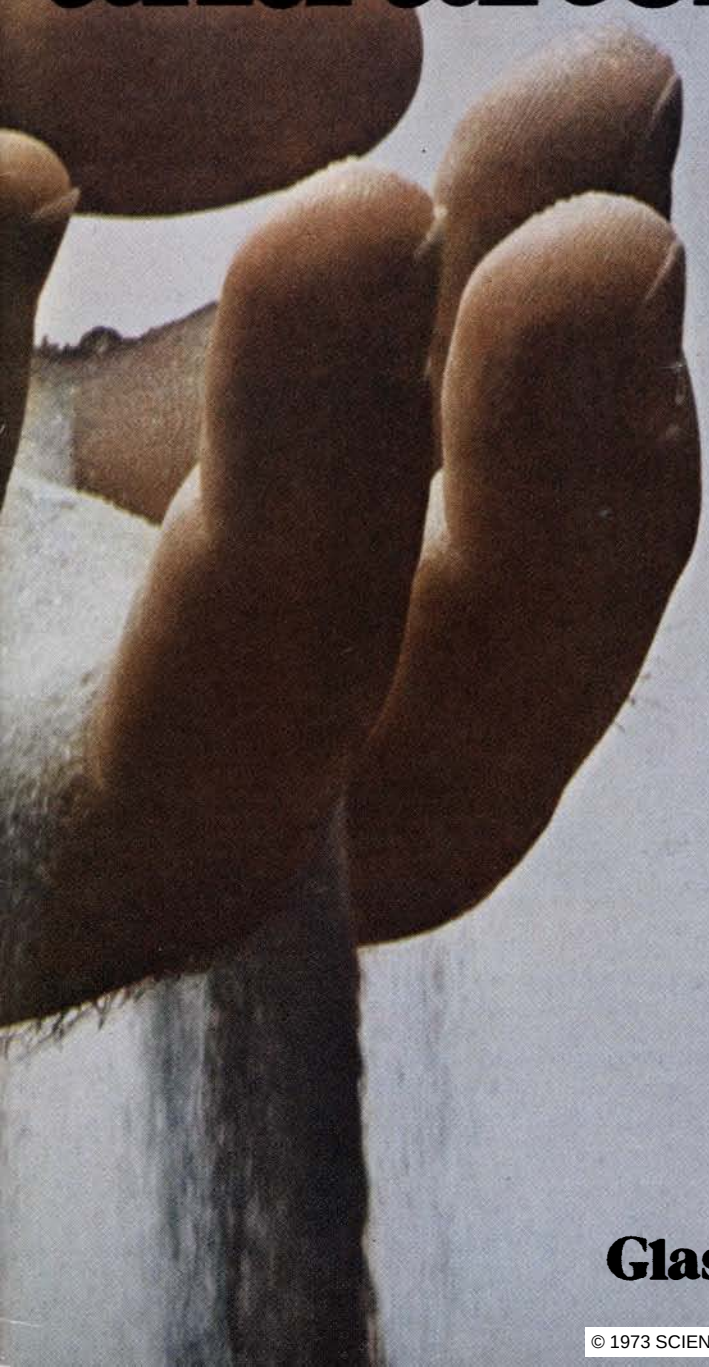
SCIENTIFIC AMERICAN

PHYSICS OF BRASSES

ONE DOLLAR

July 1973

This is a soft drink bottle before it's born and after it dies.



About 73% of glass is sand—the most common substance in the earth's crust.

It's this natural composition that appeals to the nation's homemakers. They like what glass shows them about the products they buy. They like the purity of glass and its re-sealability. They like the fact that glass protects the flavor of food and beverages.

Equally important, it's this natural composition that makes a glass container an ideal environmental package. It can be returned to the earth in virtually its original form.

Of course, not all glass goes back to the earth. But because glass is recyclable, more and more is going back to the industry for making new glass bottles and jars.

The glass industry is also continuing to develop new uses for used glass such as road paving material, construction panels, bricks, insulation, terrazzo and reflective paints.

When you add it all up, it's easy to see why consumer preference for glass containers continues to grow. In terms of its attributes, its utility and recyclability, glass is a natural. For a copy of "The glass container story," write: Glass Container Manufacturers Institute, Dept. A, 1800 K St., N.W., Washington, D.C. 20006.

Glass. It's a natural.

The Dodge Colt GT

Built across the ocean. Serviced across the street.

Up until now, when you bought a sporty little Japanese import, you pretty much had to rely on dealerships the imports themselves set up in the U.S., for parts and service.

But now, along comes the Dodge Colt GT.

Sporty as you could ask for.

But,more important, the Colt comes with a system of ready-made U.S. dealerships.

That means when your Colt gets to America, it's sold and serviced by a network of Dodge Colt Dealers that stretches from coast to coast.

So if you're thinking sporty, think of it this way: The Dodge Colt GT comes from Japan ...to America... c/o Dodge.

Dodge

Dodge Trucks

Here are just a few of the 35 things that come standard on the Colt GT: Overhead cam hemi engine ● Four-on-the-floor ● Front disc brakes ● Reclining bucket seats ● Adjustable steering column ● Rally striping inside and out ● Soft-rim sports steering wheel ● Flow-through ventilation ● Whitewall tires ● Center console with storage bin.

CHRYSLER MOTORS CORPORATION

Established 1845 **SCIENTIFIC AMERICAN** July 1973 Volume 229 Number 1

ARTICLES

17 **PUBLIC POLICY ON FERTILITY CONTROL, by Frederick S. Jaffe**
In only a dozen years U.S. policy on contraception has been completely reversed.

24 **THE PHYSICS OF BRASSES, by Arthur H. Benade**
The trumpet, for example, interacts with the player to produce musical tones.

36 **ADVANCED COMPOSITE MATERIALS, by H. R. Clauser**
The concept of combining two materials to produce a third continues to evolve.

52 **THE IMMUNE SYSTEM, by Niels Kaj Jerne**
This diffuse organ has the assignment of monitoring the identity of the body.

64 **METEORITES AND COSMIC RADIATION, by I. R. Cameron**
Nuclear reactions in meteorites provide clues to the history of the solar system.

74 **THE BRIDEPRICE OF THE SEBEI, by Walter Goldschmidt**
On the economics of how a man negotiates for a wife in an African culture.

86 **PLATE TECTONICS AND MINERAL RESOURCES, by Peter A. Rona**
Exploration for useful minerals is aided by the new picture of the earth's crust.

96 **BRAIN MECHANISMS IN MOVEMENT, by Edward V. Evarts**
In the control of muscles the highest functions are not in the cerebral cortex.

DEPARTMENTS

8 LETTERS
12 50 AND 100 YEARS AGO
14 THE AUTHORS
46 SCIENCE AND THE CITIZEN
104 MATHEMATICAL GAMES
110 THE AMATEUR SCIENTIST
117 BOOKS
122 BIBLIOGRAPHY

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. 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: $10 PER YEAR, U.S., POSSESSIONS AND CANADA; $12.50 PER YEAR, ALL OTHER COUNTRIES. EUROPEAN SUBSCRIPTION COPIES MAILED C/O INTERNATIONAL DISTRIBUTORS, KOEPOORTBRUG 1, B 2000 ANTWERP, BELGIUM.

THE COVER

The painting on the cover shows a part of the tubing of a French horn. The four silvery round objects at the top are the valves. The French horn, along with the trumpet and the trombone, is a member of one of the two main families of brass musical instruments (see "The Physics of Brasses," page 24). The instruments in this family have a considerable length of cylindrical tubing in the middle section and an abruptly flaring bell. The other main family of brasses is the conical family, in which much of the tubing increases in diameter from the mouthpiece to the bell and the flare of the bell is less pronounced than it is in the cylindrical family. Among the instruments in the conical family are the flügelhorn, the alto horn and the tuba.

THE ILLUSTRATIONS

Cover painting by Ted Lodigensky

Page	Source
18–23	Illil Arbel
25–35	Dan Todd
37	Michael F. Henry, General Electric Research and Development Center
38–39	Union Carbide Corporation
40	K. M. Prewo, United Aircraft Research Laboratories
41	Lewis Research Center, National Aeronautics and Space Administration
42	E. I. du Pont de Nemours & Company
43	St. Joe Minerals Corporation
44	John L. Walter, General Electric Research and Development Center
53	Niels Kaj Jerne
54–55	Bunji Tagawa
56	N. M. Green and Robin Valentine, National Institute for Medical Research, London
57	Bunji Tagawa
58	Bunji Tagawa (*top*), Benvenuto Pernis (*bottom*)
59	Benvenuto Pernis
60	Bunji Tagawa
64	Joseph I. Goldstein and Harvey Yakowitz
68–72	Illil Arbel
73	David J. Barber and P. Buford Price, Jr.
75	Walter Goldschmidt, University of California at Los Angeles
76	Enid Kotschnig
77	Walter Goldschmidt, University of California at Los Angeles
78–85	Enid Kotschnig
86	Earth Resources Technology Satellite
88	F. W. Goro (*top*), David A. Ross (*bottom*)
89–95	George V. Kelvin
97–100	Tom Prentiss
101	Edward V. Evarts
102–103	Tom Prentiss (*left*), Edward V. Evarts (*right*)
104–106	Illil Arbel
110	Roger Hayward
111	Aristid V. Grosse
114–117	Roger Hayward

University of Washington students and Tektronix computer display terminals are probing the secrets of sea life in an effort to find a balance between man's need and the oceans' resources.

They are research associates in the NORFISH project of the University's Center for Quantitative Science in Forestry, Fisheries and Wildlife.

They have a wealth of data to draw from, including 20 years of dusty commercial fishery statistics. Analyzing such input once meant hours of hand-drawn charting and days of drifting on paper seas. No more.

The engaging interactivity of our graphic terminals take them from data input to viewable charts in minutes. At the touch of a switch, they copy the display with a Tektronix Hard Copy Unit.

With little training they can generate orthographic projections and 3-D surfaces which describe the ocean environment far more succinctly than words alone.

Across campuses throughout the country, our terminals are adding new vision to medical diagnostic training, high energy and pattern recognition physics, biophysical waveforms and speech therapy.

At one university, library people, inspired by a cardiac diagnosis application, are developing a Library Retrieval System using our terminal.

Thus the ripple becomes a wave. Applied graphics is helping teaching and research keep more than current on our nation's campuses.

For an informative booklet on graphics and Tektronix display technologies, write: Colin Barton, Tektronix, P.O. Box 500, Beaverton, Oregon 97005.

TEKTRONIX®

Computer graphics is bringing to light the ecosystems of the deep.

Graphics. The mind's eye, for those who think tomorrow.

These calculators have such special significance for scientists and engineers that we devoted this space to describe them in some detail. Other new instruments and systems stemming from our measurement/computation technologies are being developed: look for them in the coming months.

Three alternatives (two of them brand new) to the drudgery of paper-and-pencil mathematics.

Almost 300 years ago, Gottfried Wilhelm Leibniz wrote, "It is unworthy of excellent men to lose hours like slaves in the labor of calculation." He was right, but how could anyone avoid it then?

Today, in any field of science or engineering, tedious manual calculation is neither wise nor necessary because things have changed significantly for the better, even in the last few months. We're convinced you should never again labor with slide rule, tables, scratch pads and adding machines . . . wasting your creative time getting answers that aren't as accurate as you'd like.

Any of the economical calculators that we describe in this month's message is as easy to operate as an adding machine but incomparably more powerful. All are pre-programmed to perform not only the basic arithmetic operations but also transcendental and statistical functions. All calculate positive and negative numbers throughout a 200-decade range. All automatically keep track of the decimal point and can display answers to the tenth significant digit.

One of the traits that sets them apart from the recent flood of electronic calculators is a four-register operational stack that is solidly based on computer theory. The stack automatically stores intermediate results obtained during your calculations — whether they be serial, chain or mixed chain — and brings them back to the working register when they are needed to complete the calculations. In plain English, the stack relieves you of the necessity to make scratch notes and re-enter intermediate values: it does it for you, automatically and without error.

HP-35.
The electronic slide rule

Small enough so that you can easily carry it around in your shirt pocket — it weighs only 9 ounces including rechargeable battery — the HP-35 is the original electronic slide rule introduced a little over a year ago. It has since become the constant companion of more than 75,000 scientists and engineers around the world.

Due to the economies realized in this long production run, the price of the popular HP-35 has been reduced to $295.*

HP-35

The HP-35 is easier to use, 10 times faster and significantly more accurate than the slide rule. With a single keystroke and in less than a second, it performs trigonometric (sin, cos, tan), logarithmic (log x, ln x, e^x) and other commonly used functions (x^y, $1/x$, $\sqrt{x}$, π) as well as the four arithmetic operations. It also calculates inverse trig functions.

In addition to its computer-like operational stack, the HP-35 has a constant storage register which lets you store any number and recall it as often as you want for repeat operations, without ever having to re-enter it.

The HP-35 comes with owner's handbook, battery pack, AC adapter/recharger, carrying case and travel case.

HP-45.
The scientific pocket calculator

A direct descendant of the "electronic slide rule," the new HP-45 packs nearly twice the computational power into the same package. The trick is that it has a unique gold-colored "shift" key that doubles the function of 24 of its 35 keys. Hence it does all that the HP-35 does . . . and then some.

The HP-45 is the first pocket calculator with *nine* addressable memory registers besides its operational stack. You can store data in each one — any number that appears on the display — and recall it to the working register whenever you want. (Let your imagination picture the calculating horsepower of this feature for register arithmetic, conversions, continued products, payrolls . . .)

The HP-45 also has a 14th register, called "Last X" in which the last input argument is auto-

HP-45

HP-46

matically stored. You can recall this number by pushing the "Last X" key . . . then proceed to correct it or to perform calculations with it.

There's more. The HP-45 lets you do trig calculations in any of three angular modes (degrees, radians or grads) and converts angles in any mode to degrees/minutes/seconds instantly, and vice-versa. It lets you convert polar coordinates to rectangular and vice-versa, at a single keystroke. Add or subtract vector components in polar or rectangular coordinates. Perform two dimensional accumulations for vector calculations. And convert U.S. units of length, weight or volume to metric, and vice-versa . . . to 10-digit accuracy.

The HP-45 costs $395* including owner's handbook, quick reference guide, battery pack, AC adapter/recharger, carrying case and travel case.

HP-46.
The scientific printing calculator

If you need a permanent record of your calculations and don't insist on the size and portability of our two shirt-pocket wizards, the HP-46 was designed especially for you.

It has all the computational ability of the HP-45 with the important addition of a printer. The printer's extended set of alphanumeric symbols lets it "talk back" to you as you perform your calculations. The HP-46 prints easy-to-read symbols with all operations you perform, and clearly labels the results. Its printout thus constitutes a complete permanent record.

For example, when you calculate the mean and standard deviation of a series of numbers, the printer lists each entry with a $\Sigma+$ symbol. Then, after you push the $\bar{x}$, s key, it prints the results, in order: the number of entries, the calculated standard deviation and the mean . . . each clearly labeled.

On command, the printer will also record the contents of the operational stack or the 9 addressable memory registers, each also clearly labeled. Should you make a logical error in data entry or call for an improper calculation, the printer will make an error note with a reference to an explanation contained in the operating manual.

A 15-digit LED display is available as an option.

The HP-46 costs $695* including owner's handbook, printer paper, and carrying case.

*Domestic USA prices only.

00310

Mail to: Hewlett-Packard,
1504 Page Mill Road, Palo Alto, CA 94304

I want an alternative. Send me complete information on:
( ) HP-35 Electronic Slide Rule
( ) HP-45 Scientific Pocket Calculator
( ) HP-46 Scientific Printing Calculator

Name_____

Title_____

Company_____

Address_____

City_____ State_____ Zip_____

HEWLETT hp PACKARD

Sales, service and support in 172 centers in 65 countries.
Palo Alto, California 94304. Offices in principal cities throughout the U. S.

LETTERS

Sirs:

In your April issue you published a synopsis of the United Nations' *Demographic Yearbook*'s most important results, just a month after Tomas Frejka's interesting article proposing a stable world population of about 8.4 billion. It might therefore be instructive to compare the UN's data with an empirical fit to the then available data made in 1960 by H. von Foerster, P. M. Mora and L. W. Amiot (*Science,* Volume 132, pages 1291–1295, November 4, 1960). The UN gives a current (1971) population of 3.7 billion, with a growth rate of 2 percent, or an "apparent doubling time" of 35 years. Extrapolation of the empirical curve from 1960 to 1971 would have predicted (to within 7 percent) a population of 3.32 billion, with an apparent doubling time of 39 years, so that the population has grown and is currently growing at a slightly faster rate than the empirically fitted curve. What makes these facts alarming is that in 1960 (and still) that empirically fitted curve, fitting data for the past 2,000 years, predicted populations far in excess of what any reputable demographer would risk, and in fact predicted a population growth that is faster than any exponential. It is clear from the UN data that no progress whatever has been made during the past decade in slowing the worse than cancerous growth of the human population. Mr. Frejka's clarion call to complacency seems at best somewhat premature.

THOMAS VON FOERSTER

Department of Physics
Harvard University
Cambridge, Mass.

Sirs:

I did not, and did not intend to, advocate complacency, nor did I wish to join the alarmists. The purpose of my research is to provide a realistic long-term perspective on population growth employing up-to-date knowledge. Mr. von Foerster misreads my paper as a proposal. It is an analysis "proposing" nothing. Surely one can try to say something about the future without becoming an advocate.

Whatever the reasons may be (progress and modernization, education, lack of resources, environmental degradation, "high" population density, etc.), national populations—and thus eventually the world population—will sooner or later have a tendency to approximate a stationary nongrowing state. The historical experience of the so-called developed countries over the past 50 years seems to support this statement. Individuals, families and nations strive to maximize their welfare and they shape their behavior (economic, educational and demographic) accordingly. The demographic behavior of these nations is pointing in the direction of a rough equilibrium of deaths and births, that is, stationarity.

What are the alternatives ahead for the population of the world, depending mainly on demographic trends in the developing countries?

Since presumably—in the absence of catastrophes—mortality will continue to decline, one can illustrate the type of fertility decline that is needed in order to stabilize world population in the 21st century at six, eight, 11, 15 or more billion. If an eventual world population below 10 billion is considered desirable, the fertility decline in the developing world has to proceed at a rate faster than that experienced in most western European countries during the 19th and early 20th centuries. In several developing countries there is evidence of such a rapid fertility decline: Taiwan, South Korea, Hong Kong, Mauritius, Costa Rica, Sri Lanka and others, possibly including the People's Republic of China.

Should the average fertility decline of the developing countries be at the same rate as it was in western Europe, a world population on the order of 15 billion can be expected. It logically follows that a lower rate of fertility decline or a stable current fertility would lead to a world population of more than 15 billion.

To label the 2 percent world population growth rate of the 1960's as "worse than cancerous growth of the human population" is a subjective value statement with which some will agree and others disagree. The facts, however, are that owing to the precipitous mortality decline of the past two or three decades in the majority of the developing countries, coupled with relatively high fertility, the age structure of the developing countries' population has been shaped in such a way that even if fertility (the average number of children born per woman) does decline as fast as it did in 19th-century Europe, the world population growth rate will remain around 2 percent per year for the rest of this century.

Every country has to decide how important a factor its population growth rate is in maximizing its welfare objectives. If rapid population growth is conceived as a significant obstacle to development, significant enough to create policies and programs to cause a decline of the growth rate or to accelerate the decline of a decreasing growth rate, and if the majority of the developing countries will act effectively to correct their growth rates, then the world population growth rate might fall below 2 percent before the end of the century.

The alternatives are obvious but choices are difficult to make and even harder to put into practice. Research indicates that one should expect continued high growth rates for a few decades resulting in large population numbers. How large the eventual world population will be depends heavily on the rate of fertility decline: 15 billion with a traditional rate of fertility decline, less with a faster decline or more with a slower fertility decline.

TOMAS FREJKA

The Population Council
New York

Scientific American, July, 1973; Vol. 229, No. 1. 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 Harry T. Morris, 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 25 cents.

Change of address (or other subscription correspondence) should be addressed to Circulation Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Please notify us four weeks in advance of change. If possible, kindly furnish an address imprint from a recent issue. Send both old and new addresses or complete and mail Post Office Form 3578.

☐ Please enter change of address.

☐ Please enter subscription.
U.S. and Canada:
☐ 1 year, $10 ☐ 2 years, $18 ☐ 3 years, $25
All other countries:
☐ 1 year, $12.50 ☐ 2 years, $23 ☐ 3 years, $32

NAME

NEW ADDRESS

OLD ADDRESS

This is one of the best things made in America.

The four-day weekend. With your family. Relaxing.

Or the three-day weekend. Or just Saturdays and Sundays off. And two weeks' (or more) vacation with pay. And sick leave and retirement benefits.

Are corporate profits excessive? Not according to the facts.

They all come from the same source. The free enterprise profit system.

Lately there have been some who argue that corporate profits are bloated, unjustified—even immoral.

Not so. After-tax corporate profits average four cents on every gross sales dollar, not 28 cents as the general public is reported to believe today.

Profits make possible leisure time, art, sports, welfare payments, low-cost government housing, foreign aid, the fight against cancer.

Profits make possible leisure time . . . and cancer research.

Profits allow business to accumulate surplus monies in good times so it can get through bad times, with a minimum of layoffs and other economy moves.

Profits let people start new businesses and, in the process, create new jobs.

80 million Americans owe their independence to the profit system.

Profits give working Americans more options to do something other than work. If business couldn't earn profits, we'd all have to work much longer and harder just to meet the everyday demands of living.

Not to mention doing without those long weekends, sick leave and retirement benefits.

It's as simple as that.

International Telephone and Telegraph Corporation, 320 Park Avenue, New York, N.Y. 10022.

ITT

SERVING PEOPLE AND NATIONS EVERYWHERE

Freeman Books for Psychologists

The Social Animal

ELLIOT ARONSON,
The University of Texas at Austin

"Aronson's book is sheer intellectual delight."
—Benbow F. Ritchie, University of California, Berkeley

"A superb book, clearly written, theory and research literature well selected and lucidly explained. Perhaps the best statement of T-groups within the context of social psychology theory and research available. Saul Steinberg's drawings are excellent. Aronson writes with passion, skill, and knowledge of the field. A rare book in that it bridges humanistic and behavioristic approaches."
—Clovis Shepherd, University of Cincinnati

1972, 324 pages, 10 illustrations, paper, $3.50

Orthomolecular Psychiatry: Treatment of Schizophrenia

Edited by DAVID HAWKINS,
The North Nassau Mental Health Center
and LINUS PAULING, Stanford University

The 30 articles in this volume explain a new method of treating mental illness, the guiding principle of which is that the mind's functioning is strongly influenced by the nutritional and medical state of the body. Essential reading for all who are concerned with the field of mental health.

1973, 650 pages, 81 illustrations, $17.00
Prepublication price: $12.00, until February 1973

Neuropharmacology and Behavior

VINCENZO G. LONGO
Istituto Superiore di Sanità, Rome

A concise but wide-ranging discussion of the chemical and pharmacological properties, as well as the effects, of the most important psychotropic drugs.

1972, 184 pages, 33 illustrations, cloth $6.95, paper $3.95

The Psychology of Consciousness

ROBERT E. ORNSTEIN,
The Langley-Porter Neuropsychiatric Institute,
San Francisco

This book is an attempt to reconcile two basic approaches to knowledge: the rational and the intuitive. Central to its theme is the concept of a bifunctional brain in man, with the left hemisphere normally responsible for "linear" thought and the right hemisphere for "nonlinear" or nonverbal thought.

1972, 230 pages, 13 illustrations, paper, $3.50

Introduction to Modern Behaviorism

HOWARD RACHLIN,
State University of New York, Stony Brook

"A lucid, accurate, well-organized, and intelligent discussion of the field of animal learning which my introductory psychology students can understand and appreciate."
—Martin E. P. Seligman, University of Pennsylvania, in *Contemporary Psychology*, September 1971

1970, 208 pages, 76 illustrations, cloth $6.00, paper $3.25

Selected Writings of T. C. Schneirla

Edited by LESTER R. ARONSON and ETHEL TOBACH, The American Museum of Natural History, and JAY S. ROSENBLATT and DANIEL S. LEHRMAN, Institute of Animal Behavior, Rutgers University

A collection of 34 of Schneirla's most important writings, this volume provides further insight into his observations, theories, methods, and experiments in comparative psychology and animal behavior.

1972, 1032 pages, 137 illustrations, $22.50

Development and Evolution of Behavior

Essays in Memory of T. C. Schneirla

Edited by LESTER R. ARONSON, ETHEL TOBACH, DANIEL S. LEHRMAN and JAY S. ROSENBLATT

"An important book. By maintaining the intellectual standards that were set by the scientist whose life and work they wished to commemorate, the contributors have served him well." —*Science,* January 22, 1972

1970, 656 pages, 30 illustrations, $12.95

Army Ants, A Study in Social Organization

T. C. SCHNEIRLA, Late Curator, The American Museum of Natural History. Edited by Howard R. Topoff.

This synthesis of his conclusions about the behavioral patterns of the army ants is the last work written by T. C. Schneirla. More than the work of a great scientist concerning one group of ants, it is an instrument for understanding the momentous changes that have taken place in the behavioral sciences during the past four decades.

1971, 349 pages, 89 illustrations including 8 plates in full color, $12.00

The Origins of Intellect: Piaget's Theory

JOHN L. PHILLIPS, JR., Boise State College

Jean Piaget, described by *The New York Times* as "the outstanding child psychologist in the world today," has constructed the only systematic and comprehensive theory of intellectual development. This brief book is a nontechnical, general summary of that theory.

"This is the best, most understandable short description of Piaget's work I've seen. It has the particular advantages of not oversimplifying as well as setting the theoretical work in a broader context."
—Muriel King, University of Washington

1969, 149 pages, 30 illustrations, paper, $2.95

The Antecedents of Self-Esteem

STANLEY COOPERSMITH, University of California, Davis

"What background conditions, individual characteristics, and aspects of parental treatment shape a child's level of self-evaluation? This orienting concern provides focus and structure for Professor Stanley Coopersmith's insightful and solid contribution to the growing literature linking important theoretical concepts and substantial empirical data concerning aspects of self-conception. . . . the conclusions are important in their implications for theory, child rearing, education, and therapy." —*Psychiatry,* May 1970

1967, 283 pages, 90 tables, $7.50

Readings from SCIENTIFIC AMERICAN

Collections of articles originally published in *Scientific American,* reproduced with illustrations. Selected and introduced by prominent scholars.

Perception: Mechanisms and Models

With Introductions by
RICHARD HELD and WHITMAN RICHARDS, Massachusetts Institute of Technology

390 pages, 389 illustrations, cloth $12.00, paper $5.95

Contemporary Psychology

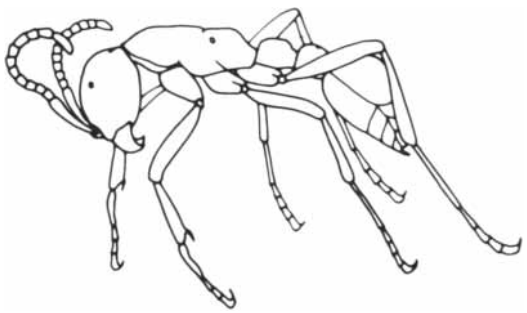
With Introductions by
RICHARD C. ATKINSON, Stanford University

486 pages, 441 illustrations, cloth $12.00, paper $5.95

Physiological Psychology

With Introductions by
RICHARD F. THOMPSON, University of California, Irvine

444 pages, 422 illustrations, cloth $12.00, paper $5.95



From your bookstore or
W. H. FREEMAN AND COMPANY
660 Market Street, San Francisco, California 94104
58 Kings Road, Reading, England RG1 3AA

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

JULY, 1923: "At the meeting of the National Academy of Sciences a few days ago Dr. W. W. Campbell made public full details of the precise and conclusive confirmation of Einstein's predictions by the observations of the Lick Observatory party at the Australian eclipse of last September. Four excellent photographic plates were secured during the eclipse, and as many more for comparison. Then, after the astronomers returned from their journey, began the laborious and tedious work of measurement and calculation. Three of the four pairs of plates have so far been worked up, and each plate shows an Einstein effect, clearly and without question. The general mean of all the results reduces to a deflection of starlight of 1.74 minutes at the sun's limb, as against 1.75 minutes predicted by Einstein."

"If we are to believe the very logical arguments of the World Metric Standardization Council, there exists no good reason at all why the meter-liter-gram system has not been adopted by the people of the United States and Great Britain, except that deep-seated quality of human nature which causes us all to put our backs up and resist changes until they are forced on us. The greatest physical obstacle to overcome is the transition from the system of machine tools. Many American manufacturers, however, already are using the metric system of measurements today for the production of export articles. What remains to be done is not so much to convince the average man of the desirability of the change on theoretical grounds as to demonstrate to him that he should contribute his share to making the change."

"The application of electric traction to the railroads of the United States is proceeding quite closely along the lines that were predicted 15 or 20 years ago. The public was quick to realize the grand scale on which electrification of the steam railroads of the country was being commenced, and predictions were freely made that within a decade or so steam would give way entirely to electricity and the steam locomotive would take its place in historical museums. The history for the past 15 years has proved the truth of these predictions. The latest development is the decision of the Virginia Railroad to electrify 134 miles of their system lying between Roanoke, Va., and Mullens, W.Va. The Westinghouse Company states that this is the largest single railroad electrification contract that has ever been placed."

"Sir James Dewar, whose death was recently announced, is popularly known to the world as the inventor of the Thermos bottle. However, he was not consciously working for what is thus known but rather for something to preserve liquid gases, with which he was experimenting. The use that his 'dewar tube' is now mostly put to came as an afterthought, not so much of his own but of the commercial world. With his invention he was able to liquefy hydrogen, and he froze it at minus 438 degrees Fahrenheit. He also isolated hydrogen, helium and neon from the air. He died at the age of 81 years."

"The recent stir in the daily press about the possibility of gasoline selling for $1 per gallon has directed public attention to the necessity of saving fuel in every way possible. Exact instructions for the most economic driving will vary with the make of the truck or car, but the following pointer will be of value to all: In ascending a hill do not wait until the last second to shift to a lower gear. If you do, not only will you lose speed and overtax your engine but also you will consume gasoline. With more knowledge of economic driving a considerable reduction can be made in the four billion gallons of gasoline consumed each year."

SCIENTIFIC AMERICAN

JULY, 1873: "At Canandaigua, N.Y., Miss Susan B. Anthony, who insisted that she had as good a right to vote as any man, and who did vote at the last election, has been tried and, we regret to say, found guilty and fined for violating the law. Judge Hunt decided that although women were entitled to the general rights of citizens, there were certain special privileges that the law of New York did not give them, one of which was the privilege of voting. The law must, in the opinion of Judge Hunt, be changed before our feminine fellow citizens can enjoy themselves at the ballot box."

"It is hardly necessary at the present day to direct especial attention to the relative proportion of advantages and defects inherent in the rotary engine as a type, nor is it here needful to draw any contrast between it and the reciprocating engine. The principal advantage of the rotary engine is a positive equalization of steam pressure on all sides of the drum, at all times and at every point in its revolution, so that frictional wear is reduced to its lowest point for the main shaft and journals. The rotary engine is cheap, as all the parts are simple lathe work, strictly circular, or else have flat and straight edges easily planed. It is compact and economical of fuel. It is simple and the working parts are readily accessible for repair. There are practical difficulties, however, particularly in the wearing of certain parts. Whereas in the future inventors will unquestionably seek to augment its advantages, at present their efforts are more particularly directed toward overcoming these difficulties."

"Robert Wilhelm Bunsen was born on March 31, 1811, in Göttingen, a town in Hanover. At the age of 17 he entered the university of his native town in order to pursue physical and chemical studies; he took the degree of doctor in 1833. In 1852 he was nominated professor of chemistry in the University of Heidelberg, which position he still holds. We owe to him important contributions relative to the combustion and diffusion of gases, etc. He is also the inventor of that wonderful instrument known as Bunsen's burner. Herr Bunsen, although now in his 62nd year, enjoys excellent health and is still increasing in the pursuit of his investigations."

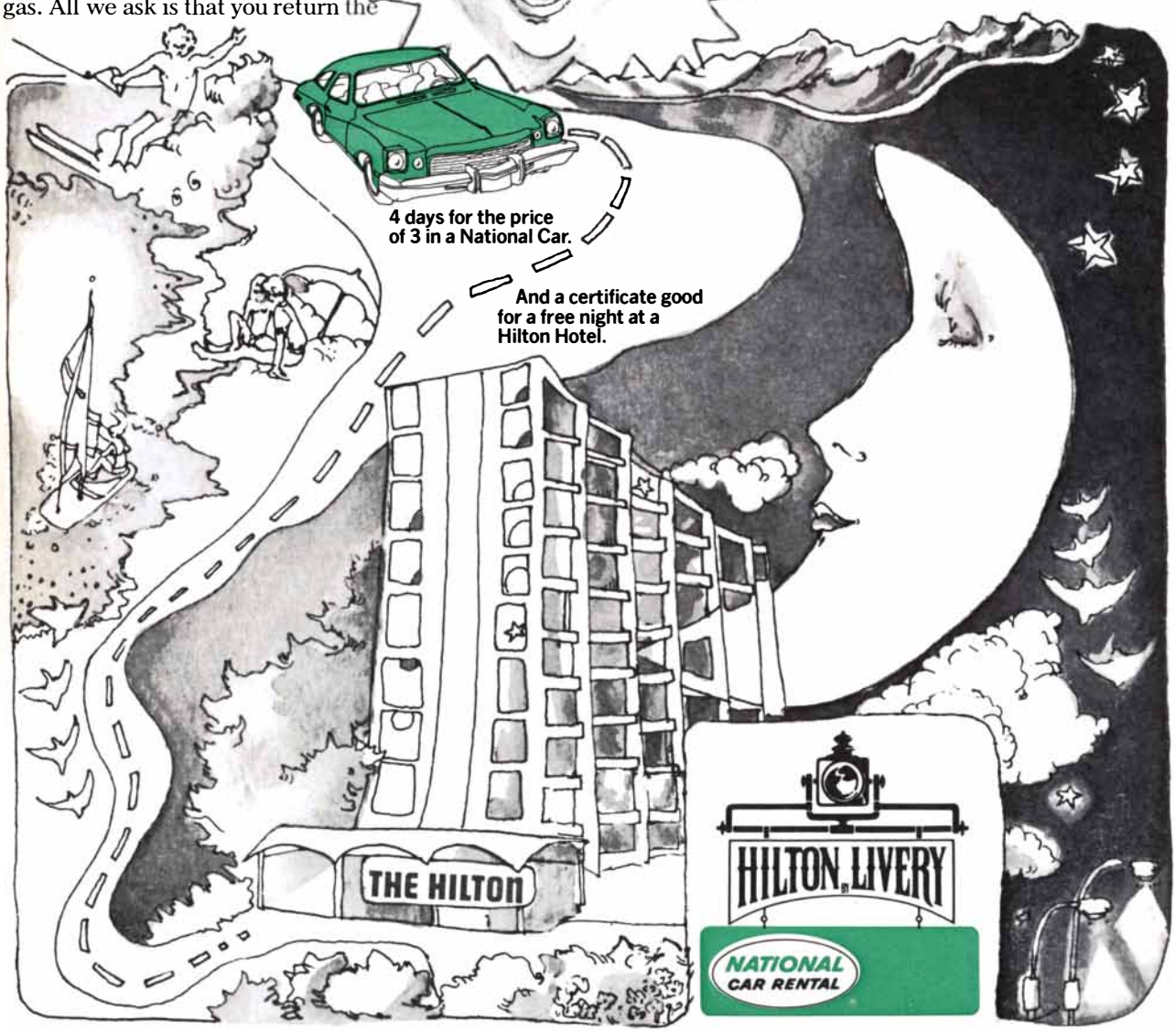
"Captain Douglas Galton, in an interesting article in the *Journal of the Society of Arts,* calls attention to the need of new inventions in devices for cooking, whereby the great waste of fuel now experienced in the best of our stoves and ranges may be prevented. The question of saving fuel for cooking purposes, he says, is even more important than economy in warming, because cooking is an operation required every day in the year, and the waste of fuel in cooking is even more considerable than in warming. Although hundreds of stove and range inventions have been made, there is still room for a new and better series of devices whose merits are to consist in their economy of fuel."

The mini-vacation that won't cost you a mini-fortune.

National Car Rental and Hilton Hotels can help you turn a weekend into a mini-vacation. One that won't cost you a small fortune.

At participating National locations† (including our Hilton Livery offices in Hilton Hotels) you'll find our 4-day weekend. We'll rent you a new Buick Century or similar-sized fine car from Thursday at 6 PM to Monday, 6 PM. You pay just $21 plus 12¢ a mile. We pay for the gas. All we ask is that you return the car to the city where you rented it.

But that's just the beginning. With every National Car 4-day weekend you get a certificate good for one free night at a participating Hilton Hotel.* (You pay for Thursday, Friday and Saturday, we'll give you Sunday night free.△) And, of course, at Hilton you'll be sure to find everything you'd expect at a great hotel. Make your next weekend trip a National-Hilton mini-vacation. You'll get a lot more vacation for your money.



To reserve your car at National call 800-328-4567 toll free. In Minnesota call 612-830-2345 collect. To reserve your Hilton Hotel room just call your nearest Hilton reservation service office. Ask for the Hilton Livery mini-vacation. Or call your travel agent to arrange for both car and room. *Certificate must be presented when you check-in. △If you wish to stay for 3 nights you pay for Thursday and Friday. We'll give you Saturday night free. © National Car Rental System, Inc., 1973. (In Canada it's Tilden Rent-A-Car.)

†Available only at participating U.S. locations, which does not include Metro New York and some other cities.

THE AUTHORS

FREDERICK S. JAFFE ("Public Policy on Fertility Control") is vice-president of Planned Parenthood–World Population and director of the Center for Family Planning Program Development. He was graduated from Queens College in New York with a degree in economics and worked in journalism and public relations for several years before he joined Planned Parenthood in 1954. In addition to writing extensively on family planning he has served as technical adviser to the President's Committee on Population and Family Planning and as special consultant to the U.S. Commission on Population Growth and the American Future. "As for other interests," he writes, "I am a tennis player and a reader, with major interests in labor, economic and social history."

ARTHUR H. BENADE ("The Physics of Brasses") is professor of physics at Case Western Reserve University. His work was originally in nuclear physics, but (he writes) "about 10 years ago my growing interest in musical acoustics, which began with reading Dayton C. Miller's *Science of Musical Sound* and blowing on everything in sight in the band room when I was in high school, swallowed up my other interests." The understanding of musical instruments he and his colleagues have gained has contributed to the design of commercial instruments. "I have made it a matter of policy to become reasonably fluent as a player on the various historical or national types of flute, clarinet, oboe and so on. In the case of instruments on which I lack fluency I can usually make sounds of good quality note by note. All of this makes for an interesting and varied life as an amateur player of chamber music and provides an indispensable insight into the relative merits of the instruments, which has proved to be an essential guide and reference point for my scientific activities and a basis for grateful appreciation of what my musical friends have to teach me."

H. R. CLAUSER ("Advanced Composite Materials") is a consultant, writer and teacher; he took up those activities in 1970 after retiring from *Materials Engineering,* where over a period of almost 30 years he was successively assistant editor, managing editor, editor and editor-publisher. His present work includes the editorship of *Research Management.* He recently completed a term on the National Materials Advisory Board and an assignment with the National Commission on Materials Policy. Clauser is the author of the book *Radiography for Industry* and is finishing a textbook on materials. In addition he is editor-in-chief of the *Encyclopedia of Materials and Processes.*

NIELS KAJ JERNE ("The Immune System") is director of the Basel Institute for Immunology in Switzerland. "All my ancestors for several generations lived in a small area of western Jutland," he writes. "My grandfather, after whom I was named, was captain of a small sailing ship that made a yearly voyage to Mexico until he went down with his ship when my father was five years old. My father did not go to sea but moved around a lot, with the result that I was born in London and received my education in Holland, where I obtained my baccalaureate when I was 16. It then took me 12 years to make up my mind what profession to choose. When I was 28, I decided to study medicine in Copenhagen, with the objective of becoming a village doctor. During my studies I took on a part-time job in a scientific laboratory, and from then on I was trapped in science." Among other things Jerne has worked for the World Health Organization, served as chairman of the department of microbiology at the University of Pittsburgh School of Medicine and spent three years as director of the Paul Ehrlich Institute in Frankfurt. "I have no interest in law, politics, walking or collecting anything," he writes, "nor in motorcars, clocks or any product of technology. Practically anything else I find interesting, including literature, music and drinking good wines in France."

I. R. CAMERON ("Meteorites and Cosmic Radiation") is associate professor of physics at the University of New Brunswick. Born in Scotland, he received his bachelor's degree from the University of Edinburgh in 1953 and his doctorate (in nuclear physics) from the same university in 1958. From 1958 to 1967 he was with the United Kingdom Atomic Energy Authority, working on the physics of various types of experimental nuclear reactor system. "After nine years of working in reactor physics," he writes, "I felt that a change to some completely different field would be stimulating The study of meteorites is particularly interesting, since it involves elements from so many diverse fields, from metallurgy to cosmology. My outside interests include classical music, history and ornithology."

WALTER GOLDSCHMIDT ("The Brideprice of the Sebei") is professor of anthropology and psychiatry at the University of California at Los Angeles and senior science fellow of the National Institute of Mental Health. His bachelor's and master's degrees are from the University of Texas (in 1933 and 1935 respectively), and his Ph.D., which he obtained in 1942, is from the University of California at Berkeley. He has served as editor of the *American Anthropologist* and is founding editor of *Ethos,* a journal dealing with the interrelation of the individual, culture and society. Goldschmidt is the author of five books.

PETER A. RONA ("Plate Tectonics and Mineral Resources") is senior research geophysicist with the National Oceanic and Atmospheric Administration; his base is the Atlantic Oceanographic and Meteorological Laboratories in Miami. He is also chief scientist of the Trans-Atlantic Geotraverse, an international cooperative venture investigating the earth's crust along a corridor between the southeastern region of North America and the northwestern region of Africa. The aim of the project is to gain understanding of continental drift, seafloor spreading and the processes of accumulation of minerals on the sea floor. Rona received his bachelor's degree in geology from Brown University in 1956 and his Ph.D. (in marine geology and geophysics) from Yale University in 1967; as he puts it, "my formal education spanned the revolution in the earth sciences." Before he joined the National Oceanic and Atmospheric Administration in 1969 he was with the Hudson Laboratories of Columbia University for nine years and worked as an oil geologist for two years.

EDWARD V. EVARTS ("Brain Mechanisms in Movement") is a member of the Laboratory of Neurophysiology of the National Institute of Mental Health. He obtained his M.D. from the Harvard Medical School in 1948 and received clinical training in psychiatry and neurology. It was while he was combining his clinical work with a year of research at the Yerkes Laboratories of Primate Biology in Florida that he became interested in the function of the brain. "My nonprofessional activities are swimming, running and learning hard languages," he writes. "My current language of interest is my wife's native tongue: Finnish."

Anatomy of a Nikon Camera (Part II)

Recently, we discussed the precision with which a Nikon is put together to make it the superb picture-taking instrument it is. But there is still another aspect of the way Nikon is put together that makes it the most versatile camera ever conceived.

We're now talking about the literally thousands of ways you can put a Nikon together yourself by way of its totally flexible, modular construction, on which the Nikon system is based.

Interchangeable Lenses

There are more than forty lenses in the Nikon system. Fourteen of them are wide angle lenses, for use at close quarters or to exaggerate perspective, including a 220° Fisheye Nikkor that actually "sees" objects behind the camera. To bring distant objects up close, you can choose any of fifteen telephotos, ranging up to 2000mm and giving up to 40X magnification compared with the "normal" 50mm lens. There are high-speed lenses for use in the dimmest available light and special lenses for extreme closeup work—even one that corrects linear distortion in architectural photos. You'll find all of these and more in the Nikon system—all sharing the exceptional quality, the surpassing sharpness and the magnificent color rendition that make Nikkor lenses the overwhelming choice of professionals.

Interchangeable Finders

Because a single lens reflex camera is built around the finder, Nikon pays particular attention to this aspect. It starts with the only 100% accurate viewing system in all of 35mm. That's right; the Nikon F and F2 are the only cameras whose finders show precisely the picture area appearing on the film—no more and no less.

To implement this unique accuracy, there are five interchangeable finders for the Nikon F and six for the Nikon F2. Brilliant eyelevel prism finders. Photomic thru-the-lens meter/finder

systems, waist-level and magnifying finders—even one for viewing accurately while wearing goggles or with the camera in an underwater housing.

Interchangeable Finder Screens

The viewing screen is where the finder image is formed by the camera lens and the finder's optical system. In most cameras it's fixed; a few others accept a handful of interchangeable screens. There are no fewer than eighteen in the Nikon system, each offering practical advantages for some specific application. Even if you never use more than one, you can choose among seven universal types to suit your individual preference.

Interchangeable Backs

This feature is mainly for the professionals, who use more Nikons than any other 35mm camera. If you happen to be one, you may want to replace the Nikon back with one holding 250 (or even 800) exposure film loads for use with a motor drive. Or, you might attach a Speed Magny back and shoot large-format Polaroid pictures (that's right, with a Nikon!).

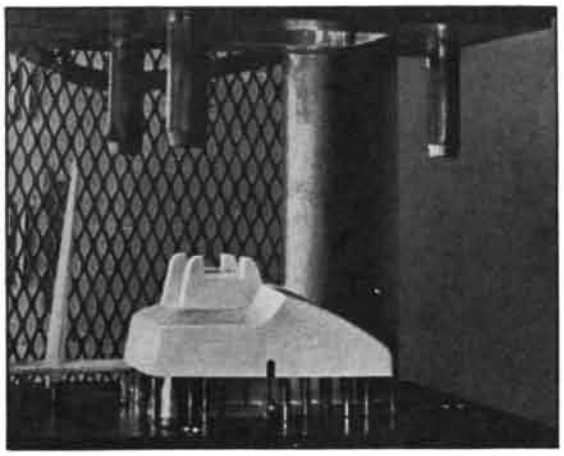
Using just the parts discussed so far, you can come up with a staggering 21,600 different combinations for your Nikon camera. And that doesn't include closeup equipment, filters and dozens of other eminently practical accessories presently in the Nikon system, and new ones constantly being developed.

Of course, no one photographer will ever need all of it. The point is, no matter how exotic or far-fetched the situation may be, you're sure to find the equipment you need in the Nikon system.

That's versatility, Nikon style. To find out more about it and how it can work for you, visit your Nikon dealer. Ask him also about the Nikon School of Photography. Or write for detailed literature folio 19. Nikon Inc., Dept. SA, Garden City, N.Y. 11530. Subsidiary of Ehrenreich Photo-Optical Industries, Inc.

(Canada: Anglophoto Ltd., Montreal 376, P.Q.)

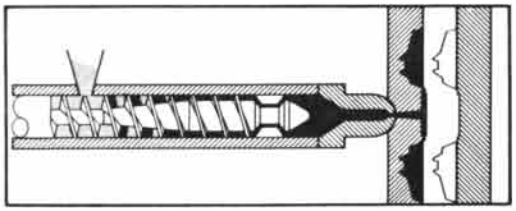
WESTERN ELECTRIC REPORTS



Molding by the millions. Western Electric people produce some 8 million phones a year. Molded plastic is used for housings and many other parts. So there is a constant investigation into the most effective way to use these materials.

$$A^*(z,t) = A_e^*(z) - [(A_f^* - A_i^*)/(1 - e^{-\beta N t_r})]e^{-\beta N t}$$

In developing the model at Western Electric's Engineering Research Center, it was found that melting behavior can be described by this formula which includes terms for shear heating and conduction heating effects. Other models were developed for temperature and pressure profiles.



End of molding cycle. At this point, the screw is stationary and heat is conducted into the plastic on the screw. After the plastic solidifies, the mold is opened as shown. The parts can then be ejected.

Solving the mysteries of molding with mathematics.

Even though plastics have been around for many years, there's still a lot to be learned about these versatile materials and their processing. So they are the subject for continuing studies by our engineers.

Some of their recent investigations have brought forth new and highly useful information about a relatively unexplored area: the melting behavior of plastics in the injection molding process.

One result of these studies is the mathematical formula, or model, above.

The model helps us predict melting behavior along the length of the injection screw molding machine used to mold telephone housings and other parts. Melting behavior is extremely important, because plastic pellets should be completely melted but not thermally decomposed before injection into the mold.

This information on melting is then used to investigate screw designs, operating conditions, machine sizes and plastic properties. All of which is aimed at obtaining optimum processing techniques.

Predictions obtained from the mathematical model have checked out closely with experimental observations. So the resulting screw designs are now undergoing evaluation by engineers at our plants in Indianapolis and Shreveport.

Conclusion: For new designs and materials, the model can help reduce the development cost for new molded parts and materials. For manufacturing current products, operating costs can be reduced.

Perhaps most significant is that we're getting information about molding temperatures not available experimentally. And many other types of information can be obtained without the use of costly, time-consuming experimental work.

The end result will be more efficient plastic molding and therefore a better product for the lowest possible cost.

Western Electric

We make things that bring people closer.

Established 1845 SCIENTIFIC AMERICAN July 1973 Volume 229 Number 1

Public Policy on Fertility Control

In the U.S. over the past dozen years policy on contraception has been almost entirely reversed. The transformation reflects the concern of the community with population size and poverty

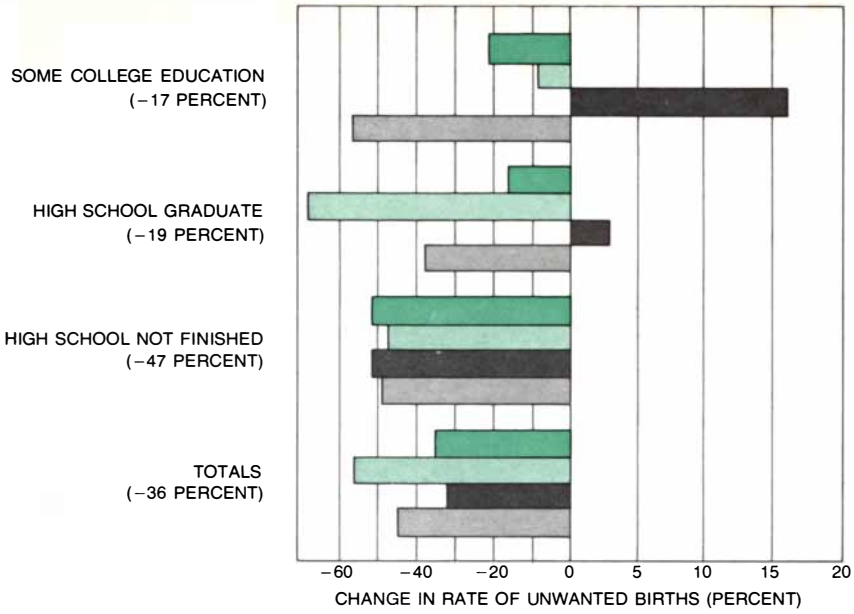
by Frederick S. Jaffe

As recently as 1971 Federal law classified contraceptives among "obscene or pornographic materials," and until a few years ago the laws of many states restricted not only their sale but also the dissemination of information about them. In other words, public policy made it difficult for health professionals to dispense and for the public to obtain advice and materials for achieving effective contraception. Now most of the laws have been repealed and the Federal Government has (in the Department of Health, Education, and Welfare) an Office of Population Affairs, a National Center for Family Planning Services and a Center for Population Research. Clearly public policy on family planning has undergone a substantial change. How did the change come about, and what effects has it had? I shall discuss these questions from the viewpoint of a participant and an observer, since my principal responsibility with the Planned Parenthood Federation of America for more than 15 years has been in the area of public policy.

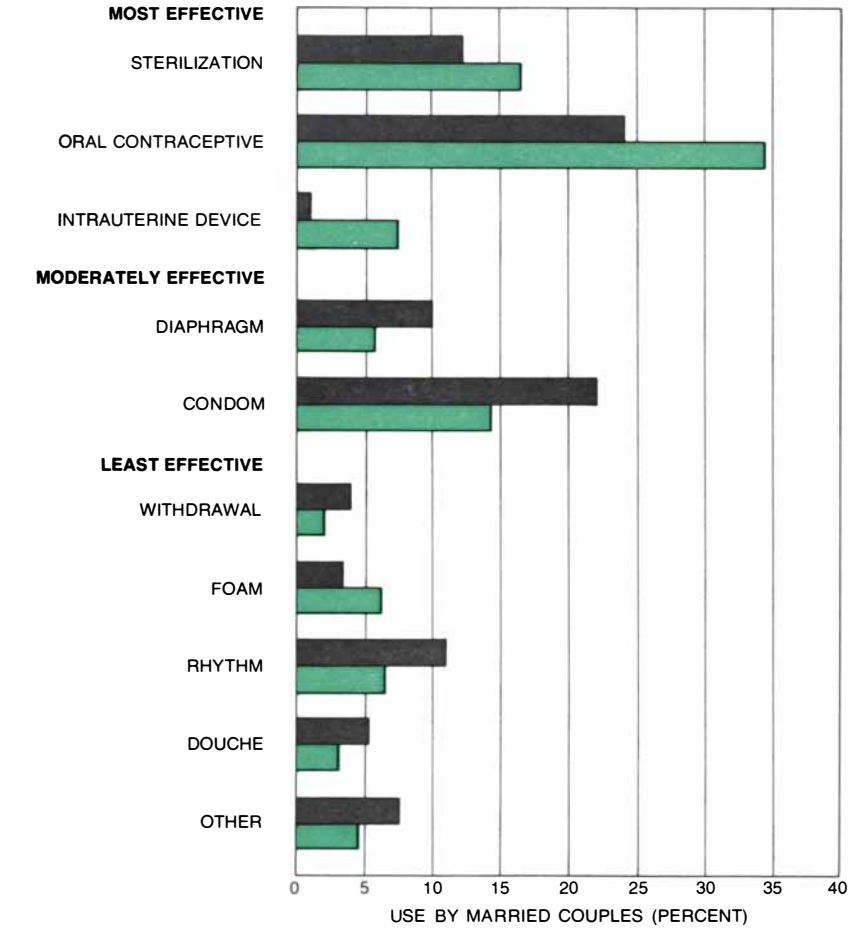
The change of policy did not transform the U.S. from a nation in which contraception was rarely practiced to one in which it is widespread; it has been widespread for a long time. What the change did do was to make it possible for increasing numbers of people to achieve their goals relating to size of family. As early as 1941, when the first large-scale study of attitudes and practices in family planning was made (in Indianapolis by the Scripps Foundation for Research in Population Problems and the Milbank Memorial Fund), most white Protestant couples were found to practice some form of contraception at some point in their lives to limit and space births. In 1955, when a group from the Scripps Foundation and the University of Michigan made a national survey, whites of all religious persuasions reported the same pattern. Both surveys revealed significant socioeconomic differences, however, in the timing of the initiation of contraception, in employing the most effective methods and in success in controlling births. Many couples relied on the least effective schemes for contraception, which were folk methods or nonprescription methods. When nonwhite couples were sampled for the first time (in 1960, by the Scripps group), the findings were similar. In short, by 1960 most Americans had a favorable attitude toward contraception, but the effectiveness of the methods they used varied widely.

In 1936 the sociologist Norman E. Himes, in his book *Medical History of Contraception,* characterized the methods employed by large segments of the married population in the U.S. and western Europe as "amateur" and "back fence" contraception. It was, he said, "learned from friends, neighbors, childhood or adult acquaintances." According to Himes, the problem "is not shall we have conception control or no conception control, but rather shall we diffuse the latest knowledge on scientific contraception or encourage by indifference and laissez-faire the use of quasi-unreliable, quasi-harmful methods with which many, including not a few of the otherwise enlightened, are in the habit of worrying along."

The problem as Himes stated it is precisely the issue that has been at stake in the evolution of public policy on family planning. The issue can be restated in ways that reveal more of its underlying social implications. Since there are significant differences in failure rates between the back-fence methods and the most effective techniques, whether or not the best methods are intentionally diffused as a matter of public policy depends on how important the society believes it is to prevent unwanted pregnancies and their social, economic, health and demographic consequences. Since there have been and still are socioeconomic differences in access to the best methods and therefore in ability to avoid contraceptive failures, the issue is also related to the society's broad policies on poverty and the prevention of dependency. Finally, since the most effective methods are the ones whose distribution is controlled by the medical profession and by health institutions, the issue becomes a test of the ability of the health system to diffuse scientific knowledge and practice—a process necessary to improve not only the control of fertility but also numerous other health practices.



UNWANTED BIRTHS, as measured by a study done by the Office of Population Research of Princeton University, have declined among almost all groups since 1960. The percentage figures compare the rate of unwanted births in the period 1961 through 1965 with the rate from 1966 through 1970. Negative percentages indicate declines in the rate of unwanted births. Educational descriptions relate to wives. Bars represent whites (*dark color*), blacks (*light color*), whites who are not Catholic (*dark gray*) and white Catholics (*light gray*).



METHODS OF CONTRACEPTION employed by married couples are compared for 1965 (*gray*) and 1970 (*color*). These Princeton findings show a shift to more effective methods.

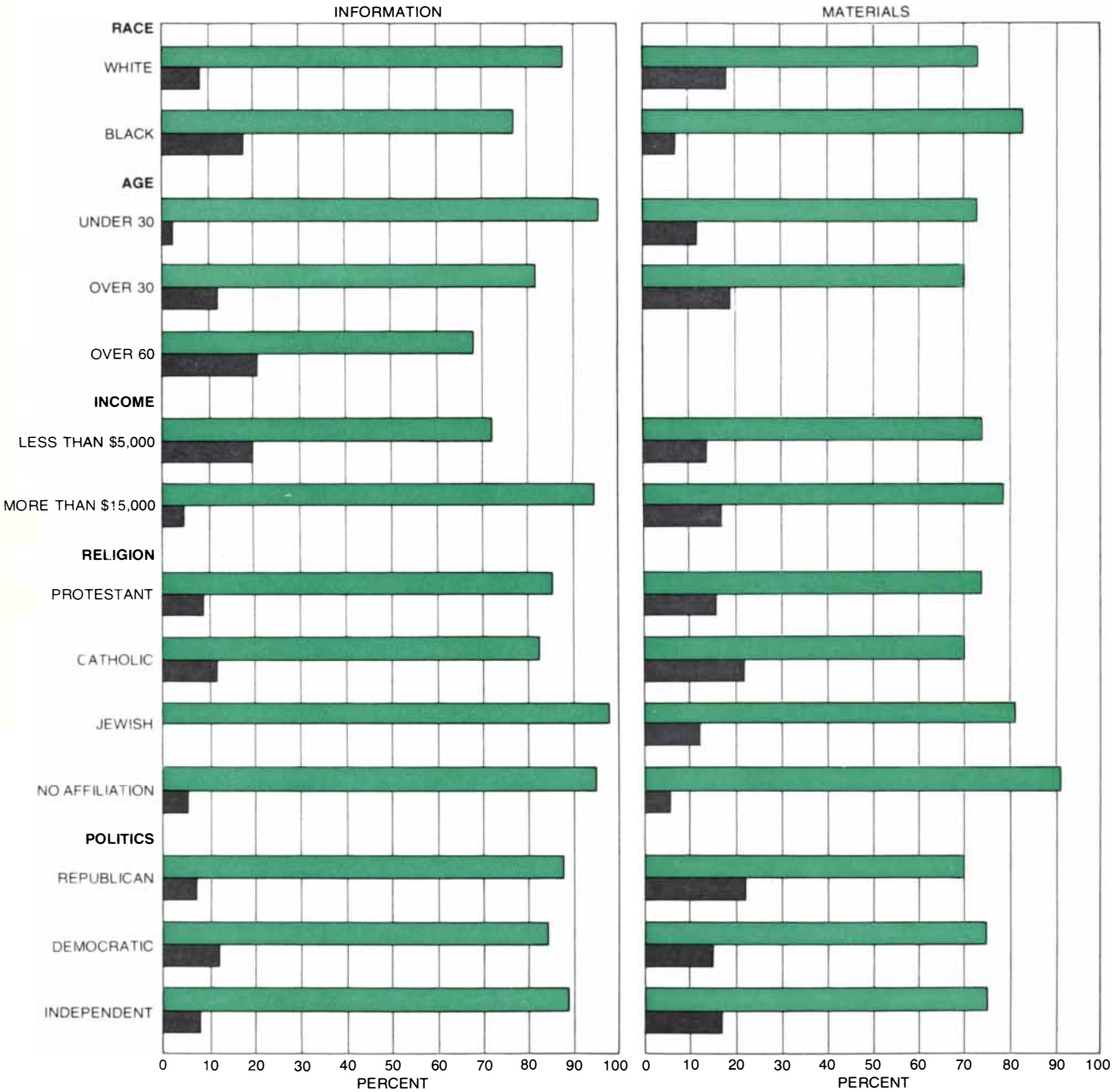
The Federal and state laws that made it difficult to dispense and obtain contraceptive advice and materials were often called Comstock laws, reflecting the far-reaching influence of the 19th-century movement personified by Anthony Comstock, who for many years beginning in 1873 was secretary of the New York Society for the Suppression of Vice. The laws, even though various court decisions helped to moderate their impact, had a chilling effect on public and professional attitudes and on the policies and practices of physicians and health institutions. For example, not until 1955 did a national magazine publish an article that for the first time named specific methods of contraception, and not until 1959 did a television network have a program that mentioned the subject. The constraints were most apparent in the administrative policies of Government hospitals and health agencies. Even for people who did not depend on publicly funded health agencies for medical care, however, public policy stigmatized contraceptive practice and created obstacles to obtaining medical contraception.

In the period from 1930 to 1950 Planned Parenthood and associated groups mounted unsuccessful efforts to repeal the restrictive laws. Paradoxically the shift in public policy began not with changes in law but with changes in administrative policies. A major turning point was reached in the late 1950's with a protracted and successful campaign by the same groups to reverse the long-standing but unwritten ban on contraceptive prescription in the municipal hospitals in New York City. (The prohibition was typical of most Government-operated health services at the time.) Although the campaign succeeded in its immediate objective of freeing hospital physicians to prescribe birth control, its major significance was in demonstrating how much public support there was for family planning.

In 1959, after a Presidential committee had recommended that the U.S. Government provide assistance on population programs to other nations requesting it, discussion of public policy on family planning became linked with discussion of the population problem. The ensuing debate was marked by the opposition of the Catholic bishops of the U.S. to the use of any public funds for birth-control programs at home or abroad and President Eisenhower's statement of opposition to a Government program. (He reversed his position four years later.) Officials of the Kennedy Administra-

tion expressed a much different view, describing in a series of speeches the Government's increasing concern over the impact of rapid population growth. The National Institutes of Health acknowledged a responsibility for supporting basic research in reproductive physiology that might lead to improved methods of fertility control. The work was assigned to the newly formed National Institute of Child Health and Human Development. (Public funds, however, played no part in the development of the oral contraceptive or the intrauterine device, which are the two methods that revolutionized contraceptive practice during the 1960's. Research on these methods was supported almost entirely by foundations, private philanthropists and industry.) By the end of 1963 attitudes in Washington had changed enough for Congress to adopt an amendment to the foreign-aid bill authorizing the use of assistance funds for "research into problems of population growth." In 1965 the National Academy of Sciences urged that family planning be made an integral part of domestic public medical programs and suggested that the appointment of an official "at a high national level" might facilitate Federal action. In the same year a Senate committee began a long series of hearings on the population problem.

The first major changes in law came at the state level in 1965 and 1966 when, with relatively little controversy, New York, Ohio, Massachusetts, Minne-



RESULTS OF POLL taken in 1971 by the Opinion Research Corporation for the U.S. Commission on Population Growth and the American Future are given. The chart shows the percentage of people in each category responding affirmatively (*color*) and negatively (*gray*) to a question about birth-control information and a question about birth-control materials. The information question was: "Do you think that information about birth control should or should not be made available by the Government to all men and women who want it?" The question on materials was whether or not the Government should make birth-control materials available.

sota and Missouri repealed certain Comstock-era restrictions on the dissemination of contraceptive information. Bills authorizing or encouraging public health departments or welfare agencies to provide family-planning services at public expense were adopted in California, Colorado, Florida, Georgia, Illinois, Iowa, Kansas, Michigan, Nevada, Oklahoma, Oregon and West Virginia. The U.S. Supreme Court, in *Griswold* v. *Connecticut,* struck down Connecticut's archaic statute—the only one in the country that prohibited the use of contraceptives—with a ruling that married couples have a constitutional right to practice contraception free of state interference. (Last year the court in effect extended the right to unmarried people.)

President Johnson referred in his State of the Union message of 1965 (and many times thereafter) to the importance of dealing with problems of population growth and providing family-planning services. The Department of Health, Education, and Welfare issued its first policy statement on family planning in 1966. The clearest statement of the objectives of U.S. domestic policy appeared the same year in the President's special message to Congress on health and education. The President cited family planning as one of four health problems requiring special attention. "We have a growing concern to foster the integrity of the family and the opportunity for each child," he said. "It is essential that all families have access to information and services that will allow freedom to choose the number and spacing of their children within the dictates of individual conscience."

The first Federal agency to move from words to action was the Office of Economic Opportunity, which as part of its antipoverty program began in 1965 to make grants to community-action agencies for family-planning projects. In the same year Federal maternity-care projects, which were designed primarily to reduce the incidence of mental retardation among children born to low-income women who were at high risk of bearing retarded infants, enabled some city health departments to start providing family-planning services on a limited basis.

Nonetheless, action tended to lag behind the change in expressed policy. The Federal role in biomedical research was implicitly acknowledged to be inadequate in 1965 when the advisory council of the National Institute of Child Health and Human Development admonished the agency to take the initiative in stimulating increased work on problems of human fertility, sterility and family planning. In 1967 an outside review of the Department of Health, Education, and Welfare found that none of the department's operating agencies "places high priority on family planning or is certain what precise functions it is expected to carry out in this field."

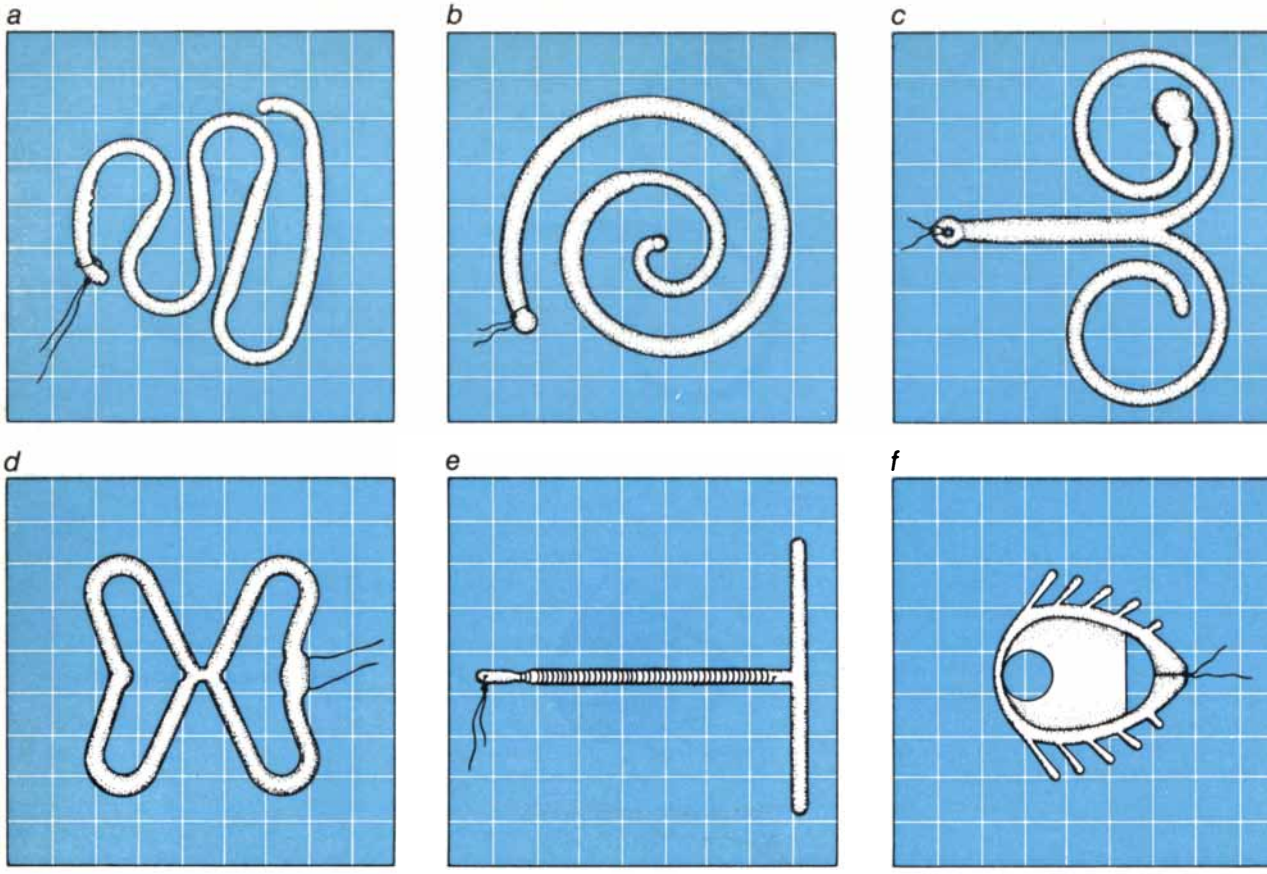
By 1967 Congress was beginning to prod the Federal agencies. In amendments to the Economic Opportunity Act family planning was designated for special emphasis in the antipoverty program. Congress also amended the Social Security Act to specify that at least 6 percent of the money appropriated for maternal- and child-health programs administered by the Department of Health, Education, and Welfare be devoted to family-planning service projects. The act also required the states to offer and provide family-planning services to all appropriate recipients of public assistance. Finally, Congress amended the Foreign Assistance Act to designate certain funds for family-planning and population programs overseas.

BIRTHRATE IN THE U.S. HAS DECLINED since 1900 except for the period following World War II. The development and rapid adoption of two of the three most effective methods of contraception, the contraceptive pill and the intrauterine device, have been factors in the recent decline. The third most effective method of contraception is sterilization.

During the same period a number of state laws prohibiting abortion except to save the mother's life came under challenge. Colorado revised its abortion law in 1967, and 17 other states did so subsequently. In January of this year the U.S. Supreme Court changed the entire abortion picture by ruling that states could not interfere with decisions made by a woman and her physician during the first three months of pregnancy.

In 1968 President Johnson appointed a Committee on Population and Family Planning to assess the adequacy of the Federal program. In its report the committee recommended rapid increases in funding and strengthening of the administrative machinery for the three main parts of the Federal program: family-planning services at home; biomedical and behavioral research to improve methods of contraception and the understanding of population dynamics, and assistance to family-planning programs in developing countries. A year later President Nixon sent to Congress a population message, calling for more Federal effort in all these directions and asking for the establishment of a population commission. Legislation creating the commission was enacted in 1970. That same year Congress passed the Family Planning Services and Population Research Act, which put into effect a modified version of the recommendations of President Johnson's population committee. The act authorized $382 million for a three-year program of services and research.

In 1971, complying with a provision of the act, the Department of Health, Education, and Welfare submitted to Congress a five-year plan; it called for a program that would provide services to an estimated 6.6 million women of low or marginal income by 1975. A year later the Commission on Population Growth and the American Future presented its final report, which appears likely to have an influence on thinking in the field for many years to come. The report recommended a broad array of health, edu-



INTRAUTERINE DEVICES, which are among the most effective methods of birth control, include the types depicted here. They are the Lippes loop (*a*), the Margulies spiral (*b*), the double coil (*c*), the Birnberg bow (*d*), the copper *T* (*e*) and the Dalkon shield (*f*).

cation, economic and social programs to enable the U.S. to achieve a stabilized population in an orderly manner as rapidly as possible.

In sum, the change in public policy was expressed in numerous legislative and administrative actions at Federal, state and local levels. It was also expressed in an enormous expansion of organized family-planning programs serving people of low or marginal income. In 1960 no more than 150 public and voluntary health agencies, most of them affiliates of Planned Parenthood, operated such programs; they served about 150,000 women. By 1972 nearly 3,000 hospitals, health departments and voluntary agencies were providing services to an estimated 2.6 million women in two-thirds of the counties of the U.S. A decade ago hardly any public money was available for such services; in fiscal 1973 the Federal appropriation alone for family-planning project grants was $128 million. Federal support for research rose from close to nothing in 1962 to $40 million in fiscal 1973. Projecting the experience of the past few years, it appears feasible to achieve the Federal Government's stated goals in this area by 1975, provided that the amount of Federal money available continues to rise. The future of the program is uncertain, however, as a result of the Administration's effort this year to cut down or eliminate support of many health and social programs, including family planning. It remains to be seen whether Congress will agree.

Whatever the outcome, the consequences of the changes that have already taken place can be measured by the increasing number of people of low or marginal income who have begun to use modern methods of fertility control, which is an aspect of medical care that was previously limited mainly to people of middle and high income. Periodic studies by the Office of Population Research at Princeton University show that between 1965 and 1970 the historic socioeconomic differences in patterns of contraceptive use were considerably reduced. Other studies indicate that fertility rates declined more rapidly in the latter part of the decade among poor and near-poor women than among women above the poverty level. Indirect effects of the policy change are less measurable but no less significant. By supporting the development of programs and facilitating public discussion of fertility control and population problems, Government actions have stimulated professional interest in family planning and the practice of effective birth control by people of all classes. As a result the change in public policy has begun to reverse the effects of the historic taboos that have surrounded contraception. The effects are suggested by the finding of the Princeton group that nearly three in five couples who practice contraception now employ the most effective methods (pills, intrauterine devices and sterilization) and that the incidence of unwanted pregnancy declined 36 percent between the first and the second half of the 1960's.

So much for the major changes. How were they brought about? Did the impetus come from within the health system or from outside it—from the political community, from citizens' groups and from the demands of individuals?

The question has no quick and easy answer. The list of people and groups that have worked to change public policy on family planning includes many physicians, health leaders and professional organizations. The health rationale for family planning, which is that

having several children in rapid succession results in abnormally high rates of infant and maternal death and illness, was essentially formulated a generation ago and has been important in the policy debate. Unfortunately it has remained only a rationale. Most physicians did not perceive family planning as a medical activity and still do not, although the best methods of contraception now known and the new methods under investigation require close medical supervision.

In my view the health institutions and the health professionals have not collectively distinguished themselves in pressing for a change in policy or in responding to the opportunities created by the change. I believe the primary impetus for change has come from outside the health system in response to two principal concerns, which are directed respectively toward the rapid rate of growth of population and toward the problems of poverty. An essential precondition for the change was the technology represented by the pill and the intrauterine device, which made large-scale programs feasible for the first time. Key roles in the process of change were played by political leaders and by voluntary organizations such as Planned Parenthood.

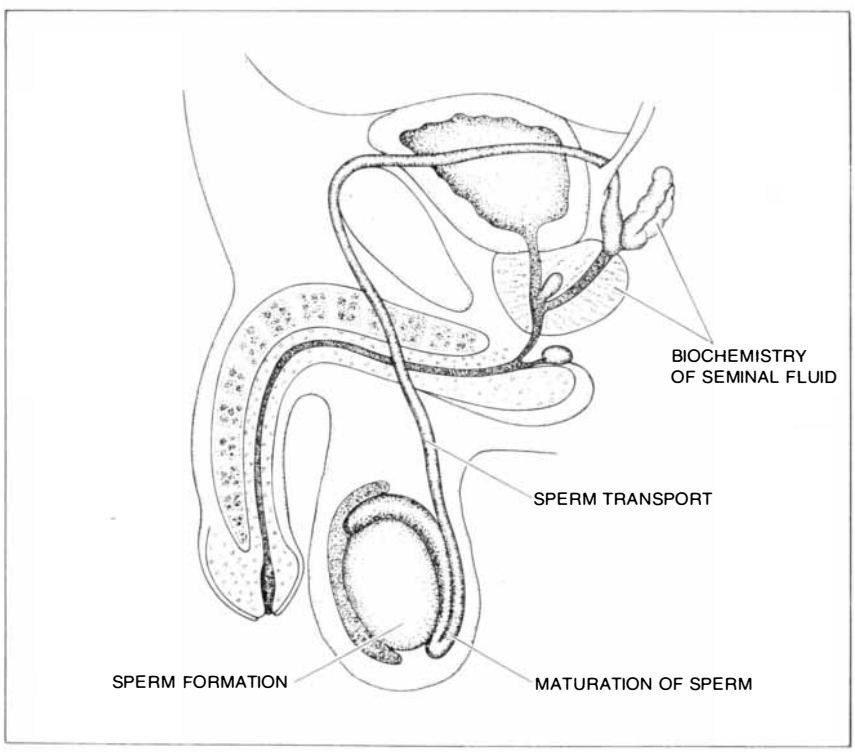
It seems to me that three basic philosophical problems are involved in the sluggish response of the health system to the opportunities presented by the policy change on family planning. The first is the protracted debate over whether medical care should be delivered by categories or comprehensively. Family planning had the misfortune to come in from the cold when the watchword in health circles was "comprehensive." Although health leaders privately conceded the difficulty, if not the impossibility, of launching a new health service without a certain amount of specialized attention and protection, they felt compelled to oppose in public any special administrative arrangements or earmarked funding and to decry any form of "freestanding" (special purpose) clinic, which is the type of clinic serving a large proportion of the people who are involved in family-planning programs.

For years the essence of the argument for comprehensive service has been that it would provide the preventive care that is now so inadequately provided and that it would reduce the need for acute, episodic care. Few people argue against the desirability of such comprehensive services, but there are many differences of opinion on how to achieve them. The problem is twofold, involving both restructuring the delivery of care and providing individuals with concrete opportunities to learn from experience the benefits of preventive services.

It is clear that little such learning has been achieved among the people who are able to pay for visits to private physicians; they continue to go to this or that physician or specialist for treatment of an existing illness. In contrast, a finding that emerged in the obstetrics and gynecology departments of some hospitals in the 1950's is of interest. It was that the introduction of family-planning services in postpartum clinics, which was necessary because the clinics were testing new contraceptive products, doubled the rate of postpartum return among ward patients. A fairly simple change yielded a considerable response from a group of patients who had previously been characterized by physicians and researchers as "unmotivated" to seek postpartum checkups.

On the basis of this experience and others Planned Parenthood suggested that if the family-planning service that low-income people appeared to need and want were offered, an opportunity would be created to provide them with other types of health care that they (like people of higher income) did not yet perceive as being needed, although health professionals had described the services as important. Data from computerized reporting systems for family-planning services show how the hypothesis has stood up. Of the estimated 2.6 million women served last year by organized family-planning programs, nine in 10 received pills or intrauterine devices, nine in 10 had annual pelvic examinations, eight in 10 had annual breast examinations and Pap tests, six in 10 had other laboratory procedures such as blood tests and screening for venereal disease and more than half had other medical examinations. In this instance a specialized approach has worked to achieve not only the primary purpose of the program but also other health goals: organized family-planning programs operated by a variety of health and other agencies have become the nation's largest providers of preventive health service to young women of low or marginal income.

The second philosophical issue is the related one of whether the focus of health policy should be on quality or on quantity. Family planning, whatever its broad social and health consequences, remains a relatively simple mass personal health service. The current health system offers few rewards for either individual practitioners or agencies to under-



MALE REPRODUCTIVE SYSTEM is portrayed schematically to indicate points where the reproductive process is susceptible to control. The process can also be approached by measures directed at the output of the pituitary gland and the production or the effect of steroid hormones. Both pituitary and hormonal processes are under the control of the brain.

take seriously the task of delivering mass personal health services.

In the light of the information that has been emerging recently about the delivery of medical care in China by "barefoot doctors" and their role in (among other things) China's extensive family-planning efforts, I suggest that the slow response of the U.S. health system to the change in family-planning policy is related to the basic issue of whether a nation should choose to emphasize quality care for the few or basic core services for the many. In the U.S. the choice has been for the former policy, although the nation has the resources to provide everyone with at least a minimum standard of health care that meets basic preventive and curative needs.

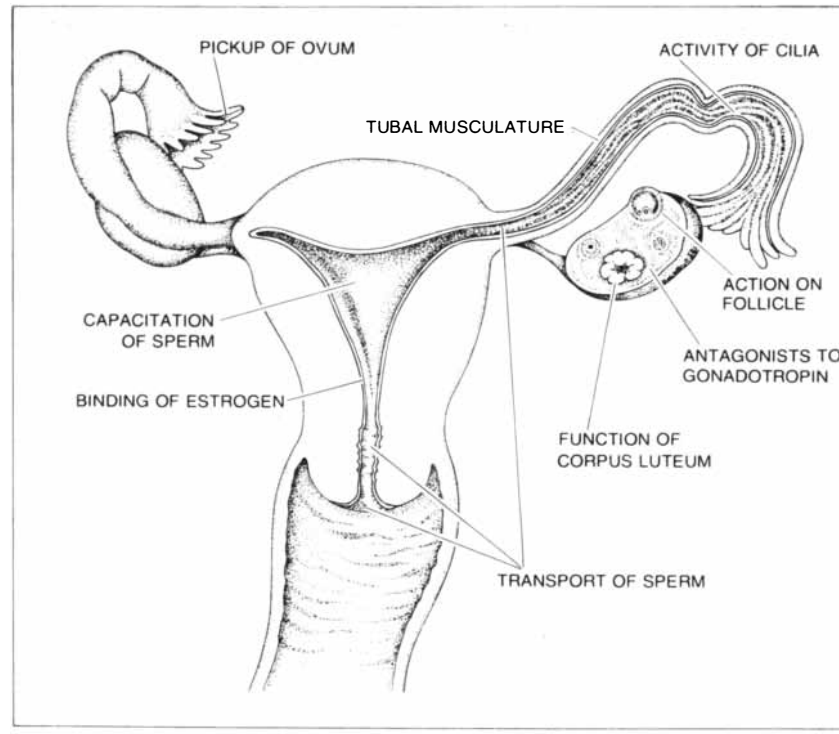
The third philosophical issue is what I see as a fundamental misperception of the family-planning problem by many professionals in the fields of health and social service. It is expressed by such frequently voiced opinions as "Everyone in the U.S. who wants birth control can get it" and "The poor are not motivated to use birth control because they want large families." Yet over the past several years acceptance of family planning has constantly outpaced even the most optimistic assessments of what was possible. Similar arguments have been raised about abortion, yet in the three years since abortion was legalized in New York the ratio of abortions to live births, particularly among blacks and people of low income, has been far higher than anyone would have predicted. Now one hears the same arguments to explain why so few sexually active teen-agers appear to practice effective contraception, although it is only recently that any organized family-planning services have been made available to these young people.

Since many of these debates have involved services for low-income people, I believe a major reason for the widespread misperceptions has been the persistence of upper-class biases about poor people and upper-class ideologies about the reasons for their poverty. The academic community has had a role in reinforcing these biases. Much of the social research done since World War II has emphasized differences in attitudes among individuals and groups and deemphasized studies of how well or how badly social institutions function. The result has been a tendency to seek remedies for social problems by attempting to change or blame the victim rather than by modifying institutions.

David Mechanic of the University of Wisconsin recently wrote: "While it is true that people's responses to health and illness are often conditioned responses to prior background and experience, the health services system has the capacity to modify such behavior patterns. It can foster dependency or encourage self-reliance. It can respect and enhance the dignity of persons or contribute to stigmatizing and humiliating them.... Barriers to medical and health care that are a product of the way health professionals and health care services function are more amenable to change than client attitudes and behavior. There is evidence that when cost and other barriers are removed from access to medical care, and a valuable service is offered, differential utilization of medical services by social class largely disappears."

That is the key hypothesis on which Planned Parenthood's approach to family planning has been based. Organized family-planning programs for people of low or marginal income have been growing at an average annual rate of 32 percent a year for five years. One can therefore argue that the family-planning experience offers a confirmation of the hypothesis and suggests its more general applicability.

In recent years there have been increasing demands for accountability in the way the health system functions. A variety of new ideas and approaches have emerged; they include the creation of provider agencies (such as free clinics) outside the health system, the concept of community control, the idea that health institutions can be and perhaps ought to be managed by professionals who are not physicians and the notion that the nation faces a health-care crisis that will require drastic changes in the mode of organizing, delivering and financing care [see "The Task of Medicine," by William H. Glazier; SCIENTIFIC AMERICAN, April]. Whatever the merits of these approaches, each one represents an attempt either to fashion a more rational and accountable system or to cope with deficits in the current system. The family-planning experience suggests that the health system can be made somewhat responsive to social goals, but it appears to lack the internal capacity to initiate social change or to generate the required modifications of institutional arrangements and practices. The achievement of such goals requires sustained, focused and systematic attention from outside the health system—a process in which political leaders and ordinary citizens can play a decisive role, particularly when their efforts can be informed by the judgments of even a few health professionals with technical skill and imagination.



FEMALE REPRODUCTIVE SYSTEM offers a number of places where the reproductive process can be controlled. They include ovulation, transport of the egg, transport of sperm, fertilization and prevention of implantation. As is the case with the male reproductive system, pituitary and hormonal processes are also subject to a certain degree of control.

THE PHYSICS OF BRASSES

A trumpet produces musical tones when the vibrations of the player's lips interact with standing waves in the instrument. These waves are generated when acoustic energy is sent back by the instrument's bell

by Arthur H. Benade

It is easy to grasp why stringed instruments make the sounds they do. When the strings are struck or plucked, they vibrate at different natural frequencies in accordance with their tension and their diameter. The energy of vibration is then transferred to the air by way of a vibrating plate of wood and a resonating air chamber, with the sound eventually dying away. The musician can vary the pitch, or frequency, of individual strings by changing their vibrating length with the pressure of his fingers on the frets or the fingerboard.

The principles underlying the acoustics of bowed-string instruments such as the violin or wind instruments such as the oboe are a good deal less obvious. Here a vibration is maintained by a feedback mechanism that converts a steady motion of the bow, or a steady application of blowing pressure, into an oscillatory acoustical disturbance that we can hear. On the violin and in the oboe different tones are produced by altering the effective length of the string or the air column.

Like the oboe and other woodwinds, the brass instruments can produce sustained tones. The question arises, however, of how a bugle, which is hardly more than a loop of brass tubing with a mouthpiece at one end and a flaring bell at the other, can produce a dozen or more distinct notes. Horns were fashioned and played for centuries before physicists were able to work out good explanations of how they worked, even though scientific attention has been directed to these questions from the earliest days. For centuries the skilled craftsman has usually been able to identify what is wrong with faulty instruments and to fix them without recourse to sophisticated knowledge of horn acoustics.

All brass instruments consist of a mouthpiece (which has a cup and a tapered back bore), a mouthpipe (which also has a carefully controlled taper), a main bore (which is either cylindrical or conical) and a flaring bell that forms the exit from the interior of the horn into the space around the instrument. Brass instruments are of two main types. Those in one family, which includes the trumpet, the trombone and the French horn, have a considerable length of cylindrical tubing in the middle section and an abruptly flaring bell. Those in the other family, called conical, include the flügelhorn, the alto horn, the baritone horn and the tuba. The generic term conical refers to the fact that much of the tubing increases in diameter from the mouthpiece to the bell and the flare of the bell is itself less pronounced than it is in the first family. Actually all the horns called conical incorporate a certain amount of cylindrical tubing in their midsection. Here I shall deal primarily with the properties of instruments in the trumpet and trombone family. The properties of the conical instruments are very similar except for being somewhat simpler acoustically because overall they have much less flare.

The acoustical study of waves in an air column whose cross section varies along its length (a "horn") goes back to the middle of the 18th century. Daniel Bernoulli, Leonhard Euler and Joseph Louis Lagrange were the first to discuss the equations for waves in such horns during the decade following 1760. Their activity was a part of the immensely rapid blossoming of theoretical physics that took place in the years after the laws of motion had been formulated by Newton and Leibniz. Theoretical investigations of fluid dynamics, acoustics, heat flow and the mechanics of solid objects took their inspiration from the workaday world outside the laboratory and the mathematician's study. The work of Bernoulli, Euler and Lagrange on horns (and their similar researches on strings) did not have much influence in the long run on the science of acoustics or the art of music. It was nonetheless a part of the initial blooming of the theory of partial differential equations underlying nearly all physics.

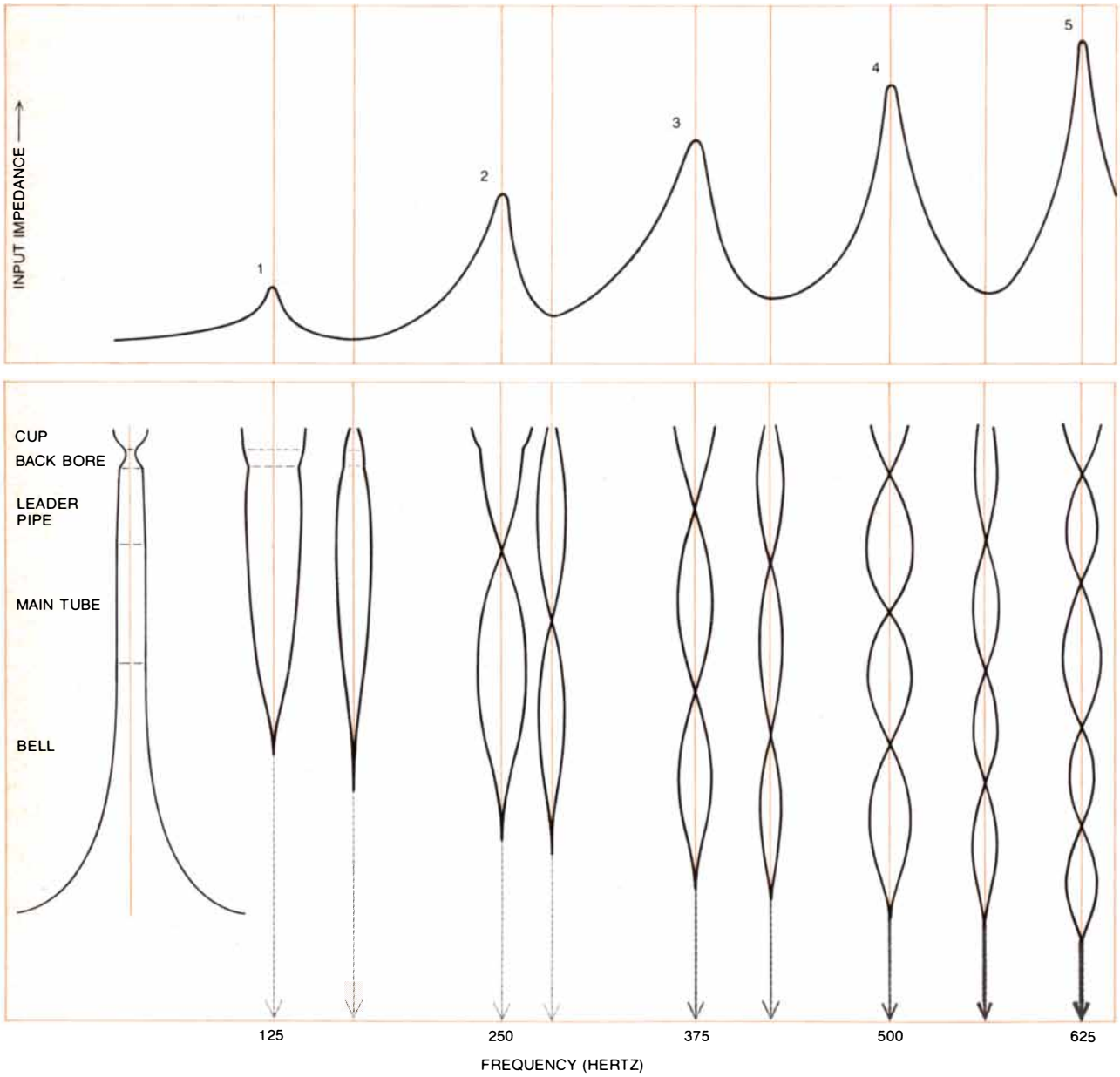
The "horn equation," as we call it today, was neglected until 1838, when George Green rediscovered it while investigating the erosion caused by waves in the new canal systems of England. Then the equation was buried again until 1876, when a German mathematician, L. Pochhammer, independently derived it for waves in a column of air and learned the properties of its most important solutions. Neither Pochhammer nor his equation was long remembered. Finally in 1919 an American physicist, A. G. Webster, published a report on the horn equation, with the result that the equation is commonly named for him.

Since Webster's time interest in loudspeakers on the part of the phonograph and radio industries, to say nothing of military demands for sonar gear to detect submarines, has kept the subject of horn acoustics in a lively state. A loudspeaker horn must be designed to radiate sound efficiently out into the air over a broad range of frequencies from a small source. A horn designed to serve as a musical instrument has quite different requirements. In a musical horn the flare of the bell must be designed to trap energy inside the horn, giving strongly marked standing waves at precisely defined frequencies.

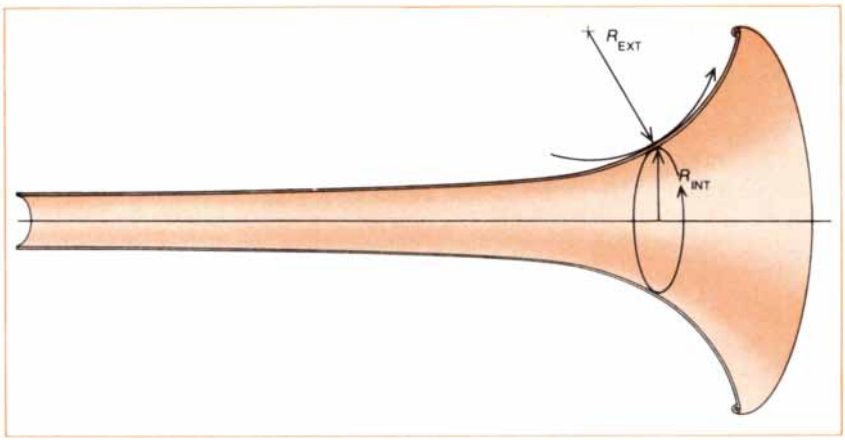
It is obvious that as a wave travels into the enlarging part of a horn its pressure

will decrease systematically, simply because the sound energy is being spread over an ever wider front. If one extracts this intuitively obvious part of the behavior of a wave in a horn from the mathematics of the horn equation, one is left with a much simpler equation that is identical in form with the celebrated Schrödinger equation of quantum mechanics. The Schrödinger equation shows that a particle of energy E has associated with it a de Broglie wavelength lambda (λ) that depends on the square root of the difference between the energy and the potential energy function V at any point in space. The "reduced," or simplified, form of the horn equation shows similarly that at any point in the horn the acoustic wavelength depends on the square root of the difference between the squared frequency and a "horn function" U that depends in a rather simple way on the nature of the horn flare [*see top illustration on next page*].

It is not difficult to show from the horn equation that sounds propagate with dif-



RESONANCE PEAKS OF A TRUMPETLIKE INSTRUMENT can be plotted (*top*) in terms of the impedance measured at the mouthpiece. Impedance is defined as the ratio of the pressure set up in the mouthpiece to the excitatory flow that gives rise to it. The impedance depends on whether the sound wave reflected from the bell of the horn returns in step or out of step with the oscillatory pressure wave produced in the mouthpiece. The shape of the air column in the trumpetlike instrument is shown at the extreme left of the bottom part of the diagram. The curves at the right are the standing-wave patterns that exist in the air column of the instrument at frequencies that produce the maxima and minima in the impedance curve. The first maximum is at about 100 hertz (cycles per second), when the reflected wave is precisely in step with the entering wave. The small irregularities in the standing-wave pattern are produced by the abrupt changes in the cross section of the instrument. The first minimum comes just above 125 hertz, where the returning wave and the incoming wave are exactly out of step with each other in the mouthpiece of the instrument. The subsequent maxima and minima are similarly explained. The number of nodes in the standing-wave pattern increases by one at each impedance peak.

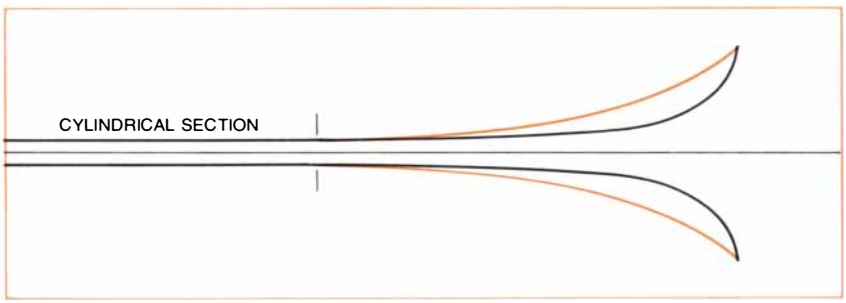


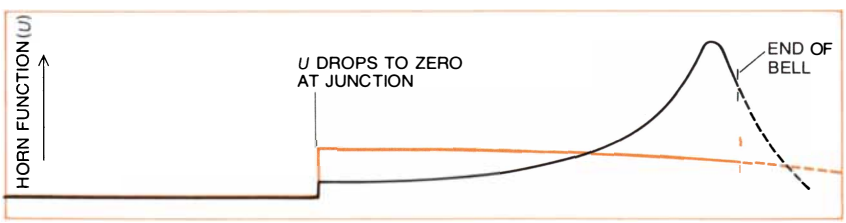
a $U \cong \frac{1}{R_{INT} \times R_{EXT}}$

b $\lambda = \frac{c}{\sqrt{f^2 - U(c/2\pi)^2}}$

c $\lambda = \frac{h}{\sqrt{E - V}}$

GEOMETRY OF HORN FLARE largely governs the pitch and timbre of sounds produced by horns of the trumpet and trombone family. As a sound wave travels into the flaring bell of the horn its pressure falls steadily as the cross section of the instrument increases. A "horn function," U, determines how much of the acoustic energy leaves the horn and how much is reflected back into the horn to produce standing waves inside the instrument. The horn function (*equation "a"*) is approximately equal to 1 over the product of the internal radius (R_{int}) of the horn and the external radius (R_{ext}) at any given point. The simplified form of the horn equation (*equation "b"*) gives the acoustic wavelength (λ) at any point in the horn, where f is the sound frequency and c is the velocity of sound. This velocity varies with U and f. The horn equation has the same form as the celebrated Schrödinger equation (*c*), which shows how the de Broglie wavelength (λ) of a particle of energy E is related to Planck's constant (h) and the potential energy function V at any point in space.





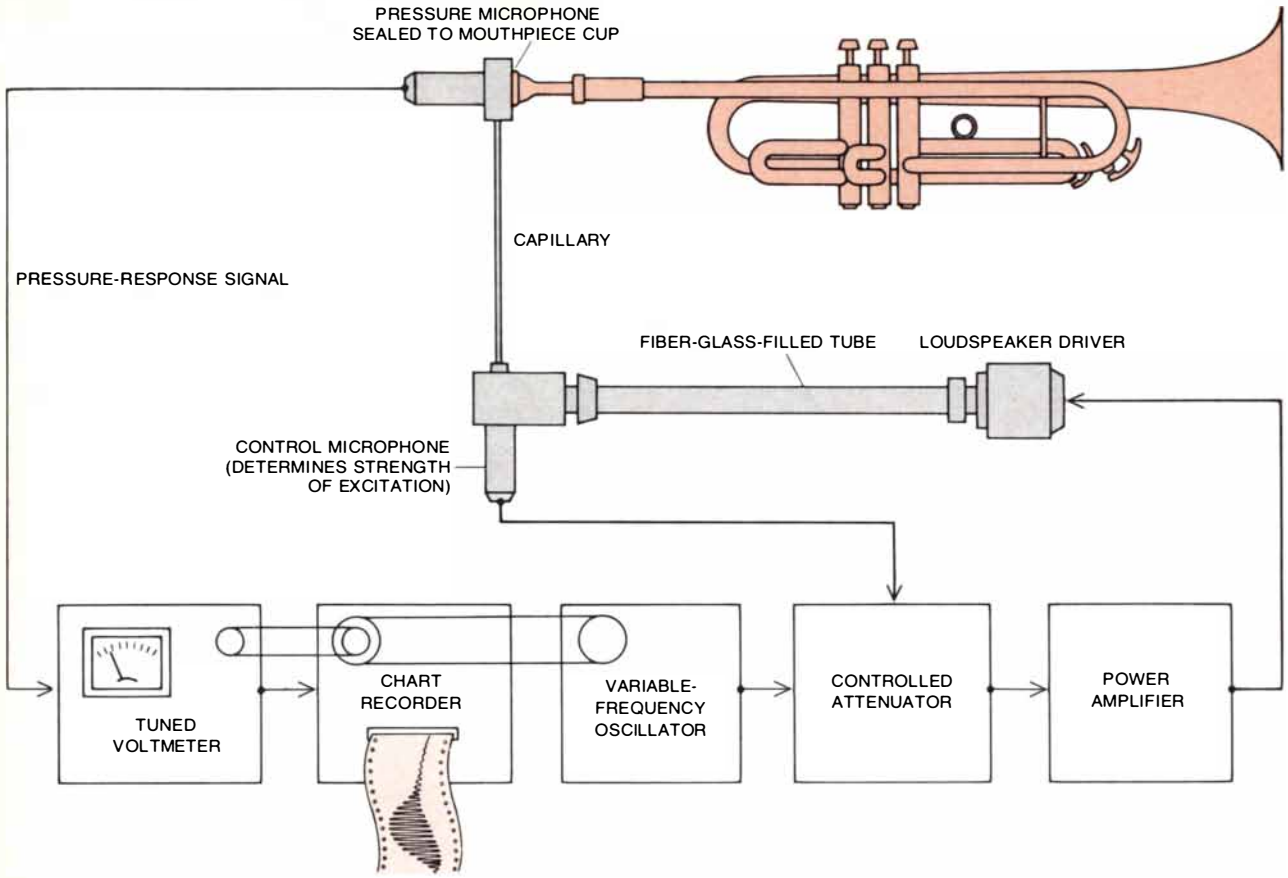
TROMBONE BELL AND LOUDSPEAKER HORN are markedly different in geometry and acoustic properties. The catenoidal shape (*black curve at top*) of the loudspeaker horn favors the efficient radiation of sound into the air. The flaring shape (*colored curve at top*) of the trombone bell is designed to save energy inside the horn, thus generating strongly marked standing waves at closely defined frequencies. Both the trombone bell and the loudspeaker horn are shown attached to a short section of cylindrical pipe. The two curves at the bottom show the horn function, U, for each horn. The catenoidal horn has a horn function (*colored curve*) that is low and nearly constant except for a slight falling off at the large end, where the sound wave fronts begin to bulge appreciably. The horn function (*black curve*) of the trombone bell rises steeply and falls. The higher the value of the function U, the higher the barrier to sounds of low frequency. Sounds of higher frequency are able to progress farther before they are reflected back by the barrier. In both cases above a certain frequency most of the sound energy radiates over the top of the barrier, so that the bell of the trombone loses its musically useful character and behaves like a loudspeaker horn.

ferent speeds as they travel through regions of differing horn function U. The speed of propagation also depends on the frequency. Another similarity between horn acoustics and quantum mechanics is that for frequencies below a certain critical value determined by the magnitude of U, the wavelength becomes mathematically imaginary, or, to put it in more physical terms, the wave changes its character and becomes strongly attenuated. In other words, regions where the horn function U is large can form a barrier to the transmission of waves and can therefore reduce the escape of energy from within a horn to the outside. The leaking of sound from the horn through the horn-function barrier is an exact analogue to the leaking of quantum-mechanical waves (and therefore particles) through the nuclear potential barrier in the radioactive decay of the atomic nucleus.

Let us look more closely at the difference between a musical horn and a loudspeaker horn. A simple example of a musical horn can be constructed by joining a trombone bell to a piece of cylindrical pipe. To a similar pipe one can join a typical loudspeaker bell, whose figure is described as catenoid. Even if the bells are matched to have the same radii at both ends, we find that their horn functions are quite different [*see bottom illustration at left*]. The catenoidal bell has a horn function that is approximately constant from one end to the other, whereas the acoustical properties of the horn function for the musical horn vary from point to point.

Five years ago Erik V. Jansson of the Speech Transmission Laboratory of the Royal Institute of Technology in Stockholm worked with me at Case Western Reserve University on a detailed study of air columns similar to those found in musical horns. In this work, which was both theoretical and experimental, we studied bells of the type found on trumpets, trombones and French horns. We unearthed a number of subtle relations between our experiments and calculations that we did not have time to clarify immediately. It is only recently that we have had an opportunity to prepare complete reports on our results. In what follows I shall lean heavily on information gained in our work five years ago and its later development, and on the earlier observations of many people concerned with acoustics or making musical horns.

In a brass musical instrument the small end of the horn is connected to the



IMPEDANCE-MEASURING APPARATUS uses the driver from a horn loudspeaker as a pump to feed a flow stimulus through a capillary into the mouthpiece cup of the instrument under study. A control microphone sends signals to an attenuator to ensure that the acoustic stimulus entering the capillary remains constant. The pressure response of the instrument, and thus its input impedance, is detected by a second microphone that forms the closure of the mouthpiece cup. The signal from the microphone goes to a frequency-selective voltmeter coupled by a chain drive to oscillator. A chart recorder coupled to the voltmeter plots the resonance curves.

player through his lips, which constitute a kind of automatically controlled valve for admitting air from the player's lungs to the horn. The opening and closing of the valve is controlled chiefly by the pressure fluctuations within the mouthpiece as they act on the lips in concert with the steady pressure from the lungs. Therefore an initial objective is to find the relations between the flow of air into the horn and the acoustical pressure set up at the input end.

Let us begin by imagining a laboratory experiment in which the horn is excited not by air from the player's lips and lungs but rather by a small oscillatory flow of air being pumped in and out of the mouthpiece through a fine capillary by a high-speed pump. This small oscillatory flow disturbance in the mouthpiece gives rise to a pressure wave that ultimately reaches the flaring part of the horn. As the wave travels down the length of the bore of the horn some of its energy is dissipated by friction and the transfer of heat to the walls of the instrument. In the flaring part of the bell a substantial fraction of the acoustic wave is reflected back toward the mouthpiece while the remainder penetrates the horn-function barrier and is radiated out into the surrounding space. The wave that is reflected back down the bore of the horn combines with newly injected waves to produce a standing wave.

If the round-trip time that the wave takes to go from the mouthpiece to the bell and back to the mouthpiece is equal to half the repetition time of the original stimulus or to any odd multiple of the repetition time, a standing wave of considerable pressure can build up and result in a large disturbance in the mouthpiece. At intermediate frequencies of excitation the return wave tends to cancel the influence of the injected wave. In other words, depending on the precise interaction between the injected wave and the reflected wave, the pressure disturbance inside the mouthpiece can be large or small. For purposes of describing such disturbances in the mouthpiece under conditions of constant flow excitation in a laboratory apparatus, engineers define a quantity termed input impedance: the ratio between the pressure amplitude set up in the mouthpiece and the excitatory flow that gives rise to it [*see illustration on page 25*].

The shape of the horn controls the natural frequencies associated with the various impedance maxima and minima by determining the penetration of the standing waves into the bell. The shape also controls the amount of wave energy that leaks out of the horn into the surrounding space. Furthermore, the kinks in the standing wave that arise from discontinuities in cross section and taper along an air column produce significant changes in both the resonance and the radiation properties of the bell. The interaction of the kinks and the primary shape of the air column can spell the difference between success and failure in the design of an instrument.

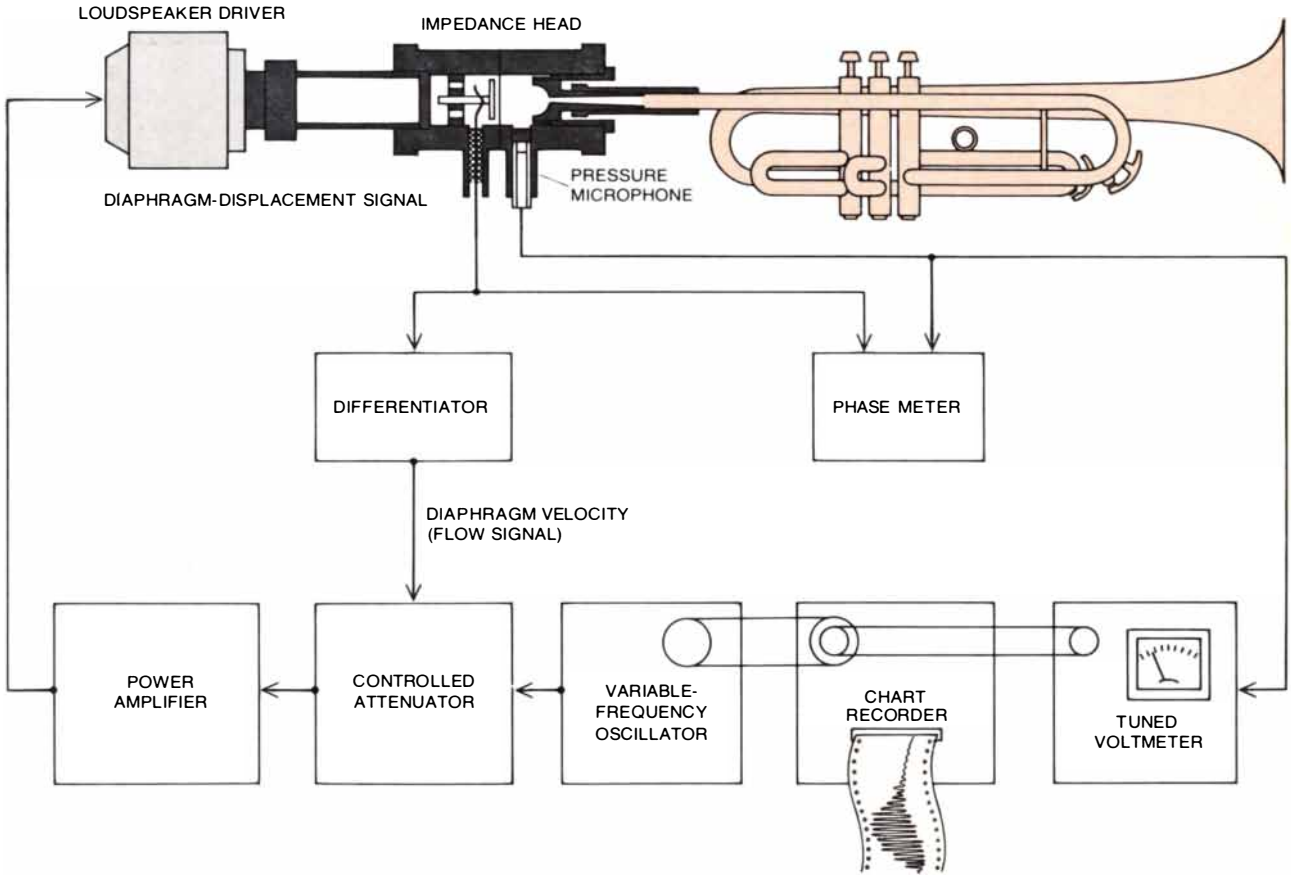
There are several ways one might measure the input impedance, or re-

sponse, of the air column. Conceptually the simplest method would be to pump air in and out of the mouthpiece through a capillary tube at some frequency and measure the amplitude of the resulting pressure fluctuations in the mouthpiece by means of a probe microphone. It is more practical, however, to use the driver of a commercial horn loudspeaker as a pump. The motion of the driver is controlled electronically by an auxiliary monitor microphone that maintains a constant strength of oscillatory flow through the capillary as one sweeps automatically through the appropriate range of frequencies. Between 1945 and 1965 Earle L. Kent and his co-workers at C. G. Conn Ltd. in Elkhart, Ind., developed this basic technique to a high degree of dependability. We often employ a modification of their technique in our work [*see illustration on preceding page*].

In Cleveland we make use of two additional methods that have special advantages for certain purposes. The first method, based on a device described in 1968 by Josef Merhaut of Prague, can be applied in measurements not only on the smaller brasses but also on bassoons and clarinets [*see illustration below*]. In Merhaut's device a thin diaphragm forms a closure at the end of the mouthpiece cup and itself serves as the pump piston. The diaphragm is driven acoustically through a pipe that connects it to an enclosed loudspeaker. The diaphragm motion is monitored for automatic control by the electrode of a condenser microphone mounted directly behind it. The second method is based on a device that was used by John W. Coltman of the Westinghouse Research Laboratories in investigating the sounding mechanism of the flute. In Coltman's device the excitatory diaphragm is driven directly by a loudspeaker coil whose motion is monitored by means of a second pickup coil that is moving in an auxiliary magnetic field [*see illustration on opposite page*].

If one attaches to any one of these excitation systems a cylindrical section about 140 centimeters long from a trumpet, one discovers dozens of input impedance peaks evenly spaced at odd multiples of about 63 hertz (cycles per second) [*see curve "a" in top illustration on page 30*]. The peaks correspond exactly to what elementary physics textbooks describe as the "natural frequencies of a cylindrical pipe stopped at one end." Because frictional and thermal losses inside the tube walls increase with frequency, the resonance peaks become smaller at higher frequencies. The energy radiated from the open end of such a pipe is only a tiny fraction of 1 percent of the wall losses.

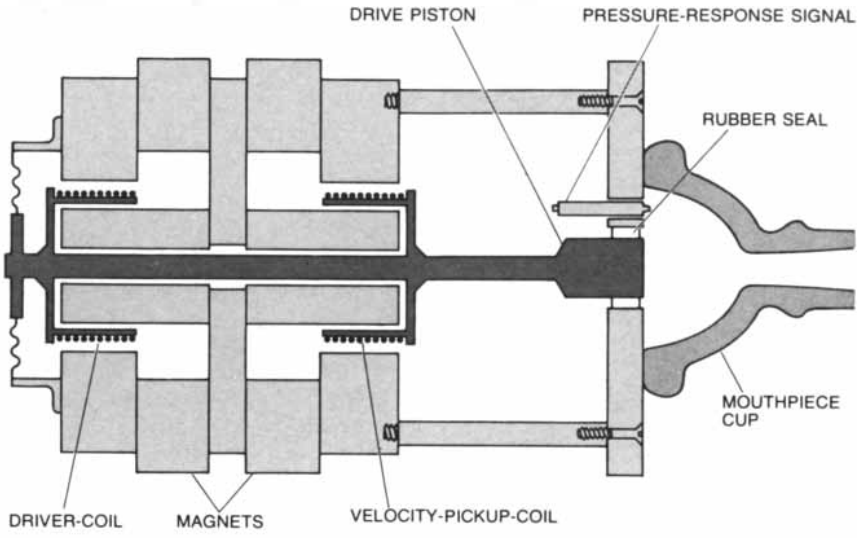
If one now adds a trumpet bell to the same cylindrical pipe, the impedance response curve is substantially altered [*see curve "b" in bottom illustration on*



SECOND TYPE OF IMPEDANCE-MEASURING DEVICE was developed by Josef Merhaut. It differs from the apparatus illustrated on the preceding page only in the way that the flow stimulus into the mouthpiece is controlled. Here the acoustic stimulus produced by a loudspeaker moves an aluminized Mylar diaphragm that in turn pumps air into the mouthpiece. The diaphragm also acts as one electrode of a condenser microphone to produce a signal proportional to the diaphragm's velocity and thus proportional to the oscillatory flow of air at the mouthpiece cup. The velocity signal adjusts the attenuator in order to maintain constant excitation at a particular frequency. The pressure response of the instrument is monitored by a microphone on the cup side of the diaphragm. A phase meter shows the relation between the phase of the input stimulus and the phase of the pressure response of the instrument.

next page]. The first peak is hardly shifted at all by adding the bell, but the frequencies of the other resonances are lowered in a smooth progression because the injected waves penetrate ever more deeply into the bell before being reflected. In addition the peaks at higher frequencies are markedly reduced in height because a growing fraction of the energy supply leaks through the bell "barrier" as the frequency is increased. In sum, the return wave in the pipe-plus-bell system is weakened not only by wall losses but also by radiation losses, particularly at high frequencies. Above about 1,500 hertz essentially no energy returns from the flaring part of the bell. The small wiggles in the impedance curve at high frequencies are due chiefly to small reflections produced at the discontinuity where the bell joins the cylindrical tubing.

By comparing these curves for incomplete instruments with the impedance curve for a complete cornet [*see illustration on page 31*] one can see at a glance that the presence of a mouthpipe and mouthpiece has a considerable effect on the overall nature of the input impedance. The resonance peaks of the cornet grow taller up to around 800 hertz, then fall away much more abruptly than the curve produced by the pipe-plus-bell system.



ELECTROMAGNETIC SOURCE for projecting acoustic waves into a test instrument was devised by John W. Coltman. The excitatory piston is directly coupled to the voice coil of a loudspeaker. The coil in turn drives the piston with an amplitude that is ultimately determined by a voltage induced in a pickup coil that is mounted on the same shaft. The mechanism is used in an overall system similar to that used with the Merhaut impedance head. The pressure response in the mouthpiece cup is detected by a miniature microphone.

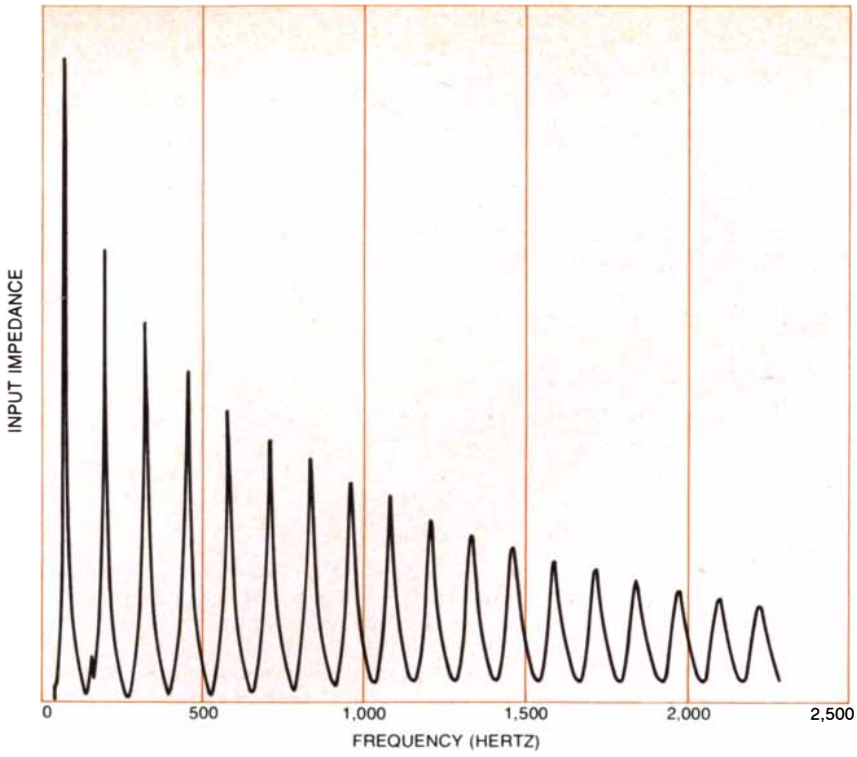
Let us now consider how the player's lips control the flow of air from his lungs into the instrument. As the player blows harder and harder, the flow increases both because of the increased pressure across the aperture formed by his lips and because his lips are forced farther apart by the rising pressure inside his mouth. Equally important is the variation imposed on the flow by pressure variations inside the mouthpiece, which tend to increase or decrease the flow by their own ability to affect the size of the lip aperture. It is this pressure-operated flow control by the lips under the influence of the mouthpiece pressure that ultimately leads to the possibility of self-sustained oscillation. Let us abstract from this rather complicated situation only the relevant part of it: the alteration in net flow that is produced by acoustical pressure variations within the cup of the mouthpiece. As long ago as the middle of the 19th century it was clearly understood that it is the flow alteration due to mouthpiece pressure that can maintain an oscillation.

In 1830 Wilhelm Weber described experiments on the action of organ reeds that led him to a correct theory for the effect of a compliant structure (the reed or, in our case, the player's lips) on the input impedance of a column of air. This effect of the yielding closure of the mouthpiece cup provided by the lips is quite separate from the lips' functioning as a valve. Hermann von Helmholtz provided the next advance. In 1877 he added an appendix to the fourth German edition of his classic work *Sensations of Tone* that gives a brief but complete analysis of the basic mechanisms by which a pressure-controlled reed valve collaborates with a single impedance maximum. He found that for a given pressure-control sensitivity (what an engineer today calls the transconductance) a certain minimum impedance value is required. Oscillating systems of the type analyzed by Helmholtz are found around us everywhere. The pendulum clock is possibly the oldest and most familiar. The wristwatch, electronic or otherwise, falls into this category. Every radio and television set has one such oscillator or more.

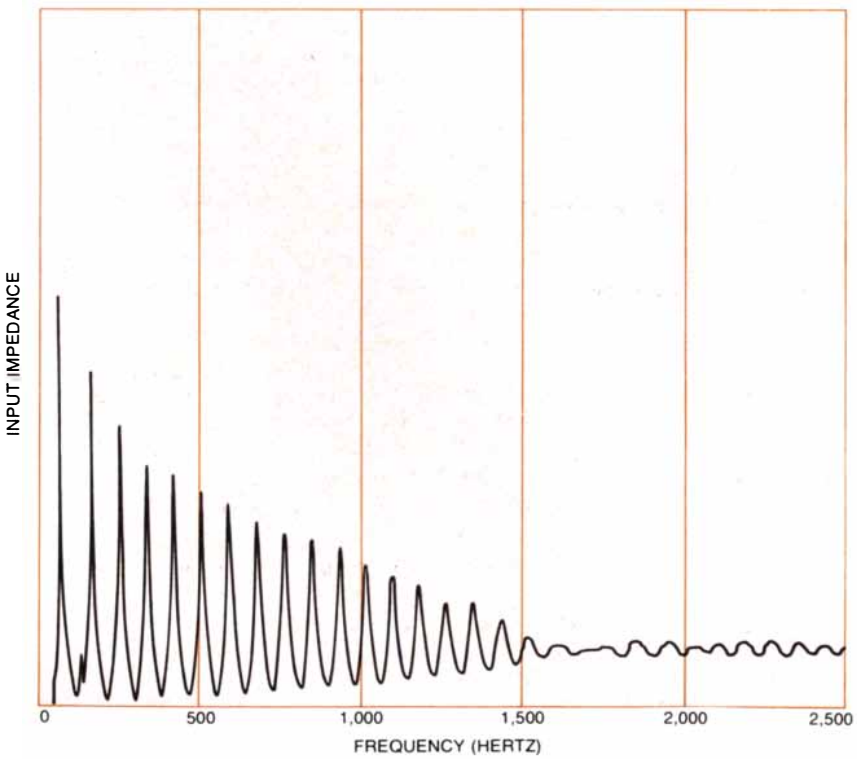
Engineers have studied oscillating systems intensively and have learned that even if the alteration in flow (of whatever kind) that results from a given pressure is not exactly proportional to the pressure (as Helmholtz assumed for simplicity in his pioneering investigation) but varies in some more arbitrary fashion, the properties of the system are not drastically altered. The presence of such nonlinearity in the control characteristics gives rise to additional frequencies at double, triple and quadruple the frequency of the basic oscillation. The net generation of oscillatory energy from the player's steady muscular effort, however, is still almost exclusively at the frequency of the impedance maximum; energy diverted in the process to other frequencies is dissipated in various ways to the outside world.

We must now try to explain how oscillations in a wind instrument can take place at not just the tallest impedance maximum but at any one of several maxima belonging to an actual air column. According to the Helmholtz theory, a wind instrument should show a strong preference for oscillations that take place at the tallest of the impedance maxima. Thus the question arises of how the bugle player finds it possible to play the notes based on lesser impedance maxima. Furthermore, one must ask how the bugler is able to select one or another of these peaks in accordance with his musical requirements.

It is not in fact difficult to deal with the problem of how the player selects one note or another. His lips are so massive compared with the mass of the air in his instrument that the influence of the air column on the lips is relatively small. The player adjusts the tension of his lips in such a way that their own natural tendency of vibration favors oscillation at the desired note, so that the



IMPEDANCE PATTERN OF SIMPLE CYLINDRICAL PIPE 140 centimeters long shows peaks evenly spaced at odd multiples of 63 hertz. The higher the frequency, the greater the loss of wave energy to the walls of the pipe through friction, hence the steady decline in the height of the peaks. Less than 1 percent of the input energy is radiated into the room.



ADDITION OF TRUMPET BELL TO PIPE lowers the overall height of the impedance peaks and squeezes them together. Whereas the pipe alone produces 16 peaks in a span of 2,000 hertz, the pipe-plus-bell system compresses the first 16 peaks into a span of 1,400 hertz. Beyond 1,500 hertz more and more of the acoustic energy leaks through the bell barrier.

air column and the lips collaborate in producing the desired frequency.

So far we have not said anything that could not have been understood in terms of 19th-century acoustics. The best account of the Weber-Helmholtz analysis and its musical consequences was made by a French physicist, Henri Bouasse, in his book *Instruments à Vent,* the two volumes of which appeared in 1929 and 1930. These volumes contain what still constitutes one of the most thorough accounts of the acoustics of wind instruments, encompassing the flute and reed organ pipes, the orchestral woodwinds and the brasses. Bouasse has left us with a gold mine of mathematical analysis, along with an account of careful experiments done by himself in collaboration with M. Fouché or selected from the writings of earlier investigators.

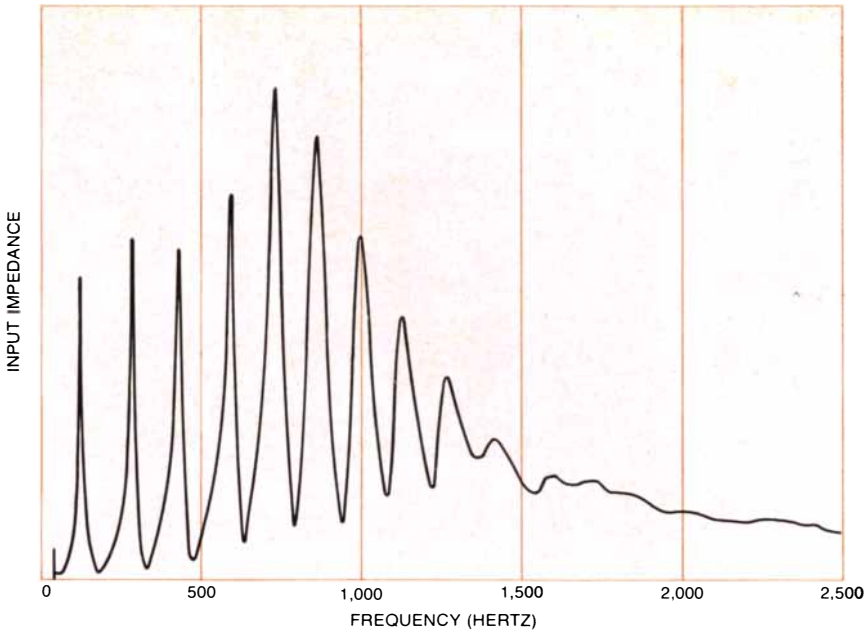
Bouasse was quite aware of the inadequacy of a theory of oscillation assuming that all the energy production is at the basic frequency of oscillation. He described many phenomena observed among the reed organ pipes and the orchestral woodwinds and brasses that underscore the limitations of this general viewpoint and that imply cooperation among several air-column resonances. Bouasse's interest in these matters was to serve both as a strong incentive and as an invaluable guide when I later undertook a close study of the subject. The first fruits of this study were described in a series of technical reports written in 1958 for C. G. Conn Ltd.

By 1964 I found it possible to deal well enough with the interaction between a reed valve and an air column having several impedance maxima that I could design and build a nonplaying "tacet horn." This "instrument" has several input impedance maxima chosen in such a way as to make them unable to maintain any oscillation in cooperation with a reed, even though the Weber-Helmholtz theory would predict the possibility of oscillation. In 1968 Daniel Gans and I published an account of this theory of cooperative oscillations. That report, based on Gans's undergraduate thesis at Case Western Reserve, included a description of the tacet horn and explanations of various phenomena discussed by Bouasse. Since that time the work has been carried much further in our laboratory, particularly by Walter Worman, who wrote his doctoral dissertation on the theory of self-sustained oscillations of this multiple type in 1971. Although his work was focused on clarinetlike systems, his results apply broadly to all the wind instruments, including

the brasses. These studies were aided by counsel from many people, in particular Bruce Schantz, Kent, Robert W. Pyle, Jr., and John H. Schelleng.

It is now time to see how the Weber-Helmholtz form of the theory had to be modified, using the trumpet as our example. When the musician sounds one of the tones of a trumpet, the air column and his lips are functioning in what we shall formally call a regime of oscillation: a state of oscillation in which several impedance maxima of the air column collaborate with the lip-valve mechanism to generate energy in a steady oscillation containing several harmonically related frequency components. Worman was able to trace out how a set of impedance maxima can work together with the air valve. The particular "playing frequency" chosen by the oscillation (along with its necessarily whole-number multiples) is one that maximizes the total generation of acoustic energy, which is then shared among the various frequency components in a well-defined way.

Experiments with instruments as diverse as the clarinet, the oboe, the bassoon, the trumpet and the French horn show that softly played notes are dominated by the impedance maximum that belongs to the note in the sense of Weber and Helmholtz. As the musician raises the dynamic level, however, the influence of the higher resonances grows in a definite way that is common to all the instruments. As he plays louder and louder, the influence of the impedance at double the playing frequency becomes more marked, and for still louder playing the resonance properties at triple or quadruple frequencies join the regime of oscillation one by one. A look at the input impedance curves for a modern trumpet will show how the peaks in a regime of oscillation cooperate so that the player can sound various notes on his instrument, including even some notes that have no peak at all at the playing frequency [*see illustrations on next two pages*]. Notes in this last category have been known to brass players since the earliest days and were a part of horn-playing technique in the time of Mozart and Beethoven. The need for such notes was reduced, however, as the instrument became more mechanized. In recent years they have returned; for example, they are sounded by musicians who want to play bass-trombone parts without resorting to a special thumb-operated valve that is otherwise required. Tuba players also find the technique useful on occasion.

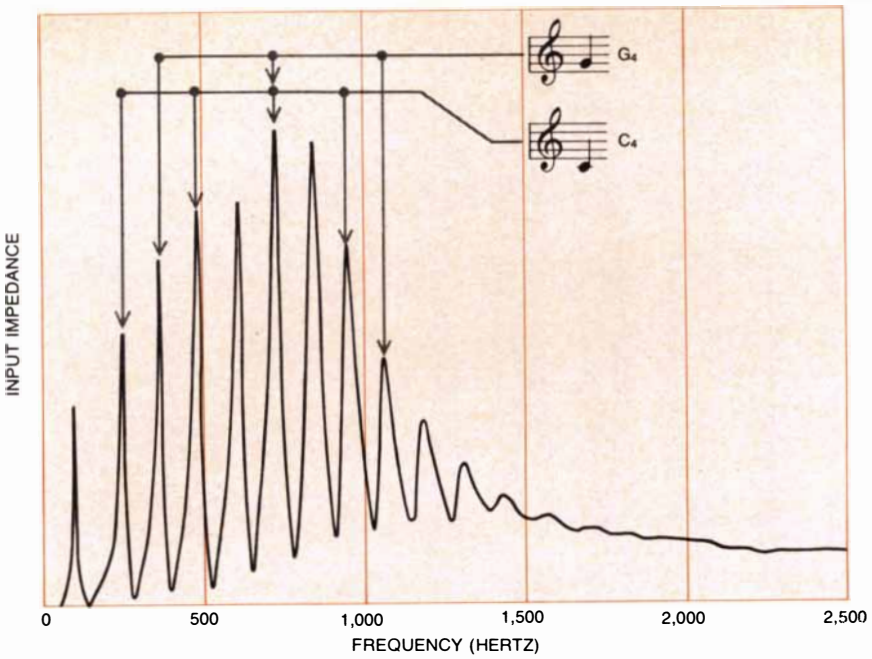


IMPEDANCE PATTERN OF A 19TH-CENTURY CORNET is typical of most of the trumpet and trombone family. The peaks grow progressively and then fall away sharply. The cornet was made in 1865 by Henry Distin. The third and fourth impedance peaks do not quite follow the smoothly rising pattern required for a genuinely fine instrument. The shortcoming is due chiefly to slight constrictions and misalignments in the valve pistons.

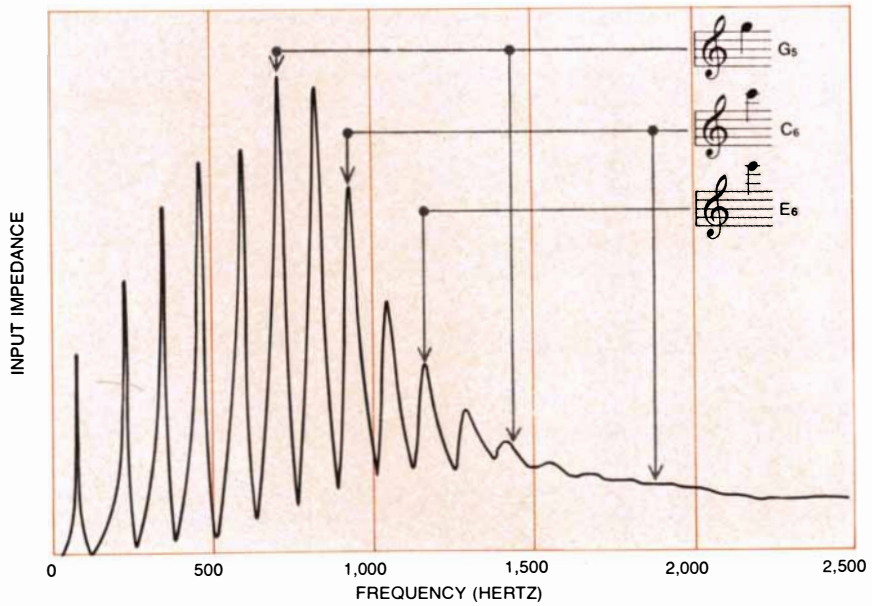
The reader may be wondering what happens when the valves on a brass instrument are depressed. Does anything radically new happen? The answer is no. The bell, the mouthpipe and the mouthpiece dominate the "envelope," or overall pattern, of the resonance curve; the pattern of peaks for a trumpet rises steadily as one goes from low frequencies to about 850 hertz and then falls away and disappears at high frequencies. When a valve is depressed, thereby increasing the length of cylindrical tubing in the middle of the horn, it merely shifts the entire family of resonance peaks to lower frequencies but leaves them fitting pretty much the same envelope.

In addition to working out the details of the regimes of oscillation in wind instruments Worman gained an important insight into the factors that influence tone color. He was able to show that in instruments with a pressure-controlled air valve (a reed or the lips) the strength of the various harmonics generated in a regime of oscillation (as measured inside the mouthpiece) has a particularly simple relation when the instrument is being played at low and medium levels of loudness. Let us take as given the strength of the fundamental component that coincides with the playing frequency. As one would expect, that strength increases as the player blows harder. Worman's striking result is that when the player blows very softly, there is essentially no other component present in the vibration as it is measured in the mouthpiece, and that as he plays louder the amplitude of the second harmonic grows in such a way that for every doubling of the strength of the fundamental as the player blows harder, the strength of the second harmonic quadruples. Furthermore, the strength of this component proves to be approximately proportional to the impedance of the air column at the frequency of the second harmonic. Similarly, the third harmonic has a strength that is proportional to the impedance at the third-harmonic frequency, and from an even tinier beginning it grows eightfold for every doubling of the strength of the fundamental component. In short, the *n*th harmonic has a strength that is proportional to the impedance at the *n*th harmonic of the playing note, and that component grows as the *n*th power of the fundamental pressure amplitude. The remarkable thing about Worman's observation is that it is totally independent of all details of the flow-control properties of the reed or the lips, provided only that the flow is controlled solely by the pressure variations in the mouthpiece [*see top illustration on page 34*].

Let me summarize what we have found out so far about how the tone



IMPEDANCE PATTERN OF A MODERN TRUMPET is annotated to show what happens when a player sounds the notes C_4 or G_4. When he blows into the horn, a "regime of oscillation" is set up in which several impedance maxima of the air column collaborate with oscillations of his lips to generate energy in a steady oscillation that contains several harmonically related frequency components. The regime of oscillation for the C_4 note involves the second, fourth, sixth and eighth peaks in the curve. When the trumpeter plays very softly, the second peak is dominant, but because this peak is not tall the beginner may produce a wobbly note. As he plays louder the other peaks become more influential and the oscillation becomes stabilized. The dominant oscillation for the G_4 note corresponds to the third impedance peak; since it is taller than the second peak, G_4 is easier than C_4 to play pianissimo. As the trumpeter plays louder the tall sixth peak comes in and greatly stabilizes the regime of oscillation, making the G_4 one of the easiest notes of all to play.



REGIMES OF OSCILLATION FOR HIGHER NOTES show why they become increasingly hard to play as one moves up the scale. G_5 is still quite easy to play because its regime of oscillation is dominated by the tall sixth impedance peak; the 12th peak makes only a minor contribution. C_6 is somewhat more difficult to play because the dominant peak of the note is lower than the peak for G_5. It takes an athletic trumpeter to reach the high E_6 and higher notes. The trumpet at this point has become virtually a megaphone: the energy production of the instrument is due almost completely to the interaction of the air column with the lips themselves, much as the human larynx operates in producing vocal sounds.

quality develops as measured inside the mouthpiece of the brass instruments. When one plays very softly, only the fundamental component associated with the playing frequency is present. As one plays louder the second, third, fourth and still higher harmonics grow progressively. If the oscillation is in the nature of a regime involving several cooperating resonance peaks, the harmonics grow in the simple fashion described by Worman's theorem; it is only at very loud playing levels that his theorem fails to give simple results. Furthermore, the theorem shows that the strength of the various components is proportional to the height of the various impedance maxima that are cooperating to generate the tone. In other words, when one plays rather loud, the strengths of the various harmonics have heights that correspond roughly to the heights of the impedance maxima from which they draw their chief sustenance. On the other hand, when a tone is generated on the basis of only a single resonance peak, as is the case in the upper part of the trumpet's range, we would be able to describe the strength of the components only if we could specify all the details of the flow-control characteristic.

Up to this point I have been discussing only the strength of the various harmonics as they are measured by a small probe microphone inside the brass instrument's mouthpiece cup. What one hears in the concert hall is, of course, very different. The transformation from the spectrum generated inside the mouthpiece, where the actual dynamics of the oscillation are taking place, into the spectrum found in the concert hall has to do with the transmission of sound from the mouthpiece into the main air column and thence out through the bell. There are many facets to the total transmission process, even without taking into account the complexities of room acoustics or the complications of our perceptual mechanism, which does a remarkable job of processing the great irregularity of room properties to give us clear-cut, definite impressions of the tone quality of musical instruments. I shall only remark that the transformation of the spectrum inside the mouthpiece to the external spectrum has the general nature of a treble boost. In other words, whatever sounds may be generated inside the instrument, it is the higher components that are radiated into the room [*see bottom illustration on page 34*].

The very fact that the bell of an instrument leaks energy preferentially at high frequencies has two important con-

sequences. On the one hand the leakage enhances the relative amount of high-frequency energy that comes out of the horn; on the other it serves to reduce the height of the impedance peaks at high frequencies that lead to the weak generation of the high-frequency part of the spectrum inside the instrument. As a result measurements made outside the instrument in a room do not show nearly as much instructive detail about the dynamics of the entire system as measurements made inside the instrument do.

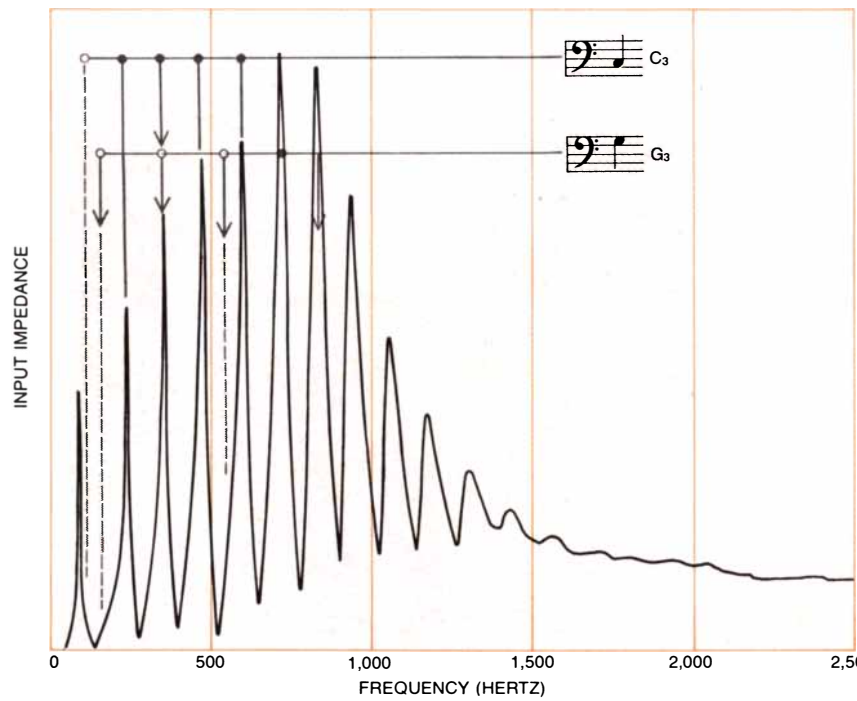
Let me conclude this discussion of the physics of brass instruments by indicating some of its implications for the musician and the instrument maker. As an illustration of the way physics can help the musician, I shall quote from an article of mine that appeared recently in the magazine *Selmer Bandwagon*. In this passage it was my intention to help French-horn players clarify and systematize their technique of placing one hand in the bell of the instrument to enhance certain frequencies.

"The player's hand in the bell is, acoustically speaking, a part of the bell. ... A properly placed hand provides... resonance peaks out to 1,500 hertz on an instrument that otherwise would lose all visible peaks at about 750 hertz [*see illustration on page 35*]. Suppose you meet a totally unfamiliar horn (perhaps during a museum visit when the curator opens the display cases) and you wish to find out quickly how well the instrument plays. Blow a mid-range note (for example concert F_3 in the bass clef) and, keeping your hand absolutely flat and straight, push it into the bell little by little until you feel a slight tingle in your fingertips. At this point (keeping the hand always perfectly straight) move the hand in and out a little until the horn sings as clearly as possible and the oscillation feels secure to your lips. Any listening bystander will agree with your final choice. Keep your hand in this slightly strained position and blow a tone an octave or a twelfth above the first one (say concert F_4 or C_5). Keeping your fingertips always in their original position, bend the palm of your hand so that its heel moves toward a position more familiar to the horn player. As you do so the tone will again fill out and get a ringing quality to it; also your lips will vibrate with a more solid feel. Your hand will now be in an excellent position for playing all notes on this horn, although an expert will be able to do even better after careful practice.

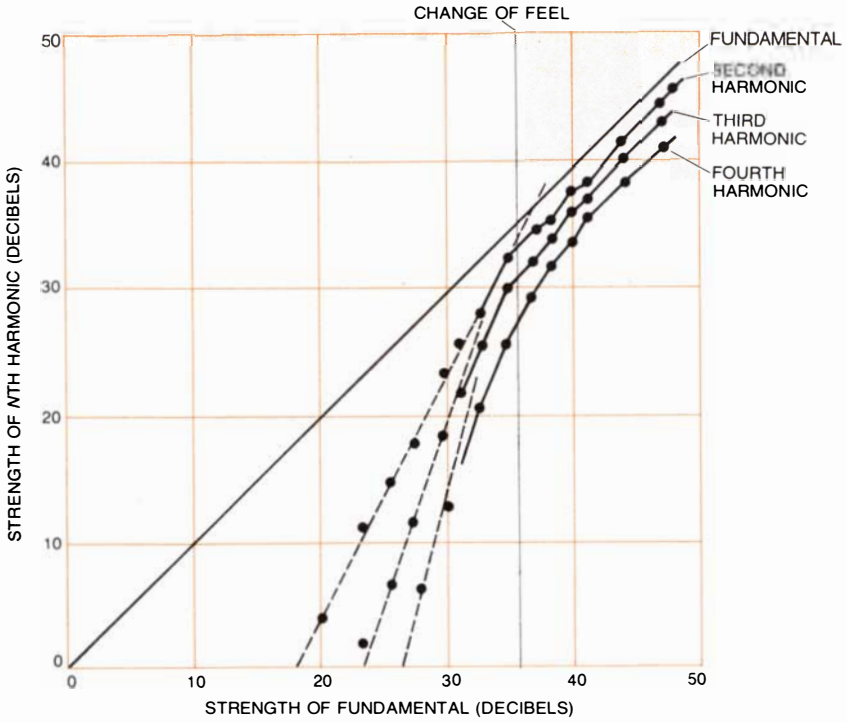
"Moving your straightened hand in and out while sounding the low F allows you to arrange to have an accurately located second helper for the tone. The unstopped horn works somewhat like a trumpet playing G_5 above the staff, while putting in the flattened hand serves to set up a regime that is analogous to the one which runs the trumpet's midstaff C_5. Bending the palm of one's hand while keeping the fingertips in place will leave the resonance peaks adjusted so far pretty much intact, but will make them taller (and hence more influential). This also gives rise to more peaks at the high-frequency end of things. The frequencies of these peaks move as the hand is bent more, so that once again the player has a means for tuning them for optimum cooperation with the other members of the regime. Trumpet players sometimes find it interesting and technically worthwhile to adapt the horn player's hand technique for their own purposes—especially for playing high passages on a piccolo trumpet."

It is only in the past few years that we have begun to have an understanding of the acoustics of mouthpieces. William Cardwell of Whittier, Calif., has provided a good theoretical basis for dealing with the relation of the mouthpiece dimensions to the tuning of the various resonance peaks. We in Cleveland, with the help of George McCracken of the King Musical Instrument Division of the Seeburg Corporation, have given attention to how the mouthpiece design controls the height of the impedance peaks. I quote again from the article for musicians to indicate the practical implications of mouthpiece acoustics.

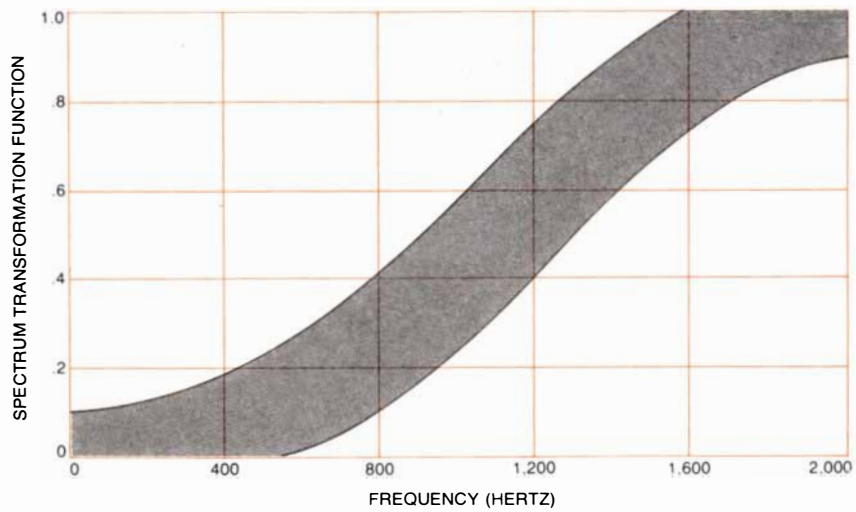
"Acoustical theory tells us that, first and foremost, a given instrument will require that the mouthpiece have a certain well-defined 'popping frequency' when its cup is slapped shut against the palm of the hand. In other words, the lowest natural frequency of the mouthpiece alone (with the cup closed) must be of the correct value. It is this requirement



UNUSUAL REGIMES OF OSCILLATION are associated with notes whose frequencies correspond to impedances that are close to minimum values. The note C_3 in the bass clef is known to musicians as the pedal tone. Its regime of oscillation is such that the second, third and fourth resonance peaks of the trumpet sustain an oscillation that lies at a frequency equal to the common difference between their own natural frequencies. Since there is actually a loss of energy at the fundamental playing frequency for this note rather than a gain, there is only a small amount of fundamental component in the sound, and even the small quantity present is converted to that frequency from the higher components by way of the nonlinearity in the flow-control characteristics of the player's lips. The situation for G_3 is even more unusual in that the second and fourth components of the tone are the chief source of oscillatory energy, whereas the fundamental component and the other odd harmonics contribute virtually nothing since the impedance is minimal at their frequencies.



TONE COLOR OF TRUMPET is related to the way harmonic frequencies make up an increasing fraction of the total sound emitted as the player blows louder. The strengths of the various harmonic components are plotted as a logarithmic scale (decibels) against the logarithm of the strength of the fundamental component. At low and medium playing levels each harmonic lies on a straight line whose slope is approximately equal to the serial number of the harmonic. As one plays pianissimo essentially no harmonics are present in the vibration as measured in the mouthpiece. For every doubling in strength of the fundamental component the second harmonic increases from an initial tiny value by a factor of four. Similarly, the third harmonic increases in strength by a factor of eight for each doubling in strength of the fundamental, and so on. This finding corresponds to a theory developed by Walter Worman at Case Western Reserve University. At the loudness where Worman's relation begins to break down the player senses a change in "feel" and listeners are aware of a change in sound. The data that are reflected in the curves were obtained with the help of Charles Schlueter, who now plays principal trumpet in the Minnesota Orchestra.



TRANSMISSION OF TRUMPET SOUND INTO ROOM is characterized by the "spectrum transformation function," which indicates what fraction of the acoustic energy at each frequency, as measured inside the mouthpiece, is emitted from the bell. Depending on the level of play and characteristics of the instrument, the energy emitted usually falls within the band plotted here. The curve has the qualitative nature of a "treble boost" because the bell leaks energy preferentially at high frequencies. Numbers on vertical scale are arbitrary.

that determines which of the peaks in the trumpet's response curve are the tallest. It also helps the peaks in this region to have the proper frequencies for good cooperation with the low-note regimes. The second most stringent requirement on the mouthpiece is that its total volume be correct (cup plus backbore). We must have this volume right in order to make the bottom two or three regimes of oscillation work properly."

So far I have discussed only the factors that contribute to favorable oscillation inside the horn and have said nothing about the tuning of instruments in the musician's sense: the relation between the pitches of the various tones that the instrument will generate. Fortunately the requirements for good tuning are almost identical with the requirements for favorable oscillation. It is for this reason that the traditional musical-instrument maker, focusing the major part of his attention on the tuning of the notes of the instrument, was able to develop instruments that would "speak" well and have good tone.

In more recent years, as our knowledge of acoustics has grown and the computer has become available, efforts have been made to design good brass instruments with the computer's help. Here the influence of loudspeaker acoustics has been great. Substantial efforts have been made to mathematically piece together a sequence of short loudspeaker-horn segments, each one intended locally to represent the shape of a workable brass instrument. This segmental approach to the problem has certain computational advantages. As we have seen, wherever the bore of a horn has a discontinuity of angle or of cross section there are anomalies in the standing-wave pattern. In spite of this fact it is always possible in principle to find suitable angles and cross sections that will place the impedance maxima of the horn with an accuracy that is acceptable by tuning standards. Although instruments built in this manner may play fairly well in tune, they can be quite disappointing in their musical value because of the neglect of the more subtle cooperative phenomena that ultimately distinguish between mediocrity and genuine excellence. Furthermore, the ability of an instrument to speak promptly and cleanly at the beginning of a tone is extremely sensitive to the presence of discontinuities, so that even though these discontinuities are arranged to offset one another in such a way as to give an excellent steady tone, it does not follow that the instrument starts well. The musician must of course

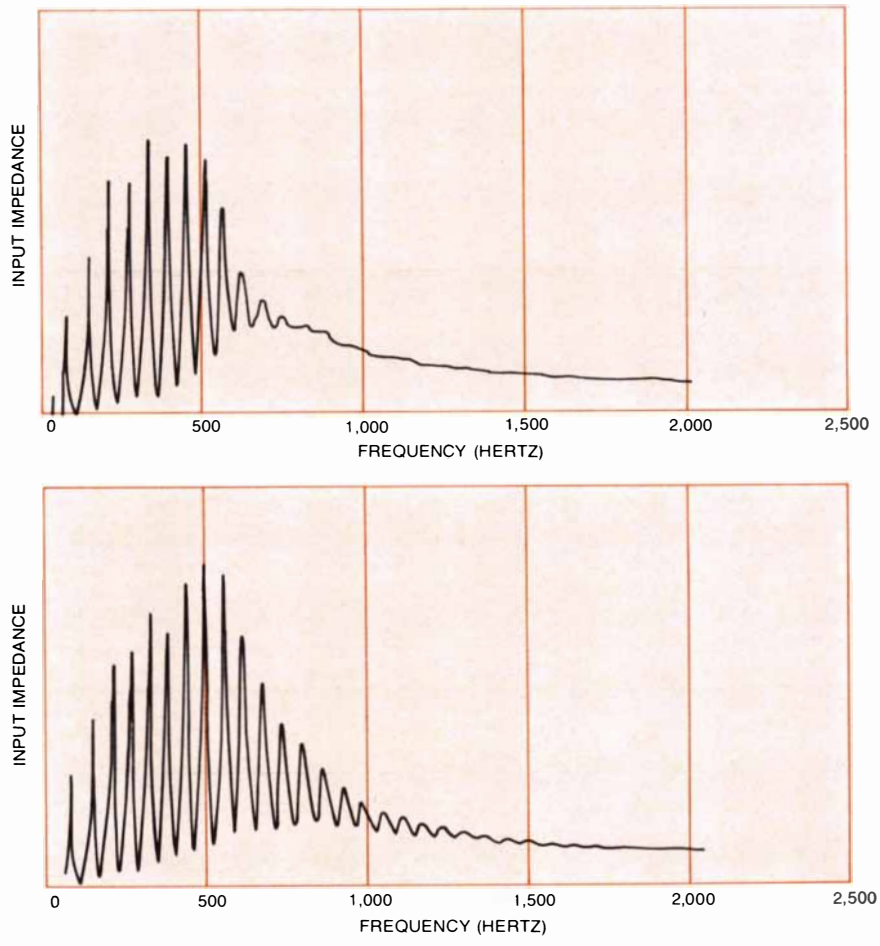
have a "clean attack" as well as a clear, steady tone.

The skillful instrument maker gradually acquires an almost instinctive feel for the subtleties of instruments, so that he can sometimes be astonishingly quick in the use of his empirical store of knowledge to find a correct solution to a tuning or response problem. Consider the problem that such a person must solve when he is asked to correct a trumpet that is faulty, with the sole error being the behavior of the tone corresponding to C_4. Let us suppose that the problem is caused by the fourth impedance peak (beginning from the peak of lowest frequency), which is somewhat high in its frequency. When the C_4 is played at a pianissimo level, the note will be in tune, but as the loudness increases somewhat the note will tend to run a little sharp as the second member of the regime (the mistuned fourth peak) begins to show its influence. The player will also notice that he can "lip" the tone up and down over a considerable range in pitch without appreciable change in tone color. He will complain that at this moderate dynamic level the tone "lacks center." If he plays louder, the influence of the still properly tuned third and fourth members of the regime becomes strong enough to partly overcome the defect of the second member. When this occurs, the player finds that the tone once again acquires what he calls a core, or center, at a certain playing level, which happens then to fall pretty well back in tune because all but one of the resonances in the regime agree on the desired playing pitch.

In the practical world of the instrument maker or designer one often meets instruments in which one or more notes are "bad" in this way. It has often proved quite difficult to correct such problems with only instinct and experience. Once one understands what is going on, however, it is often possible to bypass laboratory measurements and diagnose the errors with the help of carefully designed "player's experiments." One then uses acoustical perturbation theory to guide the alteration of the shape of the air column to give a desired correction. Such corrections are made by enlarging or reducing the cross section of the bore in one region or more of the air column. The problem is complicated by the need to preserve the locations of the correctly tuned resonance peaks while the faulty peak is being moved.

Whether one is a physicist, a musician or an instrument maker, one tries to make use of any tools at hand to provide an instrument that helps rather than hinders the creative effort of music making. At first it would seem that the computer is ideally suited to be one of these tools and that it could immediately be put to work designing the perfect instrument. As a practical matter one finds that although we have a reasonable understanding of the goals to be achieved, the complexity of the problem is such that it is very difficult to specify the problem for the computer in sufficient detail. I have found that it is much more efficient to start with an already existing good instrument developed by traditional methods and then apply the physical understanding and the technical facilities available to us today to guide the improvement of the instrument, whether it is for an individual player in a symphony orchestra or for the development of a prototype for large-scale production.

In all my work I have found it always important to keep in constant touch both with professional players and with instrument makers. They provide an inexhaustible supply of information about the properties of instruments. They also are a source of questions that have proved enormously fruitful in guiding my investigations. As the subject continues to develop it is becoming increasingly possible for the results of formal acoustical research to be translated into useful information for the player and the instrument maker.



PLACING HAND IN BELL OF FRENCH HORN is a well-known technique for extending the frequency range of the instrument. The curve at the top shows the input impedance response of a valveless prototype for the *B*-flat half of a standard French horn when measured without the player's hand in the bell. There are essentially no resonance peaks above 750 hertz. If the player tries to reach a note such as G_5 (783 hertz), all he gets is a wobbly scream because there is little or no feedback of acoustic energy from the bell of the instrument to stabilize a note of higher frequency. Notes in the octave below G_5 would also be weak and characterless for lack of a strong feedback. The curve at the bottom shows the additional resonance peaks produced when the musician points his flattened hand into the bell until he feels a slight tingling at his fingertips and then bends his palm slightly. The instrument now produces peaks well beyond a frequency of 1,000 hertz, making it possible for the musician to play the note G_5 quite dependably and even a few higher notes when he is pressed.

ADVANCED COMPOSITE MATERIALS

Materials in which the properties of one component enhance those of another are steadily evolving. As the cost of exotic components comes down, the advanced materials will show up in everyday uses

by Henry R. Clauser

Composite materials are among the oldest and the newest of structural materials. Men discovered early that when two or more materials are used together as one, the combination often performs better than each of the materials alone. Following this principle they combined clay and straw to make bricks and constructed bows of tendon, wood and silk bonded together with glue.

Then, with some notable exceptions, the further potentialities of composite structures remained virtually untapped for centuries while monolithic materials such as iron and copper served the major needs of an advancing technology. Even in more recent times, with the coming of reinforced concrete, linoleum, plasterboard and plywood panels, composite materials seemed somewhat ad hoc solutions to specific problems and thus out of the mainstream of materials development and technology.

During the 1930's and 1940's, however, lightweight honeycomb structures, machine parts made from compressed metal powders and plastics reinforced with glass fibers became commercial realities. These developments marked the beginning of the modern era of composite engineering materials. The use of composite materials has been steadily growing. The consumption of fiber-reinforced plastics, for example, has been increasing at the phenomenal rate of 25 percent annually. Nevertheless, the emergence of a distinct discipline and technology of composite materials is barely 10 years old. It has been estimated that 80 percent of all research and development on composites has been done since 1965, when the U.S. Air Force launched a major effort to make high-performance fiber composites a practical reality.

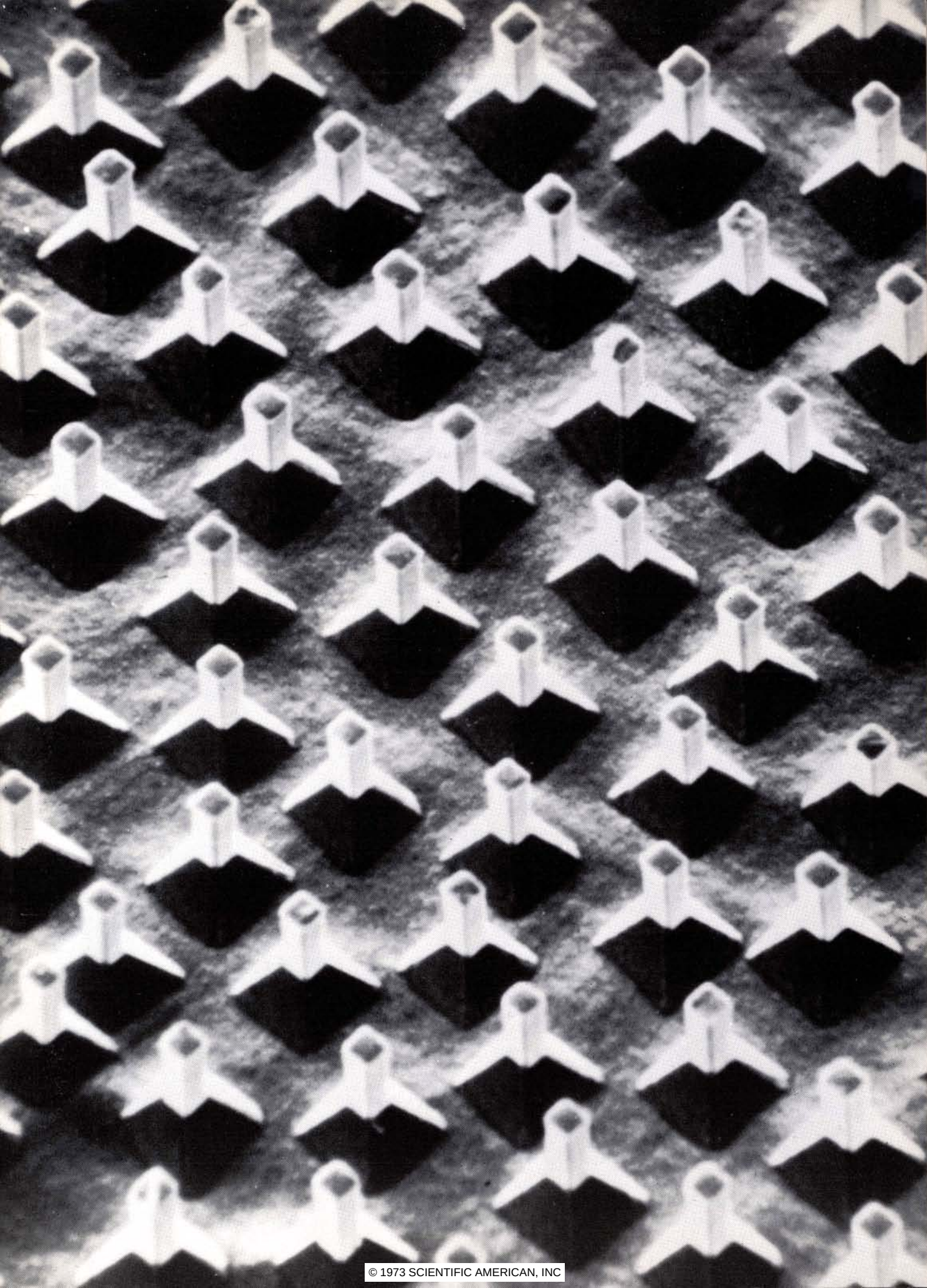
There are two major reasons for the current interest in composite materials. The first is simply the demand for materials that will outperform the traditional monolithic materials. The second, and the more important in the long run, is that composites offer engineers the opportunity to design totally new materials with the precise combination of properties needed for a specific task. Although the new composites are usually more costly, pound for pound, than conventional materials, they can be used more sparingly, because of their superior properties.

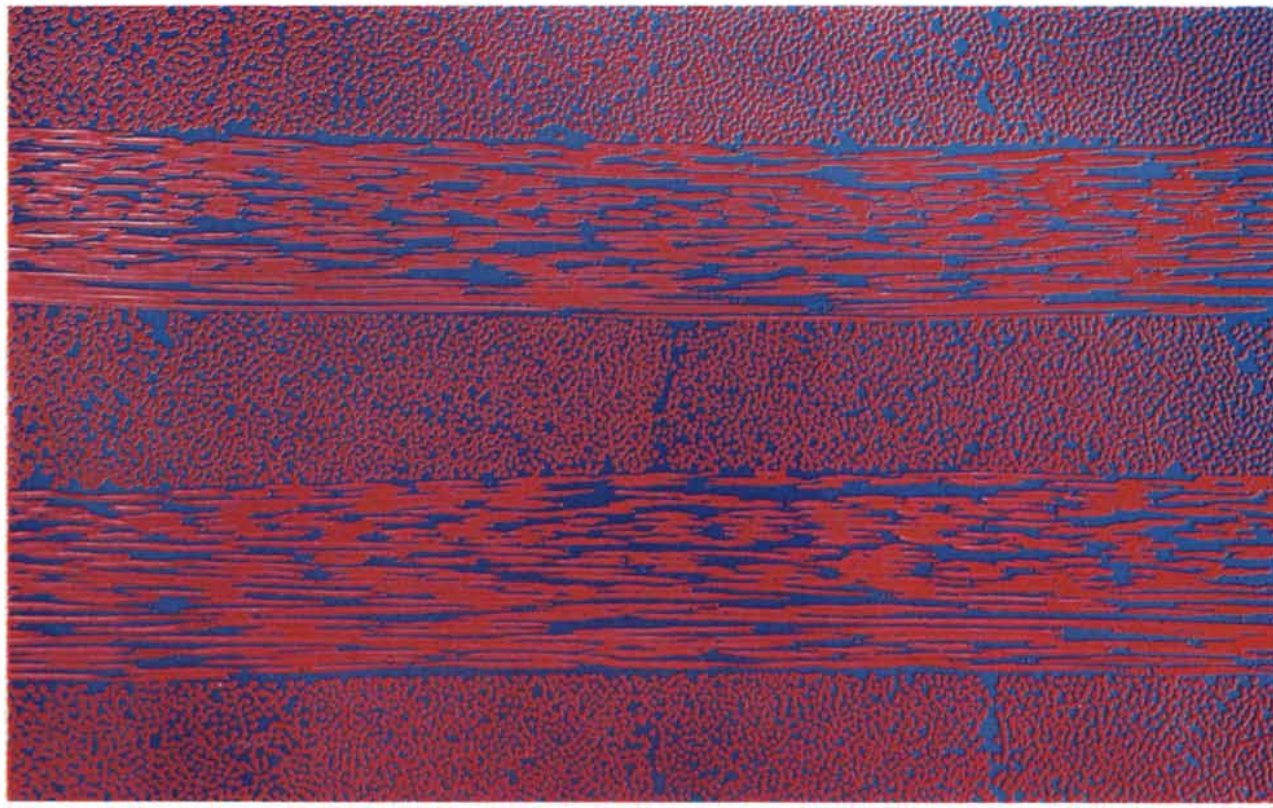
There is no all-inclusive or commonly accepted definition of composite materials. In the dictionary and in everyday usage the term "composite" refers to the concept of something made up of various parts or elements. If one tries to apply this general notion to the entire structural hierarchy of materials, one ends up encompassing most or all materials. To be meaningful, then, the definition of composites must be confined to the macrostructural level, where one deals with constituents such as glass fibers, metal particles and matrixes. On this basis a useful, although still imperfect, definition of composites is that they are a mixture of macroscopic phases that are composed of materials in a divided state that generally differ in form or in chemical composition or in both. Contrary to a widely held assumption, this definition does not require that a composite be composed of chemically different materials, although that is usually the case. The more important distinguishing characteristics of a composite are its internal geometry and the fact that its performance is the collective behavior of its constituents.

Another important distinction to be made is that the properties of composites may be either isotropic or anisotropic. In contrast, conventional monolithic materials are generally treated as being isotropic, that is, as having property values that are essentially constant from point to point. The properties of composites usually depend not only on the chemical composition of the constituents but also on their geometry and orientation. As a result the properties in planes through any point in a composite may be symmetrical or asymmetrical.

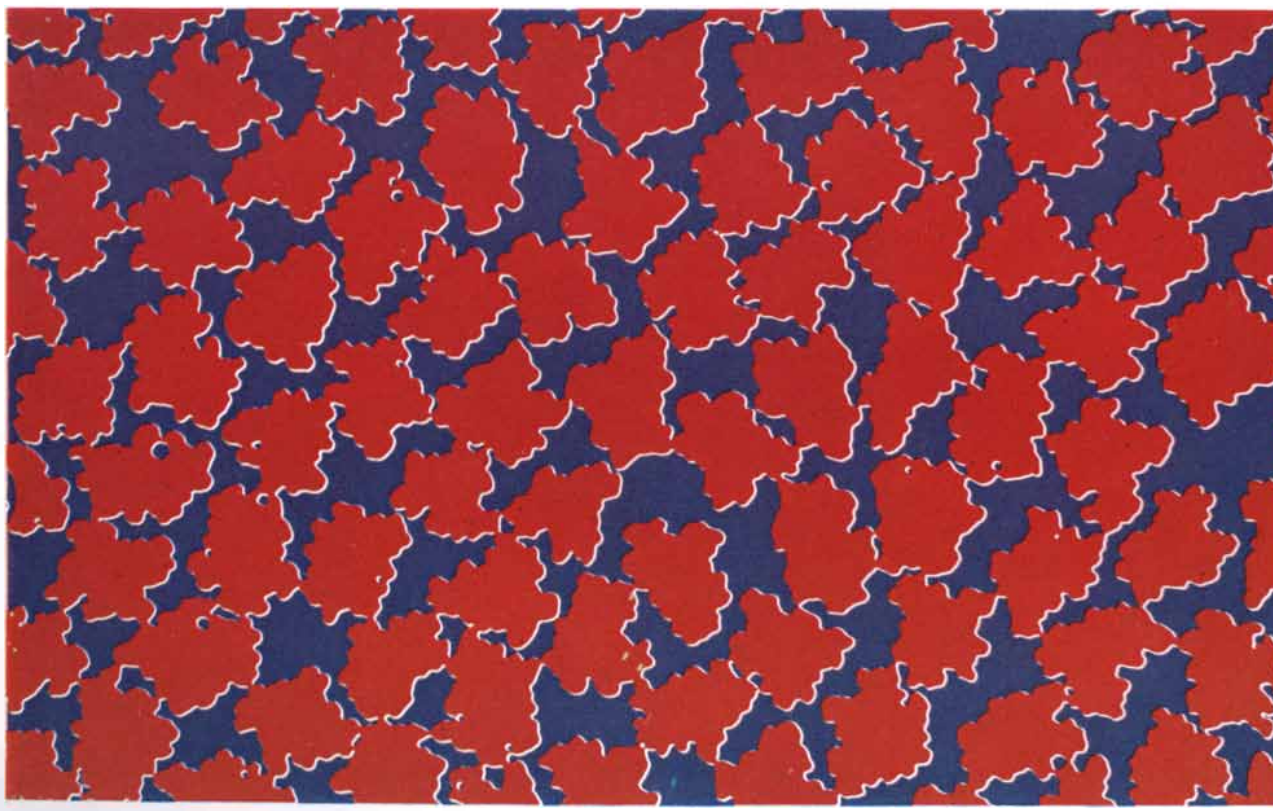
Major constituents used in composites are fibers, particles, laminae, flakes, fillers and matrixes. The matrix, which can be thought of as the "body" constituent, gives the composite its bulk form. The other five, which can be referred to as structural constituents, determine the character of the material's internal structure. The most familiar composites consist of a matrix in which is dispersed one or more types of structural constituent. For example, reinforced concrete consists of a stone-and-sand aggregate and steel fibers (rods) embedded in a matrix of Portland cement. Fiber-glass boat hulls are constructed of glass fibers supported by a plastic matrix. Not all composites have a matrix, however. Some are composed entirely of struc-

EXPERIMENTAL COMPOSITE developed by workers at the General Electric Company was produced by the unidirectional solidification of a eutectic alloy system: a system in which all the constituents crystallize at the same minimum temperature. The particular composite system shown consists of tantalum carbide fibers in a nickel-chromium matrix. The material was sectioned transversely and the matrix selectively etched away to a depth of several microns to reveal the fibers. The magnification in this scanning electron micrograph, made at the G.E. Research and Development Center, is some 2,000 diameters.





LAMINATED COMPOSITE in which the successive laminae, or layers, are themselves fiber composites oriented at right angles to one another is shown in this light micrograph made at a magnification of approximately 200 diameters. The material, which is named Thornel 50 multi-directional composite, is manufactured by the Union Carbide Corporation. It consists of graphite fibers (*red areas*) embedded in a matrix of epoxy resin (*blue areas*). The graphite fibers are produced by carbonization of rayon filaments.



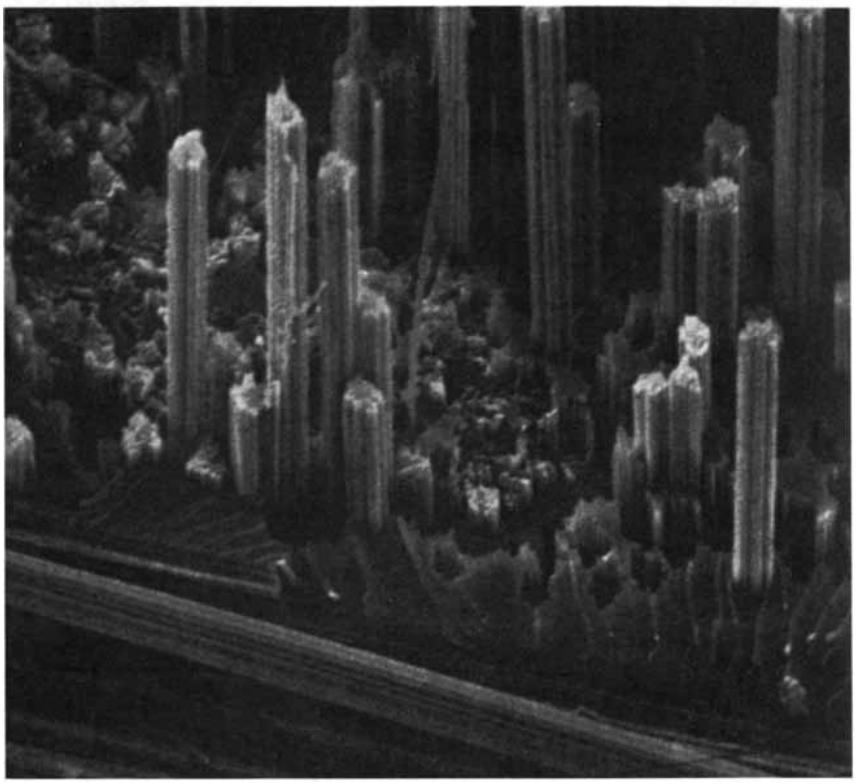
ENLARGED VIEW of a transverse layer of the Thornel 50 composite shown at the top of this page reveals the irregular cross sections of the individual fibers and the uniform distribution of the fibers in the matrix. Magnification is about 3,000 diameters.

tural constituents, usually of two or more chemical compositions. Sandwich materials (such as plasterboard) and metal laminates (such as the active element of a thermostat) are constructed entirely of laminae, or layers, which taken together give the composite its form.

Because the constituents are intermixed there are always regions of contiguity that can be considered analogous to grain boundaries in metals. The regions can be simply interfaces formed by the surfaces in contact or they can be interphases composed of a separate added phase. Examples of interphases are the coating, or coupling agent, on glass fibers in reinforced plastics and the adhesive that bonds together the layers of a laminate.

It is evident that composite structures owe their properties and behavior to two things: the properties of the constituents and the interactions among the constituents. Let us consider first how the properties of the constituents influence the performance of a composite. As in the case of monolithic materials, the chemical composition and microstructure of the constituents are critically important. They determine in large measure the range of properties a composite will have. Equally important, however, are a number of geometrical characteristics. The shape and size of the constituent elements and the relative amounts of the different constituent phases can be varied to provide specific properties or to achieve particular physical values. The way the structural constituents are distributed and arranged also helps to determine a composite's performance. Constituents can be dispersed in a uniform pattern to provide uniform physical properties or they can be distributed nonuniformly to achieve an intentional variation in properties from one region to another. If it is desired, the structural constituents can be oriented to provide directional properties. It is the almost endless opportunity to manipulate variables that makes composite structures so extremely versatile.

Going beyond the nature and properties of the individual constituents, one must also consider how the various constituents behave collectively as a result of their combination or interaction. This collective, or composite, performance has several aspects. Perhaps the simplest one follows from the mixture rule, which states that the quantitative value of a given property is the sum of the values of the constituent phases. Thus the weight of a block of concrete is the sum of the weights of the aggregate constituents (the sand and the stones) plus the weight of the set cement.



OBLIQUE VIEW of a flex-fractured sample of rayon-based Thornel 50 multidirectional composite shows the effects of the fracture on both the transverse and the longitudinal graphite fibers. Magnification in this scanning electron micrograph is about 1,000 diameters.

This summation rule applies not only to static properties such as weight but also to dynamic properties such as heat and electrical conductivity. For example, the approximate thermal or electrical conductivity of laminar composites can be obtained by simply adding the conductivities of the different laminae making up the composite.

Another kind of collective output results when each of two or more constituents contributes its particular property to the total performance. Many clad materials are composed of a low-strength surface layer with high corrosion resistance bonded to an underlayer of a stronger material that provides structural strength and stiffness. Many particulate composites have been developed to exploit in this way the properties of both plastics and metals. For example, powdered iron and iron oxides are blended with plastics to provide magnetic properties. Lead can be added to plastic structures in order to damp sound and vibration.

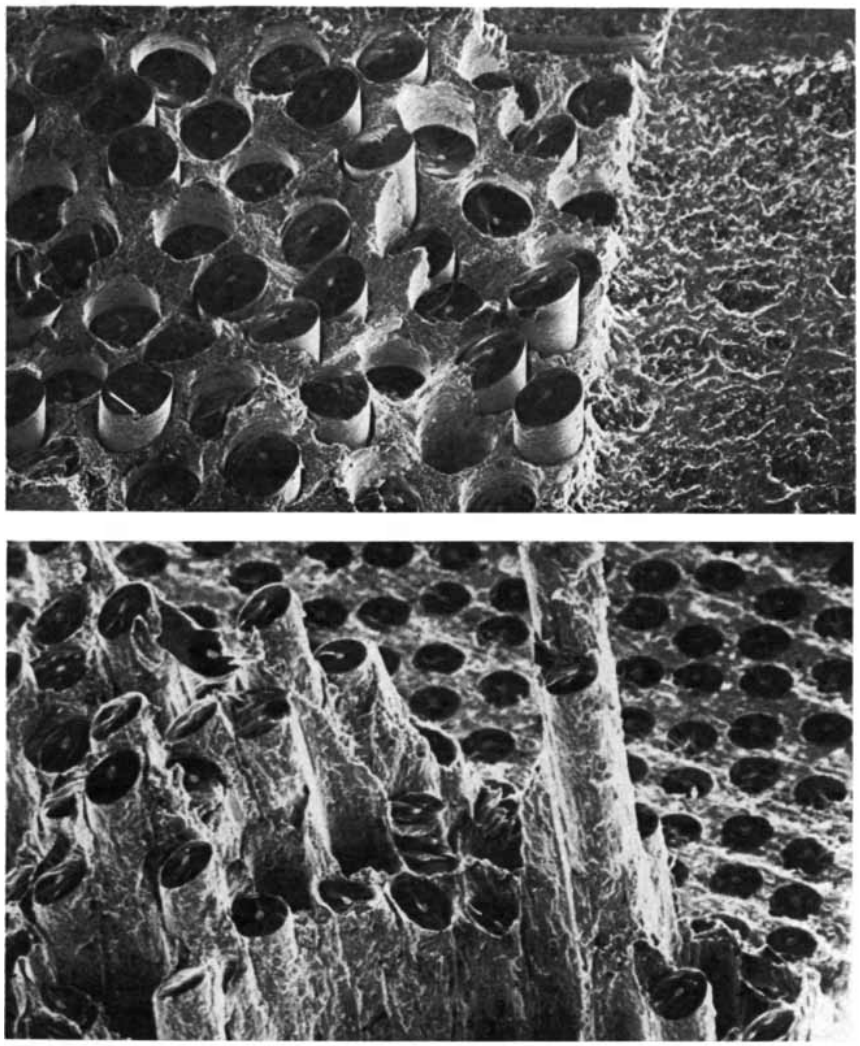
The third kind of composite output, perhaps the most important of the three, is an expression of the general systems principle that a whole can be something different from the sum of its parts. This means that the properties or functions of the constituents are not independent of one another but rather are interdependent. To gain an understanding of this principle let us see how it is used in the design of load-bearing structural composites. Many materials, including metals, ceramics, glass and graphite, have a theoretical tensile strength of up to several million pounds per square inch. Because of imperfections, however, only a small fraction of their strength is achieved when they are used in bulk form. These imperfections are of two major types. Internally they exist as microscopic dislocations in the crystal structure. When a material is stressed, the dislocations can move rapidly through it, thereby causing a shearing effect that leads to failure. Surface cracks are another source of fracture. Under an applied load the stresses tend to concentrate at tiny surface flaws, causing local overloading that can produce catastrophic cracking. Brittle materials, although they are inherently the strongest of materials, are most vulnerable to this type of failure. Because of lack of ductility even microscopic cracks on their surface will produce cracking at very low stress levels.

In the past 50 years much has been

learned about crystal dislocations and surface imperfections and how to counteract the fracture problems arising from them. In the late 1920's A. A. Griffith in England showed that threadlike forms of a material are many times stronger than the bulk form in which materials are generally used. This is particularly true of brittle materials. For example, Griffith found that glass fibers .0001 inch in diameter have a tensile strength of about 500,000 pounds per square inch compared with about 25,000 pounds for glass rods that are one millimeter thick. Since then it has been shown that even finer fibers of glass and other materials can exhibit tensile strength exceeding two million pounds per square inch. It is believed extremely small diameters inhibit the size of surface flaws and limit the directions a crack may follow.

Beyond the inherent increase in strength there is another advantage of using a material in a divided or fragmented state instead of in its bulk form. When either surface flaws or internal dislocations lead to cracking or plastic deformation, the movement is restricted to a relatively small region of the material. Moreover, the discontinuous nature of a collection of fragments tends to inhibit crack propagation and plastic flow.

By the suitable selection of constituents one can produce composite structures that resist the various fracture mechanisms. One of the oldest ways is to bond together flat layers of materials. The interfaces between the layers in the laminar composite function as crack-stoppers. Another approach is used in laminar sandwich construction. Borrowing from the *I*-beam principle, the sandwich facings, made of a high-strength material, correspond to the *I*-beam flange and carry the tensile and compressive stresses. A low-density core, such as a honeycomb structure or a foam plastic, acting like the *I*-beam web, carries the shear stresses and helps to prevent buckling of the facings. This combination of high-strength facings and a low-density core provides an extremely strong and rigid structure. For example, for the same rigidity under a bending load an aluminum-faced honeycomb sandwich weighs only a fourth as much as a solid aluminum sheet.



CONTRASTING FRACTURE SURFACES of different boron-reinforced composites near a machined notch are evaluated to reveal the causes and mechanisms of failure. In both cases shown the boron fibers are themselves composites, since they are composed of vapor-deposited boron on a tungsten-wire substrate. In the composite at top, which consists of boron fibers in a titanium matrix, the fiber-matrix bond appears to be weak and there is little evidence of matrix plasticity. In the composite at bottom, which consists of boron fibers in an aluminum matrix, the fiber-matrix bond is evidently stronger and the matrix exhibits plastic flow. Both micrographs were made at United Aircraft Research Laboratories.

The composite structures that yield by far the greatest improvement in mechanical properties are those that incorporate finely divided materials embedded in a matrix. The matrix, which is usually much more ductile than the structural constituent, can serve several functions. It can carry some or all of the applied stress. By flowing microscopically under load it can provide the medium for evenly distributing applied stresses to the structural constituents and for redistributing stresses from incipient cracks.

The structural constituents often have two functions. They are frequently the principal load-bearer. In addition they can serve as barriers to the movement of dislocations through the matrix. Particles and fibers are the main structural-constituent forms being used to exploit the advantages of materials in a finely divided state.

The geometrical characteristics of the structural constituents have a decisive influence on the properties of the composite. In the typical case, where the structural constituents are the load-bearers, the strength of the composite generally increases with the amount of constituent in a given volume, up to a point. In particulate composites the spacing of the particles, which of course involves both their diameter and their concentration, is closely related to strength properties. In fiber composites, where the fibers are the load-bearing constituent, mechanical properties in any one direction are usually proportional to the fraction of the volume of the material that is represented by the fibers oriented in that direction. As fiber orientation becomes more random, strength in any one direction decreases. Fiber diameter and length are two other critical factors.

In general, as diameter decreases, the strength properties of the composite tend to improve. Moreover, long, continuous fibers provide greater strength than short ones.

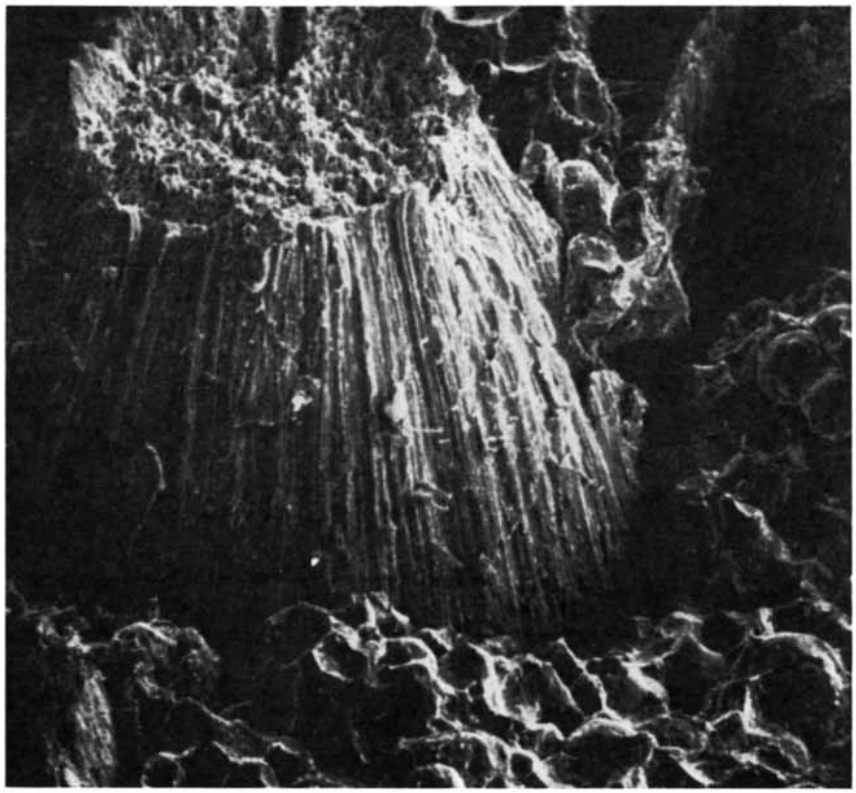
Although particles and fibers are dominant as constituents of composite materials, some interest has been shown in flakes for achieving high-strength structures. Theoretically flake composites have a higher modulus of elasticity, or stiffness, than fiber composites, and their properties are isotropic in the plane of the flakes. Moreover, when high density is desirable, flakes can be packed closer together than fibers or particles. Producing flakes in the desired shape and size, however, is difficult (except for metal flakes). Obtaining flatness and avoiding surface and edge flaws are other problems that are limiting the use of flakes for high-strength structural composites.

Let us now consider the most recent developments in particulate and fiber composites. There are two distinct classes of particulate composites, depending on particle size and the fraction of the volume represented by the particle phase. Composites in the one class, known as cermets, are composed of ceramic particles dispersed in a metal matrix. The particles are larger than one micron and their fraction of the volume ranges between 25 and 70 percent. In the other class, known as dispersion-strengthened or hardened composites, the particles are smaller than one micron and the volume fraction of the particle phase does not exceed 15 percent.

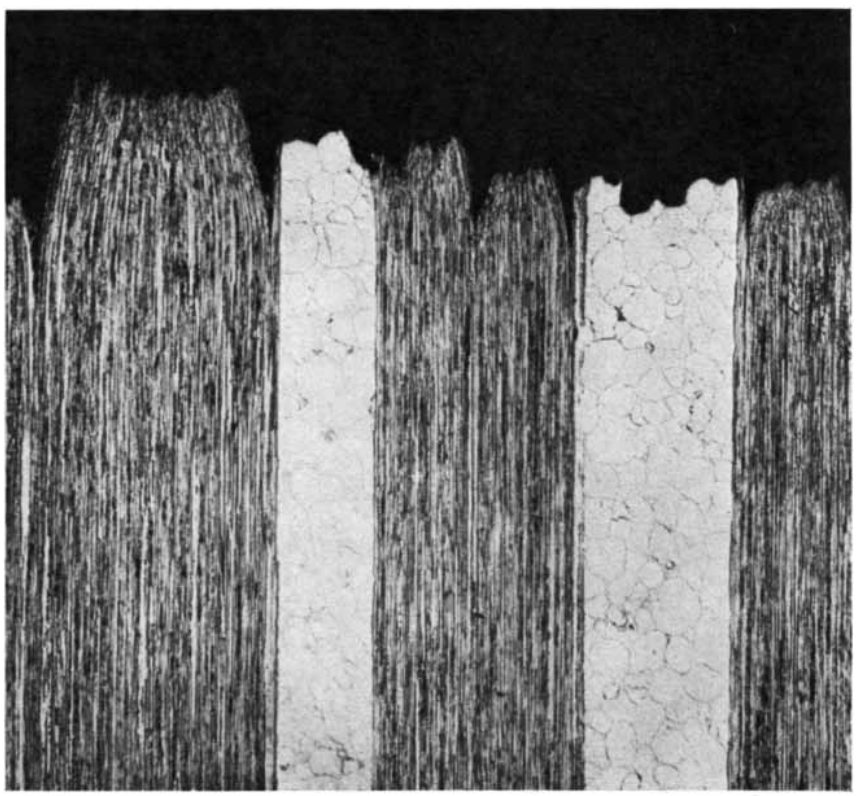
In dispersion-strengthened composites the matrix is the primary load-bearer. The fine, hard particles serve to block the movement of dislocations arising in the matrix. Thus for a given matrix material the principal factors affecting mechanical properties are the particle size, the interparticle spacing and the volume fraction of the particle phase. In general strength (particularly at high temperatures) improves as interparticle spacing decreases.

In particulate composites of the cermet type the matrix and the particles share the load-bearing function. Although the interaction of the two is complex and not fully understood, the net mechanical properties of cermets are a function of the interparticle spacing and the diameter of the particles and also of the ratio of the elastic properties of the matrix and the particles.

Cermets are produced by standard techniques of powder metallurgy in which the powder constituents are compacted by heat and pressure. Depending



IMPACT-FRACTURED COMPOSITE viewed obliquely in this scanning electron micrograph consists of tough tungsten fibers (*striated surfaces*) in a brittle nickel-superalloy matrix (*pitted surfaces*). The sample was tested at a temperature of 1,400 degrees Fahrenheit.



TENSION-FRACTURED COMPOSITE viewed longitudinally in this micrograph consists of the same two constituents: tungsten fibers embedded in a nickel-superalloy matrix. This sample was tested at 500 degrees F. Both of the micrographs that appear on this page were supplied by the Lewis Research Center of National Aeronautics and Space Administration.

on the materials involved, dispersion-strengthened composites are produced either by powder metallurgy or by techniques involving metals in liquid or colloidal form, followed by cold-working to achieve high strength levels.

Particulate composites are by no means new. Tungsten carbide, a cermet, has long been used as a cutting tool. Another cermet, tungsten thoria, has been used as a lamp filament for more than 30 years. Composites of dispersion-hardened aluminum and aluminum oxide, known as SAP (for sintered aluminum powder), are a more recent development. They are now finding wide use because they combine good resistance to oxidation and corrosion while retaining their strength at temperatures that would markedly weaken high-strength aluminum alloys. Another successful particulate composite of this type is a dispersion of thoria in a nickel matrix, called TD nickel. The composite is three to four times stronger than nickel at temperatures between 870 and 1,300 degrees Celsius. Other metals that have been dispersion-strengthened include copper, lead, zinc, titanium, iron and alloys of tungsten.

The particulate composites have important virtues and will continue to see further development, but it is the fiber composites that have evoked the greatest interest and development effort in the past five years. There are a number of reasons for this. The most important is that thin fibers seem to offer the most promising approach for exploiting the ultrahigh strengths inherent in structural materials. Moreover, materials that are normally stiff and brittle achieve their maximum flexibility when they are produced in the form of thin fibers. The fiber form also lends itself to a wide range of packing modes and geometrical patterns, so that the engineer can design structures to meet specific service requirements.

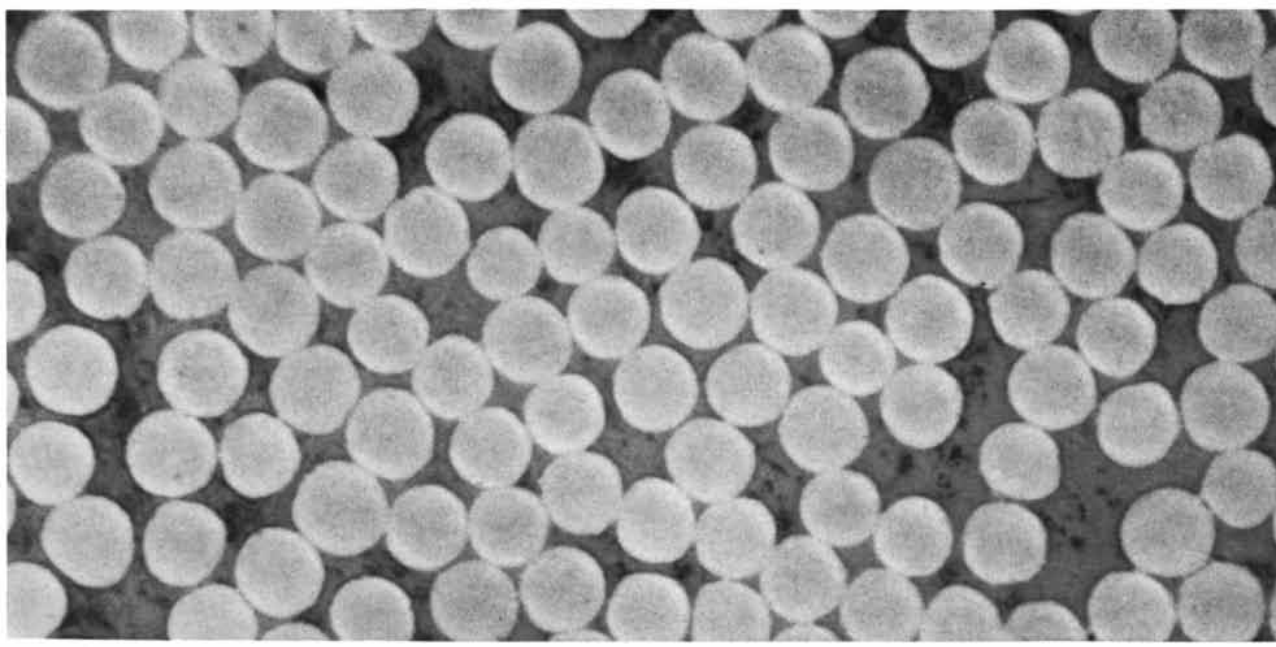
In nature perhaps the most common example of the reinforcing function of fibers is found in the structure of bamboo, a composite in which cellulose fibers are bound together in a matrix of lignin. Practically every type of man-made material is now being reinforced with fibers; fiber composites are found in such diverse objects as cafeteria trays, pleasure boats and high-performance rocket-motor cases. To anyone who has followed the steady escalation of pole-vaulting records the superiority of fiber-reinforced poles over bamboo poles is familiar.

The exceptional requirements of high-performance aircraft and spacecraft have provided a major stimulus for the development and exploitation of advanced fiber-reinforced structures. As a class the new composites combine ultrahigh strength with an ultrahigh modulus of elasticity. Since the new materials also have low density, they provide extremely high ratios of strength to weight (that is, specific strength) as well as high ratios of stiffness to weight (that is, specific modulus).

The new fiber composites extend the technology that began in the 1940's with glass fiber reinforced with plastics. Today the matrix in glass-reinforced composites may be either a thermoset plastic, such as polyester, phenolic or epoxy, or any of a number of thermoplastic resins, such as nylon, polyethylene or polystyrene. In the past few years new glasses have been developed with a modulus of around 15 million pounds per square inch as compared with 10.5 million pounds per square inch for the kind of glass used in most commercial glass-plastic composites. Even the higher figure, however, falls far short of the modulus of the new fibers used for reinforcing composites.

The two fibers that have been under the most intensive development for advanced composites are boron and graphite. They are being embedded chiefly in matrixes of epoxy resin and aluminum. Epoxy-resin matrixes have a high bond strength and toughness, but they cannot be used for service temperatures higher than 225 degrees C. Polyimide-resin matrixes extend the service temperature another 100 degrees.

The three most widely studied systems are boron-epoxy, boron-aluminum and graphite-epoxy. Their combination of strength, stiffness and light weight exceeds that of any commercially available monolithic material. Boron composites were introduced in the early 1960's, when the Air Force undertook an intensive development program. The fibers in boron composites are themselves composites. Produced by the vapor deposition of boron on a tungsten substrate, the



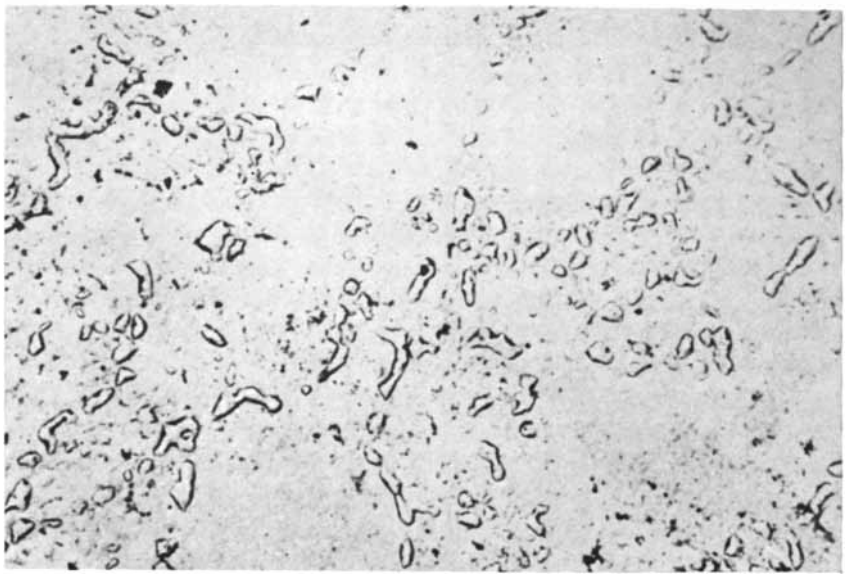
ORGANIC FIBER, known as PRD 49, has recently been introduced by E. I. du Pont de Nemours & Company for prospective use in high-strength composites. Chemically the fiber material is believed to belong to the nylon family. Magnification is 975 diameters.

fibers have a specific gravity of about 2.6 (that is, about 2.6 grams per cubic centimeter) and range from four to six mils in diameter. They have a tensile strength of about 500,000 pounds per square inch and a modulus of nearly 60 million pounds per square inch.

Graphite fibers, produced by the carbonization of rayon or acrylic fibers, average about a third of a mil in diameter and have a specific gravity of between 1.7 and 2. For use in composites they are generally made into a yarn consisting of some 10,000 fibers. Depending on the precursor fiber, graphite fibers have a tensile strength of between 200,000 and 500,000 pounds per square inch and a modulus of between 28 and 75 million pounds per square inch.

Although the properties of boron and graphite fibers are impressive, only a fraction of the fibers' tensile strength and modulus is attained when they are used as the structural constituent in a composite. Even so, the composites' specific strength and specific modulus far exceed those attainable in monolithic structural materials such as aluminum, steel and titanium. Unidirectional boron-aluminum composites and graphite-epoxy composites, with a fiber content of 55 percent, have a tensile strength ranging from 110,000 to more than 200,000 pounds per square inch and a specific strength (whose unit is inches) of between two million and 2.5 million. Their specific modulus (also in inches) lies between 350 million and 400 million. For purposes of comparison, the specific strength of high-strength aluminum is only about 750,000 inches and its specific modulus is about 100 million. It is clear that the exceptional specific strength and modulus of the boron and graphite composites can be translated into significant weight savings.

In addition to their superior mechanical properties fiber composites offer important advantages in the versatility with which they can be fabricated. One of the oldest methods of fabrication is filament winding, which involves continuously winding fibers on a mandrel of the desired shape. The most widely used method today, however, makes use of multilayer lay-up techniques. Each layer consists of composite tapes made up of graphite or boron fibers in a matrix of resin or aluminum foil. After the desired shape is formed, heat and pressure are applied to complete the manufacturing operation. It is evident that the end product is a laminar composite in which the laminae are fiber composites. The engineer can tailor the structure to meet the stress pattern of the application.



PARTICULATE COMPOSITE is represented by this electron micrograph of a sample of dispersion-strengthened lead, produced by the St. Joe Minerals Corporation. The material consists of a pure lead matrix strengthened by an extremely fine, uniform dispersion of lead oxide particles. The presence of the oxide particles inhibits the growth of crystal grains and results in a high-strength, fine-grained structure. Magnification is about 3,800 diameters.

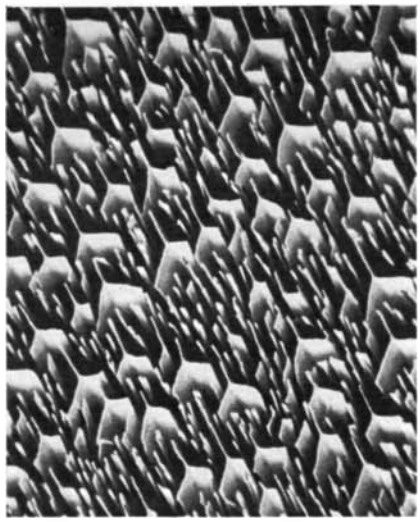
Another fabrication method makes use of conventional metal forms. It involves bonding a unidirectional composite tape to the surface of structural shapes or inserting composite rods into them. By this procedure the stiffness of aluminum shapes has been greatly increased and weight savings of between 25 and 60 percent have been achieved.

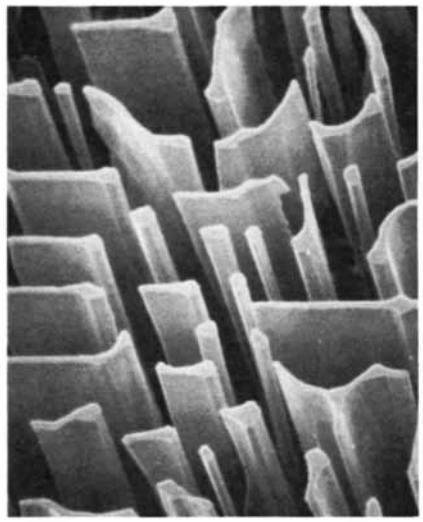
Three-dimensional weaving techniques also have been developed to produce isotropic composite structures. The three-dimensional material can be oriented in a cubic array or it can be fabricated in cylindrical form with fibers running in the radial, axial and circumferential directions. The cubic form of the material provides bulk material from which parts such as gears and bearings can be machined. The cylindrical form of the material can be directly woven into such shapes as hollow spheres, cones and ellipsoids.

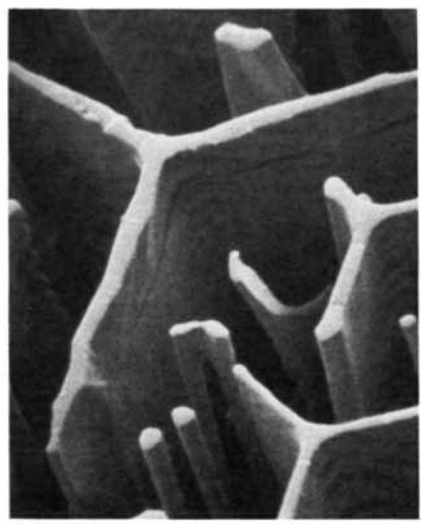
Fiber composites can now also be made by casting. They are produced by the unidirectional solidification of eutectic alloy systems: systems in which the constituents all crystallize at the same minimum temperature. Examples are alloys of aluminum–aluminum nickelide and columbium–dicolumbium carbide. One technique is quite similar to the standard foundry practice of casting into an open bottom mold resting on a cold plate that generates a vertical directional heat flow. The resulting casting contains fibers or whiskers aligned in the longitudinal direction. The second phase, with more than 30 percent of the total volume, is composed of flakes or platelets.

Although it is likely that boron and graphite will dominate the field of high-performance composites for some years to come, a variety of other fiber composites will probably find special applications. Silicon carbide deposited over boron, because of its superior compatibility with metals, shows promise for use in metal-matrix composites operating in the high-temperature region of from 815 to 980 degrees C. Boron nitride, another candidate for metal-matrix composites, combines a low precursor-fiber cost with low density and good high-temperature stability. Both its tensile strength (200,000 pounds per square inch) and its modulus (20 million pounds per square inch), however, are relatively low. Alumina, or sapphire, fibers drawn continuously from a melt are also being considered for use with metal and ceramic matrixes because of their high-temperature capability and chemical compatibility.

Recently E. I. du Pont de Nemours & Company has introduced a promising organic fiber known as PRD 49. Chemically it is believed to belong to the polyamide (nylon) family. Since its specific gravity is only 1.45 and its tensile strength is 525,000 pounds per square inch, the fiber has a higher specific strength than glass, boron or graphite. Its specific modulus is five times that of







IRREGULAR STRUCTURE of this experimental eutectic composite, developed by G.E., is evident in the sequence of successively enlarged scanning electron micrographs shown here. The fiber material, tantalum carbide, is the same as in the micrograph on page 37; the difference in the form of the fibers results from the use of a different matrix alloy, in this case pure cobalt.

glass and is as high as some graphite fibers.

Metal fibers with mechanical properties comparable to those of glass, graphite and boron will probably find increasing application in the future. For structural hardware, development is going forward on fibers of tungsten, nickel, beryllium and steel. There may be an even greater potential for metal fibers in applications such as electrically heated devices and reinforcement for tires and high-temperature conveyor belts.

It is generally recognized that the fibers that most closely approach theoretical values in strength are those prepared in the form of very fine single crystals. Called whiskers, these crystals range from three to 10 microns in diameter and have a ratio of length to diameter of between 50 and 10,000. Alumina whiskers have a tensile strength of up to three million pounds per square inch and a modulus of 62 million pounds per square inch. Other potential whisker materials are silicon carbide, silicon nitride, boron carbide and beryllia.

The large-scale use of whiskers in composite materials is still a good distance in the future. The nonuniformity of whiskers, both in size and mechanical properties, presents a major obstacle. In addition practical techniques for aligning whiskers and incorporating them into composites will have to be developed. Perhaps the biggest potential for whiskers is as an added constituent in continuous-fiber composites. Studies have shown that the whiskers greatly improve the load-transfer ability of the matrix.

As composite technology advances, hybrid composites composed of two or more different fiber types will be developed. Preliminary work in combining graphite and boron fibers in an epoxy matrix has shown that the hybrid may be as much as 30 percent superior in modulus and strength to conventional boron-epoxy composites.

What of the future? Although the great majority of current uses for advanced composites are in exotic technologies, the time is rapidly approaching when high-performance composites will appear in more mundane applications. Under serious consideration, for example, are the use of graphite composites for process equipment, for self-lubricating bearings and as strong commutators and slip rings for dynamos. In Britain graphite-fiber composites are already being used in textile machinery. In one application loom frames consisting of strips of composite bonded to aluminum have made it possible to increase the speed of the loom by 50 percent.

A major limitation to the wider use of advanced composites is the present high price of the constituent fibers. Boron fibers in tape form now cost between $150 and $280 per pound. Graphite yarn in resin-impregnated-tape form is generally available at prices of $75 to $90 per pound, depending on the volume purchased. As is normally the case, fiber prices can be expected to drop as demand increases. It is estimated, for example, that if the demand for boron fibers should reach 100,000 pounds a year (as some believe it will between 1975 and 1980), the price should drop to between $50 and $75 per pound. By 1975 the production of graphite fiber is expected to run as high as 350,000 pounds per year, at which time the price may fall to as low as $25 to $50 per pound.

When the price differentials between composites and bulk materials are being considered, a point that is sometimes overlooked is that composites often call for a smaller amount of material. In the case of one aircraft component, for example, an aluminum mill shape that weighed 4,000 pounds was machined and fabricated into a finished part weighing 460 pounds. When the same component was made of a boron composite, only 227 pounds of material was needed to produce a finished part weighing 207 pounds.

Bright as the future of composites seems to be, there is still a great deal of work to be done before they can achieve a large share of their full potential. As with all engineering materials, to make the best use of composites one must be able to characterize their properties quantitatively and be able to predict service behavior with accuracy and assurance. A fundamental difficulty is the bewildering complexity of the combinations of composite constituents. Even though one may understand the structural units individually, their combinations and interactions produce a vast array of physical values. Yet it is precisely the richness of such interactions that provides the opportunity to design composites to fit specific needs.

There is no quick, easy solution to the complexity problem. Computers, with their ability to handle lengthy calculations, will become increasingly helpful. Some feel, however, that it will be necessary to temper the strict analytical approach with the intuitions and insight of the artist. Whatever the approach, it is clear that we must develop principles for dealing with extremely complex materials involving a hierarchy of interrelated structures.

Microfilm is changing man's world.

Microfilm?

A child thinks that money put in the bank stays there physically as coins and bills. Physically, of course, your money in the bank exists as a tiny pattern of invisible magnetized spots on computer tape. No longer does a trustworthy old man make entries in the ledger with a goose quill. With a good pattern of spots you can buy a castle in Spain.

If you buy real estate in Spain or elsewhere, you probably want a piece of paper to wave at anybody who tries to chase you off your new property. The paper provides a feeling of security.

A piece of paper of another kind can give you a voice in *our* affairs and a share of *our* profits. For extra security about such securities, keep them in a safe deposit box.

You are a professor. Or perhaps you aspire to professorship. It would be well to be author of a book. Hard covers would be best. Failing that, show a list of your journal publications. Or at least one paper of which you can send out reprints, for goodness sake! Not all the recipients can—or even should—understand the content.

You are a student. You go to class. The prof lectures. That's his or her role. You take notes. That's your role. If the prof's viewpoint is not distinctive, who needs a prof? Cheaper to read a book. If everybody would read the same book, the book could be very cheap.

Ink on paper. When and where the civilized world gives it up, the reasons had better be good ones. True, goose quills don't last when more than one person in a hundred has a bank account. But the magnetic spots are *invisible*. How well do you know those cheery individuals behind the counter? For the assurance in recognizing by eye the work of the hand, bankers are still our biggest microfilm customers, and they were our first ones. A large firm of stockbrokers today uses millions of feet of microfilm a year. The securities industry still has its troubles with the archaic traffic in ornately printed paper. There may be a better way.

Microfilm needn't come in long rolls on reels. When intended more for communication than for security and proof, small sheets of it—microfiche—are more convenient. Convenience is crucial. The reading devices have to be comfortable, personal, ubiquitous. To be ubiquitous, they must cost little. The more people who need them and use them, the more the cost will drop. Our name is Kodak, and we make them and understand the principle.

Far more people are students and professors than formerly. Elitism has diminished, but the finding and fondling of new truths and new facets of old truths remain the academic obligation. Up goes the number of scholars obligated to present distinctive viewpoints. Down goes the size of each fragmented audience that appreciates the cogency, down too small to finance the ceremonial enshrinement with printer's ink and far too small to pay for color illustrations, however sadly missed.

Yet the academic enterprise will founder no more than has banking or the other enterprises founded on paper.

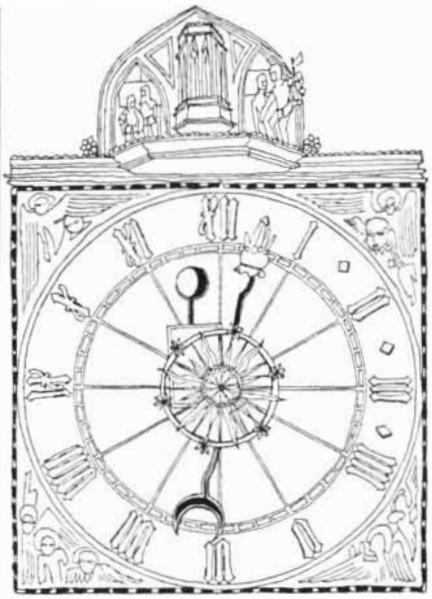
Right now arrangements can be made to send us a typed discourse with color slides or black-and-white illustrations and get back as many microfiche copies thereof as are needed to mail out to all believed capable of appreciating it. Or to be handed to students, who need assemble only to interact with the prof and each other, not for the archaic purpose of taking notes. And if the intent is to contribute to the archives of knowledge, and the content is to go through referees and editors before reaching us, that would be just fine. In fact, the detail in the "papers" in five of the principal American journals of chemical research is now published only in microfilm form.

If all this sets ideas spinning in a few heads, we can help pin down the best of those ideas. Dept. 55W, Kodak, Rochester, N.Y. 14650 can supply literature (still on paper) with titles like "Micropublishing Programs and Services," "Micropublishing for Professional Associations," "Kodak Ektalite Readers," "Basic Guidelines for the Selection of Input Material for Color Microfiche."

Kodak

A vessel of civilization

SCIENCE AND



Where the Action Is

The initial publication of the Office of Technology Assessment and Forecast, created last year in the Department of Commerce, offers a fascinating glimpse of the worldwide rivalry for technological supremacy as it is reflected in the activity of the U.S. Patent Office. The publication provides a sampling of technical areas of exceptionally rapid growth and areas where foreign inventors are particularly active. Thus one learns that the share of all U.S. patents awarded to inventors who live outside the U.S. has climbed steadily from 17 percent in 1961 to 29 percent in 1971 and is projected to average 31 percent over the next five years. Four countries emerge as the leading foreign recipients of U.S. patents: Germany, Britain, France and Japan. In 1963 they accounted respectively for 5, 4, 2 and 1 percent of the U.S. patents issued that year. Between 1963 and 1971 the patents awarded to Japanese inventors increased more than fivefold, so that today Japan is second only to Germany in winning U.S. patents and appears likely to overtake Germany before 1980.

The U.S. Patent Office has more than 11 million patents on file, divided into 300 broad classes, 8,500 "technological groupings" of roughly equal size and 85,000 subclasses. The Patent Office employs 1,200 patent examiners and last year collected $28 million in fees. Currently 80 percent of the patents are assigned to organizations, including governments; only 20 percent are retained by individuals. As measured by patents, the average growth rate for all technologies is 2 percent per year.

Although the U.S. Government is currently assigned only about 2.5 percent of the quarter-million patents issued annually, it holds between 11 and 30 percent of all the patents in six major areas: ammunition and explosive devices (30 percent), nuclear reactions and systems (22 percent), explosive and thermic compositions or charges (17 percent), radio-wave communications (17 percent), miscellaneous electron-tube systems (14 percent) and aeronautics (11 percent).

In a listing of active areas classified under "miscellaneous electron-tube systems" one finds, for example, an average growth rate of 13 percent for "electron-tube systems having means for distortion measuring and/or correction or means for unwanted wave-component elimination." Under "radio-wave communications" the report lists 22 active subclasses, of which 18 contain either the word "radar" or "antenna." These subclasses are growing twice as fast as the average of all technologies, and in them the U.S. Government has been receiving up to 38 percent of all the patents issued.

The report identifies a variety of fields in which other countries are making a heavy research investment and collectively outstripping the U.S. in obtaining patents. For example, in three broad areas of textile weaving (manipulation of warp, weft and fabric) foreign inventors have obtained 70 percent of all U.S. patents issued in the most recent three-year period (1969–1971). The most active countries are Switzerland and Britain. In the open-end spinning of yarn in the same period foreign inventors have received 72 percent of the patents, and it is projected that they will be obtaining more than 90 percent over the next five years. The leading countries are Czechoslovakia and Germany. In the production and utilization of crimped fibers foreign countries have been issued 40 percent of the recent patents and are projected to receive 80 percent; the leading country is Japan.

In the field of metal-shaping by rollers 59 percent of the patents have been going to foreign inventors; in devising tunneling methods foreign inventors have been winning 69 percent of the patents. In both areas Germany is the leading for-

THE CITIZEN

eign country. In the broad field of still photography the patents have been evenly divided between the U.S. and all foreign countries, but in the highly active subclass dealing with photoelectric controls for shutters foreign countries have been receiving 74 percent of the patents, with Germany leading Japan in the ratio of five to four. In the area of photographic emulsions foreign inventors have been receiving 57 percent of recent patents and are projected to receive 75 percent in the years ahead. The foreign leaders are Germany, Japan and Switzerland, with Japan already far ahead in the most recent year studied (1971).

Among foreign countries Germany and Japan also rank first and second in the development of polyamide resins, similar to nylon, in which foreign inventors have recently been awarded 48 percent of the patents and are projected to receive 65 percent. Germany has also been obtaining more patents than any other foreign country in the synthesis of heterocyclic alkaloid drugs (the Dutch and the Swiss are second and third) and in the design of electronically actuated fuel injectors, a field in which foreign workers have been receiving 61 percent of the patents and are projected to receive 75 percent; Japan has recently displaced France in second place.

France is the undisputed leader in the design of tracked air-cushion vehicles, in which foreign inventors have been receiving 49 percent of the patents and are projected to receive more than 90 percent. France is also second only to the U.S. in inventions related to the magnetohydrodynamic generation of power (MHD); foreign workers have been receiving 49 percent of the patents in this field and are expected to get 80 percent over the next five years.

Japan's strong push in electronics is reflected in two special fields analyzed in the report: the electronic control of musical instruments (such as electronic organs) and cartridge changer systems for magnetic recordings. Although foreign inventors have been receiving fewer than half of the patents in these two fields, the report predicts that the foreign share will grow to 85 percent and 75 percent respectively. In both fields Japan has far outdistanced all competitors except for

the U.S. The first report of the Office of Technology Assessment and Forecast is intended to be no more than a sample of the various types of analysis the new office can provide on request.

Durable Custom

In recent years it has seemed that societies of men were slowly but steadily abandoning the ancient custom of punishment by death. A worldwide survey undertaken by the United Nations indicates, however, that the death penalty is still prevalent and perhaps even becoming more so. Of 133 nations involved in the survey, only nine—Austria, Colombia, Costa Rica, the Dominican Republic, Ecuador, Finland, Iceland, Uruguay and Venezuela—describe themselves as prohibiting capital punishment. Another 16 nations have abolished the death penalty for all ordinary crimes but retain it for exceptional crimes, such as treason or killing the head of state. Those nations are Afghanistan, Argentina, Brazil, Denmark, Israel, Italy, Malta, Nepal, the Netherlands, New Zealand, Norway, Panama, Peru, Portugal, Sweden and the United Kingdom. In addition three nations—Belgium, Luxembourg and Nicaragua—still provide by law for a death penalty but have not in fact executed anyone for an ordinary crime for at least 40 years. Moreover, the study notes, "many of the retentionist countries use the death penalty so sparingly that, if one does not maintain the division between political and ordinary crimes, they may actually be executing fewer people than do certain 'abolitionist' countries."

An earlier UN report (in 1967) showed that the death penalty, which has traditionally been applied in cases of homicide, has been invoked increasingly for economic and political crimes. The current study finds that "this trend has continued." Indeed, "a few newer types of crime have begun to qualify in some countries for the death penalty." They include hijacking and trafficking in drugs.

According to the report, the death penalty "is regarded by a considerable number of governments as an efficient or at least an acceptable way of getting rid of certain types of problem—whatever the experts may have to say about the lack of deterrent effect of this penalty. Moreover, it seems clear that in most cases governments satisfy public opinion by using this sentence." The UN's position is that, in keeping with the assertions in its Universal Declaration of Human Rights that "everyone has the right to life" and "no one shall be subjected to torture or to cruel, inhumane or degrading treatment or punishment," the death penalty ought eventually to be abolished everywhere.

Fusion by Electron Beam

The possibility of using laser beams to ignite nuclear-fusion reactions appears to have stimulated much new work on a related approach: using for the same purpose an intense beam of relativistic electrons (electrons accelerated to speeds close to the speed of light). As in the laser-beam method, the electron beam would not require a "magnetic bottle" to contain the ionized heavy-hydrogen fuel of a fusion reactor. The electrons would be fired in short pulses, and the fuel would be fed into the beam with each pulse.

According to an article in *Physics Today,* work on beams of relativistic electrons for fusion purposes is in progress at the Naval Research Laboratories, Sandia Laboratories, Physics International, Cornell University, the Lawrence Livermore Laboratory, the Air Force Weapons Research Laboratory, North Carolina State University, Maxwell Laboratories and in the U.S.S.R. One problem is focusing the beam to get the energy density needed to achieve the fusion temperature of some 100 million degrees Kelvin. Another is shortening the pulses. The time needed to ignite fusion reactions is measured in nanoseconds, and existing beam machines produce pulses that are perhaps 10 times longer. Such long pulses would waste most of a fusion reactor's input power. Reviewing the various approaches to fusion power in *Physics Today,* Richard F. Post of the Lawrence Livermore Laboratory rates beams of relativistic electrons as a "dark horse."

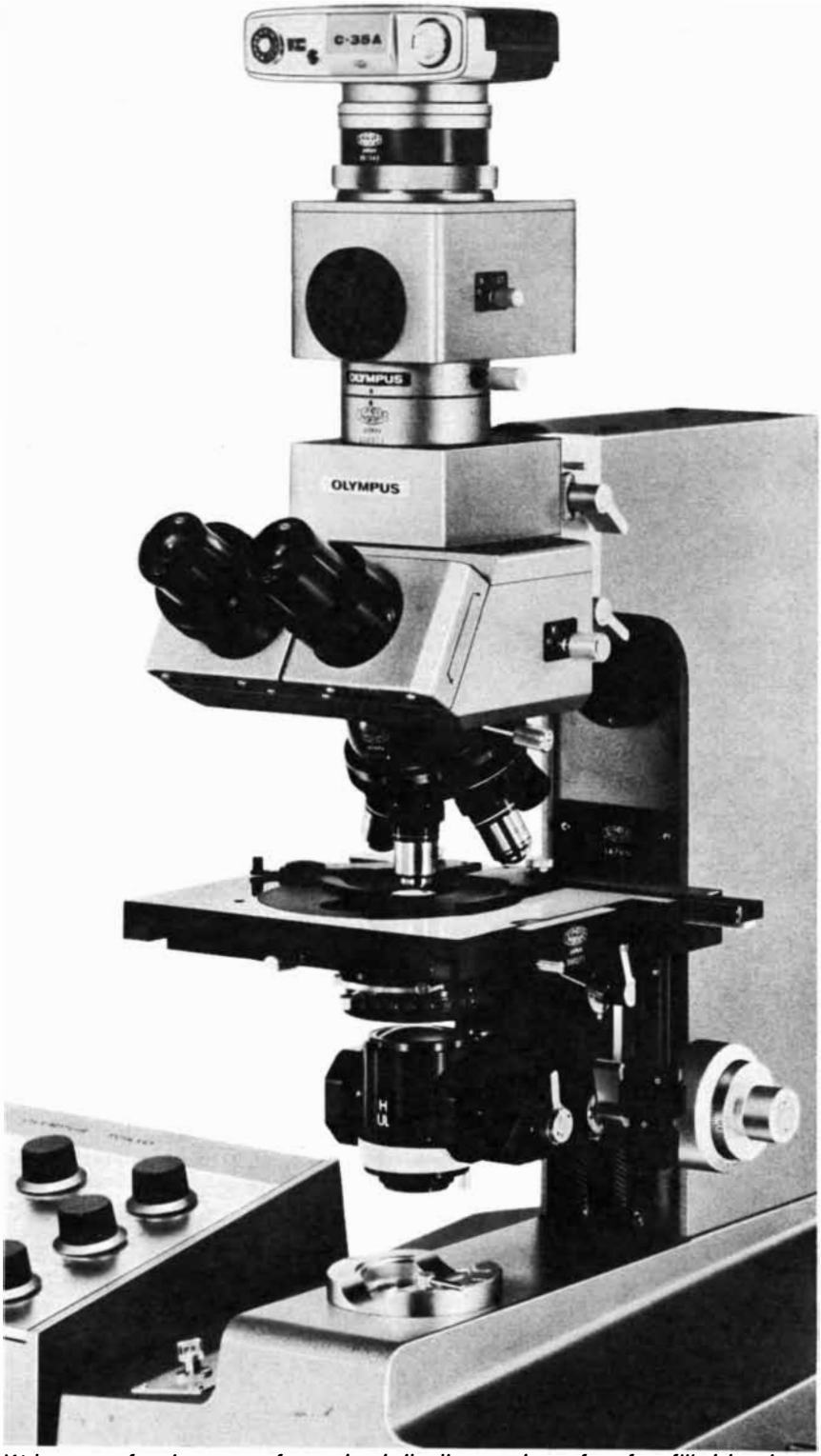
Plates, Plumes and Blobs

The plate-tectonic theory, which visualizes the outer shell of the earth as a mosaic of rigid plates in motion with respect to one another and to the earth's interior, is now generally accepted [see "Plate Tectonics," by John F. Dewey; SCIENTIFIC AMERICAN, May, 1972]. The plates are fed by magma welling up along their junctions at oceanic ridges; they are somehow propelled away from the ridges at the rate of a few centimeters a year; they converge with one another, generally at trenches along island arcs or continental margins, where one plate dives under another and is consumed. The theory is far from complete, however. Among the open questions are the sources and nature of the upwelling magma and the mechanism that propels the moving plates. In the past two years efforts to answer these questions have centered on the concept of thermal plumes: great "pipelines" of plastic rock, originating deep in the mantle below the earth's crust, that sometimes manifest themselves at the surface as oceanic islands. Some recent observations by Jean-Guy Schilling of the University of Rhode Island have made it possible to characterize these plumes more precisely.

The plume concept can be traced back to J. Tuzo Wilson of the University of Toronto, who proposed in the early 1960's that many oceanic islands were formed as volcanic eruptions, either over oceanic ridges or more localized deep sources, and were then carried away from their source by sea-floor spreading. In 1971 W. Jason Morgan of Princeton University proposed that such islands mark the breakthrough to the surface of small fractions of a few very large plumes; most of the volume of these plumes, he suggested, mushroomed out under the light crustal plates in a "thunderhead" pattern, providing the motive force with which to drive the plates. Subsequently Peter R. Vogt of the U.S. Naval Oceanographic Office reported evidence that the plumes do not flow continuously but rather seem to be turned on and off from time to time.

Until recently detailed evidence on the nature of plume material has been lacking. A group from the University of Rhode Island, working on the research vessel *Trident,* has recently dredged large amounts of rock from the sea floor along the Mid-Atlantic Ridge in the vicinity of several notable "hot spots" where major plumes are assumed to reach the surface. One of these is on and near Iceland, which sits astride the ridge at the junction of the Eurasian and North American plates. Detailed analysis of the chemistry of the rocks makes it clear that two very different sources of magma are represented, according to a report by Schilling in *Nature.* Along the ridge some 500 kilometers from Iceland the rock is from the asthenosphere, the upper-mantle layer immediately under the moving plates; it is lean in certain radioactive isotopes and in certain large ions. Over Iceland the rock is rich in these components; it comes from deeper in the mantle, the origin proposed by Morgan for his plumes. Between these two ex-

The VANOX. It's ahead of the time because we guess that's where you've got your sights.

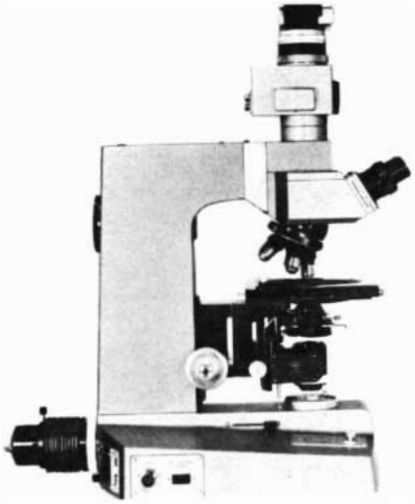


The Vanox, from Olympus, is not just another complex research microscope. It's a modular system based on an entirely new concept—a system flexible enough to adapt to your present and future research programs.

The Vanox quickly matches your application with the ability to adapt to numerous configuration. Its precision and versatility offer you nothing but convenience. Most of your requirements are met by the standard Vanox and there are numerous accessories for advanced research.

For example, it gives you opportunities to build a system around your own requirements for polarization, fluorescence, phase contrast, differential interference contrast, bright field and dark field microscopy in transmitted or reflected illumination; from achromat to planachromat. A photomicrographic system camera (PM-10), automatic, semi-automatic or manual, can be used with the Vanox. Find out how you can give your lab ultimate capabilities.

The Vanox system has so much, it takes a complete catalogue to tell all!



Write to us for the name of your local distributor, plus a free fact-filled brochure.

OLYMPUS TOKYO *Seeing beyond man's vision*

OLYMPUS

OLYMPUS OPTICAL CO., LTD.
43-2, Hatagaya 2-chome, Shibuya-ku, Tokyo, Japan
OLYMPUS OPTICAL CO., (EUROPA) GmbH.
2 Hamburg 1, Steindamm 105, Germany
OLYMPUS CORPORATION OF AMERICA
2 Nevada Drive, New Hyde Park, N.Y. 11040, U.S.A.

Photo of actual ranchland taken at Sangre de Cristo Ranches

Where the deer and the antelope play

Sangre de Cristo Ranches. Midway between Denver and Albuquerque. Where the Kit Carson Trail meets the Navajo Trail in beautiful Southern Colorado.

Big country. Wide open spaces. Majestic mountains surrounding green hillsides of the sleepy San Luis Valley. Rugged, yet serene. Crisp, clean air, scented with sagebrush and pine.

You can own land here. Leg-stretching, five-acre minimums. Land for ranches. For vacations. Retirement. For settling down and getting away from it all. For long term investment. To leave to your kids and for them to leave to theirs.

This is good land. Land you can see for yourself. Land you can stand on and dig into. It's a subdivision of the gigantic, 260 square mile Forbes Trinchera Ranch. Forbes Inc., publishers of Forbes Magazine and owners of Sangre de Cristo Ranches Inc. acquired the Trinchera Ranch some years ago. One of the oldest of the remaining big ranches in America, it ranks among the best known preserves for deer, elk, game birds and other wildlife.

Located near the intersection of U.S. Route 160 and Colorado Highway 159 just outside Fort Garland, Colorado, Sangre de Cristo Ranches are spread over thousands of rolling acres which rise to the towering mountain peaks of the Trinchera and Mount Blanca.

Since Forbes Inc. began offering this section of its big ranch for homesites, thousands of people have come from all over the nation to see this magnificent land, to purchase their own Sangre de Cristo Ranch.

Important guarantees backed by Forbes Magazine's distinguished reputation have contributed much to the Sangre de Cristo success story. Every investor is double protected: he has twelve months after his first payment during which he can—and is urged—to visit his ranch. And if it doesn't fully meet expectations, he can have all payments promptly refunded. Or if he simply changes his mind within sixty days of making his first payment, his money will be returned promptly.

You can own part of this breathtaking land for just $50 monthly. With no interest or carrying charges, the total cost of these five-acre ranches is just $5,000. Other Sangre de Cristo ranch sites are available on comparable terms at prices ranging from $3,500 to $7,500. This offer presents a wonderful opportunity for long term investment in the great southwest.

For complete details on these ranches, without obligation, simply fill in and mail the coupon below to Sangre de Cristo Ranches, Box 2000, New York, N.Y. 10011.

Sangre de Cristo Ranches Inc.

DIVISION OF FORBES MAGAZINE
Box 2000, Old Chelsea Station, New York, N.Y. 10011

Without obligation, please send me **FREE**, more information on the Sangre de Cristo Ranches.

Name_____

Address_____

City_____ County_____

State_____ Zip_____

A statement and offering statement has been filed with the Department 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 Department of State or any officer thereof or that the Department of State has in any way passed upon the merits of such offering. A copy of the offering statement is available, upon request, from the subdivider. NYA 72-833 Not available to residents of states where prohibited by law.

A 570 AD 6609 (b)

SANGRE DE CRISTO RANCHES INC., is a duly registered subdivision developer in and under the laws of the State of Colorado. This registration is not to be construed to imply state endorsement of the subdivision.

tremes there is a transitional zone, in which there is a mixture of rock from the two sources. Schilling's impression is that as the plates move apart the gap is filled by a passive flow of the depleted upper-mantle rock. The deep-mantle material of Iceland itself, on the other hand, has forced its way through the asthenosphere independently of plate separation and in excess of what is needed to fill the gap. That, indeed, is why the island is there.

Schilling has more recently begun to examine the thick piles of lava flows of the Faroe Islands and of the Greenland coast, which are now some 500 kilometers east and west of Iceland but which 50 to 60 million years ago were over the same source of upwelling material as Iceland is now. In most of the layers he finds levels of trace elements comparable to those on Iceland—evidence of a deep-mantle origin. The last few flows (upper strata), however, show an alternation between deep-mantle and asthenospheric material. These strata, Schilling thinks, may represent a weakening of the Iceland plume 50 million years ago, with consequent invasions of rock from the asthenosphere. The pattern suggests to Schilling that a plume is really a succession of "blobs," shaped like inverted teardrops, that rise through the mantle more like bubbles in a glass of champagne than like a steady jet.

The Brain Feeds Itself

Does what we eat influence the functioning of the brain? Until recently it was generally held that the brain had the ability to take from the blood the oxygen, glucose and whatever else it needed at rates that were more or less independent of the concentration of nutrients in the bloodstream. This view led to the belief that the daily metabolism of the brain is not influenced by changes in the blood following eating or fasting. Prolonged malnutrition, involving the absence of one essential nutrient or more, has, of course, been shown to cause major changes in the developing brain of the young, and most brain-nutrition studies have been concerned with such long-term effects. Now two investigators at the Massachusetts Institute of Technology, J. D. Fernstrom and R. J. Wurtman, have demonstrated that the brain also responds on a short-term basis to diet-induced changes in the blood.

In a series of experiments with rats, Fernstrom and Wurtman have shown that the rate at which brain cells synthesize the neurotransmitter serotonin from tryptophan, an amino acid, ultimately depends on the ratio of tryptophan to five other neutral amino acids in the blood (tyrosine, phenylalanine, leucine, isoleucine and valine). That ratio in turn depends on the composition of the food that has been recently ingested.

Rats given purified tryptophan showed a rapid increase in tryptophan levels in both the blood and the brain and an elevation of the serotonin level in the brain within an hour after eating. When the rats were fed a protein diet containing the competing amino acids as well as tryptophan, the level of tryptophan rose in the blood but neither tryptophan nor serotonin in the brain was elevated. It was this discovery that provided the clue that the ratio of tryptophan to the other amino acids in the bloodstream was the controlling factor. Tryptophan is far less abundant in food proteins than the other neutral amino acids are; hence consumption of protein introduces proportionately larger amounts of the competing amino acids. Paradoxically consumption of carbohydrate, which contains no tryptophan, increases the brain levels of tryptophan and serotonin because the insulin secreted in response to the carbohydrate causes the concentration of the competing amino acids to fall.

Essentially all the brain cells that contain serotonin are found in the raphe nuclei of the brain stem. These neurons are known to be involved in the control of sleep, food consumption, voluntary movements and the secretion of pituitary hormones. Fernstrom and Wurtman propose that the serotonergic nerve cells could function as "sensors" or transducers that convert information about peripheral metabolism into nerve signals. This suggests a kind of closed physiological circle in which food consumption affects brain biochemistry, which in turn affects food consumption.

Exhumed Craters

Why is the earth not covered with meteorite-impact craters as the moon is? The reason is presumably that most of the craters that were formed in the past have been obliterated by erosion and the dynamic processes of the earth's crust. There are nonetheless at least 14 good-sized craters on the Canadian Shield, some of them dating back nearly 600 million years. According to a hypothesis put forward by Brian Dent of Stanford University, the reason these craters have survived for so long is that they were made in ancient material that was subsequently covered by sediments, which were then planed off by glaciation in comparatively recent times to expose the craters anew.

Dent writes in *Geological Society of America Bulletin* that last year William A. White of the University of North Carolina proposed that the Canadian Shield had been covered with sediment for much of the 600-million-year period. If this was the case, it might account for certain puzzling features of the Canadian craters. For example, most of them seem too old and too small to have survived. Thirteen of the 14 appear to range in age from 150 million years to 550 million, and 10 of the 13 range in size from 1.9 kilometers in diameter to 26 kilometers. Yet the Gosse's Bluff crater in Australia, which is some 130 million years old and 20 kilometers in diameter, has virtually been erased by erosion. The survival of the Canadian craters is consistent with the notion that they were protected by a blanket of sediment until they were uncovered by glaciation.

Another puzzling feature of the craters is the distribution of their ages. A study of the craters of the moon presented in 1971 by R. B. Baldwin indicates that the rate of lunar cratering has decreased steadily over a period of four billion years, with the number of craters being cut in half every 250 million years. On this basis one would expect the oldest and youngest craters on the Canadian Shield to predominate; the more plentiful old craters would include large ones that would survive erosion and the less plentiful young craters would not have had enough time to be completely eroded. In fact the craters of intermediate age predominate, which is again consistent with the hypothesis that the craters had been covered with sediment. The older craters would have been erased before the sediment had been deposited, and the younger ones would have been removed by glaciation along with the sediment.

Dent points out that a sample of 14 craters does not provide a firm statistical basis for his hypothesis. Nonetheless, the hypothesis now forms a part of the effort to work out a rational scheme for the cratering of the earth and the place of meteorites, large and small, in the history of the solar system. Dent writes: "This analysis will be proved correct or not as further impact craters are discovered and dated, for it presents a consistent theory of crater preservation on the earth."

THE IMMUNE SYSTEM

This diffuse organ has the assignment of monitoring the identity of the body. Its basic constituents are lymphocytes and antibody molecules, which recognize both foreign molecules and one another

by Niels Kaj Jerne

The immune system is comparable in the complexity of its functions to the nervous system. Both systems are diffuse organs that are dispersed through most of the tissues of the body. In man the immune system weighs about two pounds. It consists of about a trillion (10^{12}) cells called lymphocytes and about 100 million trillion (10^{20}) molecules called antibodies that are produced and secreted by the lymphocytes. The special capability of the immune system is pattern recognition and its assignment is to patrol the body and guard its identity.

The cells and molecules of the immune system reach most tissues through the bloodstream, entering the tissues by penetrating the walls of the capillaries. After moving about they make their way to a return vascular system of their own, the lymphatic system [*see illustration on page 54*]. The tree of lymphatic vessels collects lymphocytes and antibodies, along with other cells and molecules and the interstitial fluid that bathes all the body's tissues, and pours its contents back into the bloodstream by joining the subclavian veins behind the collarbone. Lymphocytes are found in high concentrations in the lymph nodes, way stations along the lymphatic vessels, and at the sites where they are manufactured and processed: the bone marrow, the thymus and the spleen.

The immune system is subject to continuous decay and renewal. During the few moments it took you to read this far your body produced 10 million new lymphocytes and a million billion new antibody molecules. This might not be so astonishing if all these antibody molecules were identical. They are not. Millions of different molecules are required to cope with the task of pattern recognition, just as millions of different keys are required to fit millions of different locks.

The specific patterns that are recognized by antibody molecules are epitopes: patches on the surface of large molecules such as proteins, polysaccharides and nucleic acids. Molecules that display epitopes are called antigens. It is hardly possible to name a large molecule that is not an antigen. Let us consider protein molecules, which include enzymes, hormones, transport molecules such as hemoglobin and the great variety of molecules that are incorporated in cellular membranes or form the outer coat of viruses or bacteria.

Antigens and Antibodies

Each of the innumerable protein molecules is made up of polypeptide chains: linear strings of a few hundred amino acids chosen from a set of 20 amino acids. The number of amino acids in a large protein molecule is about equal to the number of letters in the column of text you are now reading, which is a linear string of letters chosen from an alphabet of 26 letters. Different protein molecules have different amino acid sequences just as different texts have different letter sequences. The string of letters in this column of text has been neatly "folded" into successive lines. The polypeptide chains of a protein molecule are also folded, although not so neatly. Their structure looks more like what you would obtain by haphazardly compressing a few yards of rope between your hands. There is nothing haphazard, however, about the folding of a particular polypeptide chain; the folding, and thus the ultimate conformation of the protein molecule, is precisely dictated by the amino acid sequence.

The parts of the folded chains that lie at the surface of a protein molecule make up its surface relief. An epitope (or "antigenic determinant") is a very small patch of this surface: about 10 amino acids may contribute to the pattern of the epitope. As Emanuel Margoliash of the Abbott Laboratories and Alfred Nisonoff of the University of Illinois College of Medicine showed for different molecules of cytochrome *c*, the replacement of just one amino acid by another in a polypeptide chain of a protein frequently leads to the display of a different epitope. The immune system recognizes that difference and is able to check on mutant cells that make mistakes in protein synthesis. Not only can an individual immune system recognize epitopes on any protein or other antigen produced by any of the millions of species of animals, plants and microorganisms but also it can distinguish "foreign" epitopes from epitopes that belong to the molecules of its own body. This recognition is a crucial event, since antibody molecules attach to the epitopes they recognize and thereby earmark the antigens (or the cells that carry them) for destruction or removal by other mechanisms available to the body.

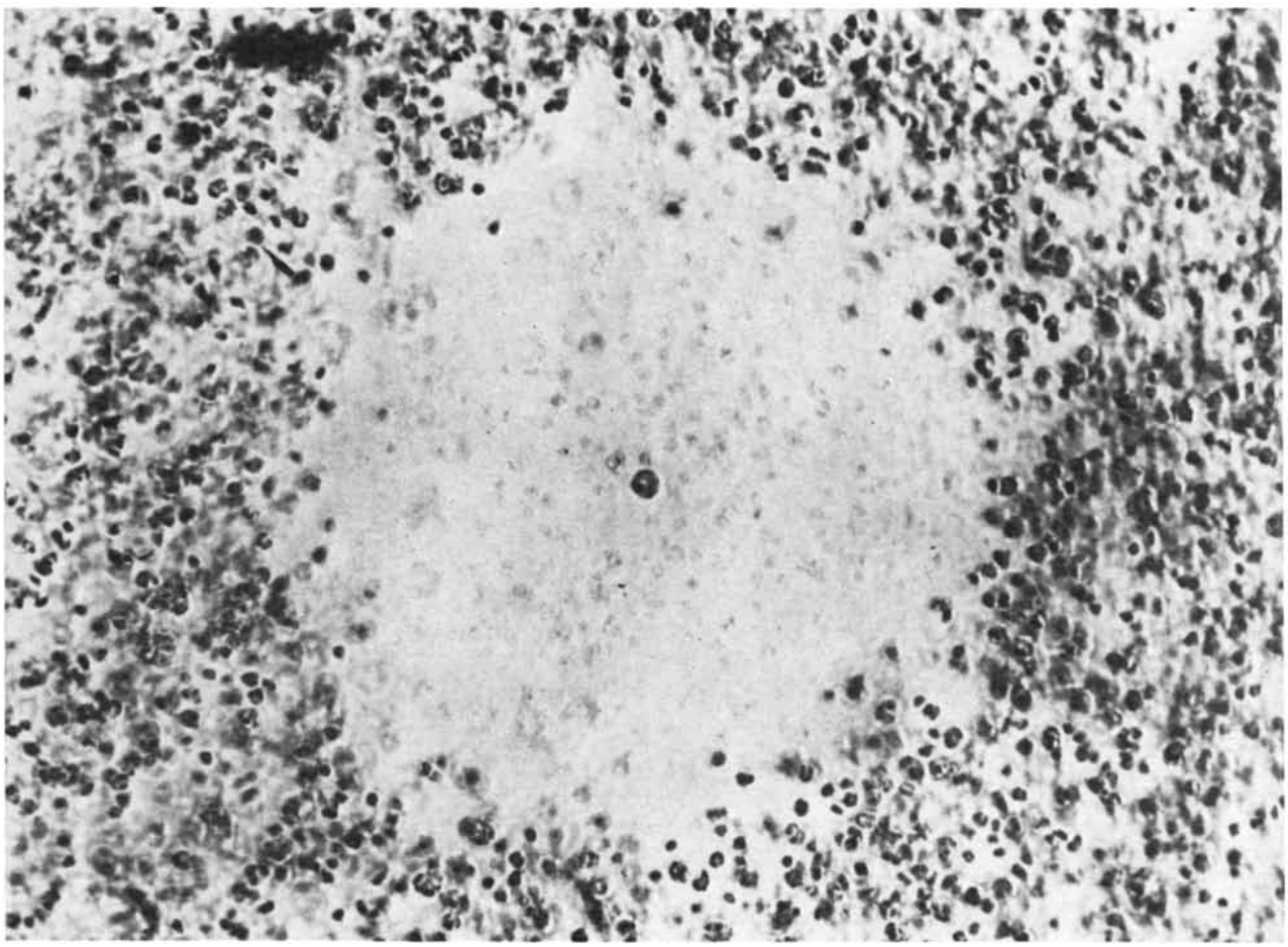
Epitopes are recognized by the combining sites of antibody molecules. An antibody is itself a protein molecule consisting of more than 20,000 atoms. It is made up of four polypeptide chains: two identical light chains and two identical heavy chains. A light chain consists of 214 amino acids and a heavy chain of about twice as many. Antibody molecules are alike except for the amino acids at about 50 "variable" positions among the first 110 positions, which constitute what is called the variable region of both the light and the heavy chains. At the tip of each variable region there is a concave combining site whose three-dimensional relief enables it to recognize a complementary epitope and make the antibody molecule stick to the molecule displaying that epitope. Whether a combining site will recognize one epitope or a dif-

ferent one depends on which amino acids are located at the variable positions. If at each of 50 positions of both chains there were an independent choice between just two amino acids, there would be 2^{100} (or 10^{30}) potentially different molecules! The situation is not that simple, however. The chains fall into subgroups, within each of which there are far fewer than 50 variable positions. On the other hand, at some of those variable positions, clustered in so-called hot spots, the choice is actually among more than two alternative amino acids. There is general agreement that the differences in amino acid sequence among antibody molecules derive from mutations that have occurred in the genes encoding antibody structure.

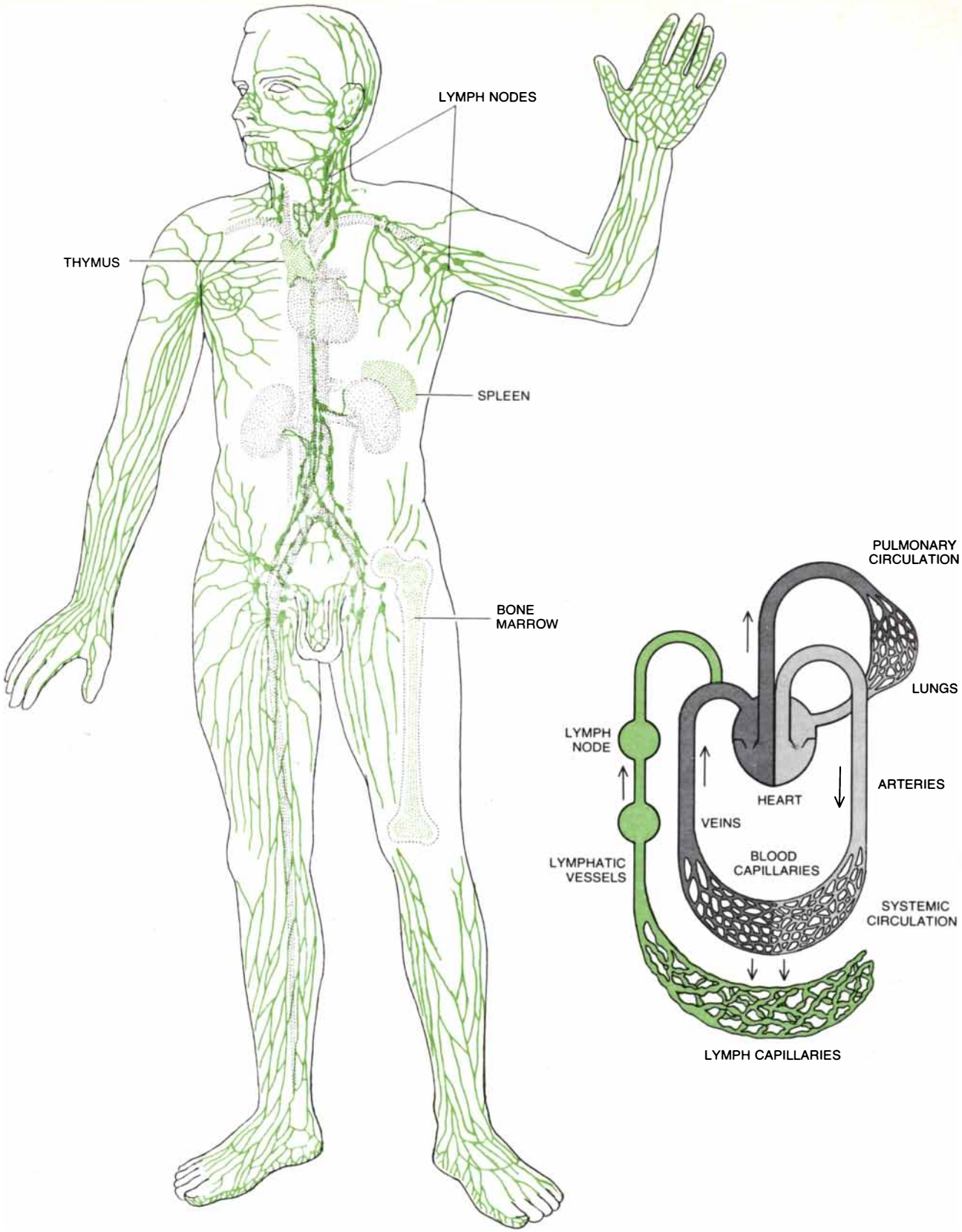
The Recognition Problem

Smallpox being the nasty disease it is, one might expect nature to have designed antibody molecules with combining sites that specifically recognize the epitopes on smallpox virus. Nature differs from technology in its approach to problem solving, however: it thinks nothing of wastefulness. (For example, rather than improving the chance that a spermatozoon will meet an egg cell, nature finds it easier to produce millions of spermatozoa.) Instead of designing antibody molecules to fit the smallpox virus and other noxious agents, it is easier to make millions of different antibody molecules, some of which may fit. By way of analogy, suppose someone makes gloves in 1,000 different sizes and shapes: he would have a sufficiently well-fitting glove for almost any hand. Now imagine that hands were a great deal more variable; for example, the length of the fingers on a hand might vary independently from one inch to six inches. By making, say, 10 million gloves of different shapes the manufacturer would nevertheless be able to fit practically any hand—at the expense of efficiency, to be sure, since most of the gloves might never find a customer to fit them. Now be more wasteful still: have a factory with machines capable of turning out gloves of a billion different shapes, but turn off 99 percent of the machines, so that the factory actually turns out a random collection of 10 million of the potential billion shapes. You would still be doing all right. So would your colleague running a similar factory. Although the two sets of gloves you and he would make would show only a 1 percent overlap, each set would serve its purpose well enough.

That is how some of us think the immune system solves its recognition problem. By a more or less random replacement of amino acids in the hot-spot positions of the variable regions of antibody polypeptide chains, a set of millions of antibody molecules is generated with



EFFECT OF ANTIBODY on an antigen is illustrated dramatically in a photomicrograph made by the author and Albert Nordin. The cell in the center is a plasma cell, an antibody-secreting lymphocyte of the immune system. It was embedded in a layer of culture medium along with millions of sheep red blood cells. The plasma cell is one that makes antibody against the sheep cells, specifically against epitopes, or small surface patches, on molecules in the surface membrane of the cells. Antibodies secreted by the plasma cell have destroyed the red blood cells in the area into which the antibodies have diffused; the radius of the area of destruction is about 1/5 millimeter. The technique illustrated here has become a standard one for measuring the immune response to an antigen.



IMMUNE SYSTEM consists of the lymphocytes and the antibody molecules they secrete. The cells and antibodies pervade most of the tissues, to which they are delivered by the bloodstream, but are concentrated in the tissues shown in color: the tree of lymphatic vessels and the lymph nodes stationed along them, the bone marrow (which is in the long bones, only one of which is illustrated), the thymus and the spleen. The lymphatic vessels collect the cells and antibodies from the tissue and return them to the bloodstream at the subclavian veins. Lymphocytes are manufactured in the bone marrow and multiply by cell division in the thymus, the spleen and the lymph nodes. The relation of the blood vessels and the lymphatic vessels is shown highly schematically in the illustration at right.

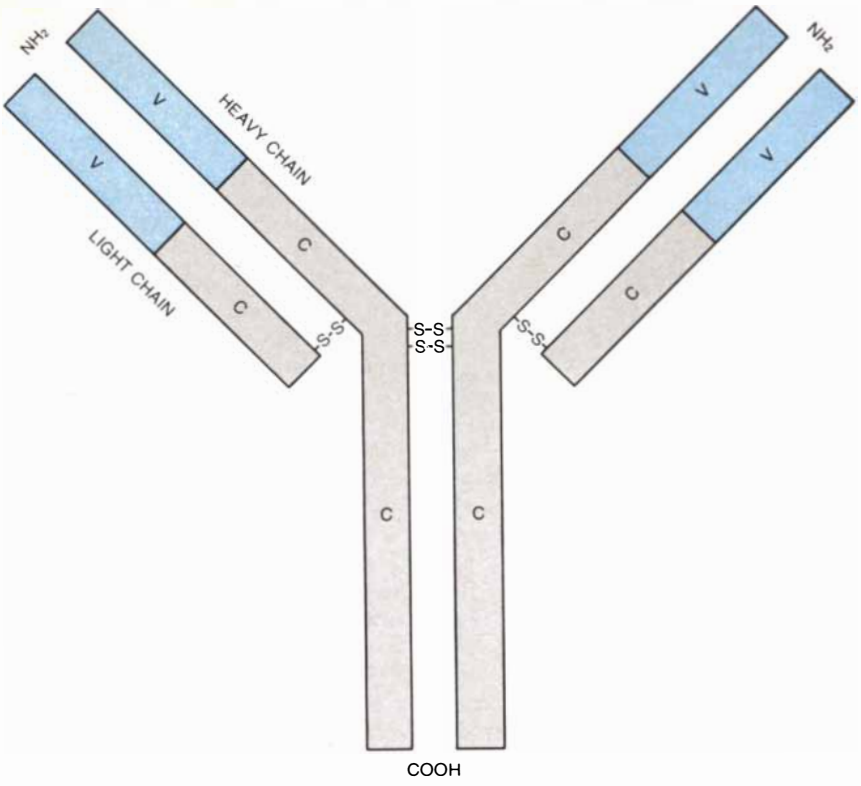
different combining sites that will fit practically any epitope well enough. As has been demonstrated by Jacques Oudin of the Pasteur Institute and by Andrew Kelus and Philipp G. H. Gell of the University of Birmingham for rabbits and by Brigitte A. Askonas, Allan Williamson, Brian Wright and Wolfgang Kreth of the National Institute for Medical Research in London for mice, individual animals make use of entirely different sets of antibodies capable of recognizing a given epitope.

There is one serious snag in all of this, to which I alluded above: one's immune system does not seem to recognize the epitopes on molecules and cells that are part of one's own body. This property, which Sir Macfarlane Burnet called the discrimination between self and not-self, is often referred to as self-tolerance. You might think that self-tolerance derived from nature's being wise enough to construct the genes coding for your antibodies in such a way as not to give rise to combining sites that would fit epitopes occurring in your own body. It can easily be shown, however, that this is not so. For example, your father's antibodies could recognize epitopes occurring in your mother; some antibody genes inherited from your father should therefore code for antibodies recognizing epitopes inherited from your mother.

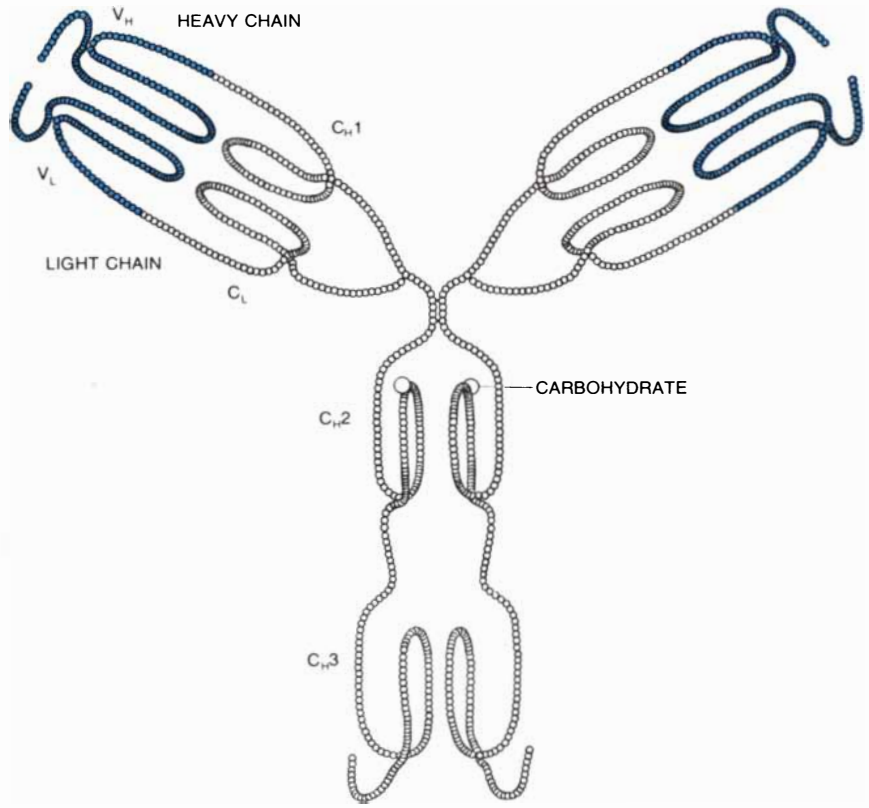
Self-tolerance, then, is not innate. It is something the immune system "learned" in embryonic life by either eliminating or "paralyzing" all lymphocytes that would produce self-recognizing antibodies. An original observation of this phenomenon by Ray D. Owen of the California Institute of Technology was generalized in a theoretical framework by Burnet and received experimental confirmation by P. B. Medawar in the 1950's, bringing Nobel prizes to Burnet and Medawar in 1960.

The Lymphocyte

Emil von Behring and Shibasaburo Kitazato discovered the existence of antibodies in Germany in 1890, but it was not until the 1960's that the structure of antibodies was determined, through the investigations initiated by R. R. Porter of the University of Oxford and Gerald M. Edelman of Rockefeller University [see "The Structure of Antibodies," by R. R. Porter, SCIENTIFIC AMERICAN, October, 1967, and "The Structure and Function of Antibodies," by Gerald M. Edelman, SCIENTIFIC AMERICAN, August, 1970]. The two men shared a Nobel prize last year for that work. Long before the structure of antibodies was



LINEAR STRUCTURE of an antibody molecule is shown schematically. The two heavy chains and two light ones are connected by disulfide bridges. Each chain has an amino end (NH_2) and a carboxyl end (COOH). Chains are divided into variable (*V*) regions (*color*), in which the amino acid sequence varies in different antibodies, and constant (*C*) regions.



FOLDING OF THE FOUR CHAINS is suggested in this drawing based on a bead model of the antibody molecule made by Gerald M. Edelman and his colleagues. Each bead represents an amino acid, of which there are more than 1,200. The variable regions are in color.

known, however, antibodies had been the subject of detailed studies. And yet it was not known that antibodies are produced by activated lymphocytes. Even 20 years ago lymphocytes were not thought to have anything to do with the immune system, something that seems odd now that they are known to constitute the immune system! It was only in the early 1960's that the involvement of lymphocytes was proved by James L. Gowans and Douglas McGregor of the University of Oxford.

Most lymphocytes (about 98 percent of them) do not actually secrete antibody. They are the "small" lymphocytes, spherical cells measuring about a hundredth of a millimeter in diameter, and they are said to be in a resting state. In order to secrete antibody a small lymphocyte must first become enlarged. In that state it can not only secrete antibody molecules but also divide and become two cells, which in turn can become four cells and so on. The offspring cells constitute the clone, or cell line, derived from one small lymphocyte.

As was originally postulated by Burnet in 1957, the antibody molecules produced by a lymphocyte and by the cells of its clone all have identical combining sites [see "The Mechanism of Immunity," by Sir Macfarlane Burnet; SCIENTIFIC AMERICAN, January, 1961]. G. J. V. Nossal, Burnet's successor as director of the Walter and Eliza Hall Institute of Medical Research in Melbourne, and his coworkers have accumulated much of the experimental evidence that now firmly supports this "single commitment" of the lymphocyte [see "How Cells Make Antibodies," by G. J. V. Nossal; SCIENTIFIC AMERICAN, December, 1964]. The cells of one lymphocyte clone are committed to the expression of two particular genes coding for particular variants of the variable regions of the light chain and the heavy chain. Already in its resting, nonsecreting state a small lymphocyte produces a relatively small number of its particular antibody molecules, which it displays on the surface of its outer membrane. These antibody molecules are the "receptors" of the cell. A small lymphocyte displays about 100,000 receptors with identical combining sites, which are waiting, so to speak, for an encounter with an epitope that fits them.

When such an epitope makes contact, the lymphocyte can either become "stimulated" (respond positively) or become "paralyzed" (respond negatively), which is to say it is no longer capable of being stimulated. Investigations in progress by David S. Rowe of the World Health Organization, working in Lausanne, and Benvenuto Pernis at our Basel Institute for Immunology suggest that the distinction between excitatory and inhibitory signals may reside in differences in the constant regions of the lymphocyte's receptor antibody molecules. Whether a lymphocyte will choose to respond positively or negatively can be shown to depend on several conditions: the concentration of the recognized epitopes, the degree to which those epitopes fit the combining sites of the receptors, the way the epitopes are presented (for example whether they are presented on molecules or on cell surfaces) and the presence or absence of other lymphocytes that can "help" or "suppress" a response. Much current experimentation aims at clarifying these complex matters.

A stimulated lymphocyte faces two tasks: it must produce antibody molecules for secretion and it must divide in order to expand into a clone of progeny cells representing its commitment. Progeny cells that go all out into the production and secretion of antibody molecules are called plasma cells. Each of them must transcribe its antibody genes into 20,000 messenger-RNA molecules that serve 200,000 ribosomes, enabling the cell to produce and secrete 2,000 identical antibody molecules per second. Other cells of the clone do not go that far; they revert to the resting state and represent the "memory" of the occurrence, ready to respond if the epitope should reappear. The immunological memory of what Stephen Fazekas de St. Groth of the University of Sydney, who is now working in our laboratory in Basel, has called "original antigenic sin" is remarkably persistent. People who are now 90 years old, for example, and had influenza in the 1890's still possess circulating antibodies to the epitopes of the influenza virus strains that were prevalent at that time.

If a lymphocyte that recognizes an epitope does not become stimulated, it may become paralyzed. Paralysis can occur when a lymphocyte is confronted by very high concentrations of epitope; this is called high-zone tolerance. David W. Dresser and N. Avrion Mitchison, who were working at the National Institute for Medical Research, have shown that paralysis can also result from the continuous presence of extremely small epitope concentrations, below the threshold required for stimulation; this is called low-zone tolerance. We need more knowledge of the mechanisms leading to paralysis, not only in order to understand how the immune system learns to tolerate self-epitopes but also to be able to induce the system to tolerate organ transplants.

ANTIBODY MOLECULES are visible in an electron microscope when they are linked to antigens and one another in antigen-antibody complexes. In this micrograph, made by N. M. Green and the late Robin Valentine of the National Institute for Medical Research in London, rabbit antibodies are enlarged 500,000 diameters. The antigen is a short polypeptide chain with a dinitrophenyl group at each end; the antibodies are from a rabbit that was immunized against dinitrophenyl epitopes. The antigens (too small to be visible) link antibodies to form polygonal complexes whose geometry derives from antibody structure.

Germ-Line and Soma Theories

The enormous diversity of antibodies raises the question of the origin of the genes that code for the variable regions of antibody molecules. Essentially two answers have been proposed to this question. They are the germ-line theory and the somatic theory. The argument of the germ-line theory is straightforward: All the cells of the body, including lym-

phocytes, have the same set of genes, namely those in the fertilized egg from which the individual arose. Therefore genes for any antibody that an individual can make must already have been present in the fertilized egg cell. They are all transmitted to the individual's children through the germ-cell line: egg and spermatozoa and their precursors.

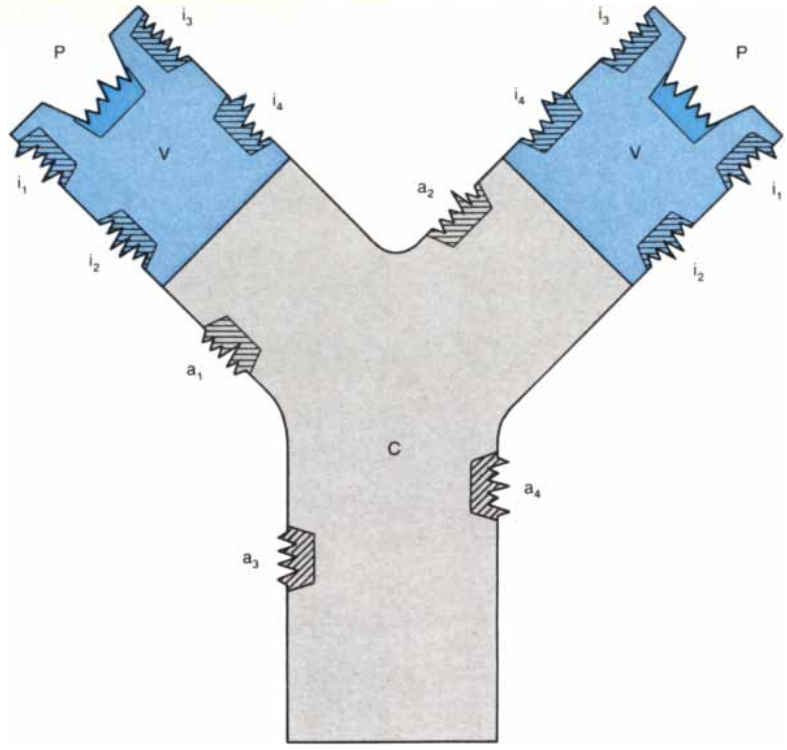
The somatic theory does not accept this approach. It is argued that the immune system needs millions of different antibodies for epitope recognition. Individual mice of an inbred strain, all having the same germ-line genes, have been shown to make use of entirely different sets of antibody molecules. The germ-line theory implies that the set of all these sets is represented in the genes of every single mouse of that strain. In that case, however, many of the genes would seem to have no survival value for the mouse, so that such a large number of genes cannot arise or be maintained in Darwinian evolution. Most antibody genes must therefore have arisen in the course of the somatic development of the individual by modification of a smaller number of germ-line genes. That is the point of departure for several variants of the somatic theory.

I have proposed that an inherited set of germ-line genes code for antibodies against certain self-epitopes. The clones of cells expressing these genes become suppressed except for mutant cells that, by an amino acid replacement, display new combining sites on their antibody receptor molecules. These mutant cells represent the enormous repertory of antibodies that recognize foreign epitopes. An organ that could breed such mutant cells is the thymus gland. More than 10^{10} new lymphocytes arise in the thymus every day; the vast majority of these cells are killed in the thymus or immediately after they leave it.

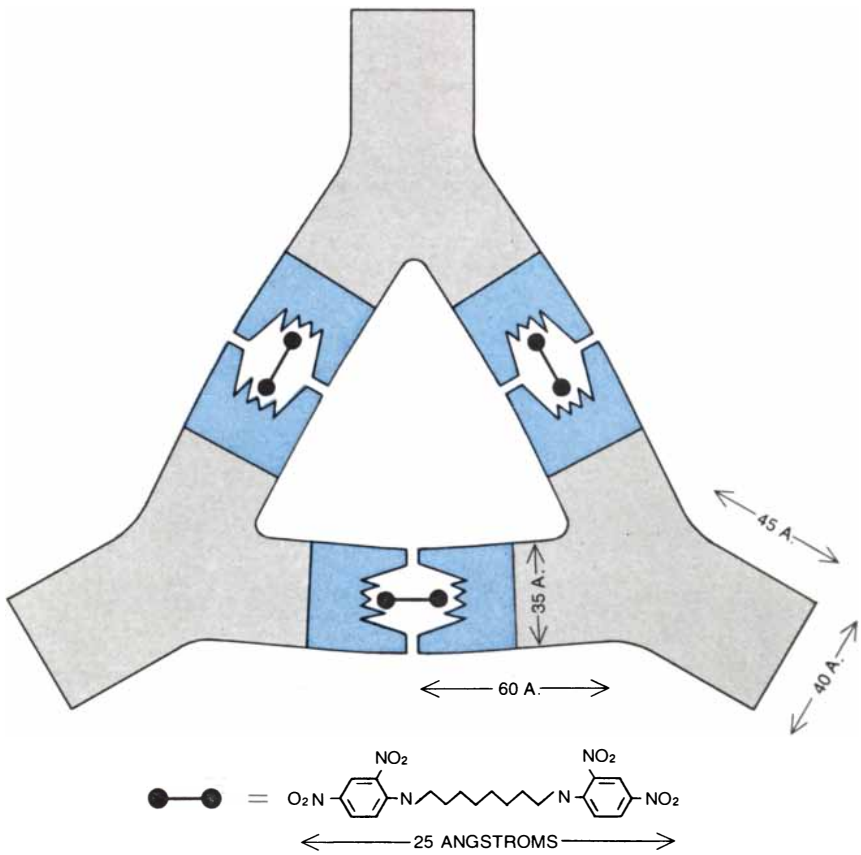
It is not possible here to discuss the merits of these theories. That would require consideration of a large body of experimental results, such as the explorations of the genetics of immune responsiveness by Baruj Benacerraf of the Harvard Medical School, Hugh O. McDevitt of the Stanford University Medical Center and Michael Sela of the Weizmann Institute of Science in Israel.

T Cell and *B* Cell

All the lymphocytes that circulate in the tissues have arisen from precursor cells in the bone marrow. About half of these lymphocytes, the *T* cells, have passed through the thymus on their way to the tissues; the other half, the *B* cells,



FUNCTIONAL TOPOGRAPHY of the antibody molecule is mapped. The end of each arm of the *Y* has a combining site (*p*) that recognizes epitopes on antigen molecules. The antibody also has its own epitopes, which can be recognized by other antibodies' combining sites. These include allotopes (*a*) in constant regions and idiotopes (*i*) in variable regions.

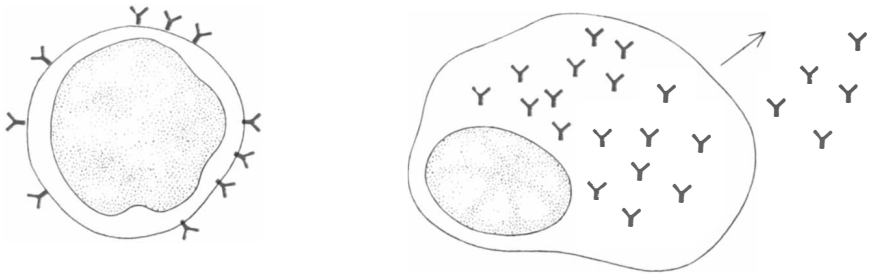


TRIANGULAR STRUCTURES in the micrograph on the opposite page are trimers, or complexes of three antibody molecules, linked by three double-ended dinitrophenyl antigens. The dimensions were worked out by Green and Valentine from electron micrographs.

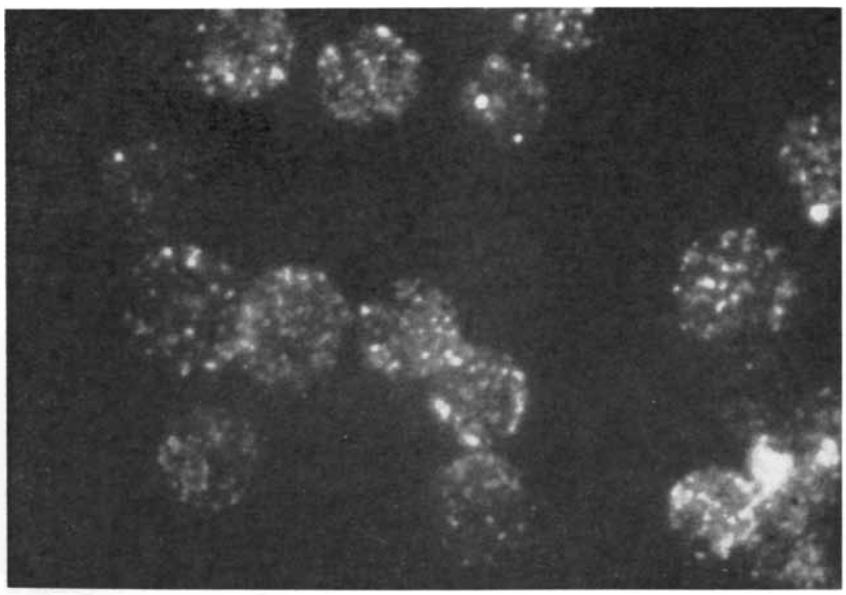
have not. This dichotomy was first discovered by Henry N. Claman of the University of Colorado Medical School and was characterized by Jacques F. A. P. Miller and Graham Mitchell, both of whom are now working with us in Basel. It has been the subject of thousands of investigations during the past five years. *T* cells and *B* cells cannot be distinguished by their form. Only *B* cells and their progeny cells secrete antibody molecules. One might think that this leaves little scope for *T*-cell function. On the contrary, *T* cells appear to be all-important. They too can recognize epitopes and must therefore, almost by definition, possess antibody molecules as surface receptors, although these receptor molecules have been much harder to demonstrate experimentally than those on *B* cells.

T cells can kill other cells, such as cancer cells, and transplanted tissues that display foreign epitopes. *T* cells can also suppress *B* cells or alternatively can help *B* cells to become stimulated by epitopes. This "helper" function of *T* cells has been repeatedly demonstrated both in animal experiments and in experiments with cells in culture. In the cell-culture experiments, based on a technique developed by Richard W. Dutton and Robert I. Mishell at the University of California at San Diego, lymphocytes taken from the spleen of an untreated animal are grown in a plastic dish together with molecules or cells that display foreign epitopes. After a few days' incubation lymphocytes that produce and secrete antibody molecules against the foreign epitopes can be shown to be present in the culture by the assay method for single antibody-producing cells [*see illustration on page 53*]. These antibody molecules are made by *B* cells, but the experiment will not work if only *B* cells are present. As soon as *T* cells are added to the culture dish, however, the *B* cells begin to respond and to produce antibody.

The dichotomy of the immune system into *T* and *B* lymphocytes adds a further dimension to the conceptual framework needed for the system's comprehension. That is not only an intellectual need but also a practical one, since the immune system is now known to be crucially involved in a vast number of diseases ranging from microbial infections and allergies to cancer, rheumatism, autoimmunity and many other degenerative disorders of aging.



LYMPHOCYTES, the cells of the immune system, produce antibodies. Each cell is committed in advance to the production of one specific antibody. In its resting state, as a small lymphocyte (*left*), the cell displays such antibody molecules (*color*) on its surface as "receptors." The advent of an antigen with an epitope that fits the combining site of this particular antibody molecule may stimulate the lymphocyte to grow, change in structure and divide, eventually giving rise to a large number of plasma cells (*right*): lymphocytes specialized for the rapid synthesis and secretion of this cell line's specific antibody molecules.



RECEPTOR ANTIBODY MOLECULES are demonstrated by a fluorescent stain in a photomicrograph made by Benvenuto Pernis. The cells are small lymphocytes from a patient with lymphocytic leukemia, in which a line of lymphocytes proliferates out of control. The receptor antibodies on the cell surfaces have epitopes in their constant regions (allotopes) characteristic of human antibody molecules. An antibody directed against those allotopes is prepared by injecting human serum into a rabbit. Rhodamine, a fluorescent dye, is coupled to the antibodies, which are added to a suspension of the lymphocytes. The bright spots on the cells represent fluorescent antibody bound to receptor molecules they "recognize."

The Lymphocyte Network

I have mentioned two striking dualisms within the immune system. One is the dichotomy of the lymphocytes into *T* cells and *B* cells, with functions that are partly synergistic and partly antagonistic. The second is the duality of the potential response of a lymphocyte when its receptors recognize an epitope: it can either respond positively (become stimulated) or respond negatively (become paralyzed). It is important to realize that the immune system displays a third dualism, namely that antibody molecules can recognize and can also be recognized. They not only have combining sites enabling them to recognize epitopes but also display epitopes enabling them to be recognized by the combining sites of other antibody molecules. That is true for the antibody molecules attached to the outer membranes of lymphocytes and serving as receptors as well as for the freely circulating antibody molecules, which can be regarded as messages released by lymphocytes.

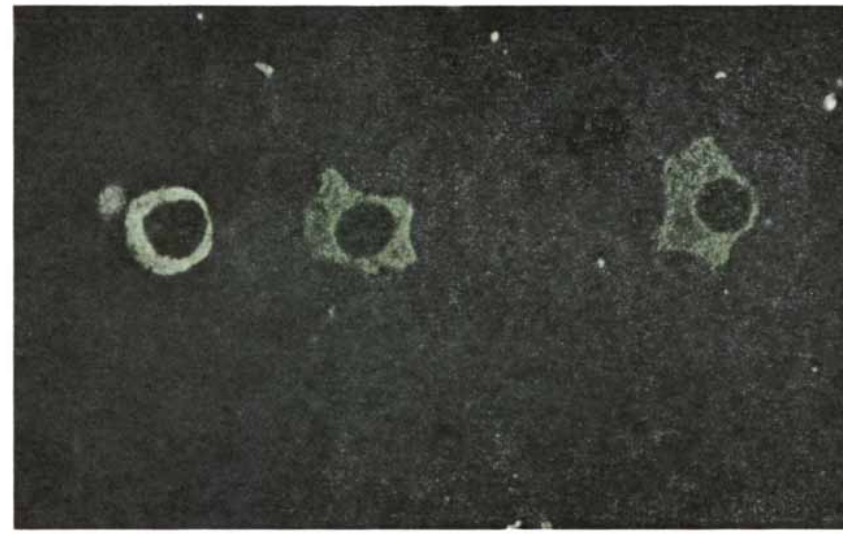
Epitopes occur on both the constant and the variable regions of an antibody molecule. Since the patterns of the variable-region epitopes are determined by the variable amino acid sequences of the polypeptide chains, there are millions of different epitopes. The set of such epitopes on a given antibody molecule was named the idiotype of that molecule by Oudin. When antibodies produced by

animal *A* are injected into animal *B*, animal *B* will produce antibodies against the idiotypic epitopes ("idiotopes") of the injected antibody molecules. That is also true when *A* and *B* belong to the same animal species and even when they are of the same inbred strain, that is, when they are genetically identical. Evidence is emerging that, within one animal, the idiotopes occurring on one antibody molecule are recognized by combining sites on a set of other antibody molecules, and that the idiotopes on the receptor molecules of one lymphocyte are recognized by the combining sites of the receptor molecules of a set of other lymphocytes. We thus have a network of lymphocytes and antibody molecules that recognize other lymphocytes and antibody molecules, which in turn recognize still others.

I am convinced that the description of the immune system as a functional network of lymphocytes and antibody molecules is essential to its understanding, and that the network as a whole functions in a way that is peculiar to and characteristic of the internal interactions of the elements of the immune system itself: it displays what I call an eigen-behavior. (Eigen in German means peculiar to, or characteristic of. Eigen-behavior is analogous to such concepts as the eigenvalue or eigenfrequency of certain physical systems.) There is an increasing body of evidence for this view.

Antibody molecules are normally present in the blood in a concentration of about 5×10^{16} molecules per milliliter. The total concentration of combining sites and idiotopes is therefore of the order of 10^{17} per milliliter. If the immune system made use of 10 million different combining sites and 10 million different idiotopes, each single variant of these elements would be present, on the average, in a concentration of about 10^{10} per milliliter. Mitchison at the National Institute for Medical Research and Nossal, Gordon L. Ada and their colleagues at the Walter and Eliza Hall Institute and the Australian National University, experimenting with low-zone tolerance, have shown that epitope concentrations ranging for different antigens from a million to 10^{12} epitopes per milliliter suffice either to suppress or to paralyze lymphocytes that can recognize the epitopes. Nisonoff and his co-workers at the University of Illinois College of Medicine and Humberto Cosenza and Heinz Köhler at the University of Chicago have shown that injecting into an animal antibodies against an idiotype suppresses lymphocytes that have receptors with idiotopes recognized by those antibodies. Leonard A. Herzenberg of the Stanford University Medical Center and Ethel Jacobson in our laboratory in Basel find that *T* lymphocytes recognizing epitopes on the receptors of *B* lymphocytes can suppress those *B* lymphocytes.

What this adds up to is that lymphocytes are subject to continuous suppression by other lymphocytes and by antibody molecules with idiotopes or combining sites that fit. Some lymphocytes escape from suppression and divide. New lymphocytes emerge. Others remain suppressed or decay. The eigen-behavior is the dynamic steady state of the system as its elements interact. As the system expands in the course of development and later life, new idiotopes and new combining sites emerge. The "self"-epitopes of other tissues impinge on the



INTERNAL ANTIBODY MOLECULES, being produced by plasma cells for secretion, are stained in this photomicrograph and the one at the bottom of the page, also made by Pernis. The plasma cells are from a rabbit that is heterozygous for the structural gene that determines the constant region of the antibody molecule and therefore its set of allotopes, or its allotype; that is, the rabbit inherited paternal and maternal chromosomes containing two different determining genes. The plasma cells have been stained with two preparations of fluorescent antibodies: one, to which a green-fluorescing stain has been coupled, is directed against the paternal allotype (call it allotype *A*) and the other is directed against the maternal allotype (*B*). The internal antibodies in some of the plasma cells bind to the green-staining preparation, indicating that they are antibody molecules carrying allotype *A*.

OTHER INTERNAL ANTIBODIES, in other plasma cells from the same field of the preparation as the micrograph at the top of the page, stain red. They are molecules with allotype *B*. Although all the rabbit's lymphocytes and plasma cells contain both the paternal and the maternal chromosomes, each cell expresses only one or the other constant-region gene.

network and cause certain elements to become more numerous and others less numerous. In this way each individual develops a different immune system.

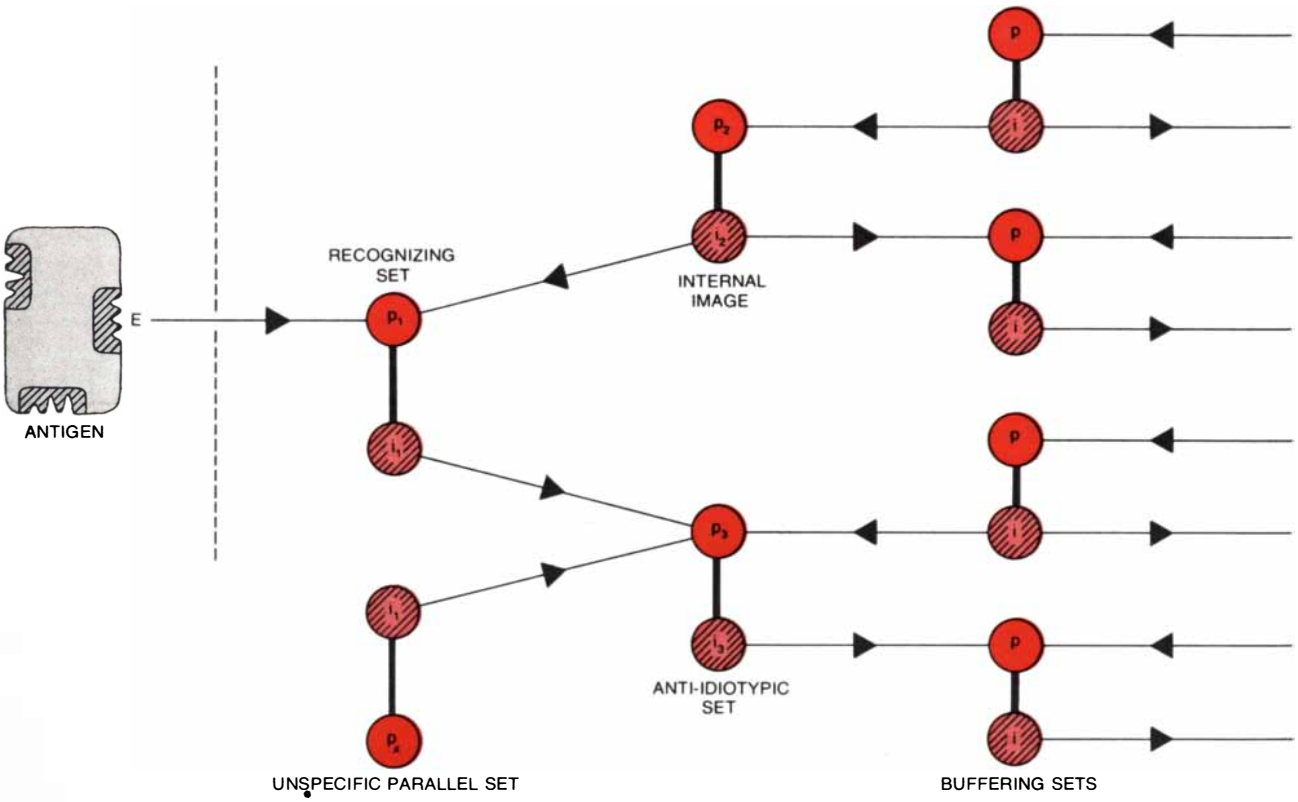
Invading foreign antigens modulate the network; early imprints leave the deepest traces. A given foreign epitope will be recognized, with various degrees of precision, by the combining sites of a set of antibody molecules, and lymphocytes that are committed to producing antibody molecules of that set are then stimulated and become more numerous. That is not, however, the only imprint made by the foreign epitope. The set of combining sites that recognized the epitope also recognizes a set of idiotopes *within* the system, a set of idiotopes that constitutes the "internal image" of the foreign epitope. The lymphocytes representing the internal image will therefore be affected secondarily, and so forth in successive recognition waves throughout the network [*see illustration below*].

The structural properties of the immune system and its eigen-behavior reside in these complex ramifications.

Immune System and Nervous System

The immune system and the nervous system are unique among the organs of the body in their ability to respond adequately to an enormous variety of signals. Both systems display dichotomies: their cells can both receive and transmit signals, and the signals can be either excitatory or inhibitory. The two systems penetrate most other tissues of the body, but they seem to avoid each other: the "blood-brain barrier" prevents lymphocytes from coming into contact with nerve cells.

The nerve cells, or neurons, are in fixed positions in the brain, the spinal cord and the ganglia, and their long processes, the axons, connect them to form a network. The ability of the axon of one neuron to form synapses with the correct set of other neurons must require something akin to epitope recognition. Lymphocytes are 100 times more numerous than nerve cells and, unlike nerve cells, they move about freely. They too interact, however, either by direct encounters or through the antibody molecules they release. These elements can recognize as well as be recognized, and in so doing they too form a network. As in the case of the nervous system, the modulation of the network by foreign signals represents its adaptation to the outside world. Both systems thereby learn from experience and build up a memory, a memory that is sustained by reinforcement but cannot be transmitted to the next generation. These striking analogies in the expression of the two systems may result from similarities in the sets of genes that encode their structure and that control their development and function.



LYMPHOCYTE NETWORK is diagrammed in an effort to indicate how its steady-state ("eigen") behavior is established and how the network responds to an antigen. An epitope (*E*) on the antigen is recognized by a set (p_1) of combining sites on antibody molecules, both circulating antibody and cell-surface receptors. Cells with receptors of the recognizing set p_1 are potentially capable of responding to the antigenic stimulus (*arrowhead*) of epitope *E*, but there are constraints. The same molecules that carry combining sites p_1 carry a set of idiotopes (i_1). These are recognized within the system by a set of combining sites (p_3), called the anti-idiotypic set because they tend to suppress (*reverse arrowhead*) the cells of set i_1. (These idiotypes i_1 are also found on molecules with combining sites that do not belong to the recognizing set p_1 but rather are unspecific with regard to epitope *E*.) On the other hand, the set p_1 also recognizes internal epitopes i_2, which therefore constitute an internal image of the foreign epitope *E*. In the steady state, molecules of the internal image tend to stimulate cells of set p_1 and thus to balance the suppressive tendency of the anti-idiotypic set. When the foreign antigen enters the system, its stimulatory effect on recognizing set p_1 allows cells of that set to escape from suppression. (The same thing happens to unspecific cells of the parallel set p_x.) The resulting immune response to the antigen is modulated by the buffering effects of many more sets of combining sites and idiotopes (*right*), which have a controlling influence on the response.

Opel: Number one seller in Germany.

Opel Manta Luxus: $3,058.50 in America.

The Opel outsold every other car in Germany in 1972. The plush Manta Luxus is one of the reasons why. For the above price you get all the standard features listed on the price sticker below. Give up a little plush and Manta is $2,849.50. Or go sporty with Manta Rallye—$3,046.50. All prices are Manufacturer's Suggested Retail Prices, including dealer new vehicle preparation charges. Optional equipment, State and local taxes and transportation charges are additional.

Manta. The personal-size sport coupe proven in more than 2 billion miles of actual driving. Manta. With rack and pinion steering. A 1.9 liter engine. Disc brakes up front. A grand-touring ride. Manta. Luxurious. Economical. Comfortable. Really ready for America.


BUICK MOTOR DIVISION
GENERAL MOTORS CORPORATION

OPEL

MANUFACTURER'S SUGGESTED RETAIL PRICE

HOMETOWN BUICK, INC.
HOMETOWN U.S.A. 00-000-000000

ARLINGTON, VIRGINIA

OPEL BALTIMORE, MARYLAND

00000 0 00000 '00 57L 2DR MANTA LUXUS CPE

00 0000000 $3,058.50

THE FOLLOWING ITEMS ARE STANDARD EQUIPMENT ON THIS 1973 MODEL OPEL: 1.9 LITER ENGINE DESIGNED FOR NO-LEAD, LOW-LEAD FUEL, WITH HYDRAULIC VALVE LIFTERS AND AUTOMATIC CHOKE • POWER BRAKES WITH FRONT DISCS, SELF ADJUSTING • CARPETED TRUNK COMPARTMENT, PACKAGE SHELF AND FLOOR • CUSTOM CLOTH TRIM WITH VINYL EDGING, AND COLOR COORDINATED INTERIOR • HINGED REAR QUARTER WINDOWS • RUBBER BUMPER GUARDS AND STRIPS • ELECTRIC CLOCK • RALLYE TYPE WHEELS AND STEERING WHEEL • ADJUSTABLE FRONT SEAT BACKS.

$ 22.00

TOTAL AMOUNT $3,080.50

German precision imported by General Motors.
Sold and serviced by 2200 Buick dealers.

There's nothing like a p
But there's a lot more to a

With a Polaroid portrait identification card, a glance tells you whether the person and the picture are the same. But an ID card should do more than match faces; it should suit the conditions of use.

For that reason we've developed all sorts of ways to make our ID cards more convenient for both the user and the issuer.

In addition to our standard cards (A and B)—carried in pocket, purse, or wallet—we have others with more sophisticated functions.

For instance, there's a card (C) which can be punch-coded. A computer reads it just as accurately as you match a picture to a face.

We have badges (D) with a clip for attachment to outerwear. They have regular or extra-large portrait areas.

We have a card (E) which can be encoded to act like a key. Slipped into an electronic device, it allows access to secure areas, parking lots, etc.

We have a new credit card (F) combining a magnetic stripe with a portrait and an embossed name and number. It's another way of encoding information securely for automatic banking transactions.

Chances are we have an ID card that will exactly fill your needs.

Security starts with the card itself.

The Polaroid ID card is highly secure. It's laminated with a clear, tough plastic which becomes integral with the card.

The card works so well that more than 6000 organizations all over the world use it. You'll find our ID cards in industry, universities, research and computer centers. They're used as credit cards and drivers' licenses.

People like to carry them. They like having their own color portrait, and they like being able to

Polaroid ®

ortrait for an ID card.
ID card than a portrait.

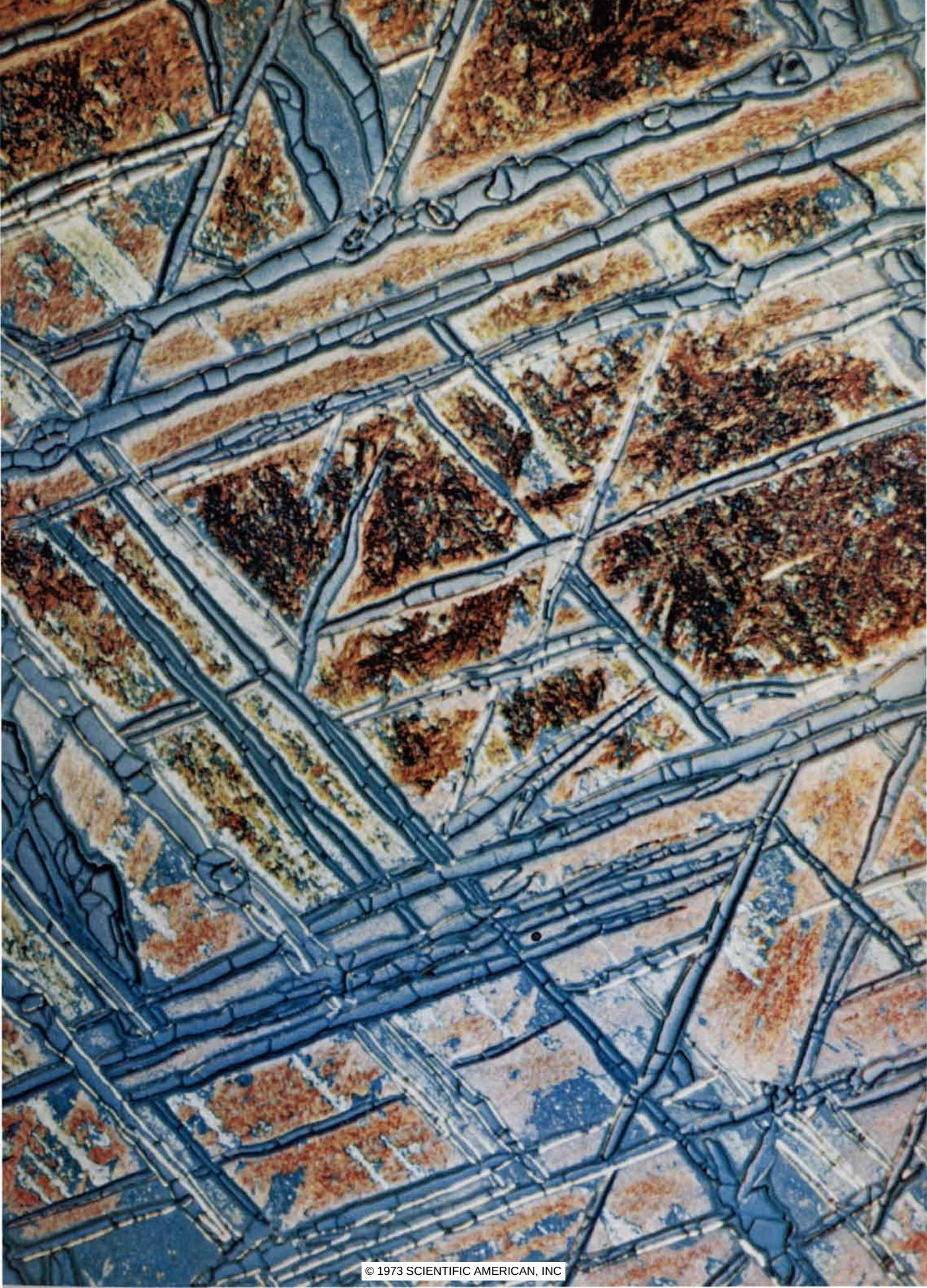
rove who they are.
They carry them with confidence.
Credit cards, for example. Before the portrait, ıany people were afraid to carry them around. ut now, with the portrait, a lost or stolen credit ırd is no longer a serious problem.
You can't forge a face.

A handy system for making cards.

Any organization can produce its own ID cards asily and quickly with a low-cost Polaroid ID-3 and Identification System.
It can be carried around like a suitcase, so you an take it from one branch of your organization) another. An operator can be trained to run it in few minutes.
With an ID-3, you can immediately issue per-ıanent ID cards to all your employees. If any-ɔdy blinks, you can make a new card on the spot.

You can use different background colors to color-code the employee's function, work area or security clearance.
Security starts with the ID-3. It's the basis of entire security systems.
We'd like to demonstrate it in your own office or factory.
We'll put your portrait on a card and let it speak for itself.
For more information or a demonstration, write to Polaroid Corporation, Dept. 57-255, 549 Technology Square, Cambridge, Mass. 02139. Or you can call collect (617) 864-6545.

The Polaroid 2-Minute ID System.



Meteorites and Cosmic Radiation

Before a meteorite falls to the ground it has been bombarded in the vacuum of space by cosmic ray particles. The nuclear products of this bombardment are clues to the history of the solar system

by I. R. Cameron

One of the most interesting questions of astrophysics is: How did the solar system form out of a diffuse cloud of interstellar dust and gas? Another interesting question is: What is the nature and origin of cosmic rays, the high-energy particles that crash into the earth's atmosphere after having traveled through space for millions of years at speeds close to the velocity of light? An intriguing link between the two questions is provided by meteorites. Meteorites are samples of the primordial material of the solar system. These samples were subsequently bombarded by cosmic rays for long periods of time in interplanetary space. The study of the effects of that bombardment has yielded a wealth of information not only about the cosmic radiation but also about the meteorites themselves.

Meteorites are solid bodies that are large enough to survive their journey from interplanetary space through the earth's atmosphere to the earth's surface. Before they encounter the earth they have circulated in interplanetary space for periods ranging from some tens of thousands of years to as much as two billion years. Until samples of the moon were brought back by the Apollo astronauts meteorites were the only extraterrestrial material available for close examination. The geological processes that complicate the interpretation of the history of the rocks in the earth's crust do not apply to meteorites, whose composition has remained essentially unchanged since the solar system came into being some 4.5 billion years ago.

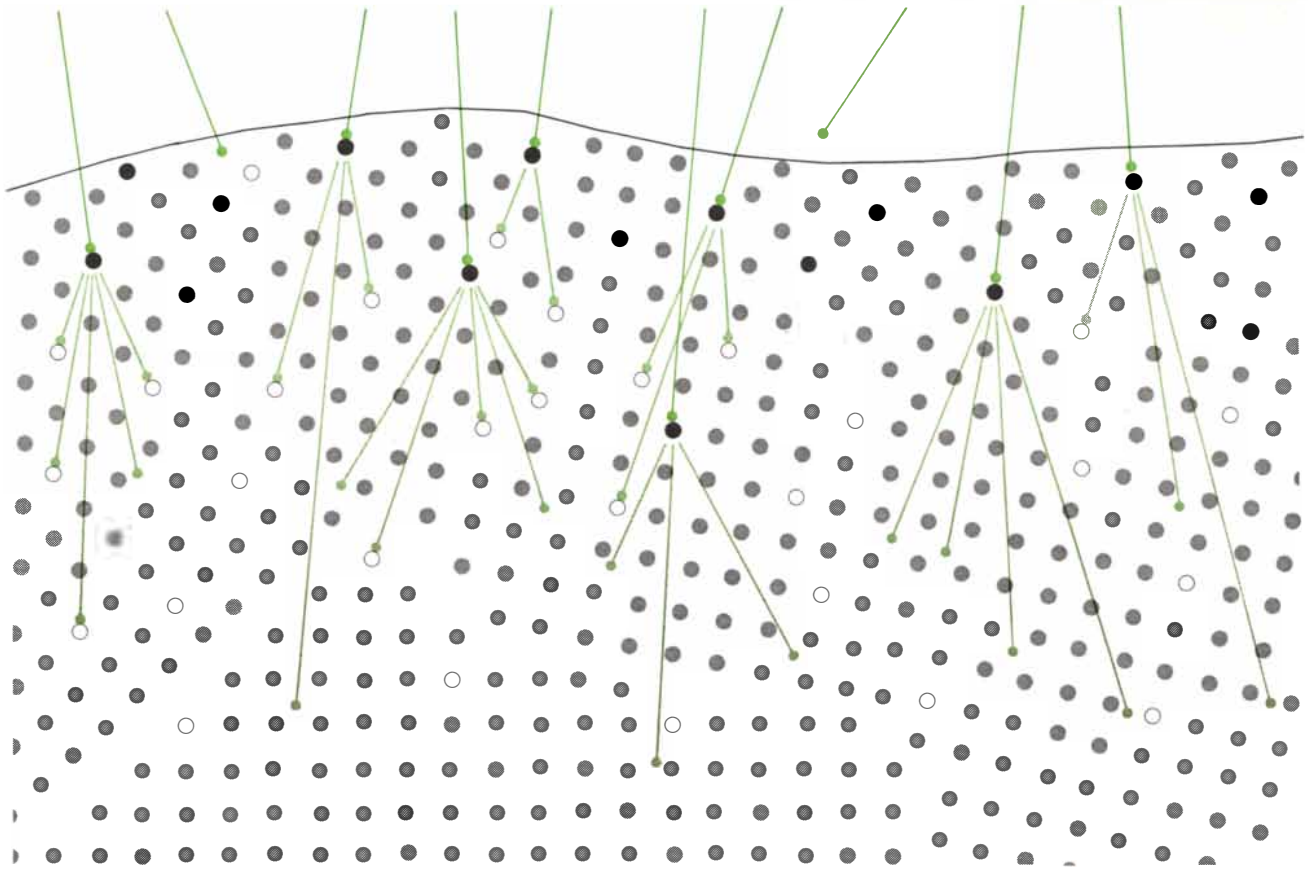
For the investigator of meteorites the basic challenge is deducing the history of the meteorites from a bewildering abundance of evidence. The richness of the problem is indicated by the sheer variety of types of meteorite. The two main classes are the stony meteorites and the iron meteorites. The stony meteorites consist mainly of silicates, with an admixture of nickel and iron. The iron meteorites consist mainly of nickel and iron in various proportions. A smaller class is the stony-iron meteorites, which are intermediate in composition between the other two. Stony meteorites are in turn divided into two groups: the chondrites and the achondrites, according to whether or not they contain chondrules, spherical aggregates of iron-magnesium silicate. Within each group there are further subdivisions based on mineralogical and chemical composition.

Early attempts to explain the origin of meteorites were of necessity based on their structural and chemical properties. The degree to which meteorite types are chemically differentiated suggests that they were formed in the interior of parent bodies that were large enough for the heat generated by the decay of radioactive elements to melt iron and allow it to sink into the core of the body. The material of the stony meteorites would then have risen toward the surface of the body, where the variety of temperature and pressure conditions would have produced the observed range of structure and composition. Since the meteorites that fall on the earth range in mass from a few grams up to 50 tons, we have to postulate that there was one catastrophic collision or more between the parent bodies, producing fragments that circulate in space for long periods before they are captured by larger bodies such as the earth.

It is possible to estimate the radius of a typical parent body from the rate at which it cooled from its original molten condition. That rate can be deduced from the detailed metallurgical structure of the nickel-iron in iron meteorites [*see illustration on opposite page*]. The cooling rate in the central region of the parent bodies at the time they solidified seems to have been between one degree and 10 degrees Celsius per million years. These rates in turn correspond to parent bodies ranging in radius from 50 to 200 kilometers.

There is much speculation about the source of the heating that gave rise to the initially high temperatures of the parent bodies. The conversion of gravitational energy, which is the basic source of heat in the early history of a massive body such as a star, can be discounted. It is quite inadequate to heat to a melting temperature a body whose radius is a mere few hundred kilometers. Moreover, the decay of the naturally radioactive elements found on the earth today is much too slow either to provide the amount of heat required or to allow the observed rate of cooling. Uranium 238, for example, has a decay half-life of 4.5 billion years. We must therefore conclude that the source of the heat that melted the parent bodies of meteorites was some radioactive element, now extinct, that was synthesized in the orig-

ETCHED SECTION OF TAZEWELL METEORITE was photographed at a magnification of 45 diameters in an investigation by Joseph I. Goldstein of Lehigh University and Harvey Yakowitz of the National Bureau of Standards. The regular linear pattern, known as a Widmanstätten figure, is characteristic of iron meteorites. The details of the figure depend on the chemical composition of the material and the rate at which it cooled. They yield information on the probable size of the parent body in which the meteorite material was initially formed.



SPALLATION occurs when a fast particle of the primary cosmic radiation (*dark color*) strikes the nucleus of an iron atom (*gray*) in a meteorite. The nucleus ejects a spray of neutrons and protons; the residual nucleus (*black*) is termed the spallation product. The spray of neutrons and protons forms a flux of secondary cosmic ray particles (*light color*). These lower-energy particles strike other nuclei deeper within the meteorite, giving rise to other spallation products (*open circles*). By measuring the concentrations of the various spallation products within the meteorite one can determine how long it has been exposed to the cosmic radiation.

inal nuclear processes that gave rise to the material from which the solar system formed. Such an element would need to have a half-life so short that no trace of it is found today. The most likely candidate is the isotope aluminum 26, which has a half-life of 740,000 years. Because of the relatively brief period for which such a short-lived radioactive element could exist in sufficient quantity to provide a useful heat source, we can conclude that the condensation process that formed the parent bodies of meteorites must also have started within a relatively brief period (less than 10 million years) after the end of the nuclear synthesis that created the elements of the solar system.

How long ago did the condensation process take place? To answer that question we can make use of the same radioactive-isotope dating techniques that have been widely applied for determining the ages of rocks on the earth. The simplest in principle, and the first to be applied to meteorites, is the uranium-helium method. The two naturally occurring isotopes of uranium, uranium 235 and uranium 238, decay successively into other radioactive elements until they finally end up as nonradioactive isotopes of lead. For both U-235 and U-238 several of the decays involve the emission of alpha particles, which are nuclei of helium. The decay of the original uranium therefore leads to a buildup of helium (provided that the mineral is sufficiently cool to retain the gas).

When the parent bodies of meteorites were first formed, their temperature was so high that the helium simply diffused away and was lost. As the parent body cooled, however, it eventually reached a temperature below which the helium could no longer escape by diffusion. The quantity of helium present in a given meteorite sample will therefore be a function of both the concentration of uranium and of the time that has elapsed since the parent body solidified. Measurement of the relative concentrations of helium and uranium thus enables us to estimate this time.

Measurements made with this method and similar ones lead to the conclusion that the parent bodies of meteorites were created about 4.5 billion years ago. The actual ages of meteorites based on the retention of helium, however, show a very large spread. Most of this effect can be explained by the fact that some helium was lost from the meteorite during its subsequent history, as a result either of collisions or of heating during a close approach to the sun. Such a loss of helium, of course, gives rise to apparent ages that are lower than the true ages. On the other hand, many of the early measurements of helium gave ages that were anomalously high, some greatly exceeding the age of the oldest terrestrial rocks known. In 1947 Carl A. Bauer of the Harvard College Observatory showed that the energetic cosmic rays bombarding meteorites would produce significant quantities of helium. It was this "cosmogenic" helium, in addition to the radiogenic helium produced by the decay of uranium, that had produced spuriously long ages for some meteorites.

The presence of cosmogenic helium in meteorites was confirmed by F. A. Paneth and his collaborators at the University of Durham in 1952. Since then a large number of cosmogenic elements have been measured in meteorites. The interpretation of the results has led to interesting conclusions about the history of meteorites after the breakup of their parent bodies. To understand the effects of cosmic rays on meteorites we must now consider some of the properties of the radiation itself.

The primary cosmic radiation consists of particles, mainly protons, that have energies extending up to the enormous value of 10^{20} electron volts. (By way of comparison, the energies of the alpha particles or beta particles that are emitted in normal radioactive decays are in the range of a few million electron volts.) In addition to the protons the radiation contains smaller numbers of heavier particles: the nuclei of elements more massive than hydrogen. The relative abundance of these nuclei decreases markedly with increasing mass.

The enormous energies of cosmic rays are a clue to their origin. They imply that the radiation is generated in cataclysmic events, such as supernova explosions, that release gigantic amounts of energy in both our galaxy and others. The lively current interest in such phenomena has led to a major effort to measure the properties of cosmic radiation. The study of the effects of cosmic rays on meteorites can thus make a valuable contribution in several areas.

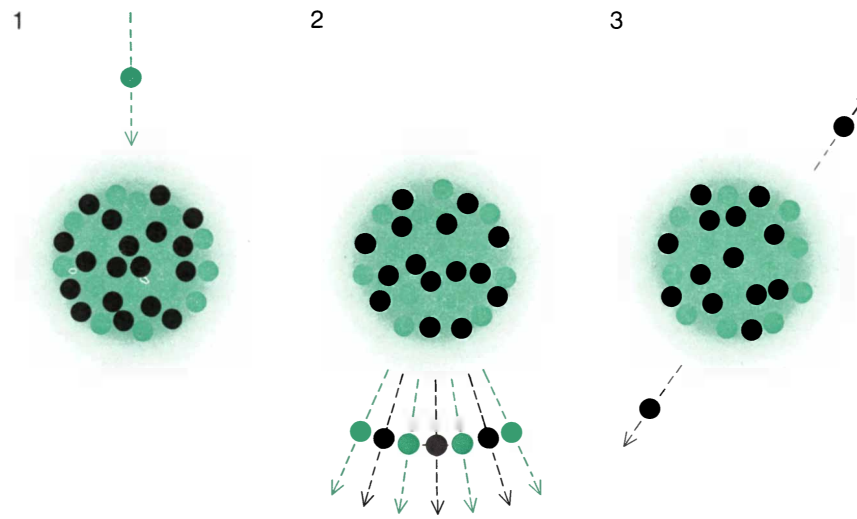
When a fast cosmic ray particle strikes the nucleus of an atom in a meteorite, it typically ejects several neutrons and protons together with some of the pi mesons representing the attractive forces that hold the nucleus together. The remainder of the energy of the bombarding particle is shared among the neutrons and protons left behind in the nucleus. One or more of these particles may subsequently "evaporate" out of the nucleus. The end result is that the struck nucleus is transformed into a nucleus of a different element, which may be either radioactive or stable. The process is called spallation, and the residual nucleus is called a spallation product. A meteorite that has been exposed to cosmic radiation as it orbited in space will contain minute amounts of a variety of radioactive elements, which can be identified by their half-lives and by the characteristic radiation they emit.

The primary particles of the cosmic radiation can penetrate matter for only a relatively short distance. Therefore the material in the interior of the parent bodies is completely shielded from the radiation; only a surface layer representing a negligible fraction of the total volume is affected. From the instant that two parent bodies collide, however, the meteorites in the debris of the collision are exposed to the cosmic radiation and the spallation products begin to build up in them. The process continues until the meteorite falls to the earth, after which it is shielded from the cosmic radiation by the earth's atmosphere. The period of time from the start of the meteorite's exposure to the cosmic radiation to the moment it is captured by the earth is known as the exposure age.

Measured exposure ages range from some tens of thousands of years up to two billion years. The most obvious way to arrive at an exposure age is simply to measure the concentration of some stable spallation product and then to divide this value by the rate at which that nucleus is produced. Such a procedure is more complicated than it sounds. Calculating the production rate of a given spallation product is difficult because the rate depends on, among other things, the composition of the particular meteorite, the difference in mass between the original nucleus and the nucleus of the spallation product, and on the intensity and distribution in energy of the bombarding particles. The energies and intensity of the radiation depend, in turn, on the meteorite's orbit and the depth to which the radiation penetrates into the object. Even when samples of meteorites are bombarded with protons in laboratory accelerators to standardize the calculations, an appreciable uncertainty remains in the prediction of the rate at which individual spallation products are formed.

We can avoid many of the difficulties involved in calculating absolute production rates by comparing the concentrations of two different spallation products of approximately the same mass. The ratio of the production rates of these two nuclei can be estimated much more accurately than the individual rates because many of the uncertainties cancel themselves. If one of the chosen nuclei is radioactive, the concentration of that element will eventually reach a constant level where the rate at which it is produced is exactly balanced by the rate at which it radioactively decays. If the half-life of the nucleus is short, this equilibrium condition will be reached fairly soon after the start of the irradiation. If the other product is stable, its concentration will build up steadily as a function of the time of irradiation. The ratio of the concentrations of the two spallation products will therefore be a direct measure of the period for which the meteorite has been exposed.

Many determinations of exposure age have been made with the mass spectrometer, which can measure the tiny amounts of rare gases produced by spallation. Another technique calls for gamma-coincidence spectrometry. Here the



TYPICAL SPALLATION REACTION involves a high-energy cosmic ray proton striking a nucleus of sulfur 34 (*1*), which consists of 16 protons (*color*) and 18 neutrons (*black*). When the cosmic ray proton collides with the nucleus, it ejects a spray of four protons and three neutrons (*2*). The remaining nucleus is left in an excited condition. It eventually rids itself of the excess excitation energy by emitting two more neutrons (*3*). The final spallation product is a radioactive nucleus of aluminum 26, which has a half-life of some 740,000 years.

meteorite sample is placed between two detectors, each of which will produce an electric pulse when a gamma ray emitted by the sample strikes it. The system is particularly well suited for measuring a radioactive isotope that decays by emitting a positron, or positive electron. The positron slows down rapidly and encounters a negative electron. The particles annihilate each other, giving rise to two gamma rays that go off in opposite directions. For a small sample placed symmetrically between the two detectors there is a fairly high probability that one of the gamma rays will be detected in the first detector just as the other is detected in the second. The detectors are connected to a coincidence analyzer that will give an output pulse only when it receives two input pulses of the right magnitude within a very short time interval. The chance of such an event's happening accidentally as a result of the two detectors' simultaneously recording unrelated gamma rays from radioactive impurities is very small. Therefore the system is highly selective for the radiation being emitted by the meteorite sample. Gamma-coincidence spectrometry has the additional advantage of not destroying the sample. In other methods the sample must be dissolved.

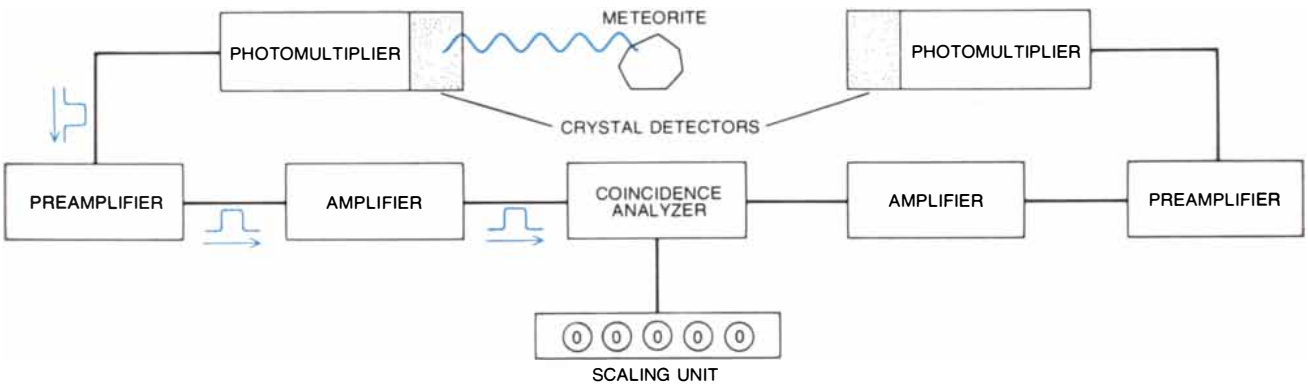
Several groups, such as the one led by Edward Anders of the University of Chicago, have measured the concentration of radioactive isotopes in a large number of meteorites. These measurements enable us to analyze the distribution of exposure ages with a statistically significant number of observations. The analysis yields some interesting clues about how the different classes of meteorites originated. The age distribution measured by H. Voshage of the Max Planck Institute for Chemistry for some 60 iron meteorites shows a "clustering" effect suggesting that the majority of them originated in three or four major collisions of parent bodies at specified times during the past billion years [*see top illustration on page 70*]. James R. Arnold of the University of California at San Diego and others, working with computer models of the solar system, have shown that exposure ages of the order of those observed for the iron meteorites are consistent with the intervals that would be expected to elapse before meteorites that had been produced by collisions happened to be captured during a chance encounter with the earth. Such models assume that the parent bodies collided within the belt of asteroids that lies between the orbits of Mars and Jupiter.

The exposure ages of the stony meteorites, however, are from one to two orders of magnitude less than those of the iron meteorites. Several ingenious hypotheses have been advanced in the attempt to explain this difference. One possibility is that the number of stony meteorites in the asteroid belt is high enough to ensure that the meteorites that do not happen to encounter the earth within a relatively short time after they are formed are destroyed by colliding with other meteorites. This mechanism would limit the lifetime of meteorites and would cut off the distribution of ages that are observed on earth. The drawback to this hypothesis is that it assumes an improbably high density of debris in the asteroid belt.

Harold C. Urey of the University of California at San Diego had suggested that some of the chondritic stony meteorites might be part of the debris ejected from the surface of the moon as the result of meteorite collisions. There are various objections to this hypothesis, among them the fact that the samples brought back by the Apollo expeditions have now shown that the lunar material is quite different in composition from the chondritic meteorites.

In 1959 Fred L. Whipple and Edward L. Fireman of the Harvard College Observatory proposed an alternative explanation for the absence of stony meteorites of long exposure age. The meteorite during its time in space is being continuously eroded by interplanetary dust and gas. Whipple and Fireman argued that stony meteorites would erode much more quickly than iron meteorites. The result would be that stony meteorites that were not captured by the earth within a certain period would be completely destroyed by erosion, whereas the iron meteorites would be relatively unaffected.

It is theoretically possible to detect the existence of such space erosion by comparing the various estimates of the exposure age of a single meteorite on the basis of the measured ratios of radioactive isotopes with widely different half-lives. For an isotope with a short half-life the radioactive atoms present at the time the meteorite fell to the earth would have been created by the cosmic rays bombarding it over a relatively short period toward the end of its exposure, since the atoms created earlier would all have decayed. For a stable spallation product or one of long half-life, however, the radioactive atoms present would have been created essentially over the total exposure age of the meteorite. If we assume that there is a significant



GAMMA-COINCIDENCE SPECTROMETRY is one technique for analyzing radioactive isotopes produced in meteorites by the bombardment of the cosmic radiation. It is particularly well suited for measuring an isotope that decays by emitting a positron (positive electron). When the positron encounters a negative electron, the particles annihilate each other, giving rise to two gamma rays that go off in opposite directions. For a small sample placed symmetrically between two crystal detectors there is a fairly high probability that one of the gamma rays will be detected in the first detector just as the other is detected in the second. The interaction of a

amount of space erosion, many of the stable or long-lived atoms present would have been produced at a period when the meteorite was much larger than just before it fell to the earth. The cosmogenic production rate in its interior would therefore be relatively low, owing to the effective shielding from cosmic ray bombardment provided by the meteorite's large volume. The result is that the measured concentration of the long-lived isotope would be lower with respect to the short-lived isotope than it would have been if the radius of the meteorite had not decreased over its exposure age, that is, if there had been no space erosion.

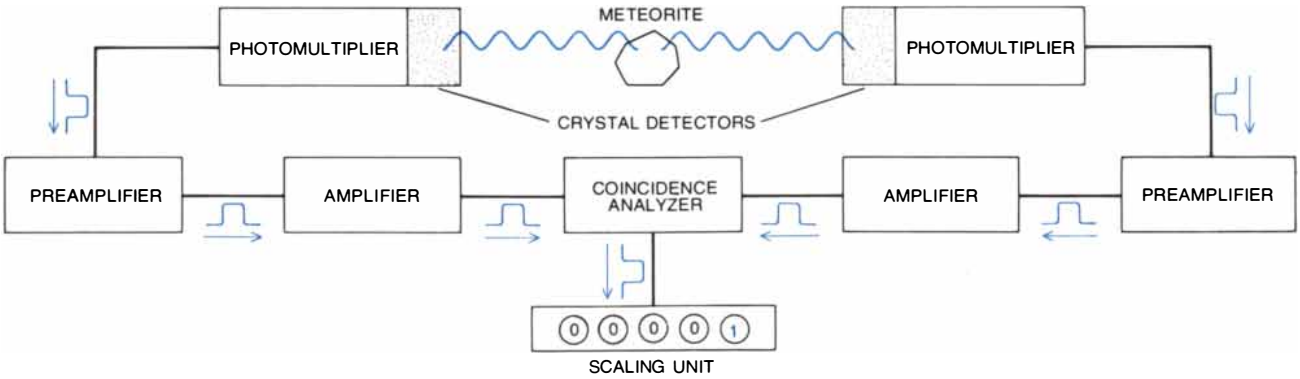
At the present time the evidence for a significant amount of space erosion seems to be inconclusive. David E. Fisher of the University of Miami has shown that an erosion rate of .2 millimeter per million years is sufficient to resolve discrepancies in estimates of the exposure ages of eight iron meteorites. On the other hand, the observed ages of stony meteorites require an erosion rate some 50 times greater: of the order of one centimeter per million years. A space-erosion rate of that magnitude would imply that many of the stony meteorites had been large enough during the early part of their exposure for their interior to have been subjected to a relatively high flux of neutrons, which build up appreciably as secondary products of the cosmic ray bombardment inside meteorites of large volume. The buildup of neutrons would in turn have led to the production of larger concentrations of isotopes such as argon 38 (which is formed by the capture of neutrons) than are observed. Moreover, evidence from the tracks of cosmic ray particles in meteorites has cast doubt on the existence of a space-erosion effect with the magnitude assumed by some. It is clear that more work is needed before the exposure-age distribution of the stony meteorites can be satisfactorily explained.

One of the factors that complicate the determination of exposure ages is that the rate at which the various spallation products build up depends on their depth below the surface of the meteorite. We must always bear this in mind when we compare the exposure ages obtained by independent measurements on separate fragments of a given meteorite. In order to create spallation products whose masses differ significantly from the mass of the original nucleus, the bombarding particles must be highly energetic. Spallation products of this type therefore tend to be concentrated in the outer regions of the meteorite, which are subject to direct bombardment by the high-energy primary particles of the cosmic radiation. The flux of high-energy primary particles decreases rapidly with depth as the particles dissipate their energy by generating a flux of lower-energy secondary particles. The secondary particles are still effective in producing spallation, but they can only form products whose mass is not very different from the mass of the original nucleus. The flux of secondary particles initially increases with depth, and the rate at which it creates spallation products is at a maximum at depths of between 10 and 20 centimeters below the surface of the meteorite. Thereafter the production rate drops off gradually with depth as the energy of the secondary particles dissipates.

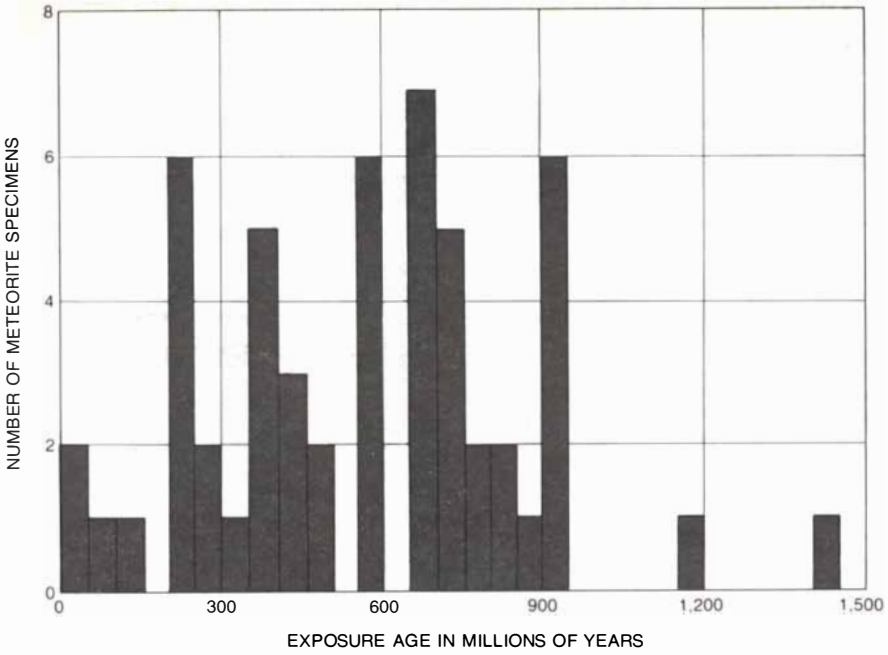
The relatively low concentration in the surface region of spallation products created by the low-energy secondary flux is counterbalanced to some extent by the low-energy cosmic rays emitted by the sun, which enhance the production of those spallation products. Both effects are difficult to measure accurately, since an appreciable amount of the meteorite's surface material is lost by heating when the meteorite travels at high speed through the earth's atmosphere. This process is known as ablation. By measuring the concentration of a given radioactive isotope and determining contours of equal concentration, it is possible to reconstruct the original shape the meteorite had before entering the atmosphere, thereby yielding information on the magnitude of the ablation.

So far I have been discussing how measurements of cosmogenic effects can be interpreted to yield information about meteorites and the history of the solar system. These same measurements can also be utilized to study possible variations in the intensity of the cosmic ray particles themselves, with respect either to a time in the past or to the meteorite's position in the solar system. For example, it is known that the intensity of the flux of cosmic rays falling on the earth varies in an 11-year cycle owing to the cyclic magnetic activity of the sun. The sun's magnetic field deflects incoming charged particles away from the solar system, and thus the minimum of the cosmic ray intensity coincides with the maximum of the solar activity.

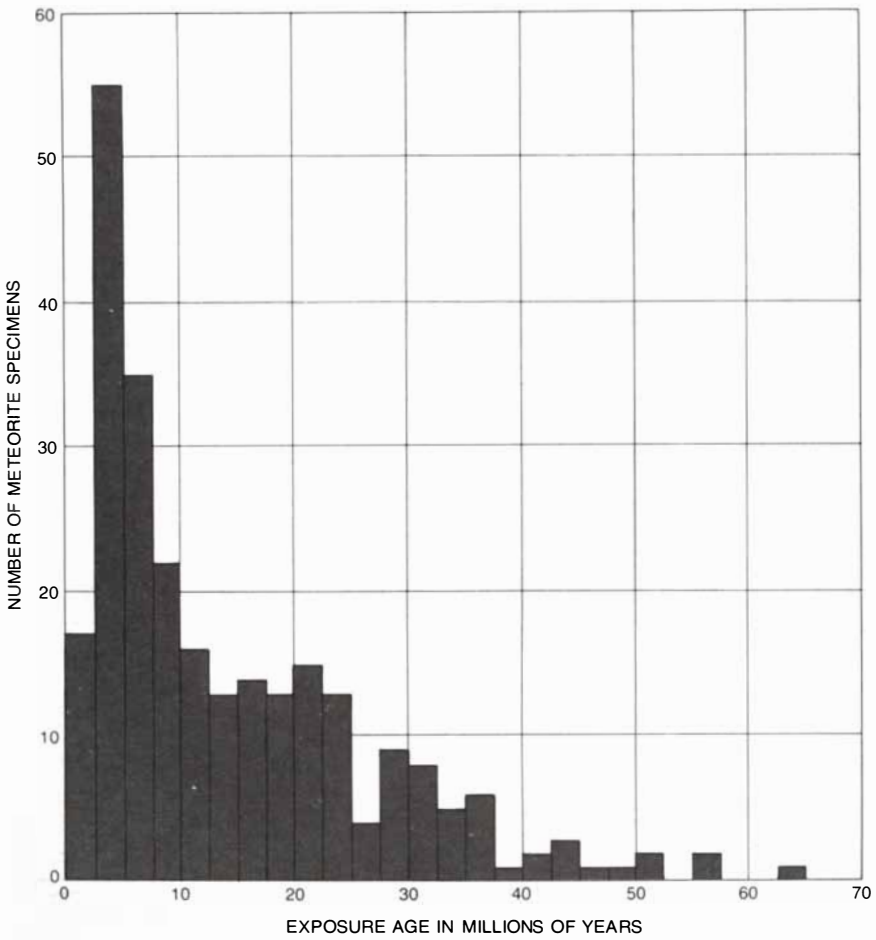
The flux to which a meteorite is exposed while circulating in its orbit will vary with the same 11-year period. The concentration of a cosmogenic radioactive isotope with a long half-life (such as aluminum 26, with its half-life of 740,-



gamma ray with one of the crystals produces a flash of light, which is amplified by the photomultiplier attached to the crystal to produce an electric pulse. The pulses travel through the circuit to a coincidence analyzer that produces an output pulse only when it receives two input pulses of the right magnitude within a very short time interval. If a random gamma ray is detected (*left*), the coincidence analyzer does not allow the pulse to pass and no count is registered on the scaling unit. On the other hand, if two gamma rays are detected (*right*), the coincidence analyzer emits a pulse and scaling unit records it as a decay event from meteorite sample.



EXPOSURE AGES FOR IRON METEORITES, that is, the lengths of time that the meteorites have been exposed to the cosmic radiation, tend to cluster around certain values. These values are assumed to correspond to the times of several major collisions between large parent bodies in the solar system. The debris from these collisions is the meteorites. This chart is based on the work of H. Voshage of the Max Planck Institute for Chemistry.



EXPOSURE AGES FOR STONY METEORITES show a more continuous distribution than the ones for the iron meteorites. Moreover, the ages themselves are much lower. One possible explanation is that older stony meteorites have been destroyed by "space erosion."

000 years) will be essentially unaffected by the 11-year cycle. The atoms present at any one time will have been produced by the exposure of the meteorite to the cosmic ray flux integrated over a period of the order of a few million years. On the other hand, the current concentration of a product with a short half-life (such as sodium 22, whose half-life is only 2.6 years) will be strongly influenced by the flux to which it has been exposed over the past two or three years, since the sodium 22 that was formed at times much earlier will have decayed away.

Fireman took a number of meteorites and compared the measured ratio of the concentration of sodium 22 with respect to aluminum 26 at the time of their fall with the theoretical values predicted on the basis of the known magnitude of the 11-year variation in the flux of cosmic rays at the distance of the earth from the sun. He found that the measured variation in the ratio of the two elements was less than the predicted variation. The finding indicates that the amplitude of the variation is less at the position of the meteorites' orbits (which are generally about twice as far from the sun as the earth is) than it is at the position of the earth's orbit.

Knowledge of the ratios of short-lived and long-lived cosmogenic isotopes in meteorites is of course much more useful if the orbit of the meteorite is well known. In recent years various recording networks have been set up to photograph the incandescent trails of incoming meteorites so that accurate estimates of their trajectories can be made. So far only two trajectories have actually been measured by this method. The more recent was the trajectory of the Lost City meteorite, which fell in Oklahoma on January 3, 1970. The photographic observations made it possible to find the meteorite very quickly, with the result that various workers could measure the concentrations of quite a number of cosmogenic radioisotopes with widely differing half-lives. By correlating these measurements with the known orbit of the meteorite, we now have more information about how the intensity of the cosmic radiation varies with distance from the sun.

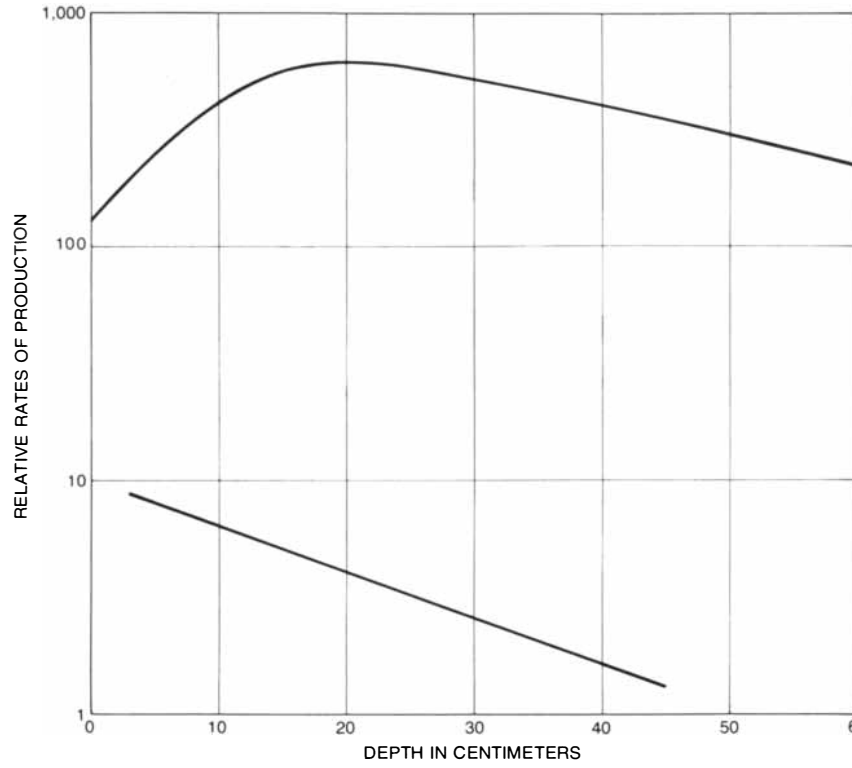
Fireman and his co-workers determined the ratio of the two cosmogenic radioactive isotopes argon 37 and argon 39, which are produced by very similar processes in the spallation of iron nuclei. The half-life of argon 39 is 270 years; the half-life of argon 37, 35 days. Therefore the concentration of argon 39 was indica-

tive of the mean cosmic ray intensity over the average radius of the meteorite's orbit, whereas the concentration of argon 37 was related to the cosmic ray intensity over the previous month or two, when the meteorite was close to the earth's orbit. The measured ratio of the two argon isotopes differed greatly from what might have been expected if the cosmic ray intensity had been the same at both locations. It was clear, in fact, that the intensity of the cosmic radiation at the mean radius of the meteorite's orbit was considerably greater than it was at the orbit of the earth. This result is particularly useful because various instrumented space probes, such as *Mariner 4*, that had directly measured the gradient of the cosmic ray intensity in space had given conflicting results.

Measurements of radioactivity induced by cosmic rays have been recently supplemented by an important new technique that can be applied to a rich variety of problems ranging from the dating of archaeological remains to investigations of sea-floor spreading. This is the technique of detecting the tracks left in solids by fast subatomic particles, originally pioneered by R. L. Fleischer, P. B. Price and R. M. Walker at the Research and Development Center of the General Electric Company [see "Nuclear Tracks in Solids," by R. L. Fleischer, P. B. Price and R. M. Walker; SCIENTIFIC AMERICAN, June, 1969]. A strongly charged particle, such as one of the heavier nuclei in the cosmic radiation, leaves a more or less permanent trail of ionization damage along its track when it passes through a dielectric (electrically insulating) material.

The region close to the track is left in a condition of high chemical reactivity. A mineral crystal in a meteorite that has been exposed for long periods to the cosmic radiation will therefore contain a fossil record of the damage left behind by the heavier particles. Fleischer, Price and Walker showed that if the crystal is immersed in a suitable etching agent, such as hydrofluoric acid, the agent will attack the crystal preferentially along the trails left by the particles and will form tracks that can be seen in a microscope.

A relatively light particle such as a proton does not cause enough ionization damage to produce etchable tracks in natural minerals. It turns out that only massive nuclei, which constitute a minute fraction of the cosmic radiation, can be detected in this way. This fact makes the method ideal for studying the heavy component of the cosmic radiation, since most other methods suffer from background interference from the vastly more numerous lighter components. (Let me note in passing that the lighter components do give rise to etchable tracks in dielectrics of lower sensitivity than mineral crystals, such as certain plastic materials. Some interesting data have come to light in the examination of, among other things, the plastic helmets of the Apollo astronauts following their exposure to the cosmic radiation during their walks on the moon.)



RATE OF FORMATION of spallation products depends on the depth to which the cosmic rays penetrate. Sodium 22 is produced from iron only by high-energy cosmic rays; its production drops rapidly with depth below the meteorite surface (*bottom curve*) as the flux of primary cosmic ray particles decreases. Cascade of lower-energy secondary particles reaches maximum between 10 and 20 centimeters below the surface. Production rate of manganese 54 (*top curve*), a spallation product of secondary particles, reaches maximum 20 centimeters below the surface, then drops off as flux of lower-energy particles is absorbed.

Analyzing the tracks found in meteorites requires great care, because the tracks from the passage of cosmic ray particles must be distinguished from the tracks arising from other sources. Meteorite crystals containing uranium, for example, show tracks from the spontaneous fission of uranium 238. In this case the two fragments of the fissioning nucleus recoil in opposite directions and produce ionization damage. When Fleischer, Price, Walker and their collaborators were investigating fission tracks, they found that the number of spontaneous fission tracks was much greater than could be explained by the concentration of U-238 in the sample. They suggested that the additional tracks had been produced by an extinct radioactive isotope: plutonium 244. This isotope, like the more familiar members of the uranium and thorium series, would have been produced in the nuclear processes that gave rise to the material of the solar system. Because of its short half-life, however, it would by now have decayed to undetectable proportions.

There are still other sources of tracks. Although the light particles of the cosmic radiation are not themselves massive enough to leave an etchable track, they can collide with nuclei in the meteorite and cause either spallation or fission. Both of these events can lead to the recoil of fairly massive nuclei, which may leave a detectable track.

All these kinds of tracks have certain characteristics that make them distinguishable from the tracks of genuine cosmic radiation. The tracks of cosmic ray primaries, for example, differ from those arising from spontaneous fission in that their density is independent of the local concentration of uranium in the sample under study. Moreover, the number of tracks from the cosmic radiation decreases rapidly with depth below the

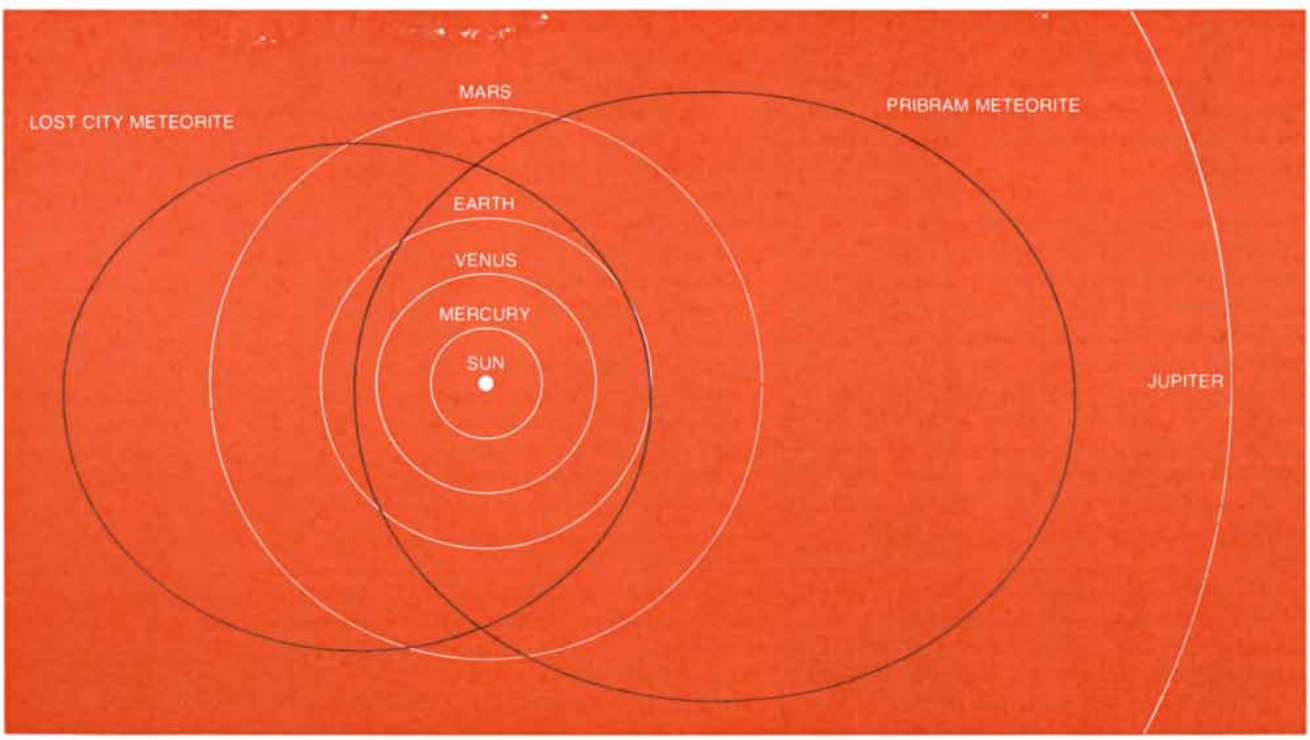
surface of the meteorite. This fact can be used along with the contours of the equal concentration of radioactive isotopes to reconstruct the shape of the meteorite's surface before it entered the earth's atmosphere and thus to yield information on the amount of ablation the meteorite suffered.

The density of the tracks in the surface region of the meteorite can also be utilized to estimate the extent of the erosion in space. If the rate of erosion is high, the density of the tracks near the surface of the fallen meteorite will be low, because during the early part of the meteorite's history its subsurface region would have been shielded by a considerable depth of material that was later eroded away. Price and his co-workers examined the density of tracks in the surface region of the Patwar meteorite. Their results imply that there is an upper limit to the rate of space erosion of about one millimeter per million years. That rate is much less than had been suggested was necessary to explain the discrepancies in the estimates of the exposure ages of stony meteorites determined from cosmogenic radioactive isotopes of differing half-lives.

Another feature of particle tracks is that the ionization damage tends to disappear through annealing when the mineral is heated above a certain temperature. Hence it is possible to estimate the rate at which the parent body of the meteorite cooled. The initial temperature of the parent body is too high to retain the potential tracks produced by spontaneous fission of uranium or plutonium within it. As the body gradually cools, however, the various mineral crystals in it will fall below the critical temperatures at which the radiation damage is eliminated by annealing, and each mineral will begin to accumulate a permanent record of tracks. Each mineral is characterized by its own critical temperature for retaining the tracks, and these temperatures can be determined in the laboratory. The minerals with the higher critical temperatures will therefore show a greater track density with respect to the concentration of uranium 238 within the meteorite. The track density can be correlated with the concentration of uranium in the different mineral crystals; it is this procedure that yields an estimate of the rate at which the parent body cooled.

Fleischer concluded from such a study that the parent body of the meteorite known as the Toluca iron meteorite cooled at a rate of about one degree C. per million years. That rate corresponds to a body with a radius of some 200 kilometers. It was also clear from his results that the great majority of the tracks had been formed by the spontaneous fission of plutonium 244 rather than of uranium 238. The very fact that this relatively large quantity of plutonium was still present at the time when the parent body solidified, in spite of its short half-life, confirms the hypothesis that the interval that elapsed between the final nuclear synthesis of the material of the solar system and the time when the planets were formed was much shorter than the total time the solar system has existed.

One of the most useful characteristics of the particle tracks is that the length of the track produced by a high-energy cosmic ray nucleus traveling through a mineral crystal is a function of the charge carried by the nucleus. The statistical distribution of track lengths therefore directly indicates the proportions of the different nuclei present in the radiation. The long exposure times of the meteorite crystals not only allow statistically significant numbers of tracks to accumulate but also afford an opportunity of discovering whether or not there have been appreciable variations in the intensity of the various components of the cosmic radiation over the exposure age of the meteorite. M. Maurette and his collaborators at Washington University concluded from measurements of track densities in



ORBITS OF THE TWO METEORITES whose trajectories have been observed by photographic recording networks show that they originated in the asteroid belt between Mars and Jupiter. The Lost City meteorite is the more recent of the two; it fell in Oklahoma on January 3, 1970, and was recovered quickly enough so that various workers could measure half-lives of many spallation products.

the Saint-Séverin chondrite that the intensity of the component of heavy particles in the cosmic radiation averaged over the 12-million-year exposure age of the meteorite was essentially the same as it is at present.

Since the length of the track in the solid depends on the electric charge of the particle, such tracks can be used in the search for "superheavy" elements in the cosmic radiation. With the exception of the recently discovered plutonium 244, no element whose nucleus has a charge greater than that of uranium has been found in nature, although heavier elements can be produced artificially in particle accelerators or nuclear reactors. The reason that transuranium elements do not exist in nature is that as the number of protons and neutrons in the nucleus increases, the positively charged protons repel one another so strongly that the nucleus becomes very unstable and decays rapidly either by spontaneous fission or by emitting an alpha particle. Superimposed on this general trend toward instability, however, is an opposing tendency toward greater stability when the numbers of protons and neutrons in the nucleus are close to certain "magic numbers." These numbers correspond to the complete closure of the shells in which the particles within the nucleus are arranged. It has been suggested that an "island of stability" may exist for elements with about 110 protons. Nuclei in this region might be expected to have half-lives of up to as much as 100 million years.

Peter H. Fowler and his co-workers at the University of Bristol found the first experimental evidence suggesting that a superheavy nucleus might exist in the cosmic radiation. They discovered a very densely ionized track in special photographic plates flown in a high-altitude balloon. Their identification of the track was not totally conclusive, however, and a number of workers have since searched for definite evidence of tracks left by transuranium nuclei in meteorites and in lunar rocks brought back by the Apollo astronauts. Several unusually long tracks, which may have been made by superheavy nuclei, have been found. The search is vigorous, since proof that superheavy elements exist would be of great interest to theoretical nuclear physicists. It would confirm their methods for predicting the properties of nuclei. Moreover, the presence of relatively short-lived nuclei in the cosmic radiation would throw light on how the cosmic ray particles diffuse through the galaxy and also help in deciding to what extent the radiation originates within our own galaxy or reaches us from more distant regions of the universe.

NUCLEAR TRACKS in a crystal of the Fayetteville meteorite, which show up as the thin, short dark lines, were formed by the nuclei of iron and some heavier elements. They were photographed at a magnification of 60,000 diameters with a high-voltage electron microscope by David J. Barber and P. Buford Price, Jr., of the University of California at Berkeley. To reveal similar tracks to a microscope working at visible wavelengths, the crystal would have to be etched with a reagent such as hydrofluoric acid or hot sodium hydroxide. The tracks resulted when the crystal was bombarded by cosmic radiation from solar flares early in the history of the solar system, before the crystal had become part of a meteorite. The two dark bands diagonally crossing the photograph are not tracks but crystallographic features.

As we have seen, an enormous range of information may be stored in a few cubic centimeters of meteoritic material. Within the past several years this information has been supplemented in a most exciting way by the results of the study of the material brought back from the moon. The lunar samples have certain advantages over meteorites. For example, interpreting the concentrations of radioactive isotopes in meteorites in terms of variations in the cosmic ray intensity is complicated by the difficulty of distinguishing between the effects of space and the effects of time, since in general the orbit of the meteorite is not known. The spatial uncertainty is removed in the case of the moon, whose orbit is known with high precision. Another advantage of the lunar material is that it contains a record of the effects of low-energy cosmic rays, which penetrate a very short distance below the surface. In a meteorite this record is destroyed by ablation during passage through the earth's atmosphere. The study of the cosmic radiation both in meteorites and in lunar material, and the comparison of the results, should yield some interesting facts over the next few years.

The Brideprice of the Sebei

In this African culture getting a bride costs a man a substantial fee in cattle, cash, clothing, pots, tobacco, milk, beer and other commodities. To what extent is the transaction strictly economic?

by Walter Goldschmidt

It is well known that marriage has its economic aspects, but in some societies those aspects seem more explicit than in others: the bride or the groom brings some form of dowry to the altar or the groom pays a substantial fee to the family of the bride. The two transactions, it should be noted, are not equivalent: the dowry assists the new couple whereas the brideprice benefits only those the bride is leaving. The prevailing anthropological view of brideprice is that the transfer of goods to the family of the bride cannot properly be regarded as a purchase. Many have even wanted to avoid the term brideprice entirely. In general anthropologists have argued that the goods are a "prestation," which has been defined as the act of paying in money or service what is due by law or custom. Pointing out that men do not buy a wife in order to make a profit (although in one society they are reported to occasionally resell a wife), they have gone so far as virtually to deny that the transaction is in any sense economic. To test the proposition that the transaction is economic an observer might undertake to see if the society he is studying perceives brideprice transactions in the same way that economic transactions are perceived among us. He would note the attitudes brought to the negotiation, the terms used, their compatibility with the terminology in other kinds of haggling, and so on. Such an approach, however, is beset by two difficulties: How in fact does the society in question actually perceive brideprice transactions? And exactly how are "economic" transactions defined in our own society?

A different approach to the question can be made. It requires only that the observer determine whether the marriage obligations of the society being studied are susceptible to straightforward economic analysis. Will or will not a body of data relating to the marriage contract yield findings that make sense in terms of economic decision making? With such a procedure it is not necessary that the investigator know if the society perceives its marriage obligations as being essentially economic ones. The question is no longer whether or not the marriage contract is an economic transaction. Instead the observer asks a simpler question: To what extent are the social choices dependent on economic considerations as distinct from matters of sentiment?

I undertook this second approach in 1961–1962 while working among the Sebei, a Southern Nilotic people who live on the northern slopes of Mount Elgon in Uganda. I collected information on brideprice contracts negotiated by 79 men in the years between 1910 and 1960. Normally my informant was the married man himself, but in a few cases it was the father of the groom. My informants were drawn from two contrasting environments in the Mount Elgon area. The data on 29 of the marriages came from Sasur, a Sebei farming community situated in the well-watered Sipi district at the western end of the Elgon escarpment. It is an area of intensive hoe agriculture largely devoted to raising bananas and coffee. The data on the other 50 marriages came from Kapsirika, a Sebei pastoral community in the drier plains district of Ngenge below the mountain to the north, where the economy is dominated by cattle herding [*see illustration at left on opposite page*]. Thus the men of Sasur are predominantly farmers and the men of Kapsirika are predominantly herders. The brides sought by the men of both communities came not only from these and other districts in the Sebei region but also from kinsmen across the border in Kenya and even from linguistically unrelated tribes.

I asked each of my informants to tell me the total amount he had contracted to pay as a brideprice: how much he had paid immediately and how much was promised for payment later. I also asked whether the union was the man's first marriage or a later one, what the bride's place of origin was, whether or not she had been married previously and who had negotiated the contract. I have no reason to doubt the essential accuracy of the responses I received. Brideprice negotiations are important social events. The contract is reached by means of formal bargaining conducted in the presence of witnesses. Large quantities of valuables are transferred, and the details of the contract are neither secret nor private. Some of my informants may, however, have understated the proportion of a brideprice that remained to be paid; a man might well be reticent on this subject.

As the total number of contracts indicates, I obtained more responses from the herders of Kapsirika than from the farmers of Sasur. This was partly because polygamous marriages are twice as common among the pastoralists. The 50 Kapsirika responses represented more than half of the married households in that community, whereas the 29 Sasur responses came from little more than a third of the married households. I do not believe the imbalance in responses has distorted the data. A more serious potential flaw is the small size of my total sample. It should be made clear that my conclusions relate only to the Sebei. Although statistical details on the actual brideprice payments elsewhere in Africa are not available, it is known that there is great variation in the attitudes, social expectations, legal rights and in-

dividual duties associated with such payments. Whether what applies to the Sebei applies to any of their immediate neighbors in closely related cultures, let alone to other African peoples, is beyond the scope of this discussion.

The central dynamic in Sebei social history stems from developments in the 19th century or earlier. Principal among them was a gradual shift from a primarily pastoral economy to one that has become increasingly involved in the raising of garden crops. Exactly when the shift began is uncertain, but it continued after the 19th-century European colonization of East Africa. This historical change is reflected in economic and social variations among the Sebei today that are largely responses to differences in environment. As I have mentioned, the area of hoe agriculture is at the western end of the Elgon escarpment. The Sebei who live to the east of this area are progressively less concerned with crops, whereas on the semiarid plains lying to the north of the mountain herding continues to be the principal livelihood.

Like many other African cattle keepers, the Sebei are organized into clans that reckon kinship in terms of male descent. This is to say that clan relatives are counted on the father's side rather than the mother's. The clan relationship is of great psychological importance to the Sebei in a number of contexts. For example, although clan authority is not involved in the marriage-contract negotiations, considerations of clanship do affect the choice of a marriage partner. Furthermore, the clan has a vital interest in the institution of marriage. Only through legal unions and the offspring they produce can clan continuity be assured, and continuity of the clan is a major social goal.

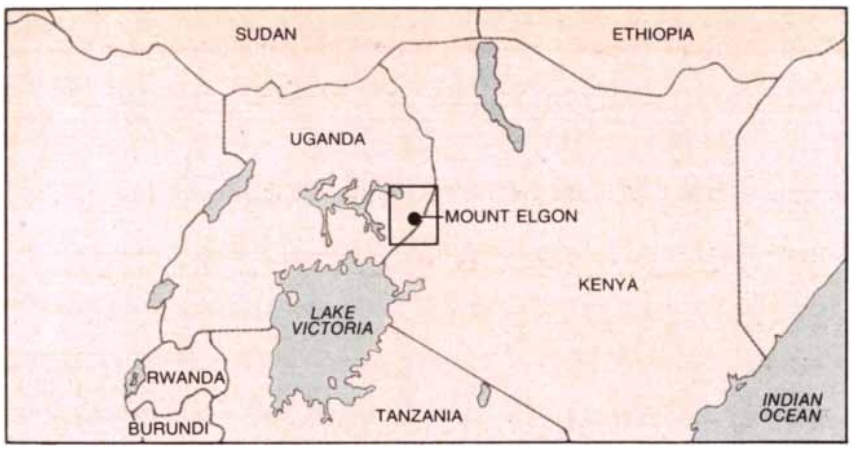
Every Sebei woman and almost every Sebei man gets married. Although today a number of divorced Sebei women live alone and get along quite well, they are viewed as being unusual. They are somewhat resented by the men and, I suspect, are envied by the other women because they have escaped male domination. The essence of Sebei marriage is the husband's concern with procreation. And it is this concern with the acquisition of descendants, reflected at the clan level, that is the essence of the brideprice contract. The payment establishes a legitimate claim by the husband's clan on the progeny of his marriage. The compelling demand for the husband to legitimize the clan's claim is the reason why, even when a man inherits his brother's widow or marries a much divorced woman, there will be at least a token brideprice payment: the transfer of a single head of cattle to the woman's family. It is a basic Sebei belief that immortality depends on having descendants. Marriage is thus a truly central element in the personal history of the Sebei, and having progeny is of the greatest importance to both marriage partners.

The socially preferred form of marriage among the Sebei is polygyny, in which a man can have more than one wife at a time. A man normally first marries when he is in his early or middle 20's; a woman, when she is younger than 20. The young man's father is expected to provide the brideprice for his son's first marriage; he also negotiates the marriage contract. The father may help with his son's subsequent marriages, but he is not socially obligated to do so, and it is the son who usually negotiates the later marriage contracts. Although Sebei men are much in favor of polygyny, the practice is far from universal. During the period of my study only 17 percent of the Sasur men had had more than one wife at a time. Polygyny was twice as common among the Kapsirika men, but more than 60 percent of them were monogamous. Even this limited number of polygynous marriages makes women scarce among the Sebei.

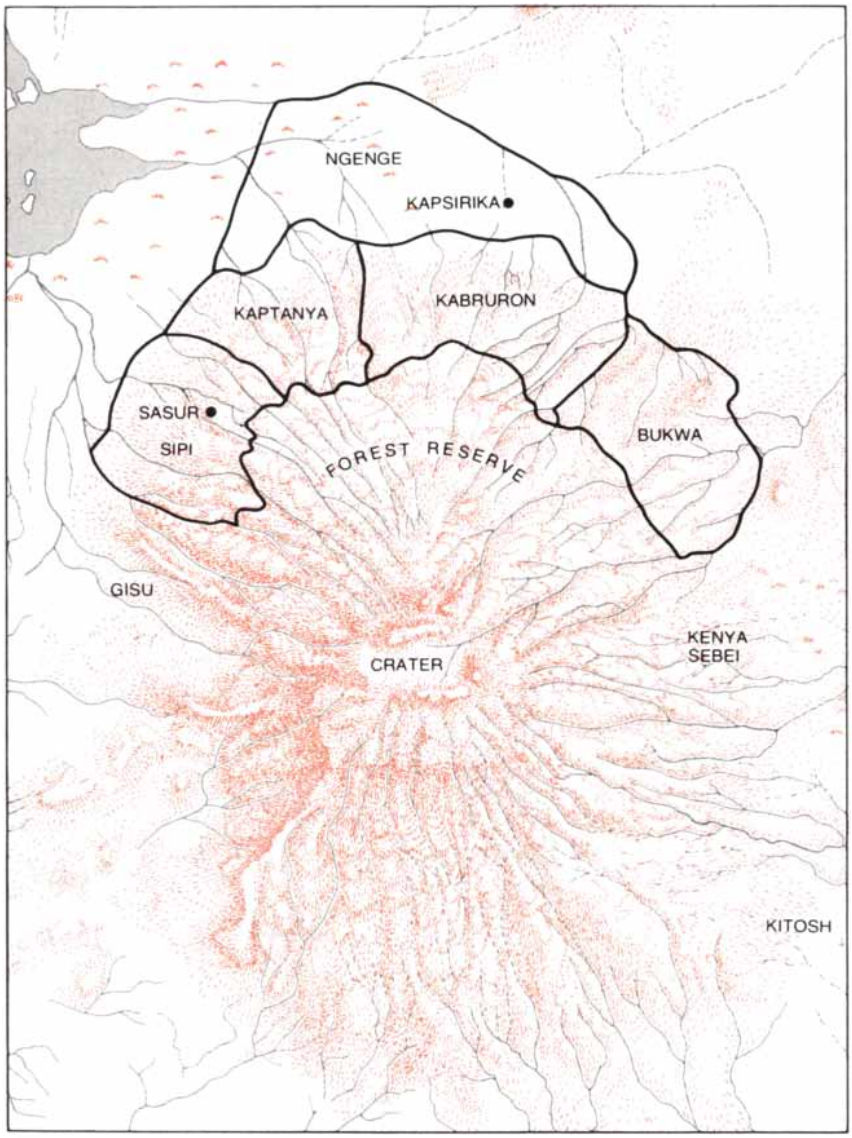
The Sebei marriage contract is essentially an agreement between two men: the prospective groom or his father on the one hand and the father of the prospective bride on the other. Other parties—particularly brothers, mothers and mothers' brothers on the bride's side—have legitimate interests in a prospective marriage and a potential stake in the outcome of contract negotiations. In principle, however, the choice of the marriage partner is entirely in the hands of the two men; even the prospective bride often has no say in the matter. I found

KAPSIRIKA HERDSMAN, a member of one of the two groups in the brideprice study, drives his cattle out to graze. In brideprice contracts a standard value for cattle is 100 shillings ($14) a head.

SASUR FARMING WOMEN, members of the other group in the brideprice study, prepare a hillside for planting. Bananas are Sasur's main crop, but corn and root crops are also cultivated.



SEBEILAND lies on the northern slope of Mount Elgon, an extinct volcano in Uganda near the Kenya border north of Lake Victoria. The Sebei population is about 35,000.



ESCARPMENT on the northern slope of Mount Elgon lies generally between 5,000 and 7,000 feet above sea level. To the west, in the well-watered Sipi region, bananas are the main crop. Eastward on the escarpment the banana crop is replaced first by grain and then by animal husbandry. The plains north of the mountain lie below 4,000 feet; there animal husbandry predominates. Data on brideprice were collected in an agricultural community, Sasur in the Sipi district, and in a pastoral community, Kapsirika in the Ngenge district.

a good deal of evidence that this not only had been true in the past but also was to a large extent true today, although efforts are now made to prevent the forcing of women into unwanted marriages.

Nowadays there is a prelude to many, if not most, Sebei marriages in the form of a period when the young man and woman simply live together, an arrangement that English-speaking Sebei call "elopement." There is no stigma attached to these liaisons, although it seems to be considered more genteel if a contract has been arranged beforehand. One way of establishing such a liaison, now on the decline but frequent in the past, is abduction. A kind of mock capture also seems to have been a way of establishing a liaison for a couple who faced opposition to the marriage by one father or both. Whatever the mode of choice, the "elopement" relationship must eventually be legitimized by a transfer of goods from the groom's side to the bride's. This means that in most instances there is a bargaining session sometime after the liaison is established.

In contract negotiations the principal spokesman for the bride's side is not her father, the person whose interests are paramount, but a neighbor who enjoys a quasikinship (called *latyet*) with the father. The spokesman for the groom may be his father, a brother or the groom himself. The interested parties meet in the bride's community, preferably at the house of the bride's mother. Relatives of both parties, including senior women, may be present during the negotiations. The women generally remain silent throughout, although they may object at the end or demand the special compensation of a goat or a cow for some past insult from the groom or his clansmen. Although a Sebei man will, if he is asked, quote an average brideprice, the brideprice reached in any one negotiation can fluctuate widely from the average. I have found, both when observing a negotiating session and when discussing the problems involved in negotiations with grooms and their representatives, that the groom's side approaches the session anxiously, fearful of "defeat" by the demands of the bride's side.

The pattern of give-and-take in a negotiating session requires that the bride's side make a demand that the groom's side either accepts, rejects or attempts to negotiate further. The procedure is for the neighbor first to make a demand for cattle, setting out a tally for each animal. When the number of cattle has been agreed on by the two parties, they put

the tallies aside and bring out others to represent the demand for sheep and goats, following the same procedure. Next they usually turn to cash, then in no set order to the various other kinds of goods that the bride's father has determined to seek. When each of the items has been agreed on, that is the final contract, although the father may also request a gift or demand a "fine" that is not part of the contract at all.

Yet the contract does not constitute a binding agreement at this stage, as it would with us. In principle either party can withdraw from the agreement at any time up to a particular point in the actual marriage ceremony. They may not, however, open negotiations to alter the agreement. Nor may either party enter into a marriage-contract negotiation with anyone else. Such an action is prohibited until the provisional contract is voided, and I have never heard a Sebei complain of either party failing to observe this rule.

BRIDEPRICE NEGOTIATION, called "breaking the sticks," is conducted at the home of a Sebei bride's mother. Tallies spread by the representative of the bride's father show how many items of different kinds the bride's family expects to receive. The groom's representative (*arm extended*) may accept or refuse the demand or seek to reduce it. Cattle are the first items to be negotiated. Originally the bride's father kept possession of the tallies, returning them to the groom one by one as each item of the brideprice was delivered.

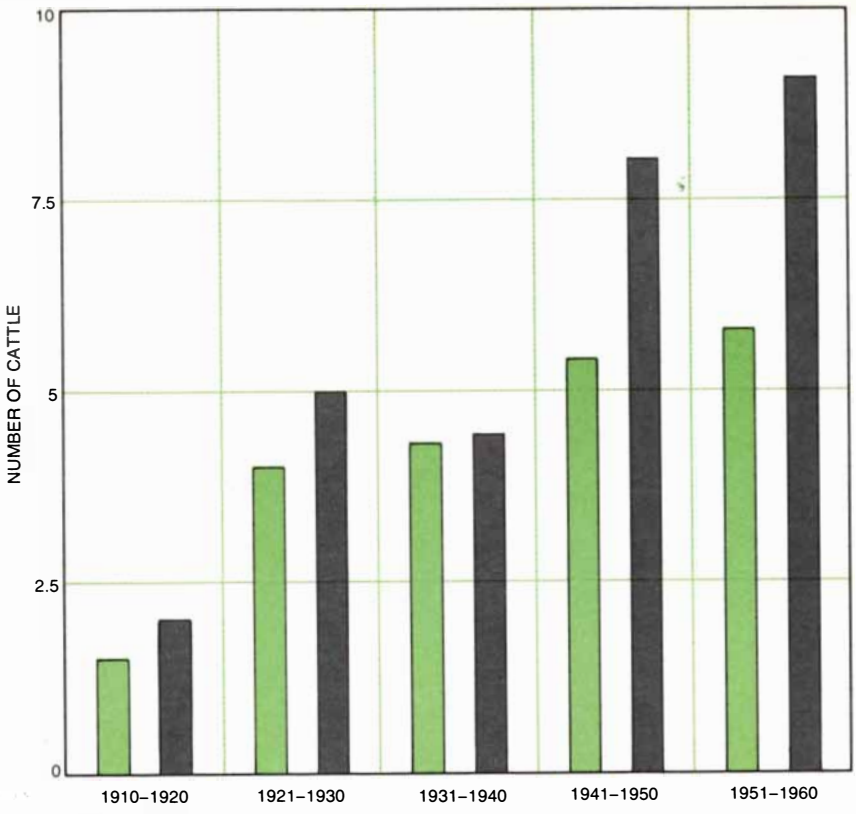
Brideprice is the largest single outlay of goods that an ordinary Sebei makes in his lifetime. In proportion to income that outlay is close to the rule-of-thumb calculation of what an American spends in buying a home, ordinarily in our society an individual's greatest single capital outlay. In terms of the specific items involved cattle were a good part, and usually the largest part, of the brideprice in all 79 marriage contracts. Other kinds of livestock also changed hands in nearly all the marriages: goats in 72 of them and chickens in more than half. Some payment of cash was included in almost 90 percent of the contracts, and 70 percent called for the delivery of beer and tobacco. Minor consumable items figured in fewer than 12 percent; among them were kerosene for lamps, home-distilled gin and milk.

The Sebei brideprice always includes at least one head of cattle. At one time it also always included in the class of personal effects an iron bracelet and an iron hoe blade for the bride's mother. This tradition is weakening today; hoe blades appeared in fewer than 60 percent of the marriage contracts and bracelets in fewer than half. Four other kinds of personal effects—blankets, coats, lengths of cloth and aluminum pots—first began to be included in the marriage contracts in the 1920's. They seem to have been immediately popular; blankets appeared in 75 percent of the contracts and coats, cloth and pots in at least half. Less popular personal items, all introduced after the 1920's, included men's robes, iron pans, mosquito nets and kerosene lamps.

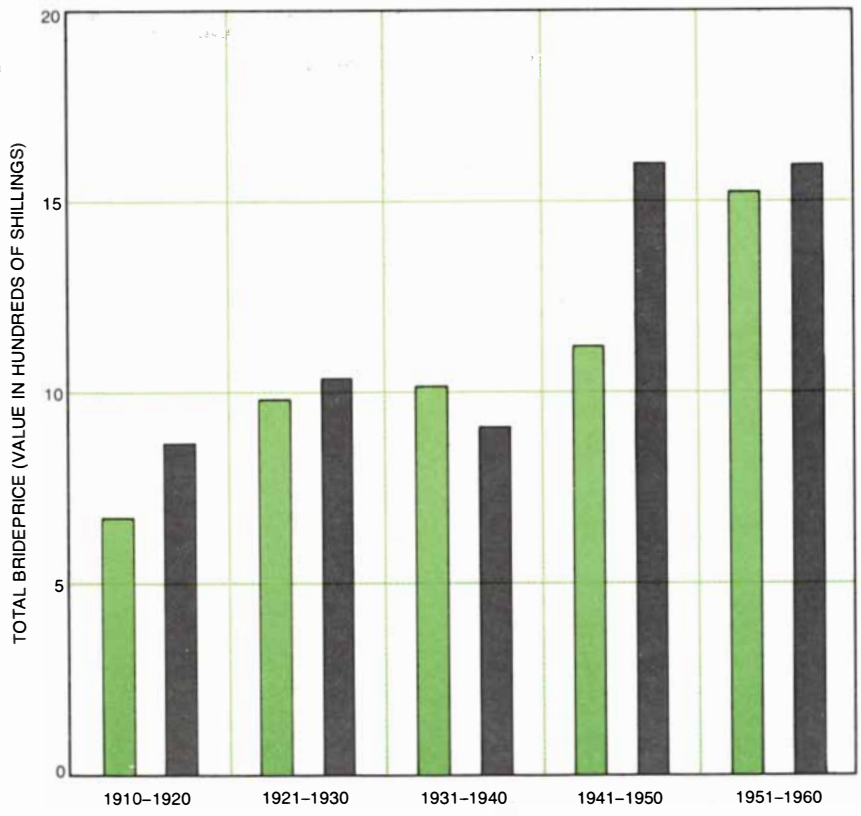
The Sebei make a distinction between the goods included in the negotiated contract and gifts. Gifts may be requested by the father of the prospective bride, and the prospective groom may feel strongly obligated to give them, but the gifts are not part of the brideprice. I have restricted my analysis to recognized brideprice items. The "fines" frequently levied against the groom for some action of his clansmen are not regarded as part of the brideprice and must be paid separately. They have therefore not entered into my calculations.

As the priority given them indicates, cattle are certainly the central element in the Sebei brideprice and were probably the original element. At least one animal should be transferred in order to make a marriage legal. Goats are also regularly included in the contract. Originally the transfer consisted of a single ram intended for the bride's mother's brother; now as many as six or eight goats are paid. Beer, milk and perhaps tobacco appear to be as traditional as the iron bracelet and the hoe blade, but the other consumables and personal effects are later additions to the list. Kerosene as a consumable, and the lamps to burn it in, first appeared in a 1951 marriage contract; home-distilled gin was first mentioned in 1955. The inclusion of some cash in the contract is by now all but traditional; it began in 1914 when the currency of East Africa was still the rupee.

To undertake an economic analysis of Sebei marriage contracts it is necessary to translate the value of the goods into money values. I have used the East African shilling (with an exchange value of seven shillings to the U.S. dollar) as my unit of currency; Uganda has its own currency today but the shilling was still in use at the time of my fieldwork. I asked my informants to estimate the value of brideprice items that the groom



NUMBER OF CATTLE in the contracts of 29 Sasur husbands and 50 Kapsirika husbands from 1910 to 1960 is seen in the graph. The Sasur (*color*) consistently gave fewer cattle.



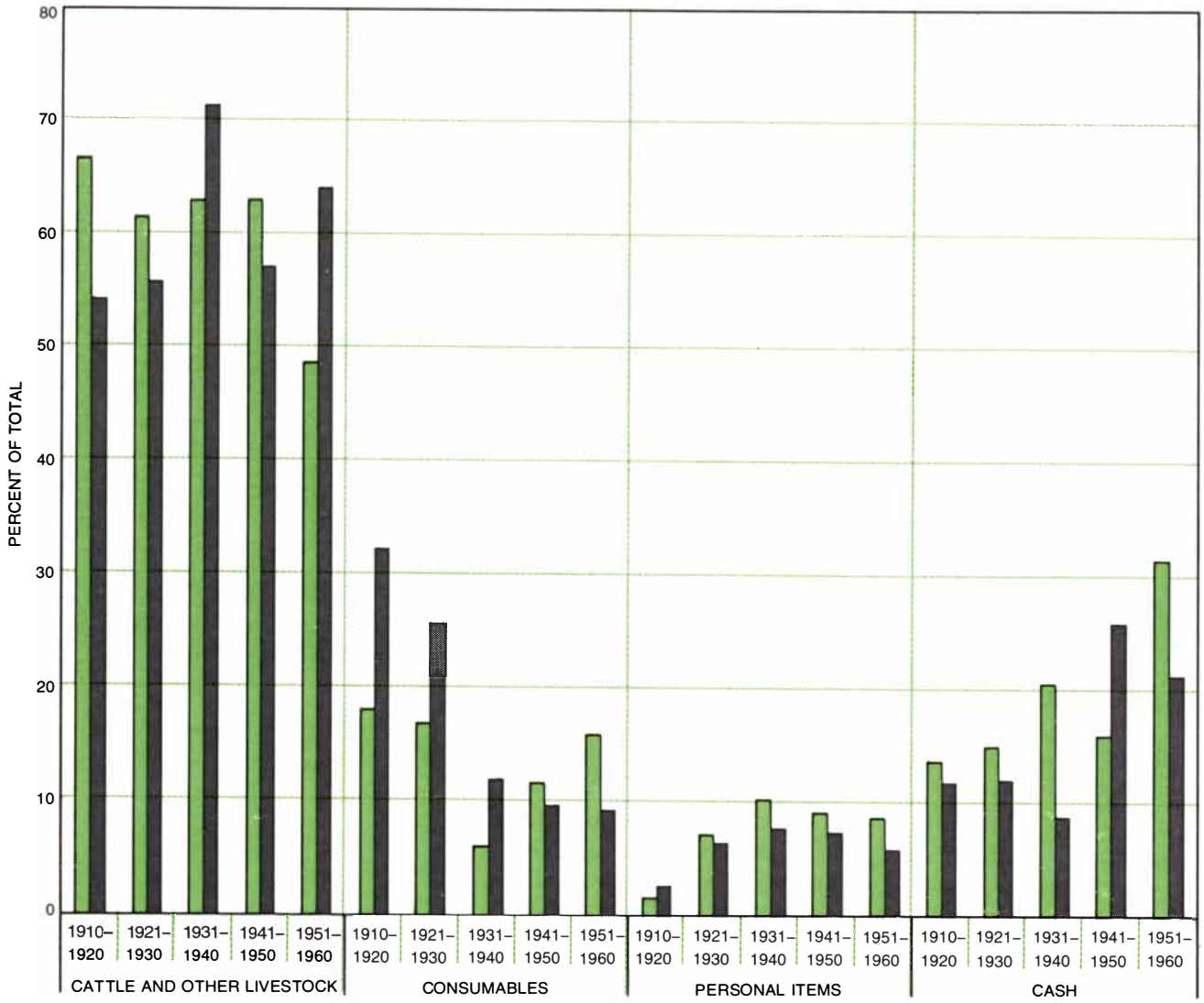
TOTAL BRIDEPRICE paid over the same 50-year period more than doubled for Kapsirika men and almost tripled for Sasur men (*color*). Slump in the 1930's reflects world depression.

had to purchase in terms of current prices. In some instances the informant might remember the exact price he had paid for some particular item; if he did, I used the actual value in my calculations. It soon became evident that a standard of expected quality existed. For example, in the course of the bargaining the bride's father might express a wish for a coat or a pot of better quality or larger size than was standard, and he would agree to pay the difference in cost himself.

I assigned to all cattle a standard value of 100 shillings per head, in spite of the fact that Sebei cattle vary widely in price and that cattle selected for transfer under a marriage contract are not all the same quality. I chose this figure because it is a substitution the Sebei make whenever they choose to make a money payment in lieu of some of the agreed-on cattle, as they occasionally do. I arrived at the shilling value of goats and sheep through the same process of utilizing standard Sebei substitutions. I find further evidence that brideprice cattle are thought to have a uniform and standard value: no Sebei man ever attempted to explain a seemingly high or low brideprice by referring to the quality of the brideprice cattle he had contracted to transfer.

The prices of consumable goods were harder to determine. For beer there was an established price of eight shillings per *debe* (four imperial gallons), so that I was able to establish money values by estimating the number of *debes* represented by my informants' counts of pots or sacks of beer. (The Sebei dry the fermented grain mash on large flat sheets in the process of beermaking. Such "dried beer" can be kept for several weeks, and sacks of it can be used for brideprice payments.) I have estimated a similar unit value for gourds of tobacco or milk. On balance I think my figures are a reasonable estimate of the 1962 shilling value of the various brideprice items.

Prices, of course, have changed greatly over the 50 years covered by the brideprices on which I gathered information. It is manifestly impossible to establish these changes with accuracy, but fortunately it is not necessary. My use of uniform prices throughout the period produces an evaluation in what economists call "constant" shillings. A steady reduction in the exchange value of the shilling over the same period, however, presents a greater problem. To adjust the values so that they reflect this inflationary factor is at best a highly



BRIDEPRICE GOODS fall into four categories: livestock (cattle in particular), consumable items such as beer, personal effects such as pots and blankets, and cash. Over 50 years the cash part of the brideprice has grown larger and the personal-effects part smaller.

technical procedure, fraught with inaccuracies. I have chosen not to attempt it, and the reader should be warned that my analyses dealing with the rise in the brideprice over a period of time and with changes in the proportion of various classes of goods are all influenced by this factor to an undetermined extent.

Over the 50-year period the lowest brideprice paid by any of my 79 informants was 288 shillings ($33) and the highest was 2,943 shillings ($420). The number of cattle stipulated in the contracts ranged from one to 17. The average value of the cattle was $192; the average number of cattle, 6.65 head. It is meaningless, however, to consider the total dollar value of the contracts or the dollar value of the cattle transferred except as these values relate to the Sebei economy. I therefore offer some measures by which the meaning of these payments can be understood. When an ordinary peasant works for others or works as a common laborer for the government, he receives two shillings per day. If he worked full time (which very few of them in fact do), his annual income would be about 600 shillings. A not very different average figure—643 shillings—was arrived at by inquiring of my informants what they had received in cash through sales and employment in the preceding year. Thus we can see that the brideprice paid by the Sebei in the later years of my analyses (around 1,600 shillings in the 1950's) is equivalent to about 30 months' income for the average Sebei, exclusive of what he produces for home consumption. The brideprice is also high when it is considered in terms of an individual's capital assets. Land is the main capital among the Sasur farmers, but because its value is difficult to calculate my capital-value comparison is confined to Kapsirika men. The value of the average Kapsirika herder's stock in 1962 (calculated at 100 shillings per head of cattle and 18 shillings per goat) was 2,159 shillings ($308). This means that an average brideprice of about 1,600 shillings ($229) equaled nearly three-quarters of a Kapsirika man's total capital value.

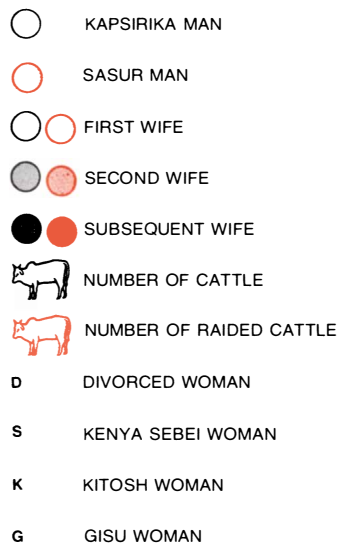
The decades between 1910 and 1960 witnessed a gradual long-range economic improvement among the Sebei. At the turn of the century they had been struck by the triple disaster of drought, disease and European rule. For example, early in the 1900's cattle were so scarce that some Sebei herders were forced to make do with goats and sheep. As time passed, however, the goods that enter into brideprice contracts gradually became more abundant. Under these circumstances economic theory calls for an increase in the average

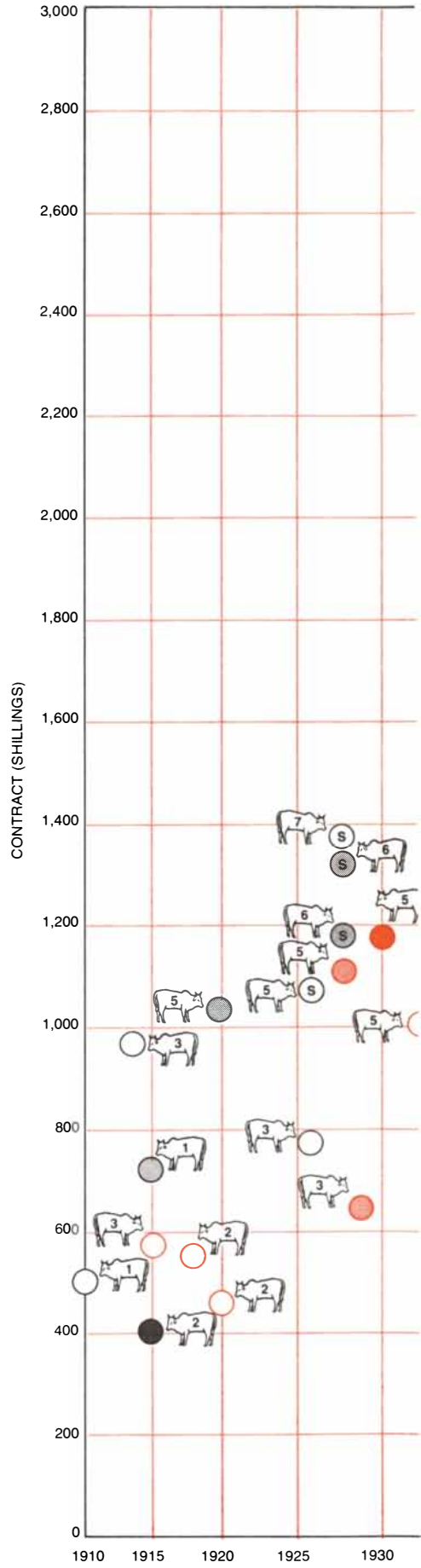
brideprice payment. A review of community marriage records decade by decade shows that the average number of cattle included in Sebei brideprice payments has more than doubled over the half-century [*see top illustration on page 78*]. This trend has been strongest among the Kapsirika herdsmen. Between 1910 and 1920 the average number of cattle paid by a Kapsirika groom was 2.4 head; by 1951–1960 the number had risen to 9.1. In the 1910–1920 period the brideprice paid by the farmers of Sasur included an average of 2.3 head of cattle. These payments too increased as the years passed, but they leveled off at about six head of cattle in 1951–1960. Like the price of cattle, the total brideprice paid was higher in the pastoral community than in the farming one. The trends, however, are parallel. The total brideprices paid by grooms in both communities by then had nearly tripled. Brideprice payments can thus be seen to reflect the growing prosperity over the half-century, as would be expected in terms of economic theory.

This relation between brideprice and prosperity would appear to be reflected in the slight decline in brideprice evident in the 1930's, particularly among the Kapsirikans. I believe this is not a statistical artifact but a response to the worldwide depression during the 1930's. For example, the proportion of cattle in the total brideprice increased during the same period. Such a shift decreased the groom's expenditure of money, in terms both of cash and of store goods purchased, and simultaneously increased his expenditure of a commodity internal to the Sebei economy: cattle. Because my presentation of brideprice data assigns a constant value to each head of cattle, the effect of this shift in proportions is masked. My informants have told me in other connections that the price of cattle was very low in the 1930's. Furthermore, if the data were adjusted to reflect the increased purchasing power of cash at the time, the decline in brideprice would be even larger.

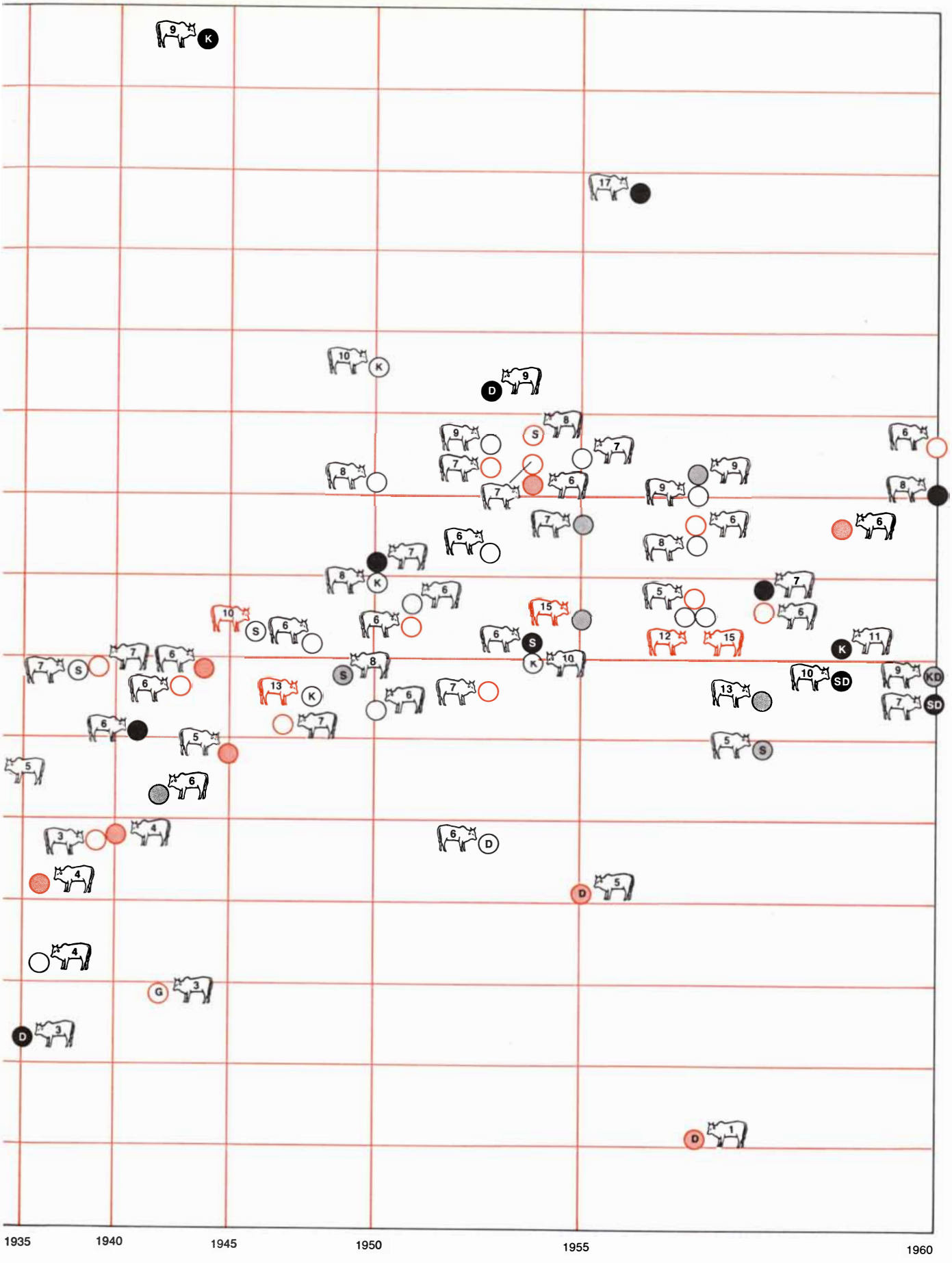
There is another way of perceiving the influence of economic factors as they are reflected in brideprice: the differences in payment demanded in the two communities. We have already seen that the price is consistently higher in Kapsirika than in Sasur. Why should this be so? It cannot be argued that the Kapsirikans are economically better off; it is not possible to obtain the information by which such a relative prosperity can be measured. It is true, and significant, that the pastoral peoples pay a larger proportion of the total brideprice in cattle, whereas the farmers, who have more cash thanks to the production of coffee, a marketable crop, pay more of their brideprice in money. That, however, would not account for the larger total payments in Kapsirika.

To answer this question we must turn to another aspect of Sebei marriage patterns: the relative prevalence of polygyny. Although the men generally feel that plural wives are desirable, only a few of them actually have more than one wife. Yet this very fact makes women scarce. Significantly there is no such thing as an old maid in Sebei. The more polygynous marriages there are, the more difficult it is to get a wife. As a measure of this scarcity it is best to think in terms of the number of women in the community who are currently married as a ratio to the number of men who are currently married. In a monogamous society the ratio should be 1:1. For Sasur the figures are 117:100; for Kapsirika they are 151:100. Thus the frequency of polygynous marriages has driven the price higher





BRIDEPRICE DATA for the 79 Sebei marriages between 1910 and 1960 appear in the graph at right. Vertical position of each circle shows the total value of the brideprice contract in shillings; numeral within the adjacent cattle symbol shows the number of cattle specified in the contract. (If the symbol is colored, cattle were seized in a raid.) The general rise in brideprice over the 50-year period is readily apparent, as is the differential between the brideprices for first and subsequent wives. Marriages of Kapsirika men and Sebei-speaking women from Kenya in 1926 and 1946 were marked by higher-than-average brideprices for those days. The Kapsirika men who married Kitosh women in 1944 and 1950, however, were leaders of a trend toward higher prices.



and the men of Kapsirika must pay a larger brideprice than the men of Sasur.

Why are Kapsirikans more given to polygyny than the men of Sasur? Essentially the reason is also economic. Polygyny is more suited to the pastoral form of economic endeavor than it is to farming. The acquisition of extra wives by the farmers of Sasur is made difficult by the limited availability of land, whereas such limitations do not apply to cattle keeping. The grazing land is open to all who have cattle.

Although I do not have information on the rates of polygyny throughout Sebei, I do know that the districts vary with respect to the relative importance of cattle keeping and farming. In general farming is more important in the west and becomes less so as one moves eastward across the escarpment; it is least important down on the plains where Kapsirika lies. In general the brideprice varies in conformity with this change. As I have noted, the men involved in these contracts all lived in either Sasur or Kapsirika, but the women came from all parts of Sebei. When I examined the brideprice paid for these women, I found that there was a significant correlation between the relative importance of cattle keeping in the area and the average payment made for women who came from these districts. The brideprice is affected by both the residence of the man and the place from which the woman comes. The more pastoral the residence, the higher the demand.

There is another effect of this difference in polygyny that points to economic expectation. Women from farming areas tend to marry men in pastoral areas much more often than in the reverse direction. The Sebei themselves recognize this in a saying: "Women marry to the east." I was never sure whether they meant by this an idealized standard of propriety or a recognition of reality, but it would seem to be the latter.

To observe this effect we must return to the marriages of the two communities. In Sasur 25 of the 29 marriages were within the Sipi district in which Sasur lies; one bride was from a neighboring tribe farther west (where farming is even more intense and a very low brideprice was demanded), and only three came from the east. In contrast, in Kapsirika only eight of the 50 marriages were within the plains district. Women married into the area from all parts of Sebei, and even from tribes outside Sebei. There were eight brides of Kapsirika men who came from the Sipi district, where Sasur is located.

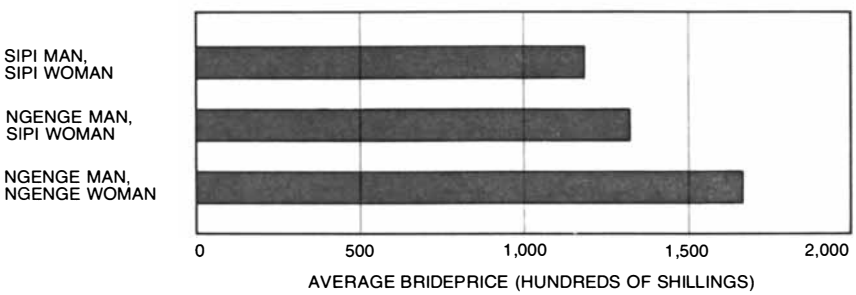
A closer look at the marriages that involve these two districts—Ngenge, in which Kapsirika lies, and Sipi, in which Sasur is located—gives further evidence that the location of both parties is relevant to the final outcome of the contract. There were 25 marriages in which a Sasur man took a wife from the surrounding Sipi district; the average brideprice of these marriages was 1,192 shillings. There were eight marriages in which a Kapsirika man took a bride from the surrounding Ngenge district; the contract called for an average of 1,677 shillings. That is a difference of more than 40 percent. There were also eight marriages in which Kapsirika men married women from Sipi; the brideprice was 1,336 shillings. That is about halfway between, as would be expected. Significantly there were no Sasur men in my sample who obtained a bride from Ngenge.

There is no standard of east-marrying propriety among the Sebei-speaking residents of Kenya, or for that matter among the Kitosh of Kenya, a Bantu-speaking group that has provided brides for Kapsirika grooms. Seven Kapsirika men have married Kitosh women and 12 have married Sebei-speakers from Kenya. The Sebei are conscious of the upward trend in brideprice over the years, and in their view one of the influences affecting this trend is the relatively high cost of wives from Kenya. The average brideprice of the seven Kitosh marriages was 1,719 shillings ($246). Kapsirika men recognize that the Kitosh price is high and that most of it is paid in cattle; their awareness suggests a rather sensitive appreciation of brideprice differentials. In the same general way they complain that the Sebei-speakers from Kenya have also contributed to the rise in brideprice. That is not true on the record; the average brideprice for a Sebei-speaking Kenyan is 1,357 shillings, or only $3 more than the average that a Kapsirika groom pays for a farmer's daughter from the Sipi area. Nonetheless, taking a longer view, the fact is that some marriages between the Sebei and their Kenyan kinsmen were among the very first to raise brideprices in general [*see illustration on preceding two pages*]. In this sense the Sebei perception is quite accurate.

What happens in the negotiations that take place after the man and woman are already living together? Put yourself in the place of the bride's father and consider what you would do. On the one hand you know that the young man is committed and perhaps you could ask for a higher brideprice; on the other your daughter is also committed to the man and therefore you are constrained not to interfere with her wishes. It offers an almost classic opposition between the rule of economics and the sway of sentiment. The Sebei assert that under these circumstances the man will have to pay a higher price.

My data support this Sebei contention to some extent, but the differences do not meet tests of statistical significance. Of the 49 Sebei marriage contracts negotiated from 1950 to 1960, 16 were arranged after an elopement. Seven of the elopements involved Sasur grooms; the average brideprice paid after the event was some 340 shillings higher than the brideprice paid by six other Sasur grooms who negotiated their contracts in advance. The nine Kapsirika men who eloped with their bride also paid more than the 33 who had negotiated in advance, although not as much more as the Sasur men. In any event the higher cost of elopement averaged 8 percent above the cost of the more conventional marriage procedure.

In the 1940's and 1950's there was a method of settling the brideprice after an elopement without recourse to con-



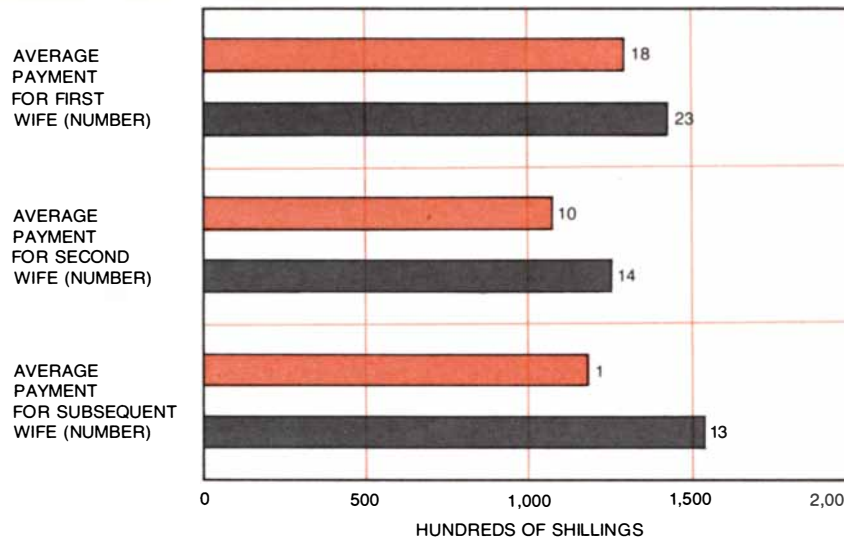
FARMERS PAY LESS FOR BRIDES and receive less for daughters, and when their children intermarry, the brideprice is low. The opposite holds for herdsmen: in such marriages the average brideprice is high. The price is intermediate when herders marry farmers' daughters. Marriages in the other direction are rare; none of them appeared in the sample.

tract negotiations. The father and brothers of the eloped girl simply raided the groom's cattle kraal and drove off what they considered an appropriate number of cattle in lieu of a formal brideprice payment. There were five such raids between 1946 and 1957. All involved Kapsirika men; this was to be expected because Sasur men do not have kraals full of cattle. It is instructive to examine the outcome of the raids in detail.

Two of the raids took place in the 1940's, when the average Kapsirika brideprice was 1,595 shillings. In one raid the number of cattle driven off was 10, or a nominal 1,000 shillings' worth; subsequent bargaining raised the brideprice to 1,476 shillings. In the other raid 13 cattle were driven off, and there were no subsequent negotiations. By the time of the raids in the 1950's the average brideprice had increased slightly. In two of the raids 15 cattle were taken; no negotiations followed. In a third raid 12 cattle were driven off; the father of the bride subsequently asked for an additional payment of 300 shillings, and his request was honored. The five incidents suggest that both parties to a marriage keep in mind a model of what is appropriate in the way of a brideprice and are quite sensitive to fluctuations in brideprice value.

Many aspects of raiding are advantageous. One is that the groom ends up paying a somewhat below-average brideprice. This, I believe, reflects the realities of the situation rather than any error in calculation. The man who raids his son-in-law's kraal gets possession of his cattle immediately rather than after what may be a considerable delay. Since the selection of the cattle is in his hands, he may also end up with superior animals. The son-in-law too may find being raided an advantage, although he has no real choice in the matter. Many grooms chafe at the numerous small demands normally involved in contract negotiations; some even say they would prefer to pay more in cattle and less in other goods, as the men who marry Kitosh women do, just to rid themselves of the nuisance of minor obligations.

On economic grounds alone, leaving aside considerations of sentiment, the brideprice of a Sebei divorcée should be lower than the brideprice of a previously unmarried woman. The reason is that the interest in progeny is the major Sasur motivation for marriage. My data support this notion. Four Kapsirika men and two Sasur men were among those who married divorcées between 1953 and 1961. None of the women was her second husband's first wife. The two who married the Sasur men were local women; both became second wives. One, who was married in 1957, had been married several times before. Her father requested only a token payment of 226 shillings ($32), the lowest payment in the entire sample. Two of the divorcées who married Kapsirika men were also second wives; the other two became third and fourth wives respectively. Only one of the four was a local woman; two were Kenyan kinfolk and the fourth was a Kitosh.



PLURAL WIVES, which the Sebei believe to be more costly than first wives, did not command consistently high brideprices. Contracts for 24 second wives averaged lower than the contracts for 41 first wives. Third or subsequent wives, however, commanded a higher average brideprice than either first or second wives, although individual prices varied widely.

Omitting the multiple divorcée, the average brideprice for the other five wives was 1,131 shillings, compared with the average Sebei brideprice of 1,568 shillings during 1951–1960. I encountered only one divorcée marriage contract with a brideprice that met or exceeded the average brideprice for a first wife, and the example was scarcely normal. A Kapsirika man had been in the process of negotiating a contract with the father of a girl from the Bukwa district, a pastoral area, when the girl eloped with another man. She later broke off the liaison and married the Kapsirika suitor; he paid her father a brideprice of 2,064 shillings. Because of her earlier elopement, the man she married considered her a divorcée. The price that her father was able to ask and that her suitor paid suggests, however, that her status as a divorcée was only nominal. On balance, then, it appears that the brideprice for divorcées is reduced among the Sebei just as economic theory would predict.

Although the incidence of polygyny raises the general level of the brideprice, it does not necessarily follow that the payment for a second wife is higher than the payment for others. Indeed, there is no economic theory that might explain the differential price for the second, third and additional wives as against first wives except the law of diminishing returns—and this quite clearly is not the Sebei attitude. Sebei men say that the price paid for second and subsequent wives is higher than the price paid for first wives. Although the data do not bear out this contention, they do suggest some interesting relations between brideprice and wife order. Of the 79 Sebei marriages contracted between 1910 and 1960, 41 were between a groom and his first wife. Of the total 23 grooms were Kapsirika men and 18 were Sasur men. The average first-wife brideprice was 1,366 shillings. Two dozen marriages, 10 Sasur and 14 Kapsirika, were with second wives; 14 more marriages, all but one involving Kapsirika husbands, were with third or subsequent wives. The average second-wife brideprice was 1,177 shillings; the average for a third or subsequent wife was 1,522 shillings. It is evident that men paid less for second wives than for first wives. The ranks of the second wives, however, include four divorcées. What is the average second-wife brideprice after the less expensive divorcées are eliminated? It rises from 1,177 shillings to 1,249, but the higher price is still more than a whole cow's value less than the price of a first bride. The Sebei contention that all plu-

ral wives are high-priced is obviously mistaken. The pattern is rather that the price of the second wife is low and the price of the third wife is high.

The misconception is explainable if we remember that it is not the husband but the husband's father who pays the brideprice for a first marriage, whereas it is usually the husband alone who meets the cost of subsequent unions. Now, there is a good deal of social pressure on a Sebei man to take a second wife. A Sebei husband therefore tends to enter a second marriage when he still can scarcely afford it. Under these circumstances it is only natural that he looks for the lowest contract he can find; at least one Sebei man candidly admitted to me that the motives of economy lay behind his marriage to a low-priced second wife.

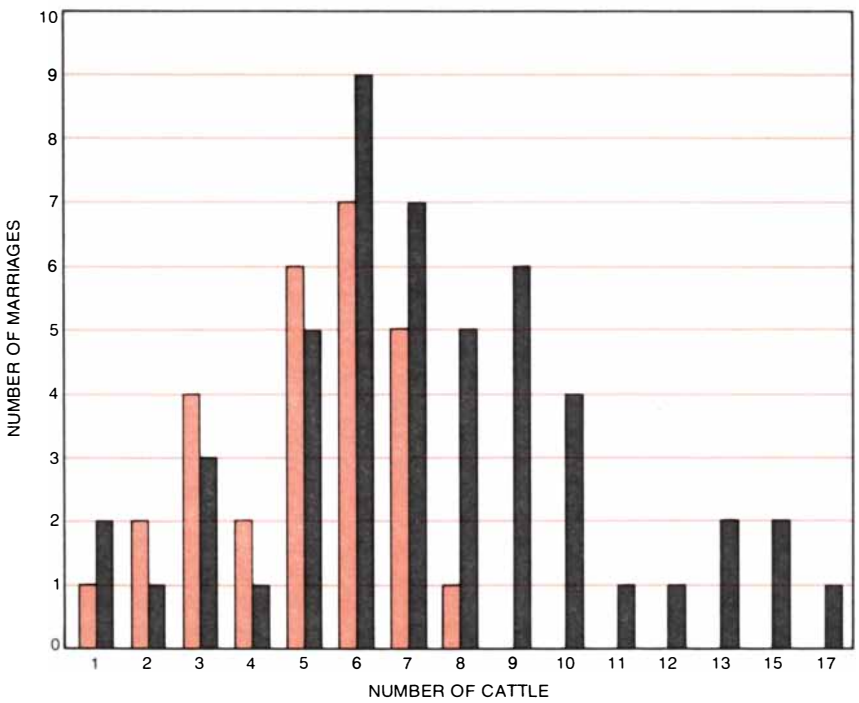
Social pressure is not so great when it comes to taking subsequent wives. Such unions usually do not occur until the husband is older and has acquired more wealth. Since he can afford a higher brideprice, the father is in a better position to demand more than he could get from a younger man. He may also be taking advantage of his daughter's reluctance to enter into such a marriage. In general—not universally—Sebei women prefer young suitors. A few polygynous suitors are more anxious simply to possess a third or fourth wife than to marry any particular woman, and such men will seek out low-cost contracts. As a result such plural marriages, although they are on the average more costly than first marriages, show a wide range of individual brideprices. Among the 14 marriage contracts for a third or subsequent wife that are included in my data are three of the four highest brideprices and two of the lowest paid by Sebei grooms during the entire 50-year period.

To test the thesis that third and fourth marriages are to a large extent prerogatives of the wealthy I analyzed the economic position of 27 Kapsirika men who had entered into marriage contracts between 1950 and 1960. The questions I asked were whether the value of the man's cattle was above or below the mean for Kapsirika men in general, and whether the size of his cash income was above or below the Kapsirika mean. Eleven of the marriage contracts were for first marriages; 10 of the men's cattle holdings and nine of the men's cash incomes were below the mean. That is of course no impediment to a union; the obligations of men contracting a first marriage are met by their fathers. Seven of the contracts were for second marriages; all seven men were close to the mean in both livestock holdings and cash income. Of the six contractors of third marriages only two were below the mean in cash income and only one was below the mean in cattle, whereas all three contractors of fourth marriages were above the mean in both economic categories. Evidently, regardless of the brideprice paid, third and fourth marriages largely involve only well-to-do Sebei men.

An aspect of negotiated agreements among the Sebei that is different from contractual obligations in Western society is that none of the contracts incorporates a time clause. Yet the Sebei are quite aware of the cost of delay; they regularly trade a bullock for a heifer that is to be delivered at some future but indeterminate date, recognizing that the heifer is much more valuable than the bullock. This failure to include a time clause is as true of contracts concerning a brideprice as it is of other exchanges, and Sebei discussions suggest that one question in the mind of every prospective bride's father during negotiations is whether or not his son-in-law will meet his contract obligation with reasonable promptness. This concern is not without foundation. I found that only four among 39 Sebei whose marriages were contracted after 1950 had discharged their contract obligations in full by 1961. The average unpaid balance was 24 percent of the total value. The Sasur farmers were worse offenders than the Kapsirika herders in this respect; the average unpaid balance of seven Sasur contracts made after 1950 amounted to more than 34 percent of the brideprice, whereas the average of unpaid Kapsirika balances over the entire decade was only 18 percent. If I had added to the 1951–1960 Kapsirika marriages the three whose brideprice had been collected on the spot by raiding the groom's cattle kraal, the average unpaid balance among the pastoralists would have dropped to 11 percent.

The Sebei brideprice shows wide variations over a 50-year period that are evidently responses to situational factors of an essentially economic kind. For example, the long-term rise in brideprice is related to a gradual increase of prosperity among the Sebei following the dark days at the turn of the century. When the growth reached a plateau during the 1950's, the rise in brideprice tended to level off.

Another example of the influence of situational factors is the finding that brideprice in areas rich in cattle is consistently higher than in cattle-poor areas. The differential is true not only when the bride's area is cattle-rich but also when the groom's area is. Cattle-poor grooms who marry within their own area pay the



COMPARISON of numbers of cattle stipulated in Kapsirika (*gray*) and Sasur (*color*) marriage contracts demonstrates that the herders barter livestock more freely than the farmers. The number of cattle granted was most often six. The greatest number in any Sasur brideprice was one contract for eight; 22 Kapsirika contracts include that many cattle or more.

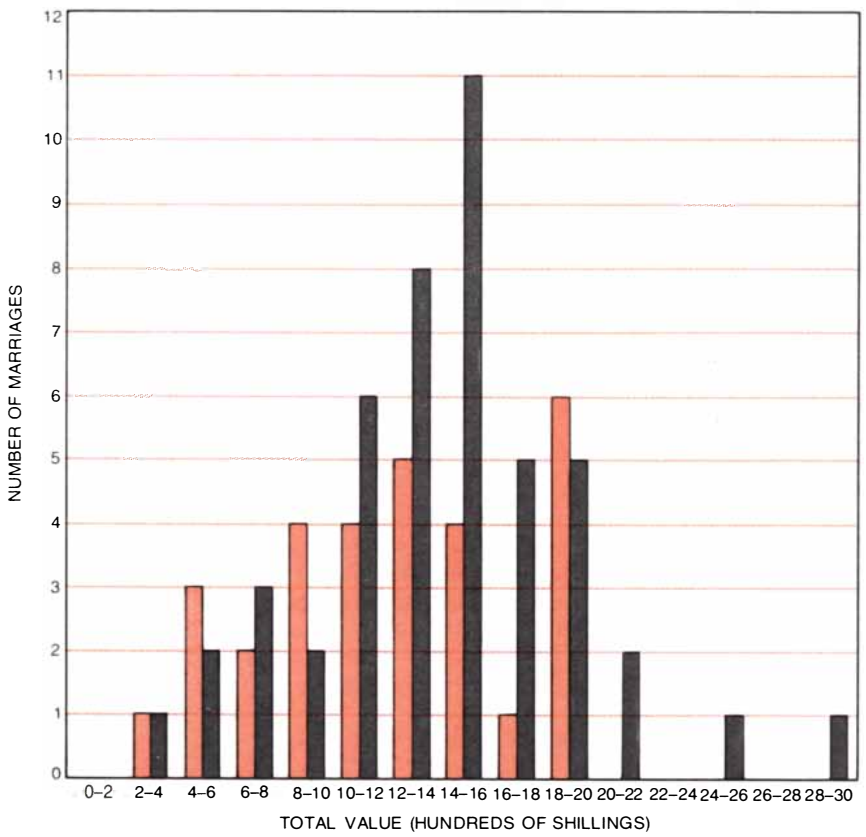
least brideprice. Grooms from a pastoral area who marry within the area or who take their brides from other herders pay the highest brideprice. Finally, men from a pastoral community who take their brides from a farming area pay an intermediate price. When one also takes into account the effect of deferred payment of the brideprice, a practice that is most prevalent in agricultural areas, the differential between the brideprice in cattle-rich communities and that in cattle-poor ones is further increased.

An earlier liaison between the bride and the groom creates a situation that reduces the husband's bargaining power. As a result an after-the-fact brideprice is higher than one negotiated in advance because the bride's father is able to take advantage of his position to increase his demands. In contrast to this state of affairs, a divorced woman commands only a reduced brideprice. That is not merely because personal items and goods for consumption are largely eliminated from a divorcée's marriage contract nor is it only because divorced women are never taken as first wives. Rather it is at least in part because childbearing is a paramount consideration in Sebei marriage, and a divorced woman, part of whose period of fertility has been expended during a prior marriage, is of reduced value to her second spouse.

The fact that polygynous men pay less for second wives than for first ones but usually pay a premium for third and subsequent wives also reflects economic circumstances. The price of a first wife is met by the husband's father, but the husband himself pays for subsequent wives. Men take second wives when their economic status is no more than average; third and subsequent wives are generally taken only by older and more affluent men.

In all these ways, then, economic considerations are important to the decision-making processes whereby Sebei men seek a wife or accept a husband for their daughter. This is a significant finding on more than one count. It demonstrates something about Sebei attitudes toward marriage and toward interpersonal relations in general. Furthermore, it indicates that when the corpus of transactions among this tribal people is subjected to economic analysis, it can yield meaningful results. The analysis not only identifies a number of the elements that enter into Sebei decision making but also demonstrates that most of their perceptions about the actions they take are realistic.

Under no circumstances should the fact that Sebei brideprice contracts are susceptible to economic analysis lead us to conclude that marriage among the Sebei is purely and simply economic. Social factors also enter into the determination of brideprice; these are difficult if not impossible to quantify, but they are nonetheless real. I still hear the echo of the Sebei man who said to me about a particularly costly contract: "It was a very high price, but I wanted her very much." As the data indicate, the Sebei brideprice may vary much more within a single year than economic considerations alone can possibly account for. I have seen Sebei men fret and worry about "what I will be charged" (to use their idiom) in a contract negotiation. I have also seen a bride's father withdraw one of the tallies set out by his negotiator, thereby refusing to ask as high a brideprice as he might. No one who has seen these things can doubt that many social sentiments enter into Sebei marriage negotiations.



RANGE OF BRIDEPRICES in Kapsirika (*gray*) and Sasur (*color*) marriage contracts indicates that a value between 1,300 and 1,400 shillings was the most common. More than half of the Kapsirika contracts were above this level and more than half of the Sasur contracts were below. Some Sasur brideprices nonetheless were high; six topped 1,800 shillings.

An even more compelling reason why the Sebei marital pattern should not be considered exclusively economic is that the pattern operates within a framework of definitions that are clearly cultural in nature. An example is the Sebei sentiment that the transfer of at least one cow to the bride's family is required to legitimize a marriage. The transfer represents the formal claim on the part of the husband's clan to possession of the progeny arising from the marriage. To be sure, a transfer of cattle is an act with a quantifiable economic outcome. Yet at the same time it is an act firmly founded on legal expectation and social definition. Similar social definitions appear again and again throughout Sebei marriage activities. Not the least of these is the mutual respect that the son-in-law and the father-in-law are bound to show toward each other and that each exhibits in a number of ways. In summary, the fact that Sebei brideprice is susceptible to economic analysis tells much about the Sebei but does not tell the whole story. To believe it does is to underestimate the complexity of human motivation. With this caveat in mind it is nonetheless fair to say that, however much sentiments or emotions may enter into a Sebei man's desire to marry, he must at the same time face the economic realities of Sebei society and make his choice of marriages within an economic context.

Plate Tectonics and Mineral Resources

The concepts of continental drift and sea-floor spreading provide clues to the location of economically important minerals such as oil and metals. These clues have already led to promising deposits

by Peter A. Rona

A scientific revolution is in progress that over the past five years has already changed our understanding of the earth as profoundly as the Copernican revolution changed medieval man's understanding of the solar system. The Copernican revolution entailed a fundamental change in man's world view from an earth-centered planetary system to a sun-centered one and led to the development of modern astronomy and the exploration of space. The current scientific revolution entails a fundamental change in man's world view from a static earth to a dynamic one and presages comparable benefits. Some of the benefits may even be economic. The implications of the new global tectonics for mineral resources, particularly the mineral resources of the ocean floor, are only now beginning to emerge.

At present the only undersea mineral resources that certainly have economic value are the vast oil and gas reserves found under many continental shelves and continental slopes, gravel, sand, shells and placer deposits on the continental shelves, various other minerals buried under the continental shelves in specific relation to adjacent continental deposits, and fields of manganese nodules that blanket large areas of the deep-sea floor. Even this limited knowledge is remarkable in the light of the difficulty that was encountered in obtaining it. Consider how much we would know about the mineral deposits of the continents if our sampling procedure were limited to flying in a balloon at an altitude of up to six miles and suspending a bucket at the end of a cable to scrape up loose rocks from the surface of the land. What are the chances that we would find the major known ore bodies, which generally underlie areas of less than a square mile?

Yet this farfetched analogy accurately describes man's present capacity for sampling the sediments and rocks of the ocean bottom, utilizing a variety of coring, drilling and dredging devices lowered from ships through the water column over an area twice as large as that of the continents. Averaged over the world's oceans, the distribution of ocean-floor rocks that have been sampled to date is only about three dredge hauls per million square kilometers!

In recent years every major discovery of a hidden mineral resource has been anticipated by a theoretical vision. For example, once field geologists realized that there was a definite association between the type of sedimentary structure termed an anticline and accumulations of oil, they knew where to drill and the rate of discovery of oil deposits accelerated accordingly. In the same way the right conceptual framework can be used to extend man's limited direct knowledge of resources of the ocean basin toward a realistic appraisal of their potential. The test of the value of such a conceptual framework is how well it explains what one sees and predicts what one does not see.

The old conceptual framework of a static earth held that the continents and ocean basins were permanent features that had existed in their present form since early in the 4.5-billion-year history of the earth. Only the most accessible continental mineral deposits were discovered, largely by trial and error, with little understanding of why or where they existed. The recent change to a conceptual framework based on a dynamic-earth model, in which continents are constantly moving and ocean basins are opening and closing, is leading toward a better understanding of the global distribution of mineral deposits in both space and time.

The basis of the new conceptual framework is the theory of plate tectonics, the essentials of which have already been reported in SCIENTIFIC AMERICAN [see "Plate Tectonics," by John F. Dewey; May, 1972]. "Tectonics" is a geological term pertaining to earth movements. The movements in question involve the lithosphere, the rigid outer shell of the earth, which is of the order of 60 miles thick. The lithosphere, which behaves as if it were floating on an underlying plastic layer, the asthenosphere, is segmented into about six primary slabs, or plates, each of which may encompass a continent and part of an adjacent ocean basin [*see top illustration on page 89*].

The boundaries of the lithospheric plates are delineated by narrow earthquake zones where the plates are moving with respect to each other. Three types of boundary are recognized. One type, called a convergent plate boundary, is where two adjacent plates move

TROODOS MASSIF on the island of Cyprus, the site of economically important mineral deposits that originated at a divergent tectonic-plate boundary, stands out clearly as the dark-colored mountainous region in the middle of the satellite photograph on the opposite page. The photograph was made recently from an altitude of nearly 600 miles by a multispectral camera system on board the first Earth Resources Technology Satellite (ERTS 1). Region is believed to be a slice of oceanic lithosphere that was formed by the process of sea-floor spreading from a submerged mid-oceanic ridge and was subsequently thrust upward.

SMALL VEIN OF PURE COPPER was discovered in a core sample of sedimentary rock obtained by the Deep Sea Drilling Project some 350 miles southeast of New York City. The copper vein, the horizontal reddish structure in this longitudinal section of a piece of the original core, is about half an inch long. It was found in sediment about 65 feet above the volcanic basement rocks under the lower continental rise at a water depth of 17,000 feet.

METAL-RICH CORE, collected from the Atlantis II Deep, one of the hot-brine pools located along the axial valley of the Red Sea at a depth of about 6,600 feet below sea level, represents the most concentrated submarine metallic sulfide deposits known. The muddy sediments containing the sulfide minerals fill the Red Sea basins to a thickness estimated at between 65 and 330 feet. The deposits are saturated with (and overlain by) salty brines considered to be the hydrothermal solutions from which the sulfide minerals were precipitated. The photograph was made by David A. Ross of the Woods Hole Oceanographic Institution.

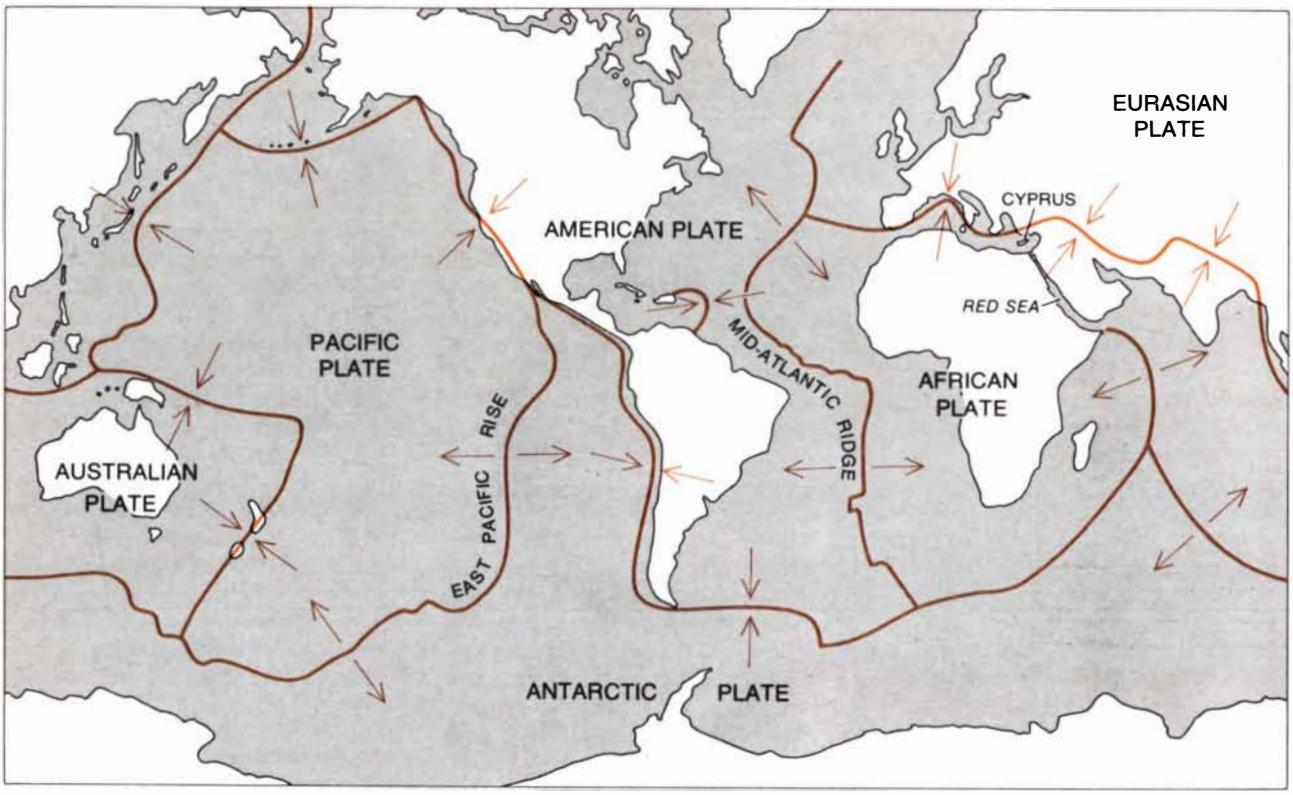
together and collide or where one plate plunges downward under the other plate and is absorbed into the interior of the earth.

The second type of boundary, called a divergent plate boundary, is where two adjacent plates move apart because new lithosphere is added to each plate by the process of sea-floor spreading. The new lithosphere, which moves more or less symmetrically to each side of the divergent plate boundary, acts like a conveyor belt, carrying the continents apart in the motion that has become known as continental drift. The dual existence of convergent boundaries where lithosphere is destroyed and divergent boundaries where lithosphere is created implies that the diameter of the earth is not changing radically.

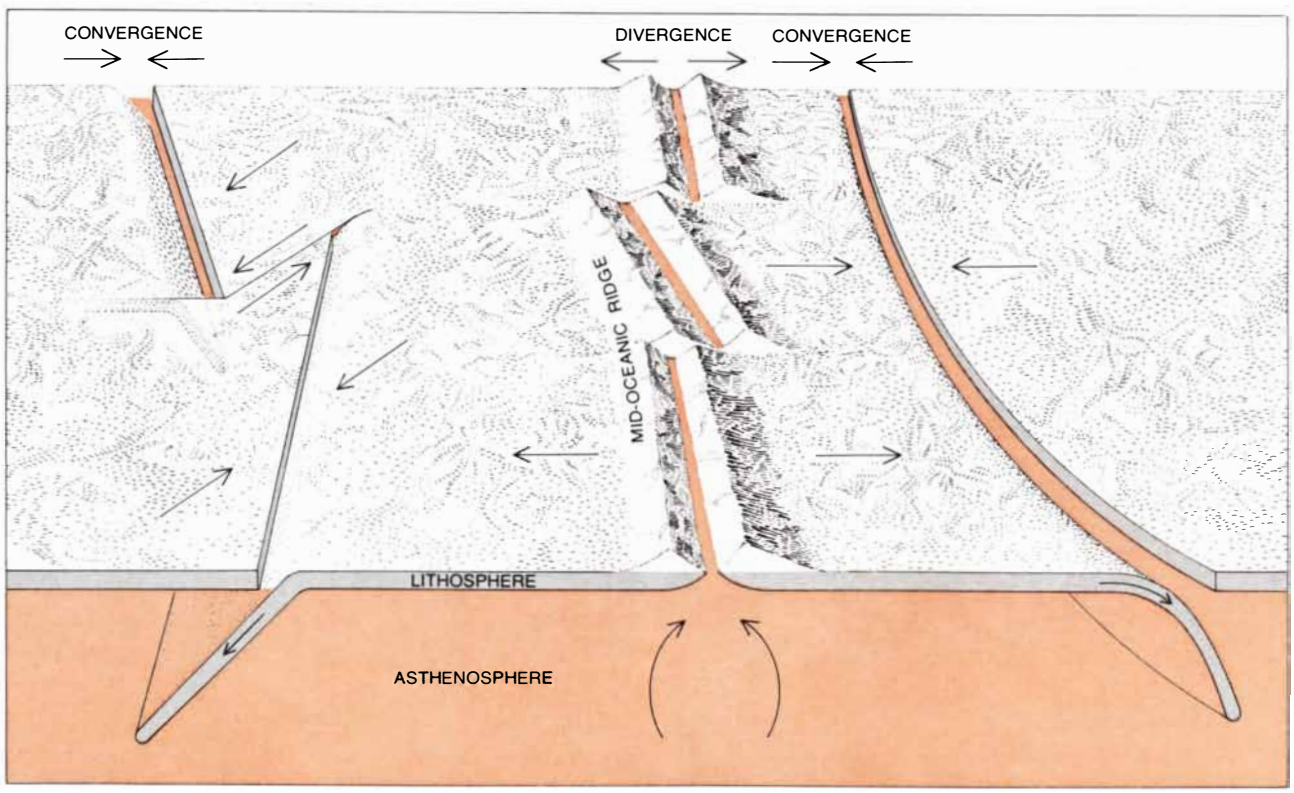
The third type of tectonic-plate boundary is the parallel plate boundary, where two adjacent plates move edge to edge along their common interface.

Hydrothermal mineral deposits, that is, mineral deposits formed by precipitation from solutions, constitute a major part of our useful metallic ores on the continents. Economically the most important types of hydrothermal deposit are the sulfides, in which various metals combine with sulfur to precipitate from the hydrothermal solution. About a year ago Frederick Sawkins, a geologist at the University of Minnesota, pointed out that most of the sulfide deposits of the world are located along present or former convergent plate boundaries where an oceanic lithospheric plate plunges under the margin of a continent (including the continental shelf) or under a chain of volcanic islands. The processes that concentrate the sulfide deposits along convergent plate boundaries, which are at present only partly understood, involve mineralizing solutions that emanate from the plunging lithospheric plate, which melts as it is absorbed into the interior of the earth.

Metallic sulfide deposits along convergent plate boundaries include the Kuroko deposits of Japan, the sulfide ore bodies of the Philippines and the deposits extending along the mountain belts of western North America and South America (the Coast Ranges, the Rockies and the Andes) and from the eastern Mediterranean region to Pakistan. Gold-bearing deposits are not sulfides but often accompany sulfide minerals. The majority of gold deposits in Alaska, Canada, the southeastern U.S., California, Venezuela, Brazil, West Africa, Rhodesia, southern India and



SIX PRINCIPAL TECTONIC PLATES of the lithosphere, the rigid outer shell of the earth, are delineated by the heavy color lines on this world map. The paired arrows indicate whether a plate boundary is convergent or divergent (*see illustration below*).



TWO TYPES OF PLATE BOUNDARY are illustrated schematically in this block diagram. The 60-mile-thick lithospheric plates move outward like conveyor belts from the mid-oceanic ridges (divergent plate boundaries) and plunge downward under the deep-sea trenches (convergent plate boundaries). The third major type of plate boundary, not shown here, is the parallel plate boundary.

southeastern and western Australia occur in rocks that can be associated with former convergent plate boundaries.

Divergent plate boundaries are formed by the spreading of lithospheric plates in the central portions of ocean basins. The Red Sea and the island of Cyprus in the Mediterranean Sea provide important clues to the potential of metallic sulfide deposits at divergent plate boundaries.

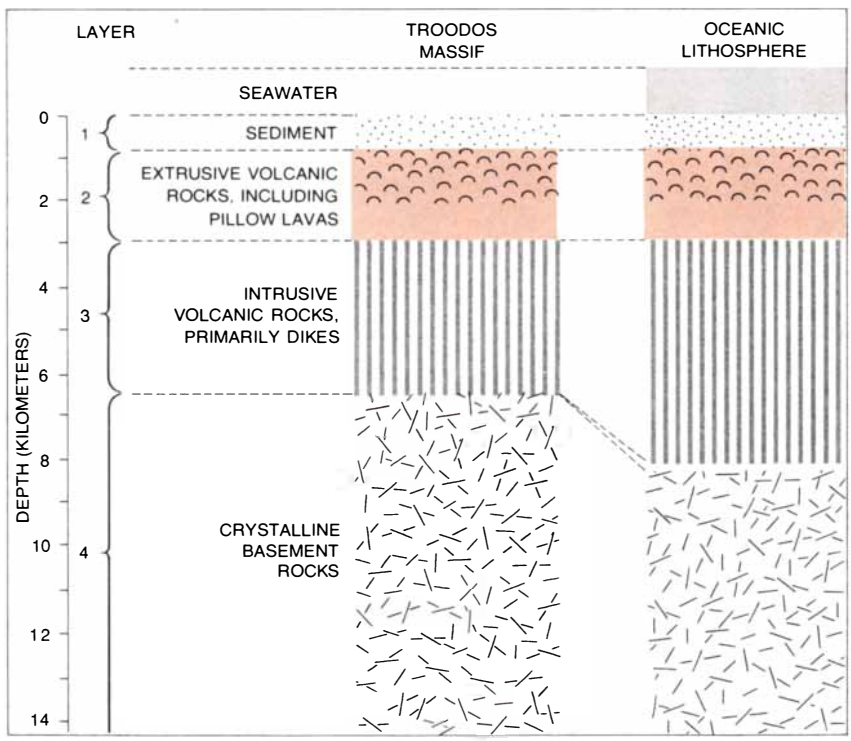
The Red Sea, the product of a divergent plate boundary developing between the African plate and the Eurasian plate, provides an accessible natural laboratory for the study of mineral processes associated with divergent plate boundaries. About five years ago the richest submarine metallic sulfide deposits known were found in three rather small basins along the center of the Red Sea at a depth of about 6,600 feet below sea level. The sulfide minerals are disseminated in sediments that fill the basins to a thickness estimated at between 65 and 330 feet. The top 30 feet or so of sediment, which has been explored by coring the largest of the basins, has a total dry weight of about 80 million tons, with average metal contents of 29 percent iron, 3.4 percent zinc, 1.3 percent copper, .1 percent lead, .005 percent silver and .00005 percent gold. The deposits are saturated with (and overlain by) salty brines carrying the same metals in solution as those present in the sulfide deposits. The salty brines are considered to be the hydrothermal solutions from which the sulfide minerals are precipitated. It remains controversial whether the brines are being charged with minerals from volcanic sources under the Red Sea or from sediments with high copper, vanadium and zinc contents adjacent to the basins where the metallic sulfide deposits are found [see "The Red Sea Hot Brines," by Egon T. Degens and David A. Ross; SCIENTIFIC AMERICAN, April, 1970].

The Red Sea represents the earliest stage in the growth of an ocean basin: the stage where a divergent plate boundary rifts a continent in two. The most advanced growth stage of a divergent plate boundary is the mid-oceanic-ridge system, a 47,000-mile undersea mountain chain that extends through all the major ocean basins and girdles the globe. The mid-oceanic-ridge system has not been adequately sampled to determine whether or not concentrations of metallic sulfides comparable to the Red Sea deposits are present at sites along its crest or in basins in its flanks. Measurements of the distribution of heat emanating from mid-oceanic ridges and of the chemical alteration of ridge rocks indicate that seawater forms a hydrothermal solution by penetrating fissures, dissolving minerals from rocks underlying the ridges and precipitating those minerals in concentrated deposits.

A limited amount of sampling indicates that hydrothermal processes are actively concentrating metals from volcanic sources underlying mid-oceanic ridges. Sediments on active mid-oceanic ridges are generally enriched in iron, manganese, copper, nickel, lead, chromium, cobalt, uranium and mercury, with trace amounts of vanadium, cadmium and bismuth. The concentrations typical of sediments covering widespread areas on mid-oceanic ridges are not economic, but much higher concentrations exist locally.

Metallic sulfides are found in rocks dredged from the Indian Ocean Ridge. In addition small veins of pure copper have been recovered by the Deep Sea Drilling Project at several sites. At the crest of the Ninety East Ridge near the Equator in the Indian Ocean, for example, veins of copper are found in volcanic rocks overlain by 1,440 feet of sediment at a water depth of 7,380 feet. Some 350 miles southeast of New York City a small vein of pure copper and clusters of copper crystals have been discovered in sediment about 65 feet above the volcanic basement rocks under the lower continental rise at a water depth of 17,000 feet [*see top illustration on page 88*].

A specimen of manganese 1.7 inches thick recently dredged from a water depth of about 12,000 feet in the median valley of the Mid-Atlantic Ridge by the Trans-Atlantic Geotraverse of the National Oceanic and Atmospheric Administration has particular significance. The composition, form and thickness of this manganese sample, which accumulated at a rate about 100 times faster than the manganese in nodules, indicates a hydrothermal origin and demonstrates that hydrothermal mineral deposits are actively accumulating at certain divergent plate boundaries in ocean basins. Because the sea floor is supposed to originate by spreading from mid-oceanic ridges, a mineral deposit on a mid-oceanic ridge would be expected to extend in a linear zone from the ridge across the ocean basin to the adjacent continental margin if the depositional process is a

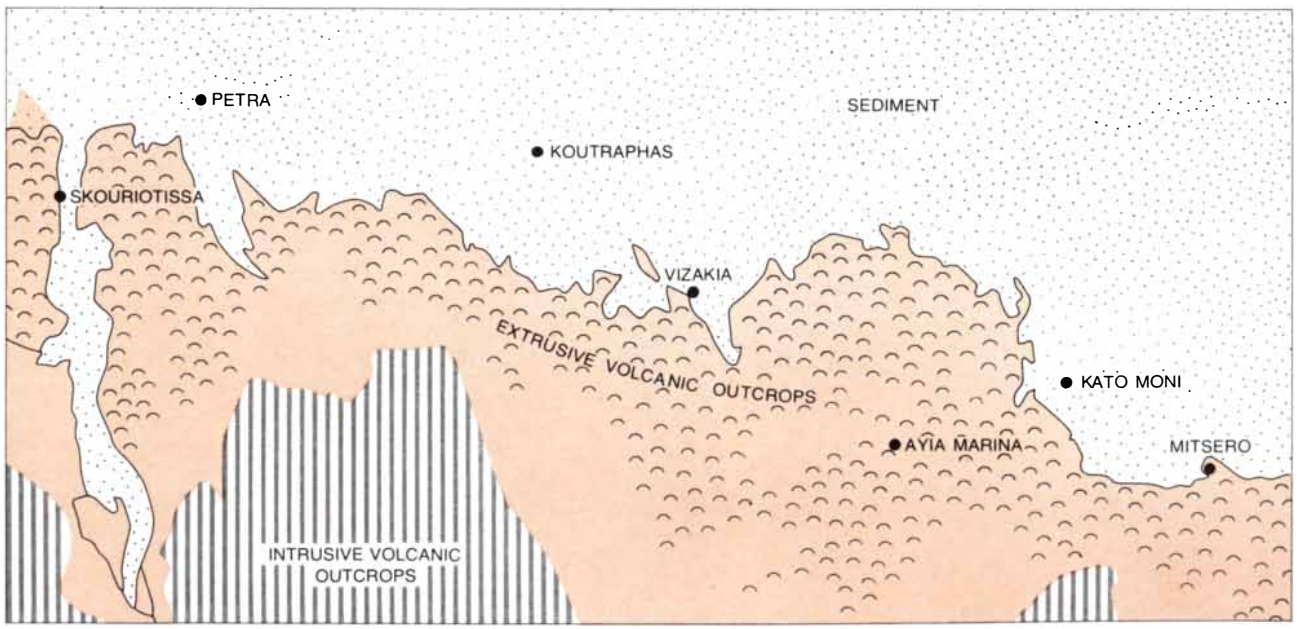


CLOSE CORRESPONDENCE between the layered sequence of rocks in the Troodos Massif (*left*) and that of the oceanic lithosphere (*right*) is evident from this comparison. The geological structure of the Troodos Massif was determined directly from rock outcrops; the structure of the oceanic lithosphere was determined indirectly by seismic-refraction techniques. The sulfide ore bodies of the Troodos Massif are in the upper portion of layer made up of extrusive volcanic rocks. Pillow shapes form when volcanic lava cools on the sea floor.

continuous one [*see illustration on next page*].

At this point in man's exploration of the oceans it would seem to be too much to expect that it would be possible to make detailed observations on an economically important metallic sulfide deposit that originated at a divergent plate boundary on a submerged mid-oceanic ridge. Yet such a deposit is known and has been extensively studied. The Troodos Massif on the island of Cyprus is interpreted as being a slice of oceanic lithosphere that was formed by the process of sea-floor spreading from a mid-oceanic ridge and was subsequently thrust upward to its present position [*see illustration on page 87*]. The composition and layered sequence of rocks that constitute the Troodos Massif are the

ISLAND OF CYPRUS has been famous for its mineral wealth since Phoenician times. The principal ore bodies are in the uppermost volcanic layers of the Troodos Massif, the total extent of which is indicated by the dark-colored area. The hatched area represents sediments, including alluvium. A geological map of a portion of the Troodos igneous complex (*small rectangle*) is shown below.

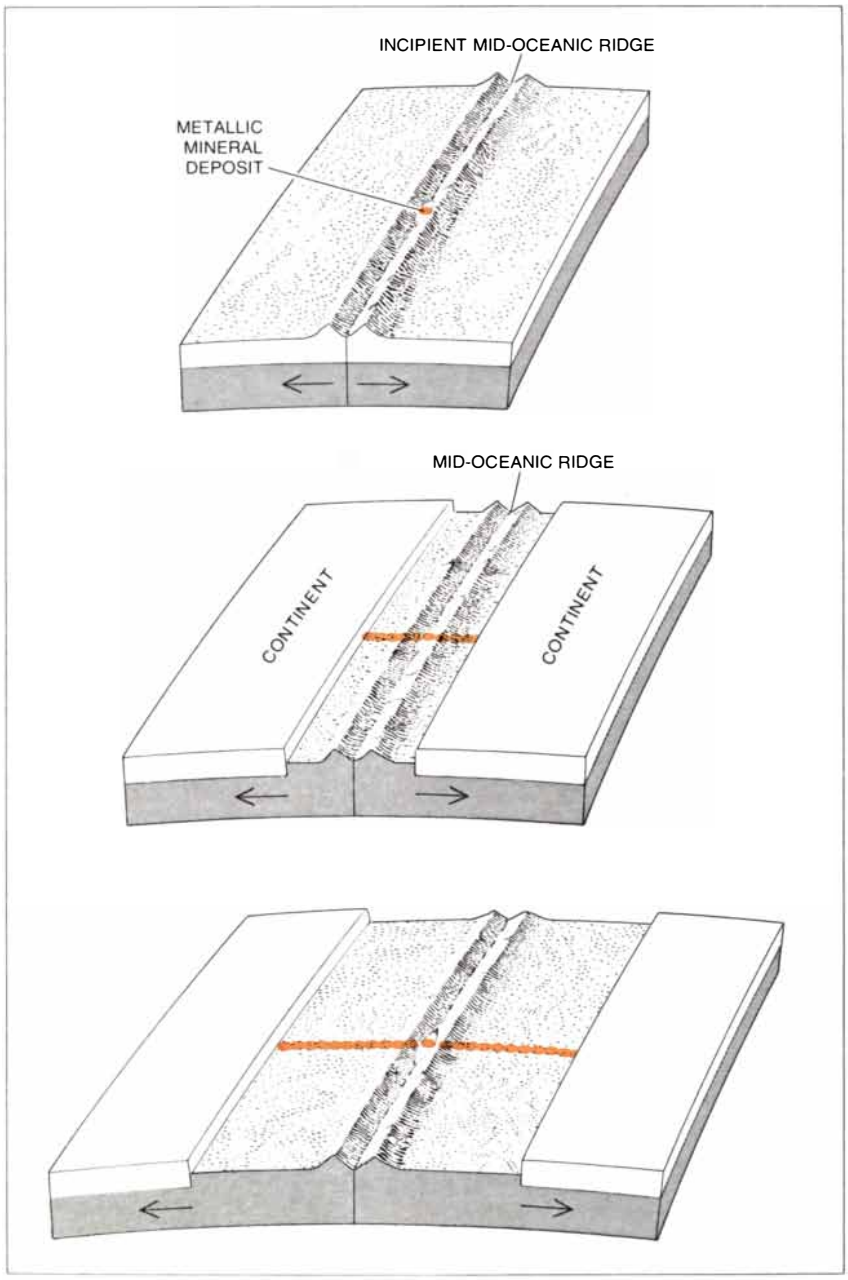


GEOLOGICAL MAP of a region that lies along the northern fringe of the Troodos Massif is based on studies that were undertaken by the Geological Survey of Cyprus. The map shows outcrops of extrusive volcanic rock that incorporate bodies of metallic sulfide ore.

same as those known to underlie the seabed.

Cyprus has long been famous for its mineral wealth. The mining of copper (for which the island is named) was an important industry in Roman and even in Phoenician times. The brilliant green stains of copper sulfides on ancient mine tailings have attracted modern prospectors. Between 1965 and 1970 the average annual exports amounted to about a million tons each of iron pyrites, chromite and gypsum, about 150,000 tons of copper pyrites and 100,000 tons of copper concentrates. The estimated value of the mineral products exported from Cyprus in 1970 amounted to $30 million.

The principal ore bodies are in the uppermost volcanic layers of the Troodos Massif. It has been uncertain whether the Troodos sulfide-ore bodies originated before the upthrust of the Troodos Massif or afterward. In the first instance the ore bodies would be representative of the seabed. In the second the ore bodies would be attributed to special conditions unrelated to the seabed. The sulfide deposits are clearly related to the volcanic rocks in which they occur. Recent studies reveal that iron-rich and manganese-rich sediments interlayered with the volcanic rocks and associated with the ore bodies of the Troodos Massif are chemically identical with those metal-enriched sediments found on active mid-oceanic ridges, indicating that both the sediments and the ore bodies were formed on the sea floor by hydrothermal processes.



HYDROTHERMAL MINERAL DEPOSIT (*color*) **formed in a hot-brine pool on the axis of a mid-oceanic ridge would be expected to extend in a linear zone from the ridge across the ocean basin to the adjacent continental margins as the ocean basin progressively widens** (*from top to bottom*) **as a consequence of sea-floor spreading from the mid-oceanic ridge.**

The Troodos ore bodies may provide the first firm evidence on the nature of metallic sulfide deposits in ocean basins. The Skouriotissa ore body, for example, is roughly elliptical in plan view, measures approximately 2,000 feet long by 600 feet wide and is lens-shaped in cross section. Its estimated mass is six million tons. The average composition of the ore is 2.25 percent copper (ranging to greater than 5 percent), 48 percent sulfur and 43 percent iron.

The Mavrovouni ore body is also roughly elliptical in plan view, measures approximately 1,000 feet long by 600 feet wide and forms a lens that attains a thickness of 800 feet in cross section. Its estimated mass is greater than 15 million tons. The average composition of the ore is 4.2 percent copper, 48 percent sulfur, 43 percent iron, .4 percent zinc, .25 ounce per ton gold and .25 ounce per ton silver.

Sediments underlying the Skouriotissa ore body, presumably a disintegration product of the pyrite in the ore body, contain 2.12 ounces of gold per ton and 12.96 ounces of silver per ton. Exposed patches of metallic oxides indicate the presence of the ore bodies under the mountainous surface of the Troodos Massif. The Skouriotissa ore body is exploited by underground shafts and the Mavrovouni ore body by strip-mining.

What kind of target for exploration would a Troodos ore body make if it were submerged under thousands of feet of water on the crest or flank of a mid-oceanic ridge? It is unlikely that any of the present exploration methods would be capable of detecting the ore body. The resolution of present geophysical exploration methods will have to be improved in order to detect such an ore body under the sea. Both the exploration methods and the engineering development involved will be costly.

The prerequisites for the accumulation of petroleum consist of a source of organic matter to generate the petroleum, a natural reservoir to contain it

and a trap to concentrate its fluid and gas constituents. Petroleum is hydrocarbons derived from the remains of plants and animals. As the progenitor of the petroleum, the organic matter must accumulate in an environment where it is preserved. The preservation of organic matter is favored by an environment that is toxic to life (so that the organic matter is not consumed as food) and deficient in oxygen (so that the hydrocarbon is not decomposed). How do conditions favorable to the accumulation of petroleum relate to convergent and divergent plate boundaries?

Convergent plate boundaries where the oceanic portion of a lithospheric plate plunges under the margin of a continent are characterized by the presence of a deep-sea trench running along their length. A system of deep-sea trenches runs along the entire western margin of North America and South America where the Pacific lithosphere is plunging under the continents. In addition to a deep-sea trench, chains of volcanic islands are present along some convergent plate boundaries; they are located between the trench and the continent. There are many such chains of volcanic islands at the western margin of the Pacific, including the Aleutians, the Kuriles, Japan, the Ryukyus, the Philippines and Indonesia. Other such chains are the Marianas, the South Sandwich Islands and the West Indies. The island chains divide an ocean basin into smaller basins partially enclosed between the islands and the adjacent continent; such basins include the Bering Sea, the Sea of Okhotsk, the Sea of Japan, the Yellow Sea, the East China Sea and the South China Sea.

Both the marginal trenches and the volcanic-island chains create a habitat that is favorable for the accumulation of petroleum in several respects. First, the trenches and island chains act as barriers that catch sediment and organic matter from the continent and the ocean basin. Second, the shape of the trenches and the small ocean basins acts to restrict the circulation of the ocean, so that oxygen is not replenished in the seawater and organic matter is preserved. Third, the accumulation of sediments and the geological structures that develop as a result of the deformation of the sediments by tectonic forces provide reservoirs and traps for the accumulation of petroleum. According to Hollis D. Hedberg of Princeton University, "these marginal semienclosed basins constitute some of the most promising areas in the world for petroleum accumulation."

AT AN EARLY STAGE of continental drift the Atlantic was a sea with its circulation restricted by the surrounding continents. As in the present Red Sea, conditions in the Atlantic Sea favored the preservation of organic matter and the deposition of rock salt, leading to the formation of petroleum accumulations under the present continental margins.

The development of divergent plate boundaries may also create a habitat favorable for the accumulation of oil, a finding that would open immense possibilities for petroleum resources in the deep ocean basin. When a divergent plate boundary develops under a continent, the continent is rifted in two and the continental fragments are carried apart on a conveyor belt of new lithosphere generated at the divergent plate boundary. As the two continental fragments move apart, a sea forms between them. The surrounding continents act as barriers to restrict the circulation of the sea. As a result organic matter is preserved and, if the evaporation of the seawater exceeds its replenishment, layers of rock salt are deposited along with the organic matter. As the continental fragments continue to move apart and to subside along with the adjacent sea floor, the restricted sea becomes an open ocean. The layers of organic matter and salt are buried under sediments. The organic matter subsequently develops into petroleum (by processes that are only partly understood) and the salt

forms into dome-shaped masses that act to trap the petroleum.

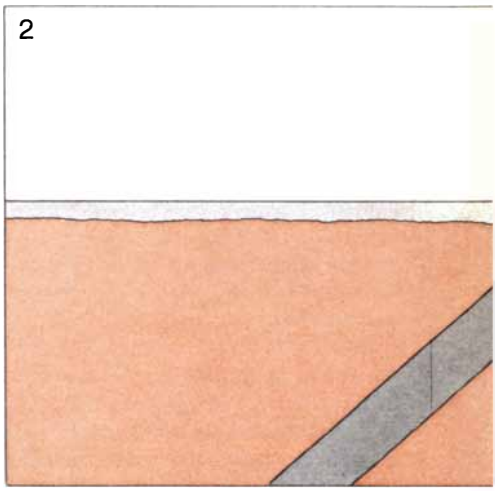
The Red Sea is an example of a restricted sea formed at an early stage of development of the divergent plate boundary along which Arabia is rifting from Africa. Layers of rock salt up to 17,000 feet thick and organic muds have been found under it. Along both the eastern and western margins of the North Atlantic and the South Atlantic apparent salt domes have been discovered extending seaward from continental shelves to continental rises in water depths of up to 16,500 feet. The occurrence of these salt domes in the deep Atlantic indicates that at an early stage of continental drift the Atlantic was a sea with its circulation restricted by the surrounding continents in their positions at that time [*see illustration on preceding page*].

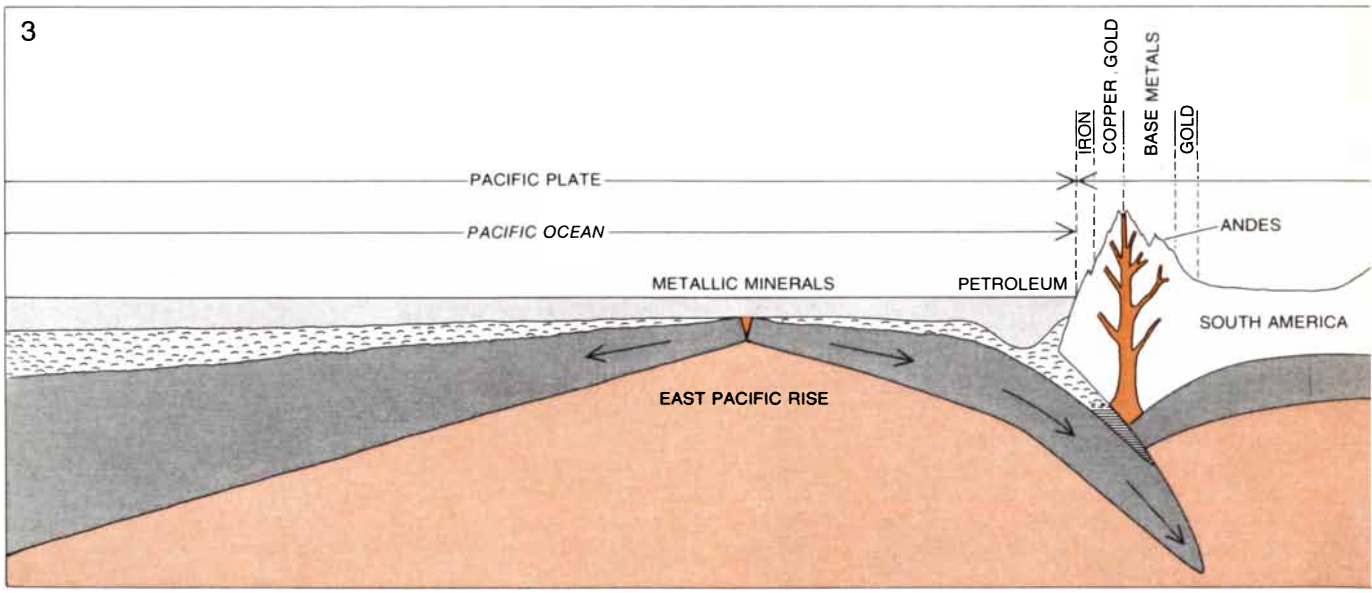
Like the present Red Sea, conditions in the Atlantic Sea favored the preservation of organic matter and the deposition of rock salt. As the Atlantic widened in response to the symmetric creation of new lithosphere by sea-floor spreading from the Mid-Atlantic Ridge, the Atlantic Sea became an ocean and the organic matter and salt were buried under sediments, forming the present margins of the Atlantic Ocean. It is reasonable to expect that petroleum accumulations will extend seaward under the continental shelf, the continental slope and the continental rise to water depths of about 18,000 feet along large portions of both the eastern and western margins of the North Atlantic and South Atlantic. Petroleum may likewise be found in other ocean basins that have grown through the stage of a restricted sea by sea-floor spreading.

In short, the patterns of mineral distribution that are emerging from the conceptual framework provided by the new global tectonics will clearly help to guide man's search for new mineral deposits. Hydrothermal processes have concentrated the majority of known metallic sulfide ore bodies along convergent lithospheric plate boundaries originally at continental margins. Hydrothermal processes are also active at divergent plate boundaries from initial stages (represented by the metallic sulfide deposits accumulating in the Red Sea) to advanced stages (represented by the metal concentration in sediments on mid-oceanic ridges and by possible metallic sulfide deposits of the Troodos Massif type). The Troodos Massif metallic sulfide ore bodies provide an actual example of the type of deposits that can be expected in sea-floor rock generated by mid-oceanic ridges. The confirmation and economic evaluation of metallic sulfide deposits of the Troodos Massif type in ocean basins await technological advances in marine exploration methods.

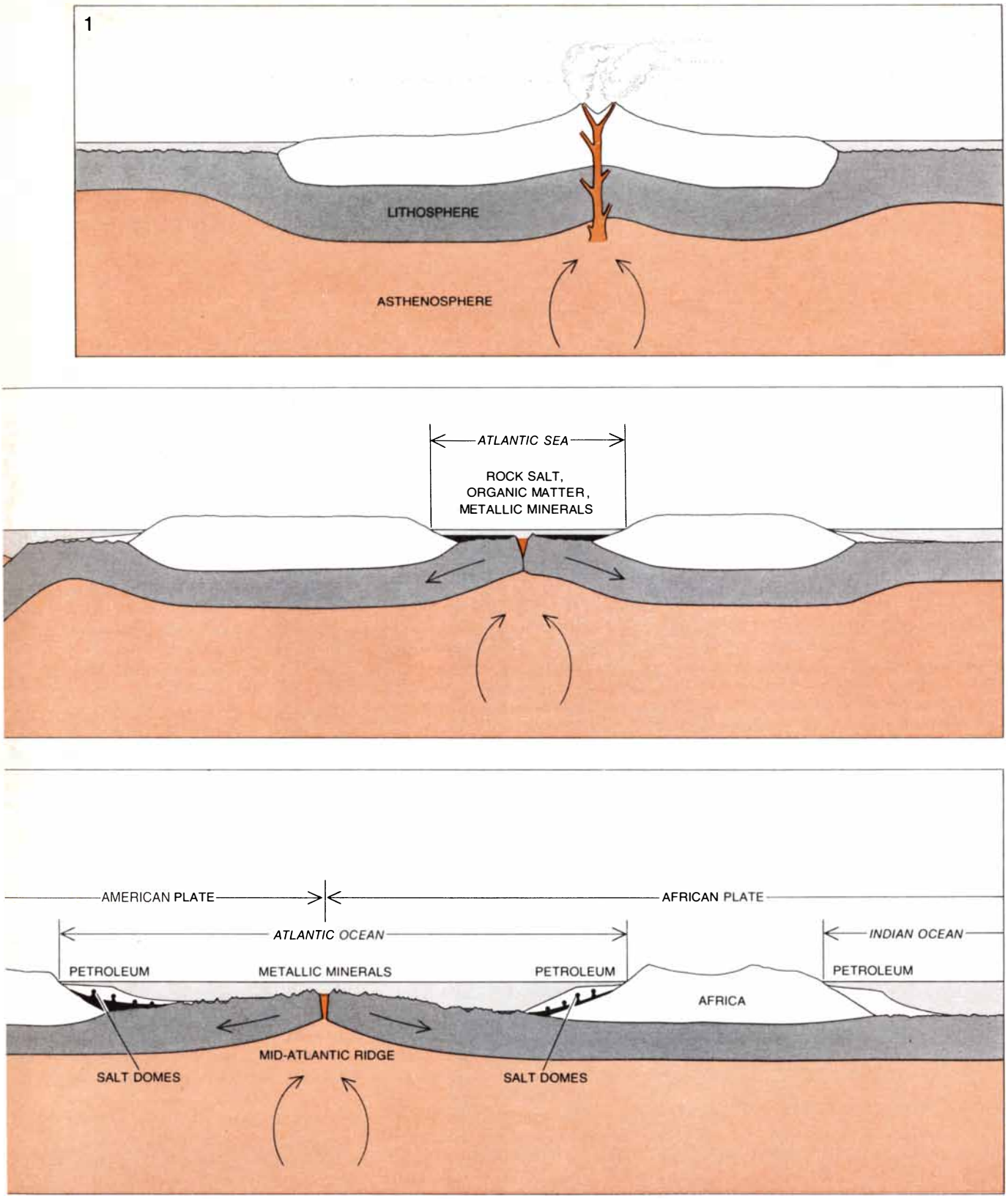
With regard to petroleum, convergent plate boundaries create conditions that form accumulations in small ocean basins and deep-sea trenches marginal to continents. Divergent plate boundaries, on the other hand, create conditions that favor the development of oil accumulations extending from the continental shelf into the deep ocean basin under the continental rise.

The global patterns of mineral distribution that are emerging from such models can be expected to accelerate the discovery of resources not only on the seabed but also on the continents.





ROLE OF PLATE BOUNDARIES in the accumulation of mineral deposits is exemplified in this sequence of cross-sectional views of the development of the South Atlantic Ocean. The position of Africa is assumed to be stationary throughout the sequence of cross sections. In stage *1* a single ancestral continent, called Pangaea, is rifted into two continents (South America and Africa) about a divergent plate boundary. In stage *2* the oceanic crust created by the process of sea-floor spreading from the divergent plate boundary (a precursor of the Mid-Atlantic Ridge) rafts South America westward and is compensated for by the consumption of oceanic crust at a trench (a convergent plate boundary) that develops to the west of South America. Thick layers of rock salt, organic matter and metallic minerals accumulate in the Atlantic Sea during this early stage of continental drift. In stage *3* continued sea-floor spreading



from the Mid-Atlantic Ridge widens the Atlantic into an ocean, rafts South America westward over the trench, reversing the inclination of the trench and producing the Andes mountain chain as a consequence of the deformation that develops at the convergent plate boundary along the western margin of South America. Metallic minerals that are melted from the Pacific plate as it plunges under South America ascend through the overlying crustal layers and are deposited in them to form the metal-bearing provinces of the Andes. Meanwhile in the Atlantic Ocean metallic minerals continue to accumulate about the Mid-Atlantic Ridge. Salt originating in the thick layers of rock salt that have been buried under the sediments of the continental margins rises in large, dome-shaped masses that act to trap the oil and gas that are generated from the organic matter that was preserved in the former Atlantic Sea.

BRAIN MECHANISMS IN MOVEMENT

The highest brain functions are generally thought to be mediated in the cerebral cortex. In the control of the muscles, however, the highest function may be served by centers deeper in the brain

by Edward V. Evarts

The traditional view of the brain is that the highest level in its hierarchical organization is in the cortex, or outer part, of the cerebrum. It turns out that this is not true for the brain's motor functions: the control and integration of muscular movements. Brain research has gradually revealed that the motor area of the cerebral cortex is actually at a rather low level of the motor control system, not far removed from the muscular apparatus itself. Structures lying deep below the cortex are at a higher functional level of the system, as judged by their position in the neural chain of command that initiates and controls movement. The implication of these findings is that the primary function of the cerebral motor cortex may not be volition but rather the refined control of motor activity.

The current era of research on how the cerebral cortex controls movement began some 100 years ago with the studies of the British neurologist John Hughlings Jackson. Reasoning from the abnormal movements present in epilepsy and from the normal movements absent in apoplexy or stroke, he proposed that the brain was a sensory-motor machine divided into different centers for the coordination of sensation and movement. From the symptoms of stroke patients and the anatomical site of the blood clot or burst vessels that caused the symptoms he concluded that the part of the cerebral cortex most directly concerned with movement lay in the territory supplied by the middle cerebral artery. Experimental evidence for Jackson's theory was provided by Gustav Theodor Fritsch and Eduard Hitzig of Germany, who in 1870 reported that the electrical stimulation of a region in one cerebral hemisphere of a dog caused the contraction of muscles on the opposite side of the dog's body. In 1874 Roberts Bartholow, an American physician, demonstrated that electrical stimulation of the cortical area proposed by Jackson as the site of motor control produced muscular contraction. That area is now called the motor cortex.

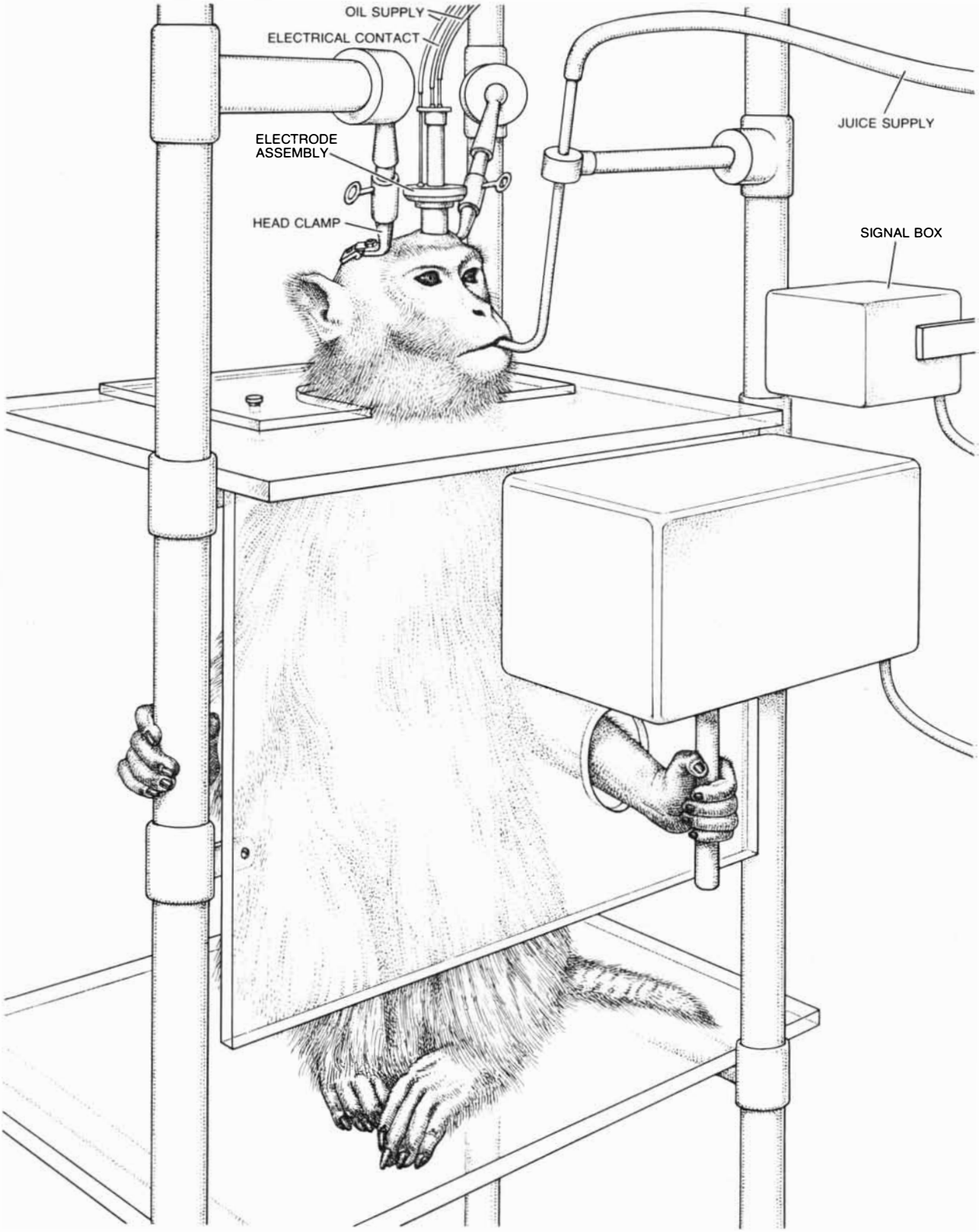
Jackson also devoted much study to focal epilepsy, a condition where convulsive movements are restricted to one part of the body, for example the thumb. He proposed that the localized movements result from excessive nerve discharges in localized areas of the cortex and that these discharges in turn give rise to localized muscular contractions without the volitional participation of the patient. Such localized epileptic attacks are now called Jacksonian epilepsy.

The discovery that muscular contraction could be produced by the electrical stimulation of a small region of the cerebral cortex came as a great surprise to the neurologists of that time. Before the work of Jackson, Fritsch and Hitzig it was generally believed that the highly convoluted cerebral cortex of man was involved in the generation of thoughts rather than of movements. The cerebral cortex was viewed as being man's highest organ of thought, and it was assumed that subcortical, or lower, centers were responsible for muscular contractions. Indeed, the intellectual climate of the day made it necessary for Fritsch and Hitzig to state that "contrary to the opinions of Flourens and most investigators who followed him, the soul in no case represents a sort of total function of the whole of the cerebrum, the expression of which might be destroyed by mechanical means *in toto,* but not in its individual parts. Individual psychological functions ... depend for their entrance into matter, or for their formation from it, upon circumscribed centers of the cerebral cortex."

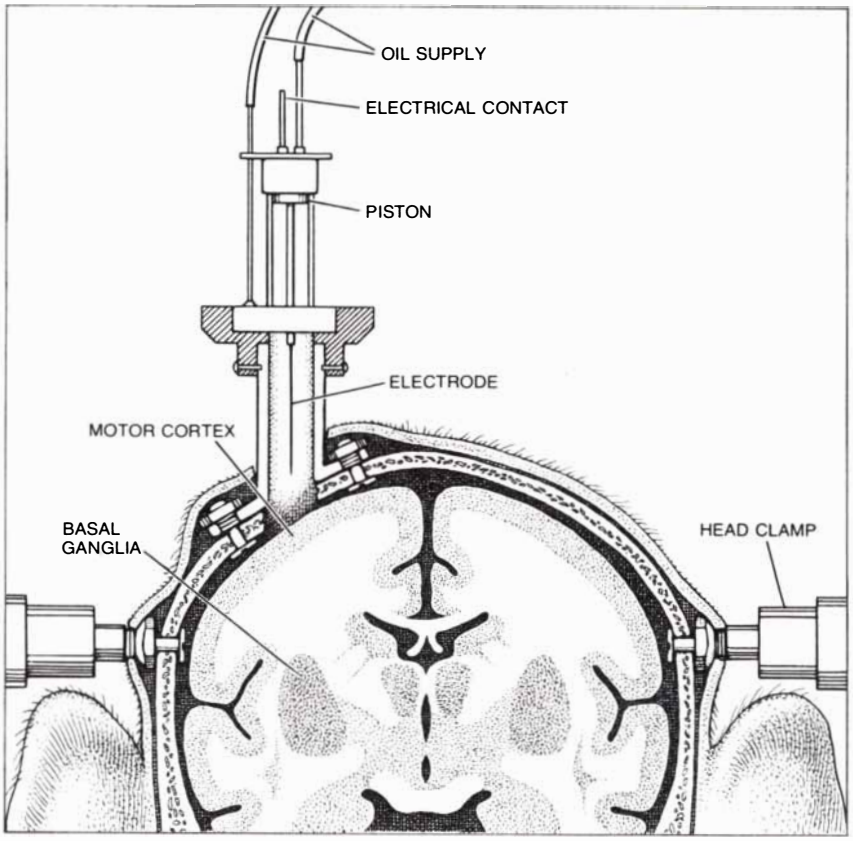
Between 1900 and 1920 Charles S. Sherrington, the foremost neurophysiologist of the time, applied the technique of electrical stimulation to study how the cerebrum controlled movement. Although he made important discoveries with this procedure, he recognized its limitations: the movements produced by the electrical stimulation of the brain are nonvolitional, resembling the movements of epilepsy more than the movements of normal motor activity. Sherrington saw the need for new techniques, and he wrote that experiments leading to an understanding of the normal functioning of the cerebral motor centers would require "combining the methods of comparative psychology with the methods of experimental physiology ... to furnish new data of importance toward the knowledge of movement as an outcome of the working of the brain."

The psychophysiological approach advocated by Sherrington was not feasible for nearly 50 years because of technical problems. As a result knowledge of the cerebral motor processes greatly lagged behind that of the cerebral sensory processes. One difficulty in studying volitional movement arose from the necessity of having the active participation of the experimental subject; that precluded the use of an anesthetized animal. Research on sensory processes moved ahead rapidly because sensory functions could be tested in such an animal. For example, the physiology of visual receptors could be studied in anesthetized animals but the physiology of eye movements could not, since such studies required animals capable of perception, attention and coordinated motor function.

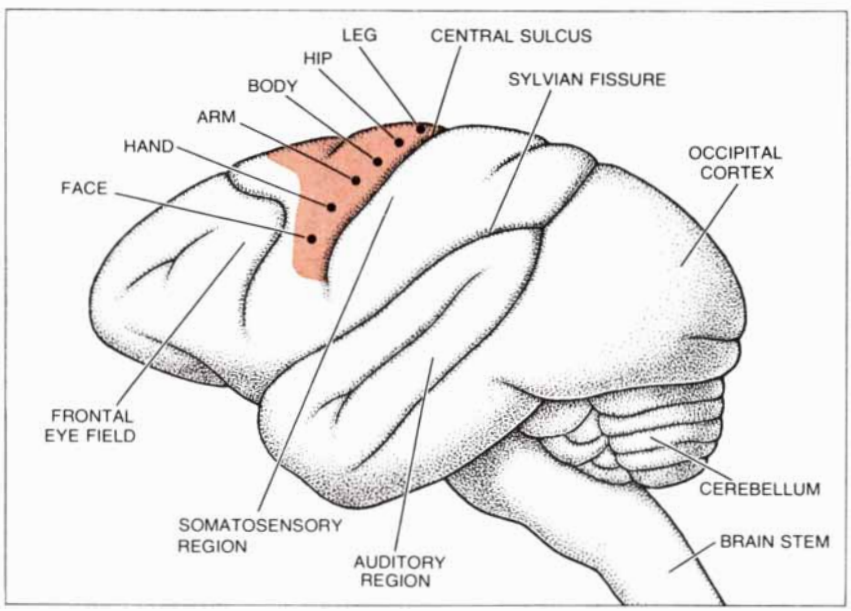
Part of the problem was solved in the 1920's when physiological psychologists



RECORDINGS OF THE ACTIVITY of single nerve cells in the brain are obtained while a monkey performs a learned task in this specially designed "primate chair" in the author's laboratory at the National Institute of Mental Health. The monkey's head is painlessly immobilized so that the microelectrode in the brain does not change position during the experiment. The monkey has been trained to move the vertical rod by flexing its wrist when a light in the signal box comes on. If it makes the required movement within a specified time, it receives a reward of fruit juice through the tube in its mouth. Signals from the microelectrode implanted in the brain, along with data from the signal box and transducers connected to the vertical rod, are fed into a computer for analysis.



MICROELECTRODE ASSEMBLY consists of a fine platinum-iridium wire attached to a hydraulically actuated piston. A stainless-steel cylinder permanently attached to the monkey's skull provides access to the brain. The bolts on the sides of the skull are also permanently implanted. They are attached to clamps during the experiment to prevent head movement. After the electrode assembly is bolted to the cylinder the electrode is lowered by pumping oil into the inlet on the right and raised by pumping oil into the inlet on the left.



CEREBRAL CORTEX of a monkey's brain is depicted with the motor cortex, which controls muscular movement, in color. Electrical stimulation of the points indicated on the motor cortex causes involuntary contraction of the corresponding group of muscles on the opposite side of the body. Damage to an area of the motor cortex usually results in paralysis of the muscles controlled by that area. The frontal eye field is involved in eye movements.

developed techniques for conditioning animals to execute certain movements that could be systematically modified and that could be readily observed and recorded in the laboratory. The greatest stumbling block was finding a way to record the electrical activity of individual nerve cells in the brain of unanesthetized animals. Cerebral nerve cells are extremely small, and in order to record their electrical discharges a microelectrode must be placed within about 50 microns of the membrane of the nerve cell. In addition the microelectrode has to remain in position even when the animal moves. Some 15 years ago Herbert H. Jasper of the Montreal Neurological Institute worked out techniques for recording the activity of individual nerve cells in animals executing learned movements. His contribution consisted in miniaturizing the system for positioning the microelectrode in the brain. The entire apparatus he developed can be attached to the animal's skull, so that head movements do not displace the recording electrode [*see top illustration at left*].

The cerebral motor cortex was long the focal point for research on how the brain controls muscular movements, but today neurophysiological studies are concerned with the cerebellum and the basal ganglia as well [*see illustrations on opposite page*]. The objective of current research is to elucidate how these three interconnected parts of the brain—the motor cortex, the cerebellum and the basal ganglia—act together to control movement.

It is known that damage to the motor cortex causes paralysis but that damage to the basal ganglia or to the cerebellum produces abnormality rather than abolition of movement. For example, symptoms of Parkinson's disease, a neurological disorder resulting from damage to the basal ganglia, include active features such as tremor and muscular rigidity and negative features such as slowness in the initiation of movement and loss of the usual facial expression of emotions. The severity of the motor disorder depends not so much on what muscles are used as on how the muscles are used. A high-velocity movement can sometimes be carried out almost normally by a Parkinsonian patient who in the next moment may have great difficulty initiating a slow movement with the same muscles.

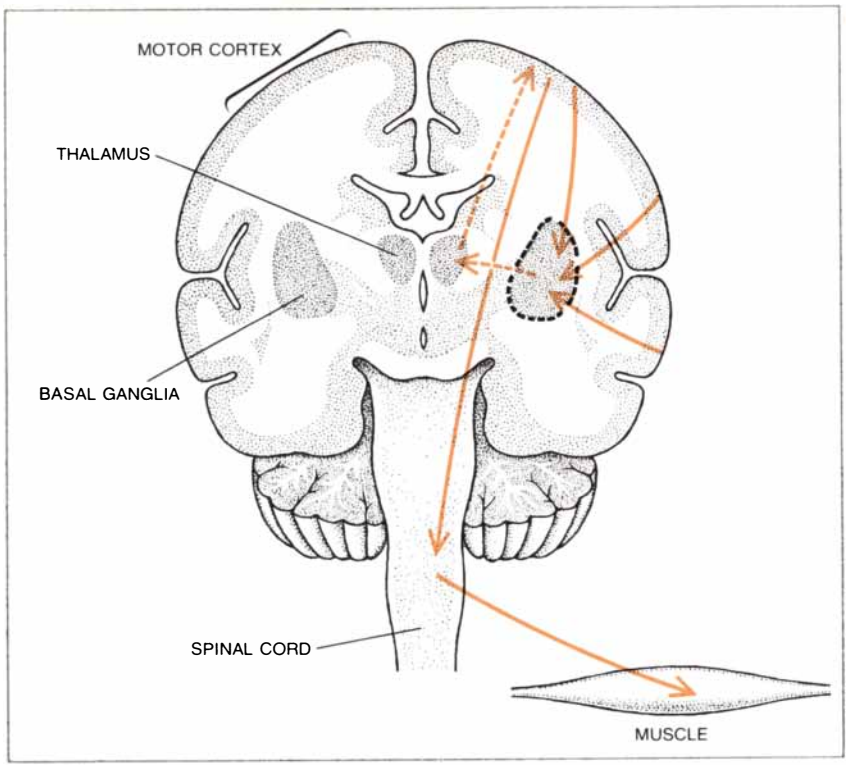
Damage to the cerebellum produces an abnormality of movement that is almost the opposite of the abnormality caused by damage to the basal ganglia. With a cerebellar disorder muscular tremor is most severe during voluntary movement and least marked when the

muscles are at rest. It seems clear that the three motor control centers are functionally interdependent. But in what temporal order do they become active, and what aspect of movement does each control? These are questions that recordings of the activity of single nerve cells in the brain during the movement of specific muscles can help to answer.

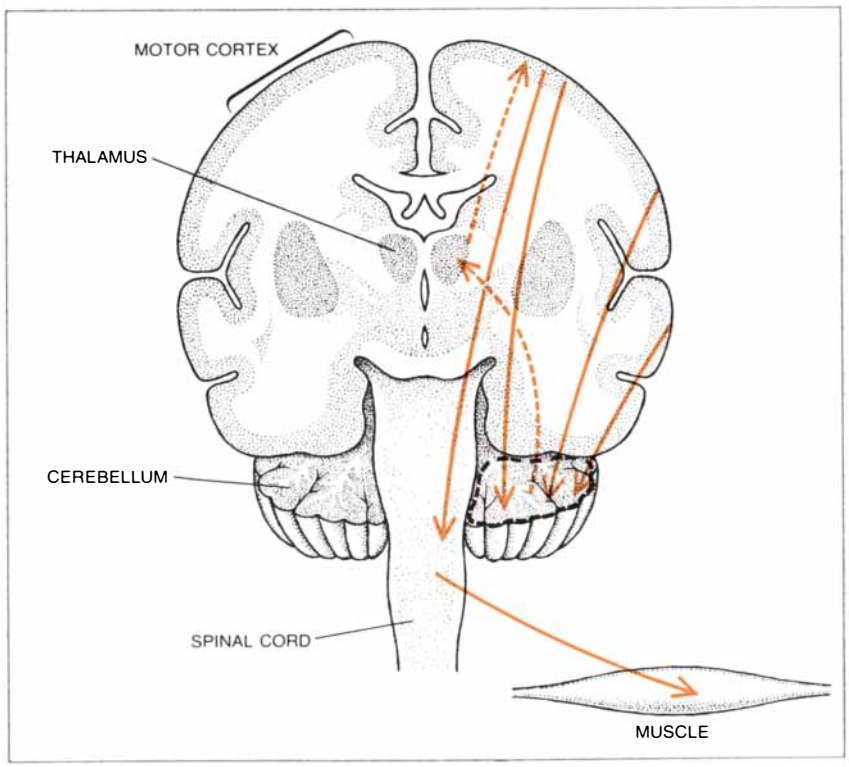
One of the first microelectrode studies involved determining the time at which nerve cells in the motor cortex of monkeys discharged when the monkey executed a simple hand movement. The monkey was trained to depress a telegraph key and to watch for the appearance of a light, which came on at unpredictable times. If the monkey released the telegraph key within 350 milliseconds or less after the light came on, it was rewarded with a few drops of fruit juice. By simultaneously recording both the brain-cell discharges and the muscle discharges, it was found that cells in the motor cortex became active prior to muscular contraction. This, together with the known anatomical connections, indicates that cells in the motor cortex are components in the circuit that initiates the motor response.

Immediately adjacent to the motor cortex is the sensory cortex, which receives inputs from nerve endings in the skin and the joints. Recordings from nerve cells in the sensory receiving area showed activity after rather than before the initial muscular contraction, indicating that although these cells may play a part in guiding movement on the basis of feedback, they are not in the circuit that initiates the first muscular contraction. The sensory cortex is not the only region of the brain with strong inputs from peripheral receptors concerned with motor control. The cerebellum, for example, receives powerful inputs from the vestibular apparatus, which senses the equilibrium of the body, and from muscle receptors. It was commonly believed that the major role of the cerebellum was regulation of movement in response to feedback from the muscles after they had begun their contraction. It therefore came as a surprise when W. Thomas Thach, Jr., of the Yale University School of Medicine discovered that changes in cerebellar activity occurred prior to movement. Then Mahlon DeLong of the National Institute of Mental Health extended the studies to the basal ganglia and found that nerve cells in that region also become active in advance of muscular contraction.

The discovery that all three motor regions discharge prior to movement has



BASAL GANGLIA receive inputs from a wide area of the cerebral cortex and send signals to the motor cortex by way of the thalamus. When the functioning of the basal ganglia is impaired, faulty signals (*broken colored lines*) pass to the motor cortex and cause postural disturbances, muscular tremor at rest and difficulty in the initiation of movement.



CEREBELLUM receives inputs from a wide area of the cerebral cortex. Damage to the cerebellum results in faulty signals to the motor cortex by way of the thalamus (*broken colored lines*). This causes muscular tremor that is most severe during voluntary movement.

led to a new notion of the functional relation of the three structures. The entire cerebral cortex sends fibers to both the basal ganglia and the cerebellum, and these two structures in turn send massive connections back to the motor cortex by way of the thalamus. Thus the basal ganglia and the cerebellum receive information from the somatosensory, visual and auditory regions of the cerebral cortex, transform this information and then send a new pattern of signals to the motor cortex. Whereas the traditional view held that the cerebral motor cortex was at the highest level of motor integration and that the subcortical structures were at a lower level, that is, closer to the muscle, it now appears that the situation is quite the reverse. The inputs going into the cerebellum and into the basal ganglia may be coded in a more abstract and complex manner than the inputs going into the motor cortex. In addition the motor cortex is more directly connected to spinal-cord motor neurons than either the cerebellum or the basal ganglia.

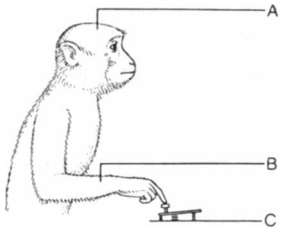
Although all three major motor regions of the brain become active prior to movement, each region is involved in quite different aspects of motor control. When studies were begun to determine the aspects of movement controlled by the motor cortex, there seemed to be two possible alternatives: control of muscle length or control of muscle tension. In other words, it was asked: Do impulses leaving the motor cortex specify the displacement to be produced or do they specify the force required to produce the displacement? Does the motor cortex control the direction and extent of the movement or does it control the direction and magnitude of the forces underlying the movement? A number of studies of motor performance in man have provided evidence for both possibilities. As one investigator, J. A. V. Bates of the National Hospital for Nervous Diseases in London, pointed out: "Force can be looked upon as the body's basic output quantity; velocity is thus the single integral of this, and displacement the double integral. To attempt a desired velocity and a desired displacement are thus in theory more complex operations than to attempt a desired force. But against this it might be emphasized that our everyday experience is a demand for accurate displacement outputs, i.e., practice in double integration."

In order to determine the primary output of the motor cortex my colleagues and I at the National Institute of Mental Health devised an experiment that involved training a monkey to carry out a task in which the direction of force and the direction of displacement could be independently varied. A panel with a vertical rod that could be grasped with one hand was mounted on the monkey's cage. The monkey received a reward of fruit juice when the speed with which the handle was moved back and forth fell within certain time limits. The limits were narrowed as the monkey gained proficiency in carrying out the task. Ultimately the monkey was trained to displace the rod in more than 400 milliseconds but less than 700 milliseconds. Two successive displacements had to be made correctly before the monkey received a reward. The rod was displaced either by a bending of the wrist followed by a straightening of the wrist (flexion followed by extension) or vice versa.

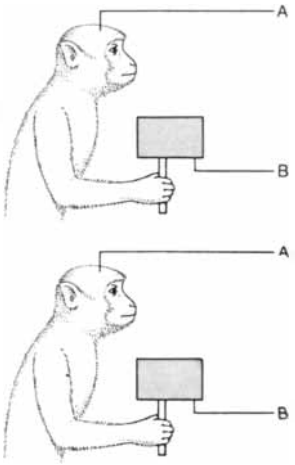
The required cycle consisted of either flexion or extension displacement within the time limits followed by either extension or flexion, also within the time limits. A weight was attached to the rod with a string, and the string was passed over one of two pulleys. When it was passed over one pulley, the load opposed wrist flexion and tended to pull the wrist into an extended position. The monkey had to exert a force in the direction of flexion; even when the load was being lowered, the flexor muscles had to exert a force to prevent it from falling too rapidly. When the string was passed over the other pulley, the situation was reversed: the load now opposed the extensor muscles and as a result the monkey had to exert a net force in the direction of extension. During training both the size of the load and the direction in which it acted were varied so that the monkey learned to make movements of the required duration independently of these variables.

When the monkey was thoroughly trained in its home cage, it was then trained to carry out the same series of wrist movements in a special chair equipped with a recording apparatus. When the monkey's performance was satisfactory, a microelectrode was implanted in its brain. Recordings of a single nerve cell in the motor cortex were then made while the monkey performed the task. The results showed that the activity of nerve cells in the motor cortex was related to the amount and pattern of muscular contraction rather than to the displacement that the contraction produced [*see bottom illustration at right*].

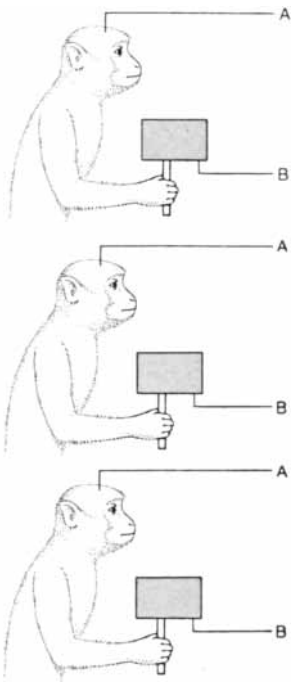
The implication of this finding may be more readily grasped by imagining what



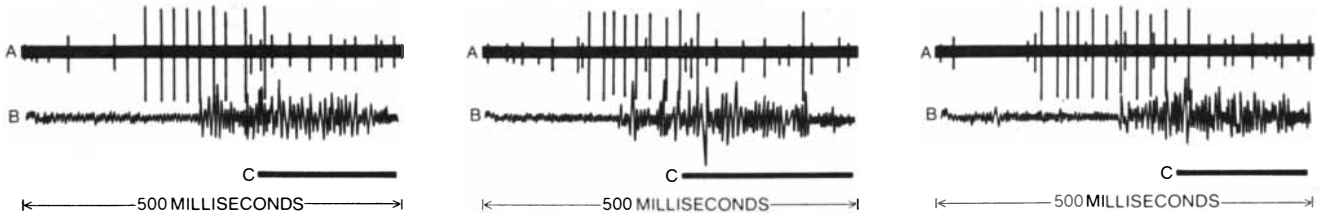
TEMPORAL RELATION between the discharge of a nerve cell in the motor cortex and a simple hand movement is shown at right. A monkey was trained to depress a



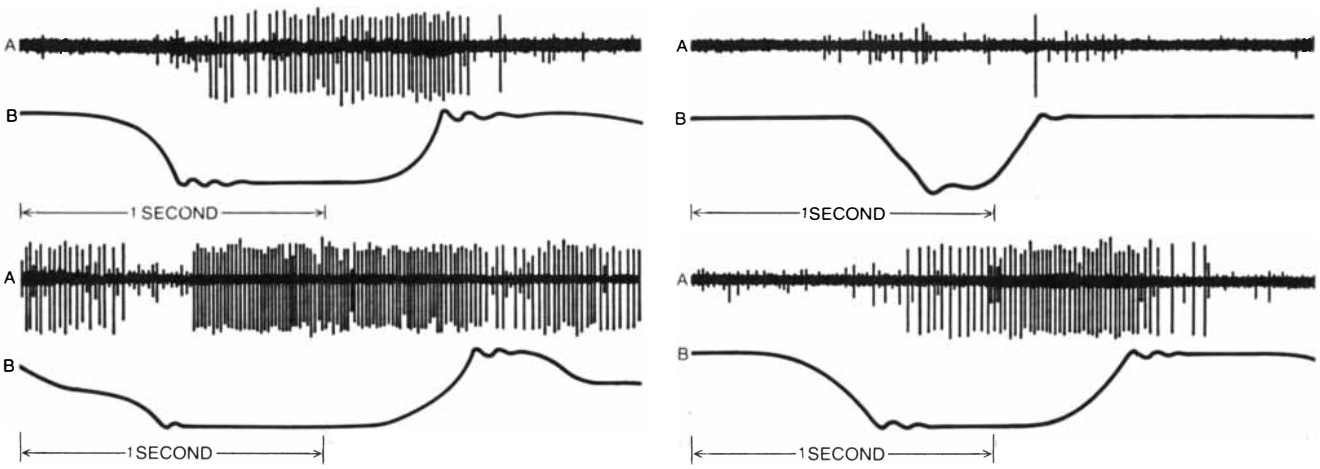
SELECTIVE ACTIVITY of a single nerve cell in the motor cortex during wrist displacement is shown at right. A monkey was trained to move a vertical rod between two stops within a certain time limit. When trace



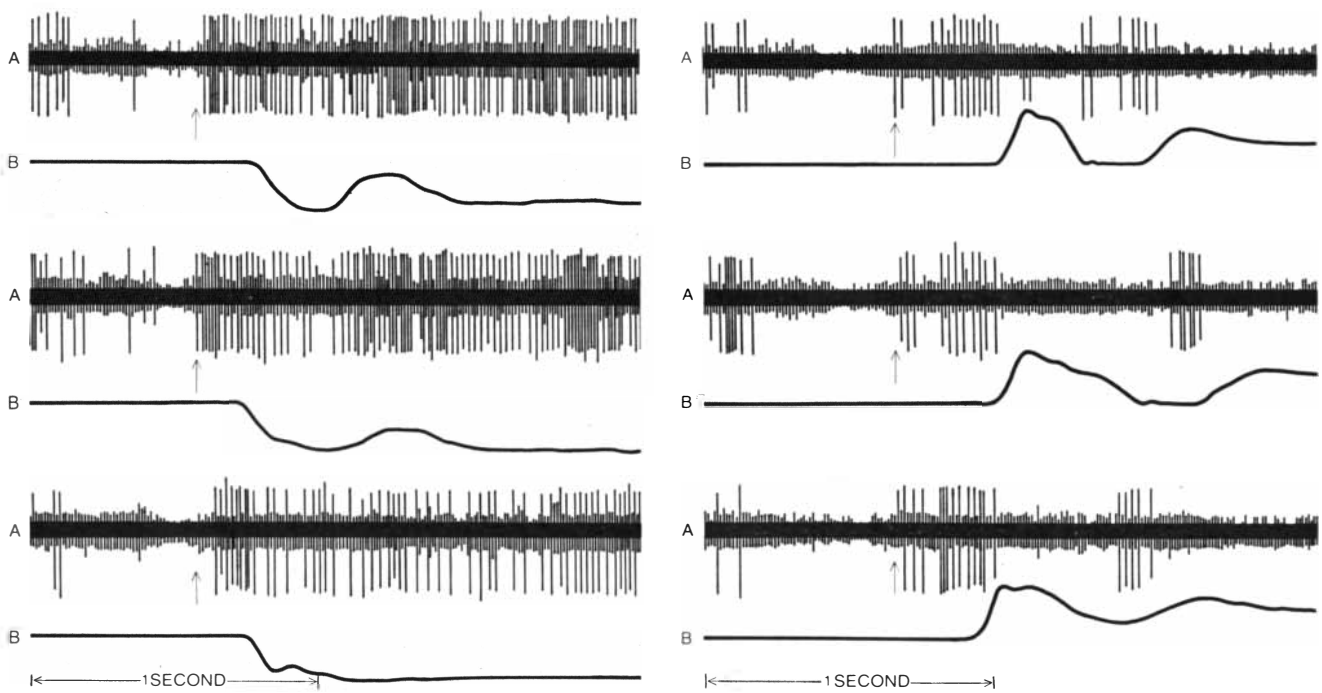
RELATION between activity of nerve cell in the motor cortex and the support of different weights by an arm is shown at right. A monkey was trained to hold a vertical handle between two stops. Weights were



telegraph key and then to release it within 350 milliseconds after a light came on. The upper traces (*A*) show the activity of a single nerve cell in the arm region of the motor cortex, which was recorded by an implanted microelectrode. The traces start at the onset of the light signal. In a series of trials the nerve cell became active first, usually within 150 milliseconds of the signal. There followed a contraction of arm muscles (*B*), which was detected by an electromyograph. Trace (*C*) shows when the telegraph key opened.



B is at its lowest, the wrist is extended. When the trace is at its highest, the wrist is flexed or bent forward. Recordings from a single nerve cell in the motor cortex (*A*) were made during the movements. When there was no weight opposing the movement (*top left*), the cell was active during flexion but not during extension. When a load of 400 grams opposing flexion was added (*bottom left*), the cell became much more active during both flexion and extension. When a 400-gram load opposed extension, however, the cell became almost totally "silent" (*top right*). With no opposing load the nerve cell again showed its initial pattern (*lower right*).



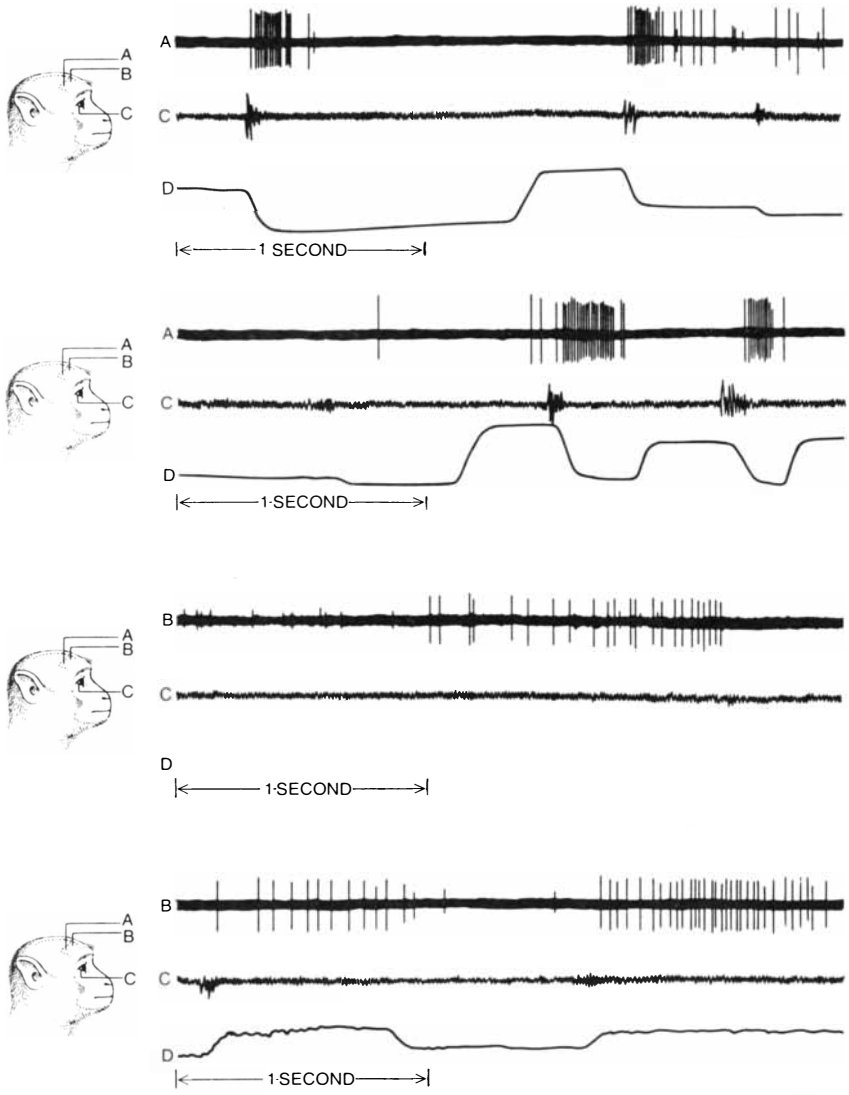
placed sometimes to oppose flexion of the wrist and sometimes to oppose extension. The magnitude of the load was also changed from time to time. Recordings from a single nerve cell in the motor cortex were made with a microelectrode (*A*). Displacement of the rod held is shown by the trace *B* and the arrow indicates when the monkey grasped the rod. The traces on the left from the top down show the activity of the nerve cell when the flexor muscles were respectively supporting weights of 400, 200 and 100 grams. The traces at right show the activity of the nerve cell when the extensor muscles supported the same sequence of weights.

happens when you lift different weights. Suppose in one case you hold a tennis ball and move your arm up and down over a fixed distance and at a fixed speed. Then you replace the tennis ball with a steel ball of the same size and repeat the arm motions exactly in the same manner. To an external observer there would be no difference between the arm movements with the two balls. Both movements would be over the same distance and at the same speed. The patterns of activity in your motor cortex would be quite different for the two movements, however, because the muscular contractions required to lift the heavy steel ball are different from the contractions required to lift the light tennis ball.

Another question about how movements are controlled is whether the pattern of activity in a motor region is related only to the physical aspects of movement or whether there are different patterns of activity that are associated with the same physical movement. For example, when the muscular contractions are exactly the same, is the pattern of activity in the motor area the same for learned movements and for innate, reflex movements? When the cerebral motor cortex was thought to be the highest level of the motor control system, it seemed logical to assume that its nerve cells would be more involved in learned movements than in reflexes. But when the activity of nerve cells in the motor cortex was recorded in association with natural movements such as scratching, eating and grooming, and also in association with highly learned movements such as reaction-time performance, it was found that the motor cortex participated in the control of movement regardless of whether the movement was innate or learned. In both types of movement the activity of the nerve cells in the motor cortex is related to what the muscles do rather than the circumstances in which they do it.

That does not mean that there is a one-to-one relation between the motor-cortex nerve cells and the spinal-cord motor nerve cells. On the contrary, there are a number of differences between the pattern of activity of nerve cells in the motor cortex and the activity of motor nerve cells in the spinal cord. It may be that the relation of the motor-cortex nerve cells to the spinal-cord motor nerve cells is similar to the relation between the cells in the lateral geniculate nucleus (a way station between the retina and the visual cortex) and the cells in the visual region of the cortex. Cells in both the lateral geniculate nucleus and the cortex respond to the location of the stimulus on the retina, but if there is a pattern in the stimulus, it will be processed by the visual cortex. In much the same way it appears that the activity of nerve cells in the motor cortex may be related to certain patterns of activity within a group of muscles, whereas the activity of a spinal-cord motor nerve cell is related only to a single set of fibers in one muscle.

In regions of the cerebral cortex outside the motor cortex, however, neural activity is sometimes more dependent on the context in which the movement occurs than on the muscular activity per se. By context is meant the mental circumstances or intentions associated with the movement of a set of muscles: whether the movement is voluntary or involuntary, learned or innate, fast or slow. For example, in the eye the same muscles serve for saccadic movements and for smooth pursuit movements. Saccadic eye movements are rapid jerks by which the gaze is shifted from one fixation point to another. Smooth pursuit movement is the tracking of a moving object so that the image remains stationary on the retina. Recordings made by Emilio Bizzi of the Massachusetts Institute of Technology from a frontal corti-

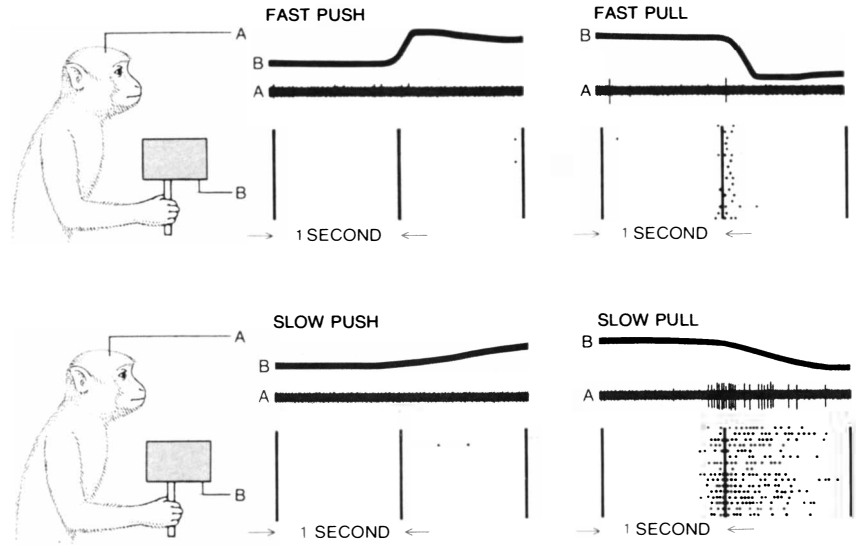


TWO TYPES OF NERVE CELL in the frontal eye field of the cerebral cortex exhibit different patterns of activity even though both are involved in controlling the same eye muscles. Recordings of the activity of single nerve cells in the frontal eye field by Emilio Bizzi of the Massachusetts Institute of Technology show that one type of cell (*A*) discharges during voluntary saccadic eye movement and a second type of cell (*B*) discharges during smooth pursuit eye movement and during maintained position. The electromyographic activity of an eye muscle (*C*) and eye movements (*D*) also were recorded. The top traces show a cell (*A*) discharging during saccadic movements. The bottom traces show the discharge of a different cell (*B*) during smooth pursuit (*upper B*) and maintained eye position (*lower B*).

cal region (outside the motor cortex) involved in controlling eye movements have revealed that one set of nerve cells participated in the saccadic movements and another set was involved in the pursuit movements even though the same muscles were serving both movements. Thus the activity of nerve cells in this frontal cortical region is not associated with the muscle activity, as it is in the motor cortex, but rather with the type of movement. In the motor cortex, on the other hand, the same set of nerve cells was found to control the contraction of arm muscles regardless of the circumstances or context of the movement.

Hans H. Kornhuber of the University of Ulm has proposed that cells in the cerebellum and in the basal ganglia are also differentially active depending on the type of movement rather than on the muscle activity. Drawing evidence both from patients with disorders of movement and from experimental studies in animals, Kornhuber in 1971 suggested that the major role of the cerebellum is to preprogram and initiate rapid saccadic or ballistic movements, whereas the major role of the basal ganglia is to generate slow movements. To test this hypothesis for the basal ganglia, DeLong carried out studies on the activity of individual cells in the basal ganglia of monkeys trained to make a quick limb movement in response to a red light and a slow, highly controlled movement in response to a green light. DeLong found that a high percentage of movement-related cells in a large portion of the basal ganglia discharged most strongly for the slow movement rather than for the fast one. Although the findings are consistent with Kornhuber's hypothesis, the exact role of these cells in the control of slow movements awaits further study. In any event it is clear that compared with the motor cortex, which is involved with both slow and fast movement, the basal ganglia are preferentially active in slow movements. The output of the basal ganglia, which presumably serves to modulate cortical output, goes to the motor cortex by way of the thalamus.

The cerebellum also projects fibers to the motor cortex by way of the thalamus, but we do not yet have direct experimental evidence on the special kinds of movement that the nerve cells of the cerebellum control. Clinical and anatomical data, however, suggest that the cerebellum may have a special role in quick, ballistic movements. Indeed, Kornhuber has proposed that the cerebellum and the basal ganglia are complementary structures, with the cerebellum controlling quick movements and the basal ganglia controlling slow movements.



ACTIVITY OF A SINGLE NERVE CELL in the basal ganglia during pushing and pulling movements of the arm shown here was recorded by Mahlon DeLong of the National Institute of Mental Health. The traces labeled *B* show the position of the arm during the trials. The traces labeled *A* show the activity of the nerve cell in the basal ganglia during the same trial, and the lower portion shows the activity of the cell in raster form during a series of successive trials. The nerve cell discharged during the slow pulling movement but not during the slow pushing movement. During fast movements the cell discharged weakly.

The results obtained so far from recording the activity of single nerve cells with microelectrodes and determining the relation between the recordings and the control of movement have been somewhat fragmentary. Their limitations serve to emphasize how much remains to be learned. A particularly promising area of investigation for the near future is the analysis of experimentally produced motor disturbances in monkeys, disturbances that are similar to the motor disturbances found in man, for example Parkinson's disease. Discovery of the neurophysiological errors in such experimental models of motor disturbances will be of value in developing and testing therapeutic drugs. Many patients with Parkinson's disease have shown a dramatic improvement after taking the drug L-dopa. With neurophysiological studies of Parkinsonism in monkeys it may be possible to determine exactly how L-dopa works. Analogous studies should be feasible for diseases involving the cerebellum.

People suffering from Parkinson's disease exhibit emotional as well as muscular disorders. Other basal-ganglia diseases, some of them genetically determined, are associated with psychological disorders. This indicates that the basal ganglia may be the region that provides the major associative link between the more specialized sensory division of the nervous system and the motor division.

The implications of the studies I have described thus extend into the areas of psychology and psychiatry. Indeed, it seems possible that understanding of the human nervous system, even its most complex intellectual functions, may be enriched if the operation of the brain is analyzed in terms of its motor output rather than in terms of its sensory input. In the past most attempts to describe the higher functions of the brain have been made in terms of how sensory inputs are processed from the receptor on up to the higher cortical centers. A strong case for an alternative approach has been made by Roger W. Sperry of the California Institute of Technology. I shall end with his comment: "Instead of regarding motor activity as being subsidiary, that is, something to carry out, serve and satisfy the demands of the higher centers, we reverse this tendency and look upon the mental activity as only a means to an end, where the end is better regulation of overt response. Cerebration essentially serves to bring into motor behavior additional refinement, increased direction toward distant, future goals and greater overall adaptiveness and survival value. The evolutionary increase in man's capacity for perception, feeling, ideation, imagination and the like may be regarded not so much as an end in itself as something that has enabled us to behave, to act, more wisely and efficiently."

MATHEMATICAL GAMES

Free will revisited, with a mind-bending prediction paradox by William Newcomb

by Martin Gardner

A common opinion prevails that the juice has ages ago been pressed out of the free-will controversy, and that no new champion can do more than warm up stale arguments which every one has heard. This is a radical mistake. I know of no subject less worn out, or in which inventive genius has a better chance of breaking open new ground.

—WILLIAM JAMES

One of the perennial problems of philosophy is how to explain (or explain away) the nature of free will. If the concept is explicated within a framework of determinism, the will ceases to be free in any commonly understood sense and it is hard to see how fatalism can be avoided. *Che sarà, sarà.* Why work hard for a better future for yourself or for others if what you do must always be what you do do? And how can you blame anyone for anything if he could not have done otherwise?

On the other hand, attempts to explicate will in a framework of indeterminism seem equally futile. If an action is not caused by the previous states of oneself and the world, it is hard to see how to keep the action from being haphazard. The notion that decisions are made by some kind of randomizer in the mind does not provide much support for what is meant by free will either.

Philosophers have never agreed on how to avoid the horns of this dilemma. Even within a particular school there have been sharp disagreements. William James and John Dewey, America's two leading pragmatists, are a case in point. Although Dewey was a valiant defender of democratic freedoms, his metaphysics regarded human behavior as completely determined by what James called the total "push of the past." Free will for Dewey was as illusory as it is in the psychology of B. F. Skinner. In contrast James was a thoroughgoing indeterminist. He believed that minds had the power to inject genuine novelty into history, that not even God himself could know the future except partially. "*That,*" he wrote, "is what gives the palpitating reality to our moral life and makes it tingle... with so strange and elaborate an excitement."

A third approach, pursued in depth by Immanuel Kant, accepts both sides of the controversy as being equally true but incommensurable ways of viewing human behavior. For Kant the situation is something like that pictured in one of Piet Hein's "grooks":

A bit beyond perception's reach
I sometimes believe I see
That Life is two locked boxes, each
Containing the other's key.

Free will is neither fate nor chance. In some unfathomable way it partakes of both. Each is the key to the other. It is not a contradictory concept, like a square triangle, but a paradox that our experience forces on us and whose resolution transcends human thought. That was how Niels Bohr saw it. He found the situation similar to his "principle of complementarity" in quantum mechanics. It is a viewpoint that Einstein, a Spinozist, found distasteful, but many other physicists, J. Robert Oppenheimer for one, found Bohr's viewpoint enormously attractive.

What has free will to do with mathematical games? The answer is that in recent decades philosophers of science have been wrestling with a variety of queer "prediction paradoxes" related to the problem of will. Some of them are best regarded as a game situation. One draws a payoff matrix and tries to determine a player's best strategy, only to find oneself trapped in a maze of bewildering ambiguities about time and causality.

A marvelous example of such a paradox came to light in 1970 in a paper, "Newcomb's Problem and Two Principles of Choice," by Robert Nozick, a philosopher at Harvard University. The paradox is so profound, so amusing, so mind-bending, with thinkers so evenly divided into two warring camps, that it bids fair to produce a literature vaster than that dealing with the prediction paradox of the unexpected hanging. (See this department for March, 1963, or the reprinted version of that piece in *The Unexpected Hanging and Other Mathematical Diversions,* Simon and Schuster, 1969.)

Newcomb's paradox is named after its originator, William A. Newcomb, a theoretical physicist at the University of California's Lawrence Livermore Laboratory. (His great-grandfather was the brother of Simon Newcomb, the astronomer.) Newcomb thought of the problem in 1960 while meditating on a famous paradox of game theory called the prisoner's dilemma [see "Escape from Paradox," by Anatol Rapoport; SCIENTIFIC AMERICAN, July, 1967]. A few years later Newcomb's problem reached Nozick by way of their mutual friend Martin David Kruskal, a Princeton University mathematician. "It is not clear that I am entitled to present this paper," Nozick writes. "It is a beautiful problem. I wish it were mine." Although Nozick could not resolve it, he decided to write it up anyway. His paper appears in *Essays in Honor of Carl G. Hempel,* edited by Nicholas Rescher and published by D. Reidel in 1970. What follows is largely a paraphrase of Nozick's paper.

		BEING	
		MOVE 1 (PREDICTS YOU TAKE ONLY BOX 2)	MOVE 2 (PREDICTS YOU TAKE BOTH BOXES)
YOU	MOVE 1 (TAKE ONLY BOX 2)	$1,000,000	$0
	MOVE 2 (TAKE BOTH BOXES)	$1,001,000	$1,000

Payoff matrix for Newcomb's paradox

Two closed boxes, B1 and B2, are on a table. B1 contains $1,000. B2 contains either nothing or $1 million. You do not know which. You have an irrevocable choice between two actions:

1. Take what is in both boxes.
2. Take only what is in B2.

At some time before the test a superior Being has made a prediction about what you will decide. It is not necessary to assume determinism, only that you are persuaded that the Being's predictions are "almost certainly" correct. If you like, you can think of the Being as being God, but the paradox is just as strong if you regard the Being as a superior intelligence from another planet, or a supercomputer capable of probing your brain and making highly accurate predictions about your decisions. If the Being expects you to choose both boxes, he has left B2 empty. If he expects you to take only B2, he has put $1 million in it. (If he expects you to randomize your choice by, say, flipping a coin, he has left B2 empty.) In all cases B1 contains $1,000. You understand the situation fully, the Being knows you understand, you know that he knows and so on.

What should you do? Clearly it is not to your advantage to flip a coin, so that you must decide on your own. The paradox lies in the disturbing fact that a strong argument can be made for either decision. Both arguments cannot be right. The problem is to explain why one is wrong.

Let us look first at the argument for taking only B2. You believe the Being is an excellent predictor. If you take both boxes, the Being almost certainly will have anticipated your action and have left B2 empty. You will get only the $1,000 in B1. Contrariwise, if you take only B2, the Being, expecting that, almost certainly will have placed $1 million in it. Clearly it is to your advantage to take only B2.

Convincing? Yes, but the Being made his prediction, say a week ago, and then left. Either he put the $1 million in B2 or he did not. "If the money is already there, it will stay there whatever you choose. It is not going to disappear. If it is not already there, it is not going to suddenly appear if you choose only what is in the second box." It is assumed that no "backward causality" is operating, that is, your present actions cannot influence what the Being did last week. So why not take both boxes and get everything that is there? If B2 is filled, you get $1,001,000. If it is empty, you get at least $1,000. If you are so foolish as to take only B2, you know you cannot get more than $1 million, and there is even a slight possibility of getting nothing. Clearly it is to your advantage to take both boxes!

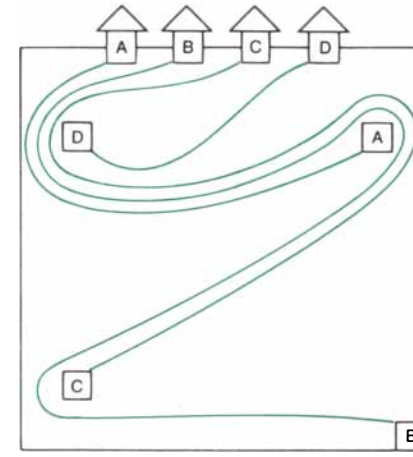
"I have put this problem to a large number of people, both friends and students in class," writes Nozick. "To almost everyone it is perfectly clear and obvious what should be done. The difficulty is that these people seem to divide almost evenly on the problem, with large numbers thinking that the opposing half is just being silly.

"Given two such compelling opposing arguments, it will not do to rest content with one's belief that one knows what to do. Nor will it do to just repeat one of the arguments, loudly and slowly. One must also disarm the opposing argument; explain away its force while showing it due respect."

Nozick sharpens the "pull" of the two arguments as follows. Suppose the experiment has been done many times before. In every case the Being predicted correctly. Those who took both boxes always got only $1,000, those who took only B2 got $1 million. You have no reason to suppose your case will be different. If a friend were observing the scene, it would be completely rational for him to bet, giving high odds, that if you take both boxes, you will get only $1,000. Indeed, if there is a time delay after your choice of both boxes, you know it would be rational for you yourself to bet, offering high odds, that you will get only $1,000. Knowing this, would you not be a fool to take both boxes?

Alas, the other argument makes you out to be just as big a fool if you do not. Assume that B1 is transparent. You see the $1,000 inside. You cannot see into B2, but the far side is transparent and your friend is sitting opposite. He knows whether the box is empty or contains $1 million. Although he says nothing, you realize that whatever the state of B2 is he wants you to take both boxes. He wants you to because, regardless of the state of B2, you are sure to come out ahead by $1,000. Why not take advantage of the fact that the Being played first and cannot alter his move?

Nozick, a specialist in decision theory, approaches the paradox by considering analogous game situations in which, as here, there is a conflict between two respected principles of choice: the "expected-utility principle" and the "dominance principle." To see how the principles apply, consider the payoff matrix for Newcomb's game [*see illustration on opposite page*]. The argument for taking only B2 derives from the principle that



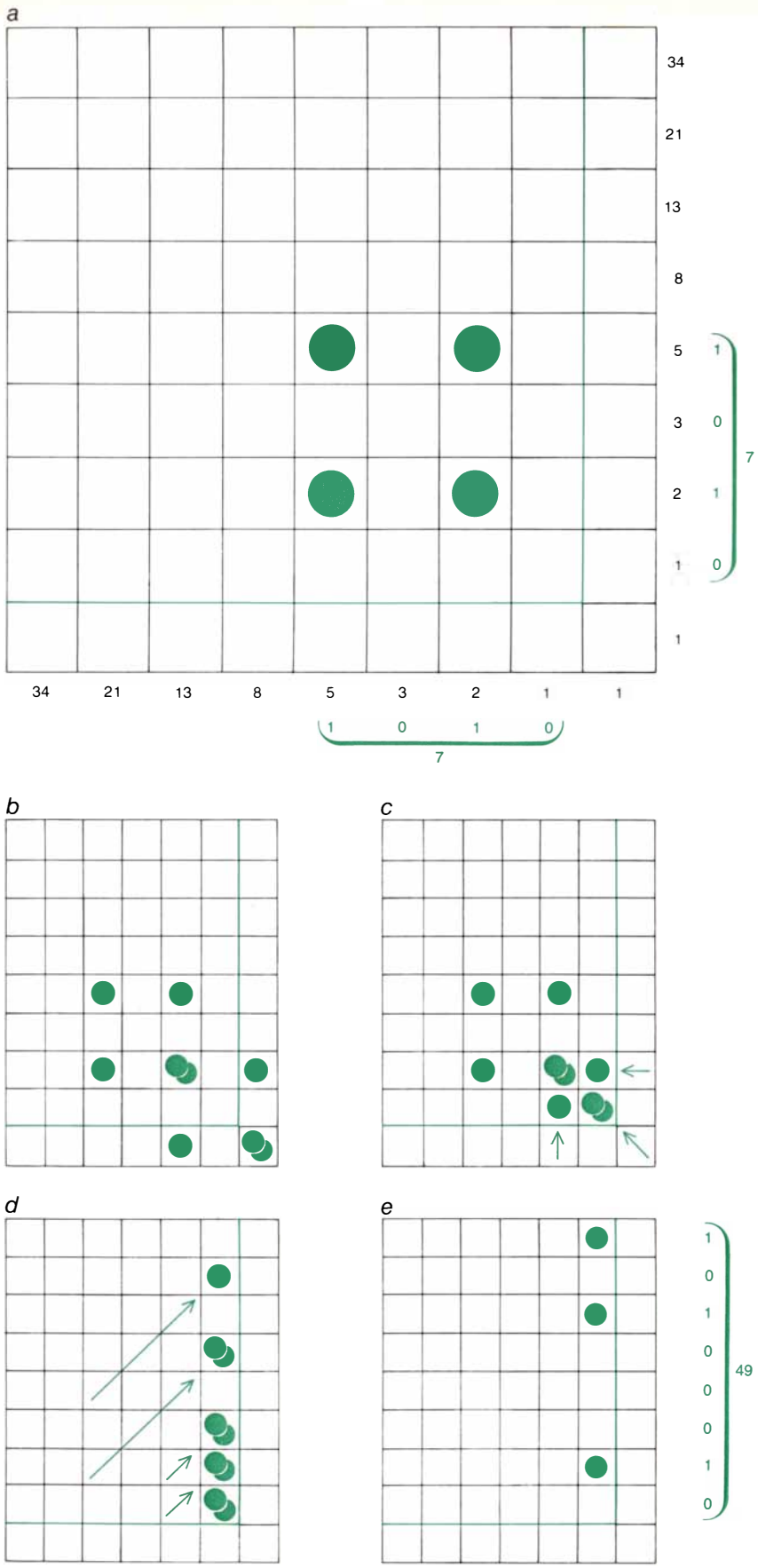
Solution to the schoolhouses problem

you should choose so as to maximize the expected utility (value to you) of the outcome. Game theory calculates the expected utility of each action by multiplying each of its mutually exclusive outcomes by the probability of the outcome, given the action. We have assumed that the Being predicts with near-certainty, but let us be conservative and make the probability a mere .9. The expected utility of taking both boxes is $(.1 \times \$1{,}001{,}000) + (.9 \times \$1{,}000) = \$101{,}000$. The expected utility of taking only B2 is $(.9 \times \$1{,}000{,}000) + (.1 \times \$0) = \$900{,}000$. Guided by this principle, your best strategy is to take only the second box.

The dominance principle, however, is just as intuitively sound. Suppose the world divided into n different states. For each state k mutually exclusive actions are open to you. If in at least one state you are better off choosing a, and in all other states either a is the best choice or the choices are equal, then the dominance principle asserts that you should choose a. Look again at the payoff matrix on the opposite page. The states are the outcomes of the Being's two moves. Taking both boxes is strongly dominant. For each state it gives you $1,000 more than you would get by taking only the second box.

That is as far as we can go into Nozick's analysis, but interested readers should look it up for its mind-boggling conflict situations related to Newcomb's problem. Nozick finally arrives at the following tentative conclusions:

If you believe in absolute determinism, and that the Being has in truth predicted your behavior with unswerving accuracy, you should "choose" (whatever that can mean!) to take only B2. For example, suppose the Being is God and you are a devout Calvinist, con-



Fibonacci notation for 7 × 7

vinced that God knows every detail of your future. Or assume that the Being has a time-traveling device he can launch into the future and bring back with a motion picture of what you in fact did on that future occasion when you made your choice. Believing that, you should take only B2, firmly persuaded that your feeling of having made a genuine choice is sheer illusion.

Nozick reminds us, however, that Newcomb's paradox does *not* assume that the Being has perfect predictive power. If you believe that you possess a tiny bit of free will (or alternatively that the Being is sometimes wrong, say once in every 20 billion cases), then this may be one of the times the Being has erred. Your wisest decision is to take both boxes.

Nozick is not happy with this conclusion. "Could the difference between one in n and none in n, for arbitrarily large finite n, make this difference? And how exactly does the fact that the predictor is certain to have been correct dissolve the force of the dominance argument?" Both questions are left unanswered. Nozick hopes that publishing the problem "may call forth a solution which will enable me to stop returning, periodically, to it."

One such solution, "to restore [Nozick's] peace of mind," was attempted by Maya Bar-Hillel and Avishai Margalit of Hebrew University in Jerusalem in their paper "Newcomb's Paradox Revisited," in *British Journal for the Philosophy of Science*, Volume 23 (1972), pages 295–304. They adopt the same game-theory approach taken by Nozick but come to an opposite conclusion. Even though the Being is not a perfect predictor, they recommend taking only the second box. You must, they argue, resign yourself to the fact that your best strategy is to behave *as if* the Being has made a correct prediction, even though you know there is a slight chance he has erred. You know he has played before you, but you cannot do better than to play as if he is going to play after you. "For you cannot outwit the Being except by knowing what he predicted, but you cannot know, or even meaningfully guess, at what he predicted before actually making your final choice."

It may seem to you, Bar-Hillel and Margalit write, that backward causality is operating—that somehow your choice makes the $1 million more likely to be in the second box—but this is pure flim-flam. You choose only B2 "because it is inductively known to correlate remarkably with the existence of this sum in

Other airlines fly as fast as we do.

And their aircraft have the same range as ours. How can we then offer faster services and fewer stops between Europe and S.E. Asia, Japan, Australia, N. America, West.

And between S.E. Asia and N. America, East.

How, indeed?

The answer is simple: *Copenhagen*. Or, more precisely, the geographical position of Copenhagen, the main gateway of our intercontinental routes.

Look at a globe. Or, better still, stretch a string over its surface between points within the areas above.

In many cases you will find that the string runs straight over Copenhagen (try, for instance, Paris-Tokyo, or London-Sydney, or Berlin-Seattle, or Glasgow-Singapore).

In other cases, Copenhagen will at least be closer to the stretched string than any other major airport in Europe (like Bangkok-New York, which makes the string run as far north as Spitsbergen).

So it's as simple as that: it's shorter via Copenhagen. And flying via Copenhagen is like running the inner track: a way to save distance and time.

And Copenhagen is the only major airport in Europe, from which nonstop flights to Bangkok, Los Angeles, Tokyo and Seattle are all within the reach of long-distance aircraft.

When you're going far east or far west you have two alternatives:

FLY VIA COPENHAGEN – OR FACE A DETOUR

SAS SCANDINAVIAN AIRLINES
General Agent for Thai International

the box, and though we do not assume a causal relationship, there is no better alternative strategy than to behave as if the relationship was, in fact, causal."

For those who argue for taking only B2 on the grounds that causality is independent of the direction of time—that your decision actually "causes" the second box to be either empty or filled with $1 million—Newcomb proposed the following variant of his paradox. Both boxes are transparent. B1 contains the usual $1,000. B2 contains a piece of paper with a fairly large integer written on it. You do not know whether the number is prime or composite. If it proves to be prime (you must not test it, of course, until after you have made your choice), then you get $1 million. The Being has chosen a prime number if he predicts you will take only B2 but has picked a composite number if he predicts you will take both boxes.

Obviously you cannot by an act of will make the large number change from prime to composite or vice versa. The nature of the number is fixed for eternity. So why not take both boxes? If it is prime, you get $1,001,000. If it is not, you get at least $1,000. (Instead of a number B2 could contain any statement of a decidable mathematical fact that you do not investigate until after your choice.)

It is easy to think of other variations. For example, there are 100 little boxes each holding a $10 bill. If the Being expects you to take all of them, he has put nothing else in them. But if he expects you to take only one box—perhaps you pick it at random—he has added to that box a large diamond. There have been thousands of previous tests, half of them involving you as a player. Each time, with possibly a few exceptions, the player who took a single box got the diamond and the player who took all the boxes got only the money. Acting pragmatically, on the basis of past experience, you should take only one box. But then how can you refute the logic of the argument that says you have everything to gain and nothing to lose if the next time you play you take all the boxes?

These variants add nothing essentially new. With reference to the original version Nozick halfheartedly recommends taking both boxes. Bar-Hillel and Margalit strongly urge you to "join the millionaire's club" by taking only B2. That is also the view of Kruskal and Newcomb. But has either side really done more than just repeat its case "loudly and slowly"? Can it be that Newcomb's paradox validates free will by invalidat-

ing the possibility, in principle, of a predictor capable of guessing a person's choice between two equally rational actions with better than 50 percent accuracy?

What does the reader think? I cannot answer letters, but in a later piece I shall report on which side got the largest vote and comment on letters of particular interest.

The first of last month's questions asked for a formula giving the maximum number of noncrossing edges that can be drawn as part of a complete graph for n points. It is $3n - 6$, for n greater than 2. The corresponding formula for complete bipartite graphs of m,n points is $2(m + n) - 4$, for m and n each greater than 1. "Odd," a friend once remarked of this second formula, "that the number is always even." Proofs of both cases are not difficult. These formulas for noncrossing edges are of no help in finding formulas for crossing numbers because there is no known way to predict the minimum number of crossings produced by the edges not drawn.

One solution to the four-schoolhouses puzzle, in which four boys have to reach their respective schools without any of their paths crossing one another or going outside the boundary, is shown in the illustration on page 105.

John Harris of Santa Barbara, Calif., discovered an ingenious way to multiply numbers in Fibonacci notation, using the Napier counting board described in April. He added an extra 1-row and 1-column outside the heavy line to the counting board [*see illustration on page 106*]. Suppose you want to multiply 7 by 7. Place the counters according to Napier's rules [*see "a" in the illustration*]. More counters are now positioned according to the following rule: On the diagonal that extends down and to the right from each counter, n, put a counter on every alternate cell, starting with the cell two cells away from counter n [b].

Each counter outside the heavy line is moved to the nearest cell inside the line [c]. Now move all counters up and to the right along their diagonals to the heavy line [d]. Clear the column according to the Fibonacci clearing rules given in April [e]. The counters, reading from the top down, give the correct product in Fibonacci notation. Readers familiar with the Fibonacci series will enjoy proving that Harris' algorithm works. Division by this method, however, seems to be hopelessly complicated.

THE AMATEUR SCIENTIST

How to blow bubbles that survive for years

Conducted by C. L. Stong

Large bubbles that display iridescence and other properties of soap bubbles can be blown from a recently developed solution of clear plastic. Plastic bubbles last for years. Unlike soap bubbles, they do not collapse when touched. Moreover, they can be colored with dye and coated with ornamental particles. Metallic powders and chemicals similarly applied can make the bubbles electrically conductive and even magnetic. The bubbles can be blown in a variety of shapes. The thickness and flexibility of the films can be altered at will.

For these reasons the new bubble solution should open fields of novel experimentation to enthusiasts of both the arts and the sciences. The plastic-bubble project was developed as a hobby by Aristid V. Grosse, president of Germantown Laboratories, Inc., an affiliate of the Franklin Institute in Philadelphia. Grosse writes:

"The new bubble solution is an outgrowth of the soap-bubble solution I described in these columns four years ago [see "The Amateur Scientist," SCIENTIFIC AMERICAN, May, 1969]. During those experiments I increased the life of soap bubbles substantially by adding a small amount of polyvinyl alcohol to the soap solution. Polyvinyl alcohol is a water-soluble plastic.

"Subsequently it occurred to me that a number of other organic polymers of high molecular weight would form bubbles. Examples are polystyrene, polyvinyl chloride, polyvinyl acetate, polyurethanes and cellulose esters. Water is specifically excluded from these compounds. The result is a water-insoluble film.

"Bubbles that are blown from appropriate solutions of plastic look much like soap bubbles. With plastic, however, surface tension plays no role in determining either the shape of the plastic bubbles or their behavior. For example, they do not necessarily expand into a sphere, nor do they shrink with time to form a flat film across the end of the blowpipe. Perhaps I should say, rather, that these properties have been observed in the solutions I have developed to date. The possibility of finding solutions with other characteristics is high. The five polymers I have named can be combined with some two dozen solvents to make more than 125 different combinations.

"Many other combinations can be devised. For example, the solvents can be mixed in various ratios and with other substances to control the rate of evaporation and alter the mechanical properties of the film, including its stiffness. Other possible variations include the temperature and viscosity of the solution, the molecular weight of the polymer and the rate at which the bubbles are blown.

"I tried only a limited number of the many variations before settling on a combination of polyvinyl acetate and acetone. It seems unlikely that I had the good luck to hit on the best possible solution after only a few experiments. Still, the combination of polyvinyl acetate and acetone is inexpensive and easy to reproduce. It can be blown into bubbles readily, and both the shape and the size of the bubbles can be controlled within wide limits.

"The vinyl polymers have been known for about 60 years, although their versatility and breadth of application was not generally appreciated before World War II. The vinyl acetate monomer is made by reacting acetylene and acetic acid in the presence of a mercury catalyst. The resulting compound can be transformed into the plastic polyvinyl acetate by heating the monomer in the presence of benzoyl peroxide.

"The resin has reasonably good heat stability. It will melt without decomposing. On the other hand, its stability of form is poor: articles molded from polyvinyl acetate tend with time to flow or flatten or otherwise lose their shape. For this reason the plastic is not generally used for molding solid articles, but it makes splendid adhesives and lacquers and is particularly good for blowing bubbles. When the resin is treated with alkali, the acetate groups are removed from the compound and replaced by hydroxyl groups to form polyvinyl alcohol, which is the water-soluble resin I used to increase the life of soap bubbles.

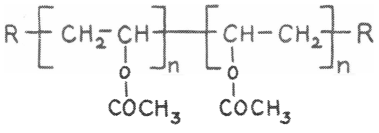
"The structural formula of the monomer is relatively simple [*see illustration on this page*]. The molecular weight of the polymer is determined by the number of times (*n* in the formula) the monomer repeats itself. The properties of the resin depend in large part on the molecular weight of the polymer. For this reason a given resin can be defined chemically only if *n* is known or if the molecular weight of the polymer is known.

"Manufacturers usually identify polymers of specific molecular weight by an alphabetical or a numerical code. For example, I experimented with two grades of polyvinyl acetate made by the Union Carbide Corporation: grade AYAT (molecular weight 167,000, *n* 1,940) and grade AYAF (molecular weight 113,000, *n* 1,315). Both grades are shipped by the manufacturer in the form of colorless, pea-size pellets that have a density of 1.18 grams per cubic centimeter at a temperature of 20 degrees Celsius. Still another grade by the same manufacturer, AYAC, has a molecular weight of only 13,000 (*n* 150). My most successful

vinyl acetate monomer

$CH_2 = CHO - COCH_3$

polyvinyl acetate



R= free radical

Structural formula of polyvinyl acetate

bubbles have been blown with the AYAT and AYAF grades. Bubbles blown with grade AYAC have quite different mechanical properties. They tend to be brittle.

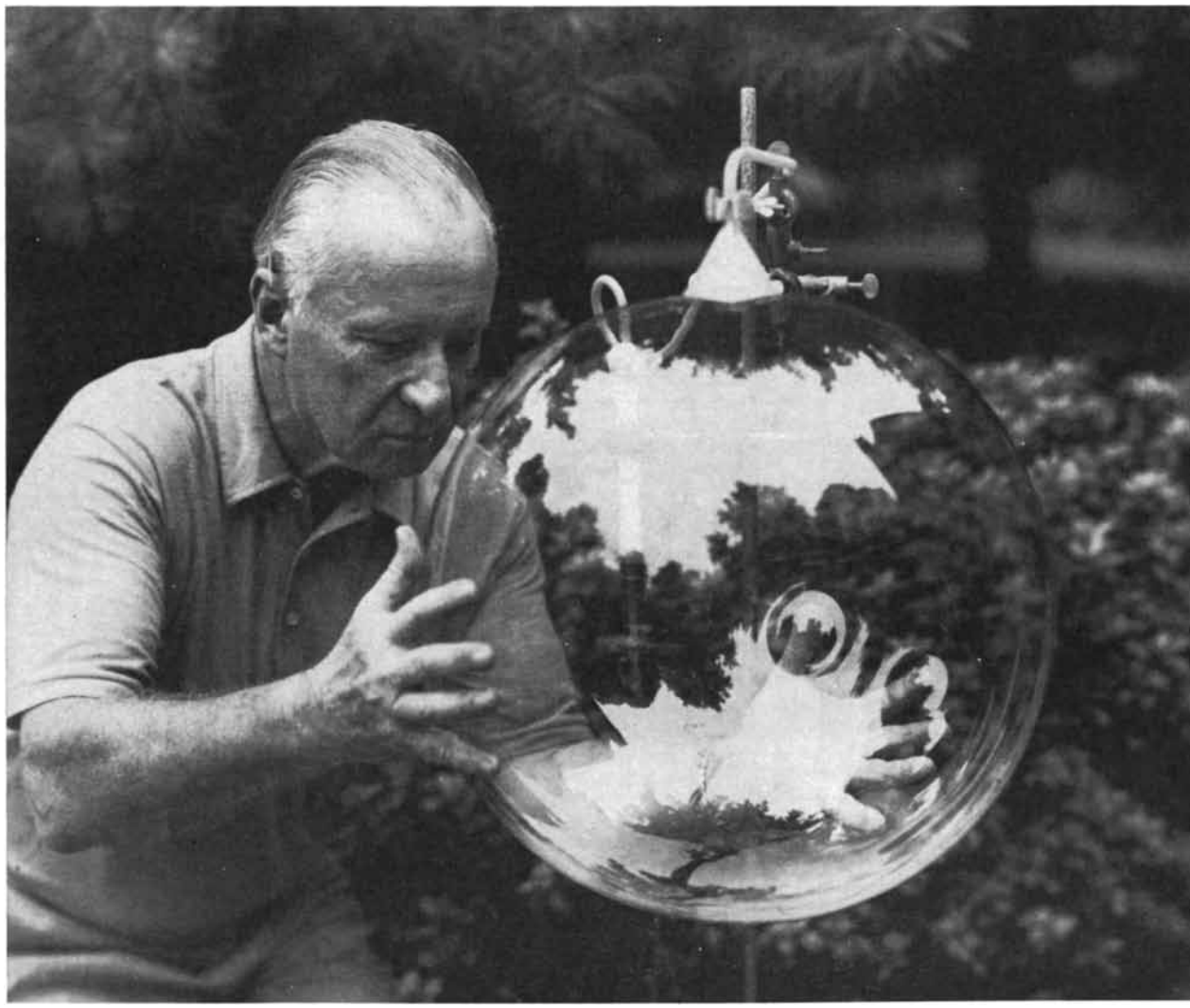
"As I have mentioned, I make the bubble solution by dissolving the polymer in acetone, which is one of the least costly solvents commonly used in chemical laboratories. In bulk it is priced by manufacturers at less than 15 cents a pound, but it may cost more than $1 a pound if it is bought in small quantities from druggists or dealers in paint supplies. Like many chemicals, acetone is a hazardous substance. The vapor is toxic and the liquid is highly flammable. Never smoke or allow flames or sparks near either a container of acetone or plastic solutions that contain the solvent. Experienced workers should handle acetone only in a well-ventilated area. Beginners are urged to work with the bubble solution either in a fume hood or, better yet, outdoors. The bubbles can be brought indoors after the acetone has evaporated.

"Nearly all my experiments were done in the basement of my house, which I have converted into a well-ventilated laboratory. Room temperature in the area is reasonably constant at 20 degrees C. Typically I make up plastic-bubble solution by putting about a pound of AYAT or AYAF pellets in a wide-mouthed jar with a cap. To the pellets I add a pint of pure acetone, which covers the polymer completely. If those proportions are maintained, the absolute quantity of the batch is not critical. A beginner can make up any smaller amount, such as four ounces of pellets in four ounces of solvent.

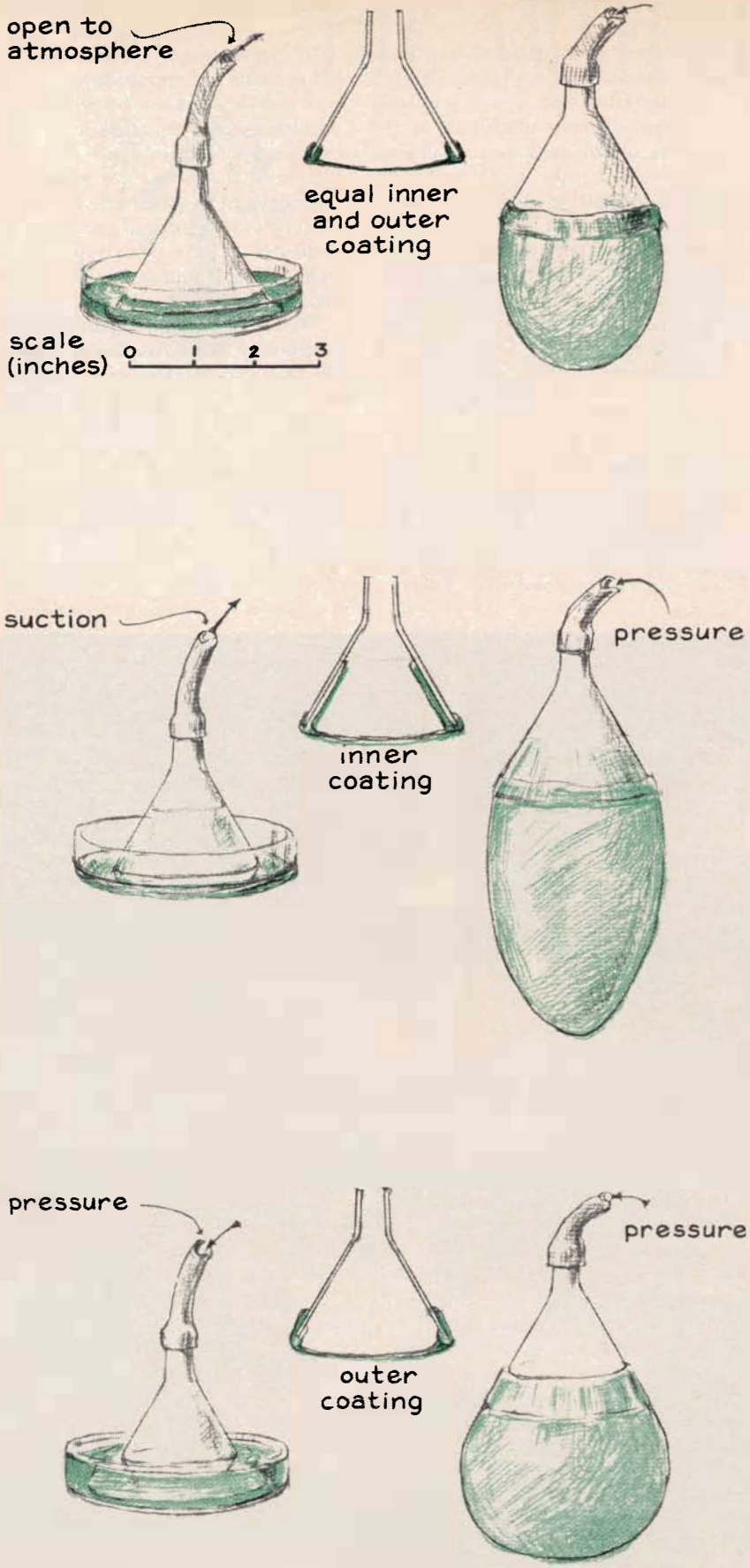
"A laboratory desiccator makes a convenient preparation vessel, particularly the type that has a serrated hose connection for coupling to a vacuum pump. I stir the solution gently but thoroughly every few hours. The pellets swell and dissolve slowly. The mixture becomes increasingly viscous with time.

"If evaporation is minimized, the solution consists of approximately 56 percent polyvinyl acetate by weight. The mixture usually contains many small air bubbles that become trapped in the solution when the pellets are stirred. The bubbles must be removed. I eliminate them by lowering the air pressure in the desiccator to two pounds per square inch, which is equivalent to a partial vacuum of about 100 torr.

"The desiccator can be pumped to this pressure by an inexpensive aspirator of the type that operates on a water tap. The solution tends to foam somewhat as the air pressure is reduced. Do not lower the pressure so abruptly that foam fills the space above the solution and overflows into the exhaust port. The foam subsides with time. I usually maintain the vessel at reduced pressure for about three hours. After air is admitted the



Aristid V. Grosse with one of his plastic bubbles



Techniques of applying plastic to the blowpipe

mixture is completely colorless, clear and free of trapped bubbles. It is fully homogeneous and ready for use.

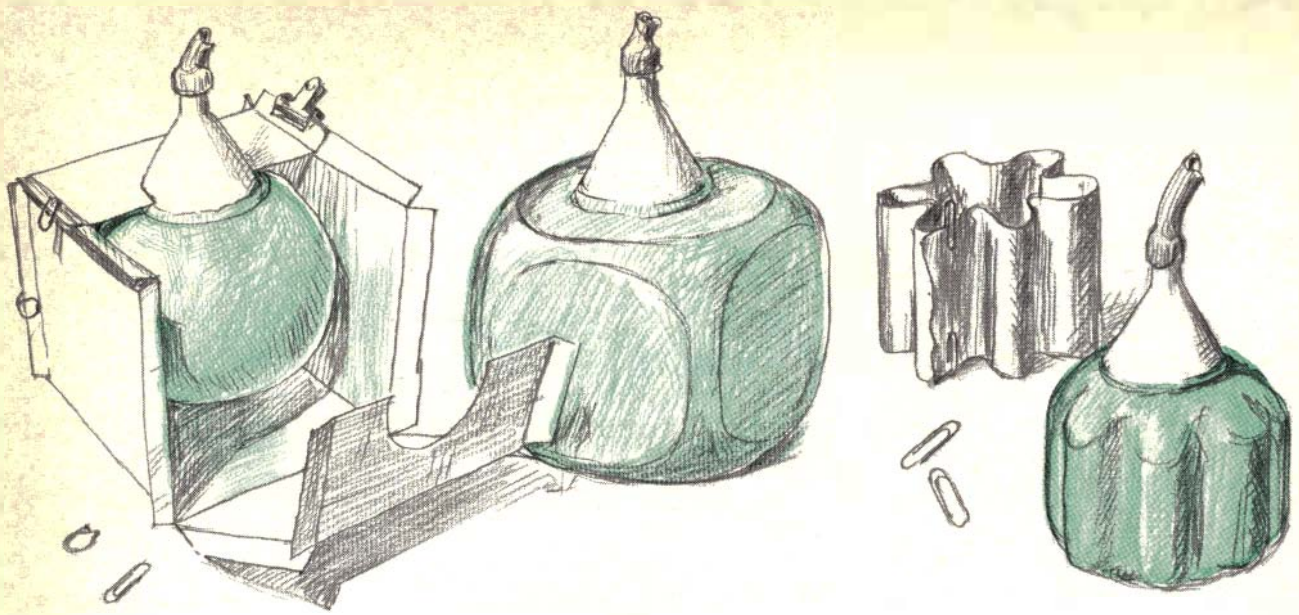
"Plastic bubbles are blown by much the same technique that I employ for blowing soap bubbles. Almost any tube can be used as a blowpipe. Of the various pipes I have tried the most convenient are the inexpensive plastic funnels that are available from dealers in scientific supplies.

"I invert the funnel and clamp the spout to a horizontal rod that is attached to the vertical rod of an apparatus stand of the kind found in chemical laboratories. A hose for blowing air through the funnel is connected to the open end of the stem. Small bubbles can be blown by mouth. A filter that contains sodium hydroxide should be inserted in the air line followed by another that contains a desiccating agent such as anhydrous calcium sulfate to absorb the exhaled moisture and carbon dioxide.

"I tend to run out of breath when I try to blow large bubbles by mouth. To inflate bubbles a foot or more in diameter I recommend working with any compressor that is capable of delivering air at a pressure of at least two ounces per square inch, which is equivalent to the pressure exerted by a column of water about four inches high. I have successfully used a bicycle pump, a vacuum cleaner and even a hair drier. An air connection was made to the vacuum cleaner and the hair drier with a perforated rubber stopper having a serrated hose connection that is inserted into the exhaust port. I now use a compressor of the rotary type made by the Gast Manufacturing Corporation (P.O. Box 97, Benton Harbor, Mich. 49022).

"The size of the funnels that serve as blowpipes varies in rough proportion to the size of the bubble to be blown. My funnels are made of either polyethylene or polypropylene. The outside diameters of the cones are 50, 78, 108, 150, 168 and 200 millimeters (two to eight inches). In addition to being inexpensive, plastic funnels make attractive blowpipes for several reasons. They are relatively light. The stems can be softened and sealed with heat. The bubble film can be peeled from the cone easily when the acetone has evaporated, after which the funnel can be reused.

"To blow a bubble put a shallow container of solution under the inverted funnel and raise the container to dip the cone into the solution [*see illustration at left*]. If the blowtube is open to the air, solution will wet both the inner and the outer surface. If suction is applied to the



Molds for shaping bubbles

blowtube at an appropriate rate, solution will be pulled into the cone and will wet only the inner surface. Conversely, solution can be excluded from the interior by applying appropriate pressure to the blowtube; solution will wet only the outer surface of the cone. In all cases a thick film of plastic clings to the wetted surfaces.

"The bubble is started by blowing into the funnel as the container of solution is lowered. A closed film of solution forms between the lower edge of the funnel and the surface of the solution. The volume enclosed by the film increases as the experimenter continues to blow and to lower the solution.

"The shape the film assumes depends on a number of variables, including the relative distribution of plastic on the inner and outer surfaces of the funnel, the velocity at which the solution is lowered, the rate at which the film is blown and the rate at which solvent evaporates from the film. Unlike soap films, which are influenced by surface tension, bubbles of plastic do not necessarily expand uniformly. The shape is determined by the distribution of plastic that adheres to the blowpipe and its subsequent manipulation by the experimenter.

"The initial film that is formed by blowing as the container is lowered must be detached from the solution, after which the film is expanded into the final bubble by additional blowing. The shape of the initial film strongly influences the shape of the final bubble. The final shapes of bubbles can range from perfect spheres to prolate and oblate spheroids. Indeed, with a bit of practice the experimenter can blow ellipsoids, hyperboloids and catenoids. A perfect cylinder can be formed by inflating a long hyperboloid.

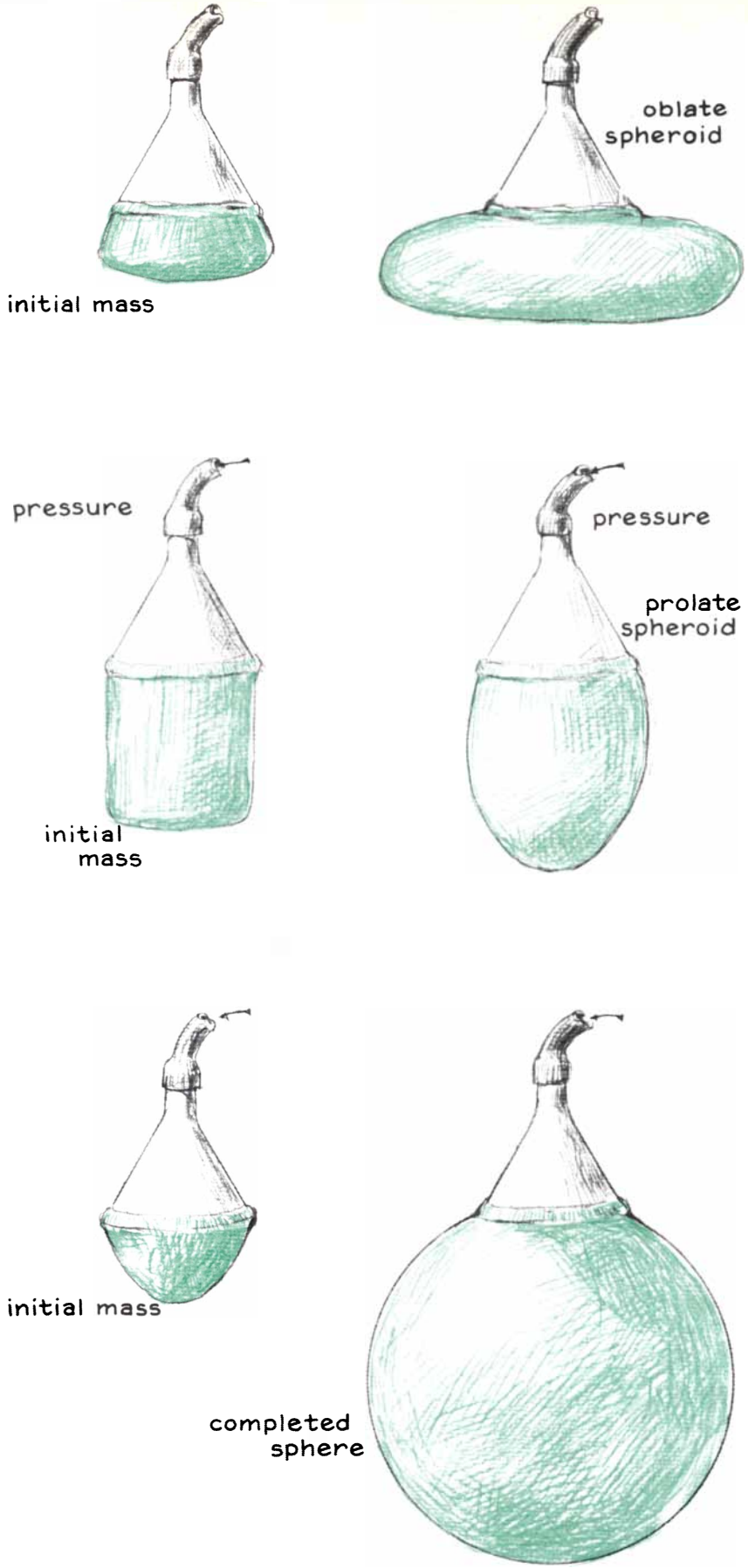
"I have blown cylinders 30 inches long by using a funnel only 15 centimeters in diameter. Polygonal bubbles—cubes, hexagons, fluted shapes and so on—can be blown with molds that are easily improvised from heavy aluminum foil [*see illustration above*]. Edges of the foil can be fastened together with tape or with a stapler. I line the molds with wax paraffin paper to prevent freshly blown plastic from sticking to the metal.

"The solution is compounded to be stiff when the solvent evaporates by adding a minimum amount of plasticizer, a material that will be discussed below. The evaporation of the solvent can be accelerated by circulating fresh air through the interior of the plastic film while it is in the mold. I do so by blowing fresh air into the bubble through one hole of a two-hole rubber stopper that fits into the stem of the funnel. Acetone vapor that evaporates from the film is swept out of the bubble through the second hole.

"The shape of the initial film that is to be expanded into the desired bubble is controlled in part by the thickness of the plastic that adheres to the inner and outer surfaces of the funnel, by the relative distance between the rim of the funnel and the surface of the solution at which the film is separated and by the shape of the cut. I make most separations by cutting the film with a pair of scissors. Freshly made solutions tend to be quite sticky, and the film may adhere to the scissors. To separate films that are blown with freshly made solutions I place strips of tissue paper below the proposed cut on both sides of the film and press the strips together with my fingers. The paper prevents the tacky material from adhering to my fingertips. The film is allowed to dry for about 30 minutes. A straight or circular cut can then be made with a pair of sharp scissors.

"Less tacky material can be separated by pressing the film together without the use of paper strips and cutting it immediately with scissors that have been lightly coated with paraffin oil. Films that are cut close to the funnel tend to expand into oblate spheroids. Those that taper toward the solution and are cut close to it expand into prolate spheroids. Cylinders expand into spheres [*see illustration on next page*].

"The film can also be detached from the solution without cutting. By adjusting the rate at which the container is lowered and the rate at which one blows it is possible to pull a conical film with the apex at the solution. The film can be twisted apart at the narrowest region. It is also possible to gather a film by dipping the blowpipe into a relatively thin solution and withdrawing it sideways, as is done with a soap solution. Such a film contains a limited quantity of plastic, however, and cannot be blown into bubbles of the largest sizes. Surplus film in the form of a thread or a tab usually remains at the place where the material



Evolution of bubble shapes

has parted. Fold the surplus against the adjacent film. It will adhere and stretch as the bubble is blown. With experience the experimenter can make seals that all but vanish.

"Beginners may have some difficulty blowing bubbles of uniform shape. Bulges tend to develop in areas where the film is thinnest. Simply stop blowing into the bubble and instead blow air across the thin area to accelerate the evaporation of solvent. The viscosity of solution in the thin area increases quickly because of the lowered temperature and the increase in the density of the plastic. Films of plastic must be inflated more slowly than those of soap solution. I often blow plastic bubbles in a series of partial expansions, allowing the plastic to flow and thus to adjust itself to local stresses between successive inflations.

"The thickness of films that have been dyed is much easier to judge by eye than the thickness of clear films. Bubbles of uniform shape can therefore be blown most easily from dyed solution. Various acetone-soluble dyes are made by Du Pont, Tennessee Eastman and many other companies. I use Du Pont's oil-blue-A and oil orange. The dye solution is made by dissolving one part dye in 99 parts acetone, by weight. I then mix 20 milliliters or more of this solution in 500 grams of clear plastic solution, depending on the intensity of the color that is wanted.

"In general bubbles can be blown to diameters about 10 times larger than the diameter of the funnel. For example, a 60-inch bubble can be blown with a 15-centimeter funnel if the experimenter blows slowly, takes care to gather an initial film of uniform thickness and maintains uniformity by accelerating local evaporation. The speed with which bubbles can be blown depends in part on the relative distribution of solution on the inner and outer surfaces of the funnel. Solvent evaporates readily from solution that adheres to the outer surface of the cone. The film stiffens quickly, which minimizes the required blowing time. Conversely, minimum evaporation occurs inside the funnel, where the atmosphere is saturated with acetone. The solution flows readily and so can 'feed' the expanding bubble for a considerable time.

"Internal coating is preferred for flowing extraordinarily thin films. Films that vary in thickness from 200 microns down to one micron can be blown with an eight-millimeter funnel that carries an internal coating of freshly made solution. (One micron is equal to about 40 mil-

lionths of an inch.) Thin bubbles must be blown quickly, as in the case of soap solution. The thinnest bubbles show rainbow colors. Hence, as Isaac Newton pointed out, the thickness of the films must approach the wavelength of light.

"The funnel can also be detached from the bubble, although frequently I just uncouple the blow hose and plug the stem with a stopper. The stopper can be removed to reinflate the bubble from time to time. To detach the funnel I make a tool by cutting a hole slightly larger than the diameter of the funnel in a sheet of either polyethylene or Teflon that is mounted on a flat backing of stiff cardboard or plywood. The diameter of the sheet should be several times the diameter of the hole. The plastic surface surrounding the hole is lubricated with a film of paraffin oil.

"The stoppered funnel is dropped through the hole so that the fully inflated bubble rests on and is supported by the upper surface of the lubricated plastic. I then exert a light, downward pull on the stem of the funnel and simultaneously rotate the stem two or three revolutions. This maneuver twists the film near the funnel into a short cord. I cut the cord with scissors close to the bubble and brush the tip lightly with acetone to ensure an airtight seal.

"Bubbles that have been sealed completely can be reinflated by thrusting a large hypodermic needle through a thick area of the film. The resulting hole can be sealed with a dab of viscous plastic solution. After drying for a few weeks bubbles tend to retain their initial shape even if they are not sealed.

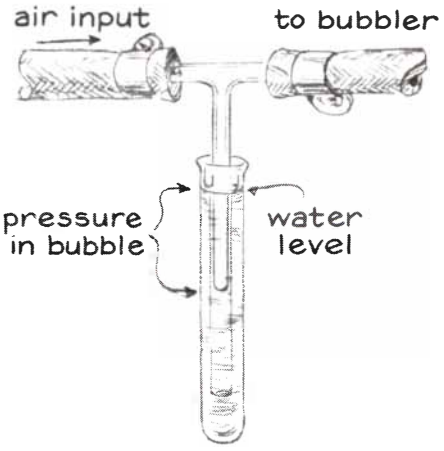
"Bubbles blown from solutions that consist solely of polyvinyl acetate dissolved in acetone tend to become rigid and brittle when dry. This may be a desirable property if the bubble is not subjected to stress. The plastic can be made as flexible and tough as one wishes, however, by adding to the solution one or more plasticizers. A plasticizer is essentially a liquid that boils at about the same temperature as a heavy oil and is a good solvent for the material it is to plasticize.

"The substances known as phthalates (particularly dibutyl, diethyl, butyl benzyl and butyl octyl) have been found to be quite effective plasticizers for the vinyl plastics. After some experimentation I settled on dibutyl phthalate (which is also known as *n*-butyl phthalate) as a good plasticizer for bubble solution. I dissolve the oily liquid in an equal part by volume of acetone and, depending on the desired flexibility, add from 2 to 15 percent (by weight) of the mixture to the bubble solution. The optimum percentage of plasticizer must be determined by experiment. In general, however, a 2 percent solution results in bubbles that approach brittleness after they have fully dried, whereas bubbles blown with solutions that contain 15 percent plasticizer remain almost as flexible as rubber balloons.

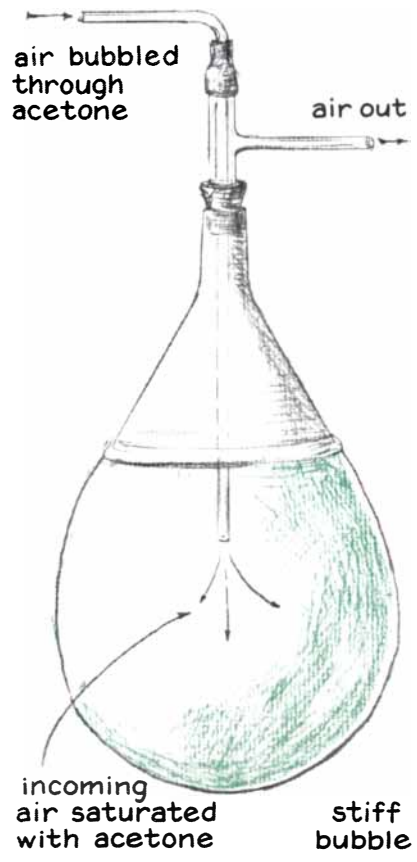
"Occasionally the experimenter may wish to soften a bubble that has become too stiff. Such softening can be accomplished easily. Just as the film of the bubble hardens as the solvent evaporates, so can it be made pliable by allowing the plastic to absorb acetone. Do not apply liquid acetone to the bubble, however. The liquid will dissolve the plastic and make a hole. To soften the film displace the air inside the bubble with an atmosphere that is saturated with acetone vapor. I achieve such an atmosphere by bubbling air from a compressor through a closed container of acetone. Vapor from the container enters the bubble through a glass tube in the stem of the funnel. The spent air is exhausted from the bubble through a *T* fitting that surrounds the inlet tube coaxially [*see bottom illustration on this page*].

"With practice and patience the experimenter can blow plastic bubbles in a remarkable range of sizes. My largest plastic bubble measured just slightly less than 11 feet in horizontal diameter. In contrast, my largest soap bubble was just over four feet. I have also had a lot of fun blowing what I call 'microbubbles,' which are tiny spheres that show rainbow colors. To blow them I file the sharp end of a large hypodermic needle to a square tip. After dipping the tip in plastic solution I blow the bubble by gently pushing the lubricated plunger of the syringe. The bubbles range from .5 millimeter to 10 millimeters in diameter. They dry much more quickly than large bubbles and last for months. Incidentally, plastic bubbles will float in water indefinitely without damage.

"Polyvinyl acetate is available from a number of U.S. companies, including the Borden Chemical Division (Borden, Inc.), the Celanese Corporation, E. I. du Pont de Nemours & Company, Goodyear Tire & Rubber Company, Polyvinyl Chemicals Inc., the Sherwin-Williams Company and the Union Carbide Corporation. The material currently costs about $1.50 a pound, but it is available from manufacturers only in a minimum quantity of 100 pounds. For a limited period experimenters can buy half-pound packages of grade AYAT polyvinyl chloride from our laboratory for a handling charge of $2.50. Make checks payable to Germantown Laboratories (4150 Henry Avenue, Philadelphia, Pa. 19144). In addition the laboratory has assembled bubble-blowing kits in three sizes. The kits contain various dyes, plasticizers and related materials. Written material describing them will be forwarded on receipt of a self-addressed, stamped envelope. Acetone, which is a highly flammable solvent, cannot be mailed. It is available in small quantities from most druggists."



Bubble-pressure gauge



Putting acetone vapor in bubble

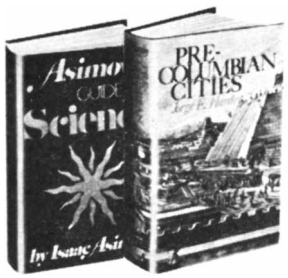
Summertime reading spree... from The Library of Science

Add any three to your science library for $3.95 (values to $55.40)

64500. A NEW PHOTOGRAPHIC ATLAS OF THE MOON. *Zdenek Kopal.* More than 200 unique photographic plates, a lunar Baedeker. *Counts as 2 of your 3 choices.* **$20.00**

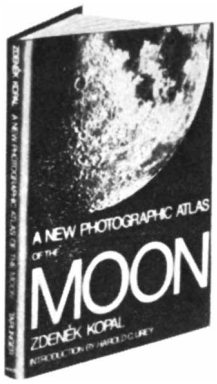
48090. THE FACE OF THE DEEP. *Bruce C. Heezen and Charles D. Hollister.* Superb, scientific account of the strange and fascinating portion of the world's surface that is deeply covered by the sea. Beautifully illustrated with more than 600 photographs. *Counts as 2 of your 3 books.* **$25.00**

34630. ASIMOV'S GUIDE TO SCIENCE. *Isaac Asimov.* The incomparable Isaac Asimov has done it again with this exciting general account of modern science—an encyclopedic work of 945 pages. **$15.00**



70240. PRE-COLUMBIAN CITIES. *Jorge E. Hardoy.* A landmark volume that vividly describes the little-known but bustling cities of pre-Columbian America—El Tijan, Machu Picchu, Chan Chan of the Chimu, and others —how they grew and decayed and how they hold a clue to the fate of our cities today. "When early American history is written as it should be, the primary reference will be this book."—*Louis A. Brennan.* Lavishly illustrated. **$19.95**

39731. COMMON INSECTS OF NORTH AMERICA. *Lester A. Swan and Charles S. Papp.* This first major one-volume survey of American insects to appear in nearly half a century will delight the naturalist, prove indispensable to the amateur collector and fascinate the interested browser. **$15.00**



(retail prices shown)

61210. MATHEMATICAL THOUGHT FROM ANCIENT TO MODERN TIMES. *Morris Kline.* An enormous (over 1,200 pages) and hugely entertaining and enlightening history, from the ancient beginnings of mathematics to concepts that acquired significance *after* 1700 (two-thirds of the book). Rich in fact and anecdote, and written with clarity and eloquence by the distinguished author of *Mathematics in Western Culture. Counts as 3 selections.* **$35.00**

36010. THE BIG, FERTILE, RUMBLING CAST-IRON, GROWING, ACHING, UNBUTTONED BELLY-BOOK. *James Trager.* Rollicking, fact-filled, often startling book about foods and fads, diet and digestion, vitamins, metabolism, nutrition and health, by the author of *The Foodbook.* Plus vital data on the factors behind heart disease, the relationship between mental illness and nutrition and much more. **$15.00**

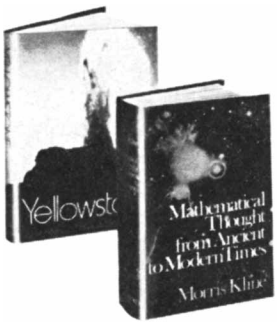
88120. YELLOWSTONE: A Century of the Wilderness Idea. *Ann and Myron Sutton.* A lavish celebration — in evocative words and breath-taking photographs (72 of them in color) — of deer, bison and bears, awesome canyons, the spectacular lakes and forests of this great century-old national park. *Counts as 2 of your 3 books.* **$25.00**

by starting your trial membership this summer in

THE LIBRARY OF SCIENCE

Take any 3 books now all for only $3.95

If you will join and simply agree to accept 3 more books (at substantial member discounts) over the next 12 months.



81800. SYMMETRY, ORBITALS, AND SPECTRA. *Milton Orchin and H. H. Jaffé.* The unity of modern chemistry stems directly from the integrating concepts of symmetry, molecular orbital theory, and absorption spectroscopy. Here, two renowned chemists blend the three into a total approach to chemical theory. **$16.50**

64910. NUMBER WORDS AND NUMBER SYMBOLS. *Karl Menninger.* A delightful cultural history of numbers: how man learned to count, how it affected his language and culture, and how his number systems changed through the ages. A big book with over 300 illustrations. **$14.95**

78190. SEXUAL SELECTION AND THE DESCENT OF MAN. *Edited by Bernard Campbell.* In one seminal volume, ten of the world's outstanding biologists, anthropologists and sociologists assess some of the problems originally posed by Darwin in 1871. Unanimously hailed as a modern masterpiece. **$14.75**

57280. KINGDOM OF THE SEASHELL. *R. Tucker Abbott.* Expert introduction to the fabulous world of marine mollusks—how they grow, reproduce and defend themselves. Magnificent illustrations, many in color. **$14.95**

45490. ENVIRONMENTAL GEOLOGY. *Peter T. Flawn.* Lucid exposition of the relationship between man and the natural earth processes. **$14.50**

36760. BORANES IN ORGANIC CHEMISTRY. *Herbert C. Brown.* One of the world's leading researchers in the important field of organic synthesis now reviews the work of a lifetime in an unusual and stimulating volume that will appeal to scientists of every stripe. *Counts as 2 of your 3 selections.* **$24.50**

56191. INTEGER PROGRAMMING. *Robert S. Garfinkel and George L. Nemhauser.* Brings together the theory, methodology and applications of one of the most exciting and rapidly developing areas of operations research. Comprehensive sets of chapter-end notes supply a complete reference guide to integer programming. **$18.95**

The Library of Science 2-58Y
Riverside, New Jersey 08075

Please accept my application for membership and send me the three volumes indicated, billing me only $3.95 for all three. I agree to purchase at least three additional Selections or Alternates during the next 12 months at special members' prices. Savings range up to 30% and occasionally even more. My membership is cancelable any time after I buy these three books. A shipping and handling charge is added to all shipments.

3 BOOKS FOR $3.95 **Indicate by number the 3 books you want.**

Some expensive books count as 2 or 3 selections.

Send no money. Members are billed when books arrive.

Name _____

Address _____

City _____ State _____ Zip _____

Book selections purchased for professional purposes may be a tax-deductible expense. (Offer good in Continental U.S. and Canada only. Prices slightly higher in Canada.)

HOW THE CLUB OPERATES

- The Library of Science **News,** describing the coming Main Selection and Alternate Selections, will be sent to you 15 times a year at three to four week intervals.
- If you wish to purchase the Main Selection, do nothing and it will be shipped to you automatically.
- If you prefer one of the Alternates, or no book at all, simply indicate your decision on the reply form always enclosed with the **News** and mail it so we receive it by the date specified.
- The **News** is mailed in time to allow you at least 10 days to decide if you want the coming Main Selection. If, because of late mail delivery of the **News,** you should ever receive a Main Selection without having had the 10-day consideration period, that Selection may be returned at Club expense.
- After completing your trial membership, you will be entitled to take advantage of our fabulous bonus plan which offers savings up to 70% and more.

BOOKS

Yukawa's reflections on life and science, Szilard's papers and a tribute to Dirac

by Philip Morrison

CREATIVITY AND INTUITION: A PHYSICIST LOOKS AT EAST AND WEST, by Hideki Yukawa. Translated by John Bester. Kodansha International Ltd. ($8.95). THE COLLECTED WORKS OF LEO SZILARD: SCIENTIFIC PAPERS, Bernard T. Feld and Gertrud Weiss Szilard, editors, with Kathleen R. Winsor. The MIT Press ($17.50). ASPECTS OF QUANTUM THEORY, edited by Abdus Salam and E. P. Wigner. Cambridge University Press ($23.50). It was given to Einstein to weave once and for all into a single fabric the strong fibers—space and time, matter and energy, weight and inertia—that had since the Renaissance come to span the physical world. The first half of this century witnessed, we can now see plainly, the composition and development of three other physical themes, even more unexpected than the rewriting of Newton's *Principia*. The deepest of these is doubtless the rise of quantum mechanics: new laws of motion, a powerful set of statutes whose writ runs further even than the Newtonian laws, into the recesses of all matter. Cause and chance are here united opposites, fused with a subtlety that exceeds that of relativity physics. The new mechanics forced a new recognition of the formal, of patterns of wholeness and of symmetry, that still rules the work of the physicists. Cause and chance struggled for being, not just around the quantum but in new readings given the century-old laws of thermodynamics. These too were put on a fresh footing before 1950, and that reexamination yielded up the idea of physical information (in turn the essential concept so brilliantly exploited as the metaphor that rules contemporary molecular biology, the domain of the informed helix). Finally, drawing on all these conceptual transformations but by comparison matter of fact, all but free of the new epistemology, nuclear physics arose out of ingenious experiment: the study of the atom's core. Its importance transcends any philosophy; by its work we know how the stars shine, how old our earth is, above all how the cities of men can come to live or to die.

The generation of physicists born in the decade or so during which the new physics itself was born—the years between the discovery of Röntgen rays and the spread of special relativity—ran through a newly opened gate. Their teachers were Einstein, Bohr, Planck, Rutherford; theirs were new paths to map. Their names too are lustrous: we may list Schrödinger, Fermi, Pauli, Heisenberg, Joliot, Bethe, Rabi....

The first of the books under review is the work of Hideki Yukawa. It was Yukawa who saw in late 1934 that under the quantum laws all the forces of nature must arise from the exchange of specific particles. And so, for example, photons mediate the electromagnetic forces. Yukawa first proposed that a then unknown "heavy photon" (we call them mesons now) of a special kind must exist to carry between the protons and the neutrons in the nucleus the momentum flow that we now recognize provides the strongest glue in all matter.

Yukawa's book is a collection of personal essays almost without technical content (there are one or two equations in 200 pages) outlining his views of the enterprise of science and more intimately the state of his own heart and mind, as recalled from his youth or as he marks the first years of reflective retirement. The son of a university professor, Yukawa remembers himself as a clumsy, bookish youth, ashamed that he did not share the common dexterity of the "average Japanese." To this he added a shut-in sense of isolation ("I'm poor at dealing with other people") and an early literary taste for the melancholy and fatalism of some Japanese novels and indeed of the Chinese philosophical masters Lao-tzu and Chuang-tzu, whose works he eagerly heard and read from the age of five. He learned German in high school and conceived the ambition to read a book in German. Then he found in a bookshop the first volume of Max Planck's introduction to theoretical physics. For him it was intelligible, fascinating, logically reasoned. His way was fixed; he was plainly a brilliant student.

After graduation from the University of Kyoto he determined "to get moving," to do research on the atomic nucleus and cosmic rays. The European flowering of quantum mechanics was fast, too fast for a distant Japanese observer far from the stimuli of Göttingen and the Copenhagen group to hope to compete directly; best to seek unopened fields where almost nothing was known. In five years, by 1934 ("perhaps I was... stupid for taking so long to produce results"), he saw the necessity of the meson, estimated its mass, connected it with the theory of the beta decay of radioactive nuclei and became the leader of an isolated but remarkably original and devoted group of young Japanese theorists who worked even in wartime and whose students now form a large and powerful school within the world community of physics.

The most striking aspect of these moving brief personal essays and interviews is their general independence from the familiar categories and parables of a Western education. The natural world is modular, we say; it consists of "atoms and the void." For Yukawa the idea is the same but the figures differ. The laws of physics are an "invisible mold" striking out innumerable copies of electrons, protons, mesons, all the same. Did not Chuang-tzu speak of "heaven and earth as a great crucible" molding human beings to be refashioned when the time came into something else? The Greek way included a powerful capacity for abstraction, for seeing the way of numbers, of atoms and the void. Yukawa's classics largely lacked that power; theirs was the unifying strength of intuition, of the sense of analogy. ("The Way is ... everywhere ... in insects, plants, tiles and walls.") It was the harmony between the abstraction of the atomists and the intuitional powers of philosophers whose thought could join science with literature and the arts that was the glory of Greece. It is that harmony we have lost today; it

is a just balance between the novel patterns we can grasp through the creative aspect of our minds and the logical, even inductive, arguments of the scientist that now eludes physicists who confront, too abstractly, the world of particles today. This juxtaposition sets the tone of the entire book, as it supplies the title.

It was in the mind of Professor Yukawa that modern science first conspicuously expressed itself no longer as a Western development but as a universal product of human thought. Mesons owe more to the *Way* than to the *Metaphysica.* His book, agreeable to read if less than simple to grasp, will place the thoughtful in touch with past and future; with a gentle, generous and humane mind; with a philosophy on the one hand resigned and on the other hand hopeful, a mixture of the elegant and the energetic not to be found where they write with only a few dozen letters. "Resignation may be important for human beings; the time may come when I have to give up. But I have not given up yet." And "I do hope that some really bold new ideas will come out sooner or later from among young physicists."

The Collected Works of Leo Szilard, very different in form and style, displays "bold new ideas" in plenty. They are the ideas—in 30-odd scientific papers, a number of memorandums and once secret reports, a spate of patents, applications and proposals—of one of the most prescient and original of all the minds that built today's physics. This volume, thick and rather stringently technical, holds Szilard's scientific legacy. (A volume yet to come will bring together his diverse writings on society, particularly on war and peace.)

Szilard added nothing to quantum mechanics. Indeed, although the physics of slow neutrons depends utterly on the wave nature of matter, most of the physics of nuclear chain reactions appears to be understandable in classical and statistical terms. Szilard wrote his conceptually most striking paper in 1922, and it is presented here both in a facsimile of the published German text and in a clear English translation. It is nothing less than the first working out of the thermodynamics of a Maxwell demon: an analysis of how, even given the "intervention by an intelligent being" who can choose among moving molecules to suit himself, there is still no way to make a perpetual motion machine of the second kind, making heat flow uphill, lowering the entropy at no cost in energy. "A simple inanimate device can achieve the same essential" choice, but in so doing it generates just that entropy required by thermodynamics. Choice itself here becomes a thermal quantity, the first step toward the modern theory of information. Young Szilard, a generation before Claude Shannon, looked down that extraordinary road.

Alone, such a result would signal Szilard's originality. But here is a design for cyclotrons before Ernest Lawrence, one for linear accelerators before Rolf Wideröe, for microfilm in 1934, for electromagnetic pumps (with Einstein) in 1927. (The pump was meant to be quiet, since it had no moving parts, but it "howled like a banshee" from cavitation, one eyewitness reports.)

The kernel of the book is that famous insight ("As the light changed to green and I crossed the street"): an element that, on being split by neutrons, released two neutrons whenever it absorbed one could sustain a nuclear chain reaction, confound Lord Rutherford by liberating nuclear energy on an industrial scale—and yield atomic bombs. This thought occurred in the fall of 1933 to a man prudently come to London and generously busy trying to find posts for his less-prepared colleagues who had lost job and home in the Nazi takeover. ("The thought became an obsession with me.") In the spring of 1934 he applied for a patent on the neutron chain reaction, complete with the idea of a critical mass, years before any such reaction as fission had been discovered. The only way to keep the patent from public view was to make it secret, and so he assigned it to the British Admiralty.

Szilard turned his obsession into the obsession of all of us after the discovery of fission. It was he who prepared the draft for Einstein's famous letter to F. D. R., and a more detailed letter over Szilard's own name went with it. It was he who applied for a patent on the uranium chain in early 1939; it was he who inspired and joined Enrico Fermi in developing the idea of a graphite pile with a lattice of uranium spheres in 1940. By 1944 Fermi and Szilard together applied in secret for a patent on "neutronic reactors" in modern form. That patent, held by "the United States of America as represented by the United States Atomic Energy Commission," became public only in 1955. Szilard said in 1946: "While the first successful alchemist was undoubtedly God, I sometimes wonder whether the second successful alchemist may not have been the devil himself."

The story of how Szilard clear-sightedly opposed the diabolical force of atomic warfare from early 1945 to the end of his life will be more explicit in the second volume of his works. This volume displays rather his postwar entry into biology by way of microbial metabolic control and into essays toward a theory of aging and a theory of memory; these remain anticipations of the greatest tasks of modern molecular biology, still incomplete.

Let the Paris microbiologist Jacques Monod speak of his late friend; he warms an acute appraisal with an intimacy of spirit. "Szilard . . . was as generous with his ideas as a Maori chief with his wives. Indeed, he loved ideas, especially his own. . . . He had close friends, mostly among younger men, whom he would visit, or summon by phone, at frequent erratic intervals, any time he had an idea to communicate, an experiment to suggest, or . . . felt he needed . . . more facts to abate temporarily his huge intellectual appetite. . . . Leo Szilard was too rich in ideas . . . too joyfully familiar with all of them . . . ever uniquely to pursue only one of them, aggressively reiterating, illustrating, and defending it, as most of us do. . . . Szilard was different . . . his memory will remain as a unique image of a man to whom science was more than a profession, or even an avocation: a mode of being."

Aspects of Quantum Theory is the most conventional of these evocative books. No personal expression by the man it honors at 70—Paul Adrien Maurice Dirac—appears. It holds rather the tribute of his scientific friends and colleagues. Perhaps a fifth of the text is a fascinating set of reminiscences of the "golden age of theoretical physics," as Jagdish Mehra calls it, and the rest is a set of reflective discussions, all but one or two at the level of the professional quantum-field theorist, seeking to appraise Dirac's work and its outcome. It was P. A. M. Dirac who first of all perceived the antiworld, that conjugate universe where all the particles have their curious reflection. It was implied by the joint acceptance of quantum mechanics and relativity. He was a gifted student who trained as an electrical engineer but plainly loved mathematics. He found his way to Cambridge and to science on a grant in 1923, when he was 21. He never looked back.

He was a quiet, somewhat withdrawn (although never solitary) man, remarkably guileless, direct and free of pretense. In 1968 he recalled, "I have the best of reasons for admiring Heisenberg. He and I were young research students at the same time, about the same age, working on the same problem. Heisenberg succeeded where I failed." The problem Heisenberg succeeded at, of course, was inventing the quantum me-

chanics! Dirac did not fail again. His remarkable insight and analytic originality led to the most persuasive and complete formulation of the new quantum mechanics and its extension into relativity and to the positron, first harbinger of the antiworld. By 1933 he "had himself largely created the language of theoretical physics and become the personification" of that golden age. Anyone with some appreciation of the quantum theory can read with pleasure the historical chapters of this book. The more technical ones are for the cognoscenti, with the exception of two. One of these (by Edoardo Amaldi and Nicola Cabibbo) describes the long search for the fundamental free magnetic poles Dirac conjectured long ago, unknown particles whose absence mars the symmetry of the antiworld. Perhaps they are coyly revealing themselves in the mysterious bursts of pure gamma rays we see in cosmic rays from time to time. A newly made pair of monopoles might separate for a brief while but they interact so avidly that in their motion they emit many photons, thus losing all relative energy, so that they can recombine and mutually annihilate again—all within some 10^{-21} second! The other piece, by Freeman Dyson, reviews the status of another idea of Dirac's, that the constants of nature may change in cosmic time. Only so, Dirac felt, can we hope to explain the large pure numbers of cosmology. They started as numbers near unity, he proposed, in principle calculable, and only became huge—10^{40}—as time passed. The case for this looks poor at the moment; if all five forms of the hypothesis fail, writes Dyson, "it will be up to the speculative cosmologists, and up to Dirac in particular, to think of something new."

GEOGRAPHICAL ECOLOGY: PATTERNS IN THE DISTRIBUTION OF SPECIES, by Robert H. MacArthur. Harper & Row, Publishers ($13.95). The species is the atom of this biological analysis, and the statistical mechanics of complex and diverse natural samples of such "atoms," weaving together the role of chance with that of deterministic laws, is its subject matter. Experiment joins observation. Ingenious and satisfying argument, based on a far-flung net of observations in the manner of Darwin, is bolstered by mathematical models as Darwin never could have done. The mathematics is made graphical and general as often as it remains in the specific language of coupled equations. The book is highly original, "a nontraditional outline" by a master in the field, the equations caged in appendixes, so that the reader unarmed with calculus or probability theory can still browse safely on text and graphs.

The experimenters cited can be both elegant and bold. A few years ago one team chose four "very small islands of red mangrove" in the Florida Keys, recorded carefully all the species—insects and spiders—and hired a professional exterminator who enclosed each island in a polyethylene sheet and fumigated the entire fauna. Then, for up to two years, they recorded step by step the colonization of the room-size islands by new species. The experimenters confirmed the facts about island fauna already seen off southern California, New Guinea and Panama: islands hold an equilibrium number of species, fewer than the mainland nearby. New species enter and established ones become extinct all the time, but the overall number remains about the same, generally surprisingly well predicted by area alone. (A straight line connects the number--on a log-log plot—of West Indian reptilelike species with island area, over a range from Redonda, at one square mile, up to Cuba!)

The theory is compelling; it fits all the facts. New species enter randomly, although just which they are depends on the nature of organism and habitat. They persist until extinction; the more intense the competition, the smaller the population of each species. Immigration and extinction balance at a stable equilibrium. Small numbers guarantee eventual extinction by random causes; the descendants of the founders must become numerous if the line is to be safe. Competition can inhibit every species, of course: simple plausible models show that reducing the maximum size of the population from 20 individuals to 15, say, by taking into account one reasonable competitor for *Lebensraum,* shifts the expected species survival time from 25,000 years to 700 years!

Competition is no simple relation: it includes predation, parasitism, disease, how "*A* can reduce *B*'s food supply" and much more. Condors do not compete with codfish; there must be some common or at least sequential use of space. The allowed principle that two species cannot coexist unless they do different things is of little use until one determines how different "different" must be. Flour beetles of two species cannot coexist in laboratory bottles of pure flour. Add some broken wheat kernels or pieces of broken glass tubing, however, and the two happily find distinct "niches" in one bottle. In the diversity of natural environments the species blend is rich. Competition for resource use must go before aggressive competition; surely there is no point to fighting unless some commonly useful resource is to be gained. For this main reason aggressive behavior —*pace* Robert Ardrey and other hard-shell romantics—is "far from universal in the animal kingdom."

Professor MacArthur has made a book that uniquely bonds strict calculation with the naturalist's eye and links the personal with the abstract. It is at once demanding and rewarding. His untimely death robs us of much, but we can celebrate what he gave us in some 40 years of life. Let him speak with the feeling of the first sentences of this book:

"To do science is to search for repeated patterns, not simply to accumulate facts, and to do the science of geographical ecology is to search for patterns of plant and animal life that can be put on a map. The person best equipped to do this is the naturalist who loves to note changes in bird life up a mountainside, or changes in plant life from mainland to island.... But not all naturalists want to do science; many take refuge in nature's complexity as a justification to oppose any search for patterns. This book is addressed to those who do wish to do science. Doing science is not such a barrier to feeling or such a dehumanizing influence as is often made out. It does not take the beauty from nature.... No one should feel that honesty and accuracy guided by imagination have any power to take away nature's beauty."

FLINT: ITS ORIGIN, PROPERTIES AND USES, by Walter Shepherd. Faber and Faber. Distributed in the U.S. by Transatlantic Arts, Inc. ($18.50). The great seas of the Cretaceous period, a round 100 million years ago, left less dry land than at any other known period in the history of the world; they smothered the continents with an algal ooze that hardened into the soft white limestone called the Chalk. It is widespread: around London and Paris, in Kansas and Kenya, there is plenty of chalk.

Where there is chalk there is flint. The more than glass-hard black nodules of this mineral, a nearly pure silica, are embedded in the chalk or freed by erosion to lie like pebbles along the stream beds or on the banks and beaches of later times.

Flint is the best-known material used by Stone Age man. It remains a valuable economic mineral, particularly in England; there it supplies silica today for pottery makers. (Flint glass, although once made from crushed flint, now de-

rives from silica sand.) The oldest mines we know are flint mines, which Neolithic miners worked for "first-quality tabular flint" down to 40 feet below the surface. From about 2200 B.C. to 1900 B.C. the big workings at Grime's Graves in Suffolk were active, meeting the flint-ax demand. Polished igneous rock and the newfangled bronze imports reduced the trade greatly, but it did not come to an abrupt end. Men still recall the last flint miner at Brandon, "Pony" Ashley, who worked by candlelight "using a single-tined pickaxe roughly resembling a Neolithic antler-pick." They still work flint there, just as the old people did, and a fortunate visitor might well strike up conversation with one of the Edwards family, craftsmen-knappers for two or three centuries at least, in a pub such as the Knapper's Arms. Gun flints still went out in 1950 in numbers to American collectors, and to Thailand and West Africa for the locally made flintlocks. The Turkish flint-knapping industry was described in these pages in 1969; there the knappers use some 500 tons of flint a year for blades set into heavy planks to make threshing sledges. Drawn over the grain on the threshing floor by teams of horses or oxen, the flint edges chop up the stalks and release the wheat.

Flint has been the firestone since the Iron Age. Until 1970 the paschal candle had to be lighted for a Catholic Easter Vigil with "new fire" struck from flint. Exegesis had allowed most such fires to be lighted with the cerium alloy of the cigarette lighter, vulgarly called flint, for the past couple of generations, but until then true flint and steel had been used throughout Christendom. Even in Virgil's Rome the slaves made fire daily from the white-hot particles shocked by the hard flint from iron or steel. It is in the language firmly; even if one has never seen flint, "none needs to be told what is meant by a 'skinflint,' a 'flinty eye' " or requires an explanation of Captain Flint's character.

The book is replete with the lore of flint and its users, but it is by no means a volume of curiosities. Its center of gravity is dual: how flint was formed, and how it was worked as a tool for most of the history of humankind.

Flint nodules are 98 percent silica, scattered through a great thickness of chalk, itself 98 percent calcium carbonate. The black flint is embedded in the white chalk at a good flint site at the rate of about a ton of flint to every 130 tons of chalk. That silica was once dissolved in the waters of the Chalk Sea, a part or two per million. How the flint was segregated is a battlefield of theories and inferences still unsettled after a century of work. Did the silica deposit out in a geological instant as lumps on the sea floor? It has happened, but it cannot be common. Perhaps the flint gathered before a million years went by in layers under the sea floor, where the ooze was becoming chalk. Or perhaps, tens of millions of years after the sea had receded, the coarse reeds and grasses that grew on the overlying soil extracted the silica, to deposit it as tiny opaline globules in their tissues. Water leaches the soluble material out of the plant debris down to the rocks below, where various chemical and physical processes over millions of years form the nodules, nucleating on an occasional sponge spicule or the like in the rock layers. All these schemes have some success at explaining the flint; all encounter difficulties with the structure of particular nodules or with features of the banded distribution of the flint in depth. Perhaps statistical studies will one day apportion the theories, each to its moiety.

"Considered together, flint and chert simply represent the natural end-state of all random drifts of soluble silica in the earth's crust. There is nothing 'special' about them. . . . It is the other forms of silica—rock crystal, agate, opal—which are extraordinary. But if flint has this undistinguished status in the mineral kingdom, its importance to the animal kingdom can hardly be exaggerated." Shepherd, experimental mineralogist, collector, scholar, folklorist and a man who has his way with words, has done his share to celebrate flint for us primates in this unusual, engagingly broad study.

ATLAS OF CULTURAL FEATURES, by Benjamin F. Richason. Hubbard Press ($6.95). WORLD FACTS AND TRENDS, by John McHale. Collier Books ($2.95). Useful and unusual references, well worth the time of any interested reader, these are two letterhead-size thin volumes. The *Atlas* presents 36 aerial photographs, each paired with a detailed contoured photographic map of the region shown, usually a few square miles and all but two within the 50 states. With each pair there are a crisp page of explanatory text and a few questions, answerable after more than superficial study of the two images. Professor Richason, a geographer, seeks to display "impressions of man's activities" on the land. We see the mark of the surveyors themselves on field and road; we see the village and the farm, mines, mills, railroads, canals and the patterned cities. The work is, of course, an open door to

the knowing use of maps and aerial photographs, but it is more than that; the places we see so clearly are unusual, and some are unforgettable.

There is the Point in Pittsburgh, a city nucleated around the sharp confluence of rivers that forms the Ohio. Multiple rail lines and wide highways thread the very site of Fort Pitt, strategic during the French and Indian War. Or look at Abiquiu on the Rio Chama, where the old Spanish village still lies above the ditch; the rounded bones of those valley walls are known to us in the paintings of Georgia O'Keeffe. Or stare down into Bingham Canyon, Utah, where the copper men have carried away a porphyry mountain to leave a stepped bowl 1,500 feet deep—two billion tons of ore gone. Or view Scottsbluff, where the Oregon Trail, the Pony Express and the Overland Trail all together leave map and photograph by one pass, past high Eagle Rock out of the low Platte valley.

In *World Facts and Trends* McHale, once an associate of R. Buckminster Fuller, has compiled an interesting, meticulously documented set of well over 100 graphs and flow charts (visually mediocre) set off by a clear text. There is a cogency to the choice that distinguishes the book. The aim is not only to make ready to hand but also to insert into the mind the facts that gauge the relations between human beings and the planet they dwell on: population and its changes, the natural and the man-maintained flows of material and energy, food, industry and its tools, communications, minerals, all on a world basis and many with national and regional breakdowns. The sources are usually recent monographic studies, not mere statistical tables. Most of the data are the quantitative stipulation of important commonplaces such as per capita use of fuels, copper and so on. Some give a view into the unexpected, like a flashlight beam into the dark: In 1850 as much energy came from burning dung as from muscle power, although coal gave rather more. (Wood was then the king of fuels.) The leaders—the U.S., the U.S.S.R. and Japan—now consume about 40 times as much steel per capita per year as India does. We live in an age of metals. Total material product per person (excluding fuels, food and water) is more than 350 pounds of metals per year; synthetics and natural products add up to only about 50 pounds. Refined metals now cycle, taking a grand weighted average, about once every 22 years. Close all the loops! Only a few errors mar the work; beware of a mislabeled graph on marine energy sources.

INDEX OF ADVERTISERS

JULY 1973

BRITISH LEYLAND MOTORS INC. 46
Agency: PKL Advertising, Inc.

BULLY HILL VINEYARDS AND WINE COMPANY 120
Agency: Wine Country Advertising Agency, Inc.

CHRYSLER CORPORATION, DODGE DIVISION 2
Agency: Batten, Barton, Durstine & Osborn, Inc.

EASTMAN KODAK COMPANY 45
Agency: Rumrill-Hoyt, Inc.

FORBES SANGRE DE CRISTO RANCHES INC. 50
Agency: Compton Direct Marketing

GENERAL MOTORS CORPORATION, BUICK MOTOR DIVISION/OPEL 61
Agency: McCann-Erickson, Inc.

GENERAL MOTORS CORPORATION, CHEVROLET MOTOR DIVISION ..... Back Cover
Agency: Campbell-Ewald Company

GENERAL MOTORS CORPORATION, GMC TRUCK & COACH DIVISION ..... 124, Inside Back Cover
Agency: McCann-Erickson, Inc.

GLASS CONTAINER MANUFACTURERS INSTITUTE ..... Inside Front Cover, 1
Agency: Marsteller Inc.

HEWLETT-PACKARD 6, 7
Agency: Richardson Seigle Rolfs & McCoy, Inc.

HONEYWELL INC. 123
Agency: Batten, Barton, Durstine & Osborn, Inc.

INTERNATIONAL TELEPHONE AND TELEGRAPH CORPORATION 9
Agency: Needham, Harper & Steers, Inc.

LIBRARY OF SCIENCE, THE 116
Agency: Henderson & Roll, Inc.

McARTHUR'S SMOKEHOUSE 47
Agency: Clifford A. Rathkopf

NATIONAL CAMERA, INC. 47
Agency: Langley Advertising Agency

NATIONAL CAR RENTAL SYSTEM, INC. 13
Agency: Richard K. Manoff Inc.

NIKON, INC., ASSOCIATE OF EHRENREICH PHOTO OPTICAL INDUSTRIES 15
Agency: Gilbert, Felix & Sharf Inc.

NIPPON KOGAKU K. K. 15
Agency: K & L Advertising Agency

OLYMPUS OPTICAL CO., LTD. 49
Agency: Fuji Agency, Ltd.

POLAROID CORPORATION 62, 63
Agency: Doyle Dane Bernbach Inc.

QUESTAR CORPORATION 109

SCANDINAVIAN AIRLINES SYSTEM 107
Agency: Sormlands Grafiska AB

TEKTRONIX 5
Agency: Young & Roehr Incorporated

TRANS WORLD SERVICES, INC. 120
Agency: Stackig, Sanderson & White, Inc.

UNITED AUDIO PRODUCTS, INC. 4
Agency: Ries Cappiello Colwell, Inc.

WESTERN ELECTRIC COMPANY 16
Agency: Foote, Cone & Belding Advertising, Inc.

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

PUBLIC POLICY ON FERTILITY CONTROL

FAMILY PLANNING, PUBLIC POLICY AND INTERVENTION STRATEGY. F. S. Jaffe in *Journal of Social Issues,* Vol. 23, No. 4, pages 145–163; July, 1967.

REPRODUCTION IN THE UNITED STATES: 1965. Norman B. Ryder and Charles F. Westoff. Princeton University Press, 1971.

REPORT OF THE SECRETARY OF HEALTH, EDUCATION, AND WELFARE SUBMITTING A FIVE-YEAR PLAN FOR FAMILY PLANNING SERVICES AND POPULATION RESEARCH PROGRAMS. Senate Committee on Labor and Public Welfare. October, 1971.

THE FUTURE OF FEDERAL SUPPORT FOR FAMILY PLANNING SERVICES AND POPULATION RESEARCH. Jeannie I. Rosoff in *Family Planning Perspectives,* Vol. 5, No. 1, pages 7–18; Winter, 1973.

THE PHYSICS OF BRASSES

HORNS, STRINGS AND HARMONY. Arthur H. Benade. Doubleday & Company, Inc., 1960.

THE TRUMPET AND TROMBONE. Philip Bate. W. W. Norton & Company, Inc., 1966.

COMPLETE SOLUTIONS OF THE "WEBSTER" HORN EQUATION. Edward Eisner in *The Journal of the Acoustical Society of America,* Vol. 41, No. 4, Part 2, pages 1126–1146; April, 1967.

ON PLANE AND SPHERICAL WAVES IN HORNS OF NON-UNIFORM FLARE. E. V. Jansson and A. H. Benade. Technical Report, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, March 15, 1973.

ADVANCED COMPOSITE MATERIALS

ADVANCED FIBROUS COMPOSITES. Proceedings of the Society for the Advancement of Material and Process Engineering. Los Angeles, 1966.

COMPOSITE MATERIALS. Edited by Leslie Holliday. American Elsevier Publishing Co., Inc., 1966.

MODERN COMPOSITE MATERIALS. Edited by Lawrence J. Broutman and Richard H. Krock. Addison-Wesley Publishing Co., Inc., 1967.

HANDBOOK OF FIBERGLASS AND ADVANCED PLASTIC COMPOSITES. Edited by George Lubin. Van Nostrand Reinhold Company, 1969.

THE IMMUNE SYSTEM

THE CLONAL SELECTION THEORY OF ACQUIRED IMMUNITY. Sir Macfarlane Burnet. Cambridge University Press, 1959.

ANTIGEN SENSITIVE CELLS: THEIR SOURCE AND DIFFERENTIATION. J. F. A. P. Miller, G. F. Mitchell, A. J. S. Davies, Henry N. Claman, Edward A. Chaperon and R. B. Taylor in *Transplantation Reviews,* Vol. 1, 1969.

INDIVIDUAL ANTIGENIC SPECIFICITY OF IMMUNOGLOBULINS. John E. Hopper and Alfred Nisonoff in *Advances in Immunology,* Vol. 13, pages 57–99; 1971.

THE TAKE-HOME LESSON—1971. Melvin Cohn in *Annals of the New York Academy of Sciences,* Vol. 190, pages 529–584; December 31, 1971.

THE REGULATORY INFLUENCE OF ACTIVATED T CELLS ON B CELL RESPONSES TO ANTIGEN. David H. Katz and Baruj Benacerraf in *Advances in Immunology,* Vol. 15, pages 1–94; 1972.

ANTIGEN DESIGN AND IMMUNE RESPONSE. Michael Sela in *The Harvey Lectures 1971–1972,* Series 67. Academic Press, 1973.

THE PROBLEM OF MOLECULAR RECOGNITION BY A SELECTIVE SYSTEM. Gerald M. Edelman in *The Problem of Reduction in Biology,* edited by F. Ayala and T. Dobzhansky. The Macmillan Company, in press.

METEORITES AND COSMIC RADIATION

METEORITES. Brian Mason. John Wiley & Sons, Inc., 1962.

ORIGIN, AGE AND COMPOSITION OF METEORITES. Edward Anders in *Space Science Review,* Vol. 3, pages 583–714; 1964.

METEORITES AND THEIR ORIGINS. G. J. H. McCall. David & Charles, 1973.

THE BRIDEPRICE OF THE SEBEI

SURVEY OF AFRICAN MARRIAGE AND FAMILY LIFE. Edited by Arthur Phillips. Published for the International African Institute by Oxford University Press, 1953.

SEBEI LAW. Walter Goldschmidt. University of California Press, 1967.

ECONOMICS IN EAST AFRICAN ABORIGINAL SOCIETIES. Harold K. Schneider in *Economic Anthropology: Readings in Theory and Analysis,* edited by Edward E. LeClair, Jr., and Harold K. Schneider. Holt, Rinehart and Winston, Inc., 1968.

PLATE TECTONICS AND MINERAL RESOURCES

EXPLORATION METHODS FOR THE CONTINENTAL SHELF: GEOLOGY, GEOPHYSICS, GEOCHEMISTRY. P. A. Rona. National Oceanic and Atmospheric Administration Technological Report ERL 238-AOML 8, U.S. Government Printing Office, 1972.

SULFIDE ORE DEPOSITS IN RELATION TO PLATE TECTONICS. F. Sawkins in *Journal of Geology,* Vol. 80, pages 377–397; 1972.

HYDROTHERMAL MANGANESE IN THE MEDIAN VALLEY OF THE MID-ATLANTIC RIDGE. Martha R. Scott, Robert B. Scott, Andrew J. Nalwalk, P. A. Rona and Louis W. Butler in *EOS: American Geophysical Union Transactions,* Vol. 54, No. 4, page 244; April, 1973.

BRAIN MECHANISMS IN MOVEMENT

MOTOR APPARATUS OF THE BABOON'S HAND. C. G. Phillips in *Proceedings of the Royal Society of London: B,* Vol. 173, No. 1031, pages 141–174; May 20, 1969.

THE BASIS OF MOTOR CONTROL. Ragnar Granit. Academic Press, 1970.

CENTRAL CONTROL OF MOVEMENT. Edited by Edward V. Evarts, Emilio Bizzi, Robert E. Burke, Mahlon DeLong and W. Thomas Thach, Jr., in *Neurosciences Research Program Bulletin,* Vol. 9, No. 1; January, 1971.

MATHEMATICAL GAMES

THE DILEMMA OF DETERMINISM. William James in *The Will to Believe: And Other Essays in Popular Philosophy,* by William James. Longmans, Green and Co., 1897.

THE LOGIC OF DECISION. Richard C. Jeffrey. McGraw-Hill Book Company, 1965.

DECISION THEORY. D. J. White. Aldine-Atherton, Inc., 1969.

THE AMATEUR SCIENTIST

TRUE PLASTIC BUBBLES AND HOW TO BLOW THEM. A. V. Grosse. Plastic Bubbles Report No. 1, GL 1970–4, Germantown Laboratories, Philadelphia, 1970.

Premise: Pollution control begins with effective sensing and measurement.

To the burning question of waste disposal, Honeywell suggests an answer.

Growing quantities of solid waste today pose growing disposal problems for most municipalities.

Incineration is a proven method of disposal. But the question is, can incinerators control one source of pollution without creating another? Without spewing noxious gases, fly ash and other particulate matter into the air?

Honeywell believes that clean, economical incineration is both possible and practical.

The answer lies in automated systems of sensing and control. Systems well within current Honeywell technology.

Although refuse varies widely in BTU values, complete combustion can be achieved consistently. Sensitive thermocouples and draft gauges can read changing furnace conditions and signal prompt compensating adjustments of air flow to maintain proper combustion and temperatures.

Similar sensors and components also control gas cooling and dust collection. Stack emissions are checked continuously by smoke density monitors and stack gas analyzers. Finally, all control components are integrated in one total control system which keeps all individual parts in balance for the cleanest possible combustion.

Could Honeywell's specialized products and systems solve a pollution problem for you? Ask us details about automated incineration control systems for solid wastes, sewage sludge, recalcining and carbon reactivation, or for other information. Write Mr. Richard Boyle, Director of Environmental Programs, Honeywell, 2701 Fourth Avenue South, Minneapolis, MN 55408.

We'd like to help.

Honeywell

Introducing the GMC MotorHome. It doesn't ride like a truck. It doesn't look like a box.

1. A NEW ERA IN MOTORHOMES IS BEGINNING.

Before we started building our MotorHome, we studied every other type of motorhome that exists.

We found that the simpler the basic construction, the fewer the problems. So we started with a strong, durable, steel perimeter frame and attached to this a cage of heavy aluminum ribs.

On top of this, we bonded both aluminum and fiberglass panels molded to a smooth finish. It's the same construction people are flying all over the world in.

Except now you'll be driving.

Then we sprayed the interior with a thick, rigid polyurethane foam for thermal insulation and noise suppression.

Next we assigned a 3-man team to install the interior. From beginning to end.

If that sounds like a return to old-fashioned pride of craftsmanship, you're absolutely right.

2. MORE POWER TO YOU.

To give you the excellent road performance, we installed a 455-cubic-inch V8 engine up front and coupled it to a 3-speed Turbo Hydra-matic transmission. We coupled that to a front wheel drive unit with a 3 to 1 ratio and put it all on top of torsion bar springs and stabilizer bar.

With our low overall body weight, it all means getting up to highway speeds quickly. Excellent traction. Excellent weight distribution.

3. A MOTORHOME IS NO FUN IF IT'S NO FUN TO MOTOR IN.

We took our basic construction and raised it only 15" from the ground. This puts the center of gravity only 37 inches above the ground. For easy handling.

See the rear wheels. We put one behind the other for four reasons: To give you a wider base. More room inside. Greater stability than you'd have with dual wheels. And so we could place a special air spring between the two wheels to pass the bumps from one to the other

instead of to you!

These air springs are the only ones of their kind on motorhomes.

To keep the weight and balance of the interior within design limits, we fed all the data into a computer. It fed back what we needed to put things where they belong.

4. ABOUT OUR SIX-WHEEL BRAKING SYSTEM.

In addition to power steering, there's a six-wheel braking system with power disc brakes up front and four large finned-drum power brakes in the rear, plus an available leveling device operated from the driver's compartment for parking on uneven ground.

Incidentally, the parking brake grabs all four rear wheels.

5. CHOOSE FROM 15 DIFFERENT FLOOR PLANS. TWO LENGTHS.

The GMC MotorHome is available in 23- and 26-foot lengths. The standard 26′ floor plan includes a dinette that converts to a double bed opposite a sofa that turns into double bunks. In the middle, a double sink, 6-cubic-foot refrigerator (it's electric so there's no pilot light that'll blow out), a range and oven with exhaust hood. There's also a bath with all the necessities plus ample cabinet space.

That's one floor plan. There are 14 more available.

6. WE INCLUDED TOP INTERIOR DESIGNERS IN OUR PLANS.

To put the finishing touches on the inside, we had *House and Garden* magazine's interior designers help us.

The driver and passenger seats are high, contoured seats with built-in arm rests. This high-level seating arrangement, combined with the big, wide-angle windshield, offers you panoramic visibility.

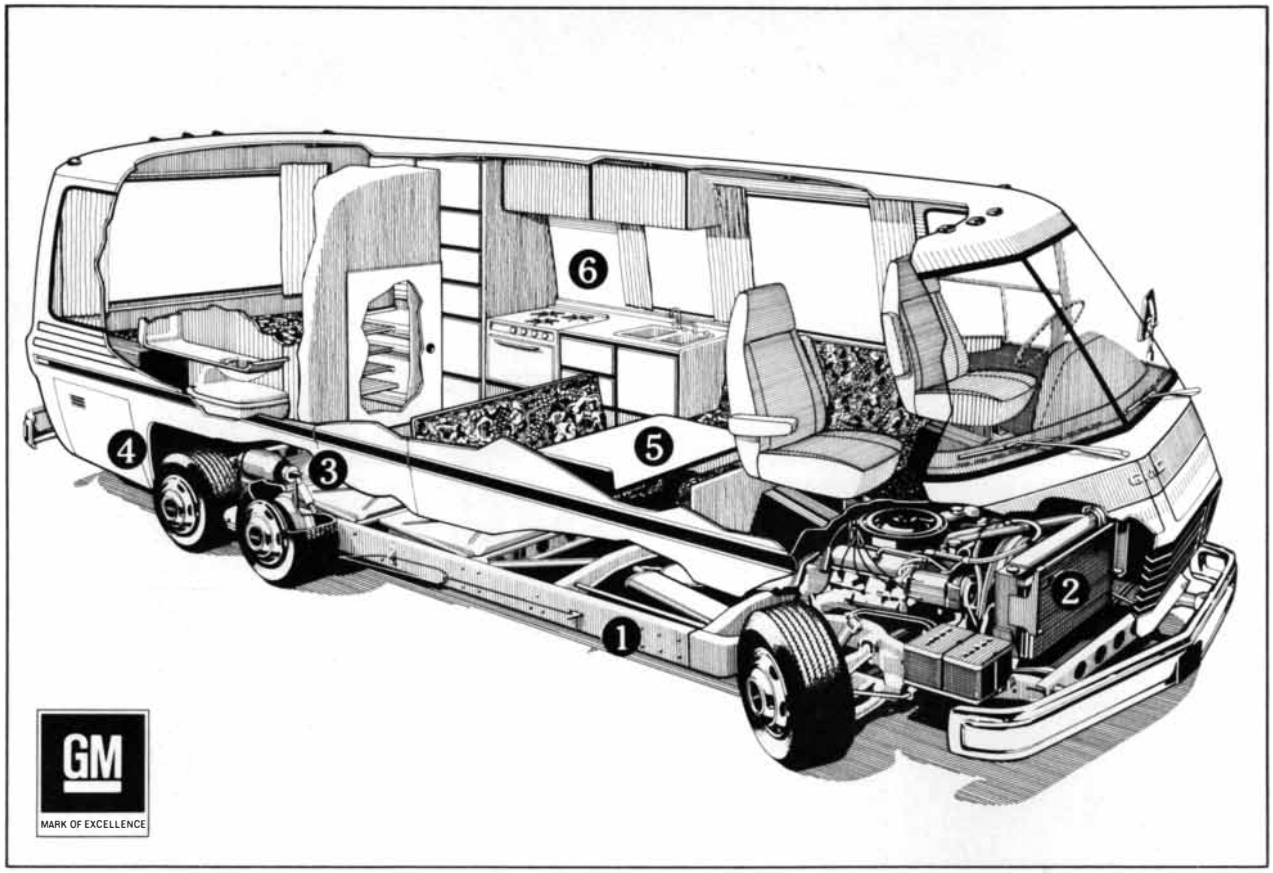
Every counter top has rounded corners. All cabinet knobs are eliminated. Every hinge is concealed.

There are thick, shag or cut pile carpets. And wood-grained vinyl on the walls and cabinets.

You also get a choice of four color-coordinated interior decors.

7. ONE-STOP SERVICE. AFTER-HOUR ASSISTANCE.

Your GMC MotorHome dealer services everything he sells. Inside and out. From the engine to the air conditioner and furnace. And there's a toll-free number you can call and immediately get the number of the nearest MotorHome dealer representative available for after-hour assistance.



For our 28-page, four-color catalog, write GMC MotorHome Headquarters, Drawer Y, Dept. 120, Lansing, Mich. 48909.

Better yet, see your GMC MotorHome dealer. He'll be glad to show you around the house. Have a good life.

GMC

The MotorHome from General Motors.

SEE WHAT IT'S LIKE TO DRIVE A WINNER.

See what it's like from the driver's seat.

You'll feel Vega's wide stance and low center of gravity ease you through turn after turn.

Available GT equipment includes special instrumentation: tach, amp and temp gauges, electric clock, and sport steering wheel.

Vega's available Custom Interior looks attractive, sits comfortable, feels extra nice.

Vega has won six major honors in less than three years. Motor Trend Car of the Year in 1971 and Economy Car of the Year in 1973. Car and Driver Readers' Poll winner in the "economy sedan" class in 1971, 1972 and 1973. Motor Service and Service Station Management survey winner of easiest to service designation in the subcompact class. Drive the most honored little car built in America.

VEGA
CHEVROLET

The award winning little car.