES OF IMPOSSIBLE FIGURES. Thaddeus M. Cowan in *Perception,* Vol. 6, No. 1, pages 41–56; 1977.

THE GAS CENTRIFUGE

TECHNICAL BASIS OF THE GAS CENTRIFUGE. Donald R. Olander in *Advances in Nuclear Science and Technology: Vol. 6,* edited by Ernest J. Henley and Jeffery Lewins. Academic Press, Inc., 1972.

URANIUM ENRICHMENT BY GAS CENTRIFUGE. D. G. Avery and E. Davies. Mills & Boon Limited, London, 1973.

CENTRIFUGATION. Stelio Villani in *Isotope Separation.* American Nuclear Society, 1976.

STELLAR WINDS

THE CIRCUMSTELLAR ENVELOPE OF ALPHA HERCULIS. Armin J. Deutsch in *The Astrophysical Journal,* Vol. 123, No. 2, page 210; March, 1956.

THE SOLAR WIND. E. N. Parker in *Scientific American,* Vol. 210, No. 4, pages 66–76; April, 1964.

THE FAR-ULTRAVIOLET SPECTRA OF SIX STARS IN ORION. Donald C. Morton in *The Astrophysical Journal,* Vol. 147, No. 3, pages 1017–1024; March, 1967.

THE HOMINIDS OF EAST TURKANA

EARLIEST MAN AND ENVIRONMENTS IN THE LAKE RUDOLF BASIN: STRATIGRAPHY, PALEOECOLOGY AND EVOLUTION. Edited by Yves Coppens, F. C. Howard, Glynn Ll. Isaac and Richard E. F. Leakey. University of Chicago Press, 1976.

HUMAN ORIGINS: LOUIS LEAKEY AND THE EAST AFRICAN EVIDENCE. Edited by Glynn Ll. Isaac and Elizabeth R. McCown. W. A. Benjamin, Inc., 1976.

KOOBI FORA RESEARCH PROJECT, VOL. 1: THE FOSSIL HOMINIDS AND AN INTRODUCTION TO THEIR CONTEXT 1968–1974. Edited by Mary Leakey and Richard E. F. Leakey. Oxford University Press, 1978.

THE TOMATO

THE HISTORY OF THE USE OF THE TOMATO: AN ANNOTATED BIBLIOGRAPHY. G. A. McCue in *Annals of the Missouri Botanical Garden,* Vol. 39, pages 289–348; 1952.

POTENTIAL GENETIC RESOURCES IN TOMATO SPECIES: CLUES FROM OBSERVATIONS IN NATIVE HABITATS. Charles M. Rick in *Genes, Enzymes, and Populations,* edited by Adrian M. Srb. Plenum Press, 1973.

VARIETAL INFLUENCE ON NUTRITIONAL VALUE. M. A. Stevens in *Nutritional Qualities of Fresh Fruits and Vegetables,* edited by Philip L. White. Futura Publishing Company, Inc., 1974.

THE TOMATO. Charles M. Rick in *Handbook of Genetics: Vol. 2,* edited by Robert C. King. Plenum Press, 1975.

NEGATIVE ABSOLUTE TEMPERATURES

A NUCLEAR SPIN SYSTEM AT NEGATIVE TEMPERATURE. E. M. Purcell and R. V. Pound in *The Physical Review,* Vol. 81, No. 2, pages 279–280; January 15, 1951.

THERMODYNAMICS AND STATISTICAL MECHANICS AT NEGATIVE ABSOLUTE TEMPERATURES. Norman F. Ramsey in *The Physical Review,* Vol. 103, No. 1, pages 20–28; July 1, 1956.

SPIN TEMPERATURE. A. Abragam and W. G. Proctor in *The Physical Review,* Vol. 109, No. 5, pages 1441–1458; March 1, 1958.

THE THERMOSTAT OF VERTEBRATE ANIMALS

TEMPERATURE REGULATION IN MAMMALS AND OTHER VERTEBRATES. J. Bligh. Elsevier North-Holland Publishing Company, 1973.

PROBING THE ROSTRAL BRAINSTEM OF ANESTHETIZED, UNANESTHETIZED, AND EXERCISING DOGS, AND OF HIBERNATING AND EUTHERMIC GROUND SQUIRRELS. Harold T. Hammel, H. Craig Heller and F. R. Sharp in *Federation Proceedings,* Vol. 32, pages 1588–1596; 1973.

THERMOREGULATION DURING SLEEP AND HIBERNATION. H. Craig Heller and S. F. Glotzbach in *International Review of Physiology,* Vol. 15, pages 147–187; 1977.

SLEEP AND HIBERNATION: ELECTROPHYSICAL AND THERMOREGULATORY HOMOLOGIES. H. Craig Heller, J. M. Walker, G. L. Florant, S. F. Glotzbach and R. J. Berger in *Strategies in Cold: Natural Torpidity and Thermogenesis,* edited by Lawrence Wang and Jack W. Hudson. Academic Press, Inc., 1978.

THE TEKTITE PROBLEM

TEKTITE STUDIES, I–X. G. H. R. von Koenigswald in *Proc. K. Ned. Akad. Wetens., Amsterdam,* Series B, Vols. 63–78; 1960–1975.

ON THE LUNAR ORIGIN OF TEKTITES. Dean R. Chapman and Howard K. Larson in *Journal of Geophysical Research,* Vol. 68, No. 14, pages 4305–4358; July 15, 1963.

THE CHEMISTRY, ORIGIN AND PETROGENETIC IMPLICATIONS OF LUNAR GRANITE AND MONOZONITE. M. J. Rutherford, P. C. Hess, F. J. Ryerson, H. W. Campbell and P. A. Dick in *Proceedings of the Lunar Science Conference, Seventh, 1976.* Pergamon Press, 1976.

TEKTITES AND THEIR ORIGIN. John A. O'Keefe. Elsevier Scientific Publishing Company, 1976.

ORGANISMS THAT CAPTURE CURRENTS

INTERFACIAL ORGANISMS: PASSIVE VENTILATION IN THE VELOCITY GRADIENTS NEAR SURFACES. Steven Vogel and William L. Bretz in *Science,* Vol. 175, No. 4018, pages 210–211; January 14, 1972.

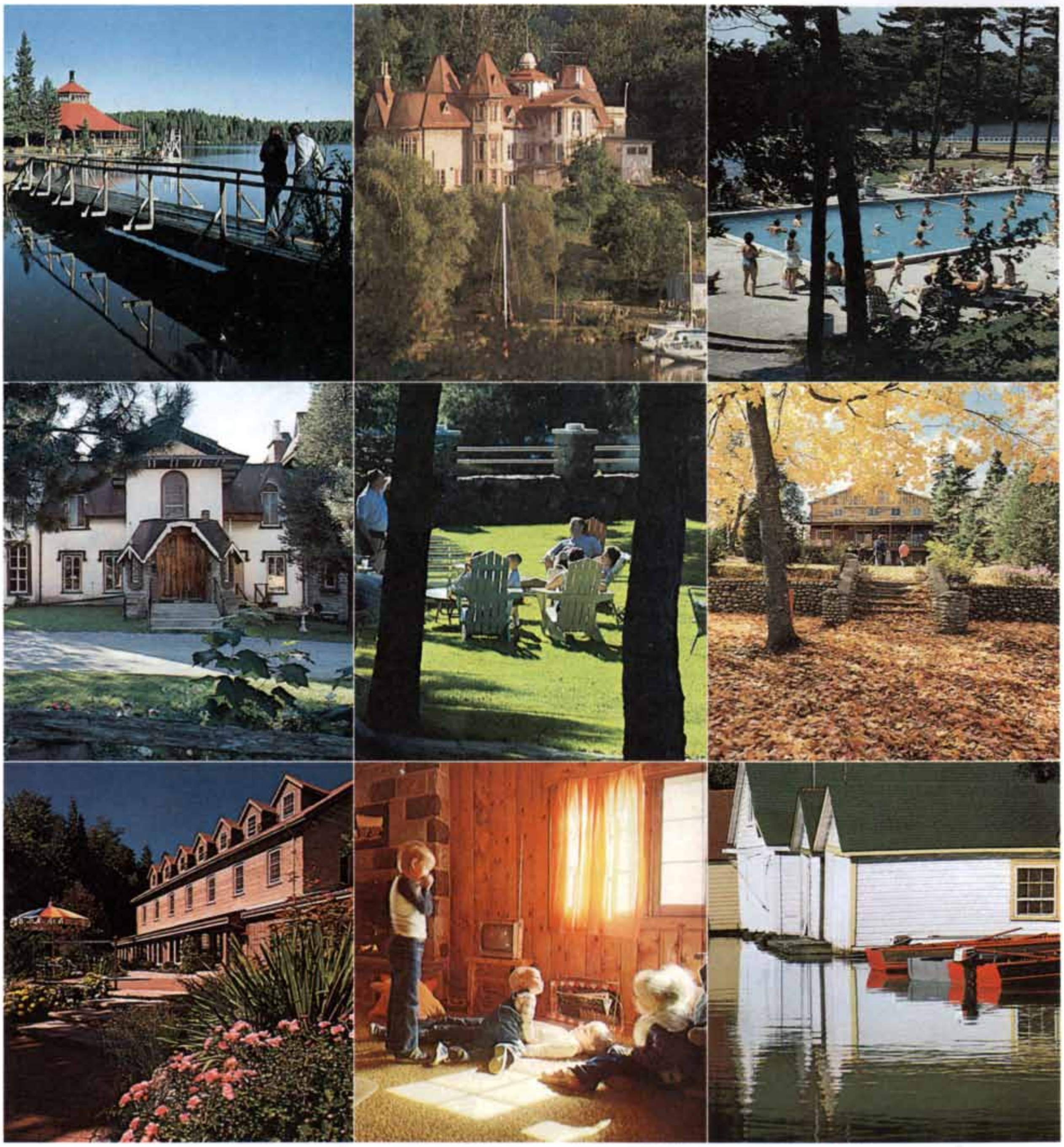
WIND-INDUCED VENTILATION OF THE BURROW OF THE PRAIRIE DOG, *CYNOMYS LUDOVICIANUS.* Steven Vogel, Charles P. Ellington and Delbert L. Kilgore, Jr., in *Journal of Comparative Physiology,* Vol. 85, No. 1, pages 1–14; 1973.

CURRENT-INDUCED FLOW THROUGH LIVING SPONGES IN NATURE. Steven Vogel in *Proceedings of the National Academy of Sciences,* Vol. 74, No. 5, pages 2069–2071; May, 1977.

BRACHIOPOD ORIENTATION TO WATER MOVEMENT, 1: THEORY, LABORATORY BEHAVIOR AND FIELD ORIENTATIONS. Michael C. LaBarbera in *Paleobiology,* Vol. 3, No. 3, pages 270–287; Summer, 1977.

THE AMATEUR SCIENTIST

BASIC MECHANISM IN FREE ABRASIVE GRINDING. Edward J. Saccocio in *Applied Optics,* Vol. 14, No. 11, pages A224-A227; November, 1975.



Jameson Irish

"~~Scotch~~ on the rocks."

If you like Scotch, you'll love light Jameson Irish.

Try a glass of Jameson Irish the way you would your favorite Scotch.

You'll notice how much it tastes like fine Scotch—only lighter and more delicate. Not smoky tasting like Scotch.

The dedicated Scotch drinker will instantly appreciate this flavor difference.

Though it may take a little time getting used to saying, "Jameson Irish on the rocks, please."

Jameson. World's largest-selling Irish Whiskey.

80 & 86 PROOF • CALVERT DIST. CO., N.Y.C.