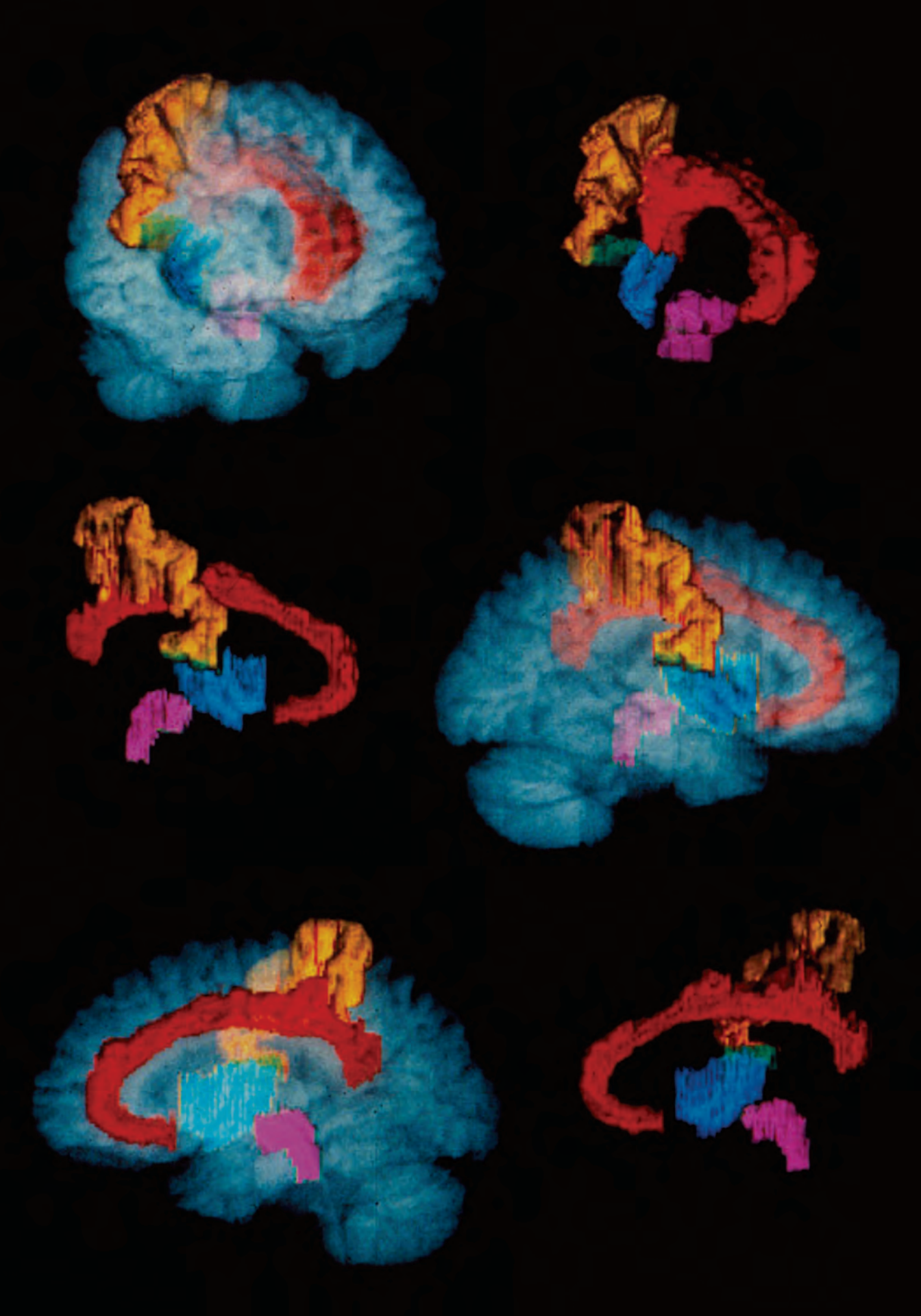
Dædalus

Journal of the American Academy of Arts & Sciences

Summer 2006

AMERICAN ACADEMY
OF ARTS & SCIENCES



Inside front cover: Some of the main human brain structures involved in the representation, simulation, and modulation signals from the body-proper. The structures are shown in color in a 3-D brain reconstruction obtained with magnetic resonance imaging (MRI) and the BRAINVOX technique. The surface of the cerebral cortex (light blue) has been made transparent to reveal hidden structures (the insular cortex, in darker blue; the cingulate gyri, in cardinal red; the hypothalamus and brainstem tegmentum, in magenta; and the somatosensory cortices, in orange yellow). See Antonio Damasio & Hanna Damasio on *Minding the body,* pages 15–22: "The body in mind helps us construct our selves and then allows us to understand others, which is nothing short of astounding." Photograph courtesy of Hanna Damasio, MD, and the Dornsife Cognitive Neuroimaging Center, University of Southern California.

James Miller, Editor of *Dædalus*

Antonio Damasio, Consulting Editor
for essays on body in mind

Phyllis S. Bendell, Managing Editor
and Director of Publications

Esther Yoo, Assistant Editor

Board of editors

Steven Marcus, Editor of the Academy

Russell Banks, Fiction Adviser

Rosanna Warren, Poetry Adviser

Joyce Appleby (U.S. history, UCLA), Stanley Hoffmann (government, Harvard), Donald Kennedy (environmental science, Stanford), Martha C. Nussbaum (law and philosophy, Chicago), Neil J. Smelser (sociology, Berkeley), Steven Weinberg (physics, University of Texas at Austin); *ex officio*: Patricia Meyer Spacks (President of the Academy), Leslie Cohen Berlowitz (Chief Executive Officer)

Editorial advisers

Daniel Bell (sociology, Harvard), Michael Boudin (law, U.S. Court of Appeals), Wendy Doniger (religion, Chicago), Howard Gardner (education, Harvard), Clifford Geertz (anthropology, Institute for Advanced Study), Carol Gluck (Asian history, Columbia), Stephen Greenblatt (English, Harvard), Thomas Laqueur (European history, Berkeley), Alan Lightman (English and physics, MIT), Steven Pinker (psychology, Harvard), Diane Ravitch (education, NYU), Amartya Sen (economics, Harvard), Richard Shweder (human development, Chicago), Frank Wilczek (physics, MIT)

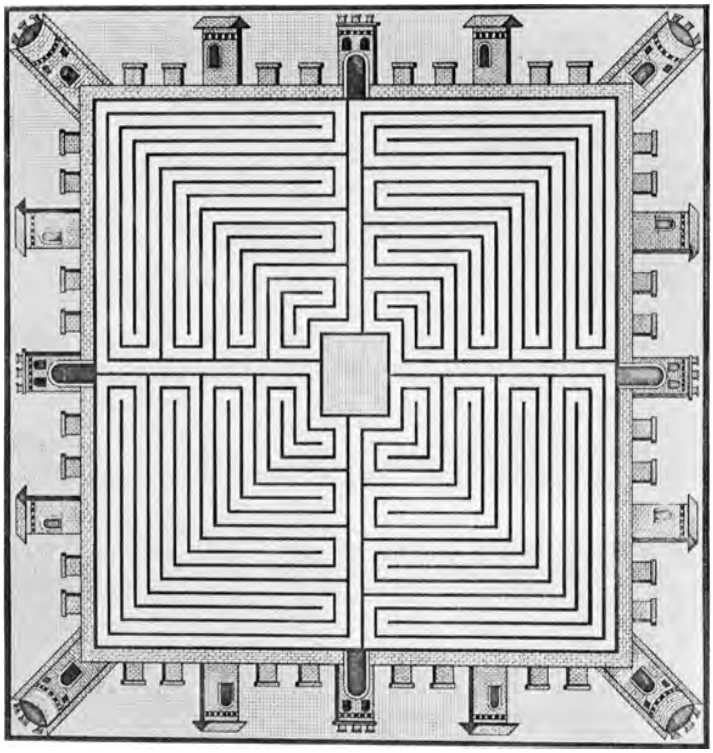
Contributing Editors

Robert S. Boynton, D. Graham Burnett, Peter Pesic, Danny Postel

Dædalus is designed by Alvin Eisenman

Dædalus

Journal of the American Academy of Arts & Sciences



Nineteenth-century depiction of a Roman mosaic labyrinth, now lost, found in Villa di Diomede, Pompeii

Dædalus was founded in 1955 and established as a quarterly in 1958. The journal's namesake was renowned in ancient Greece as an inventor, scientist, and unriddler of riddles. Its emblem, a maze seen from above, symbolizes the aspiration of its founders to "lift each of us above his cell in the labyrinth of learning in order that he may see the entire structure as if from above, where each separate part loses its comfortable separateness."

The American Academy of Arts & Sciences, like its journal, brings together distinguished individuals from every field of human endeavor. It was chartered in 1780 as a forum "to cultivate every art and science which may tend to advance the interest, honour, dignity, and happiness of a free, independent, and virtuous people." Now in its third century, the Academy, with its more than four thousand elected members, continues to provide intellectual leadership to meet the critical challenges facing our world.

Dædalus Summer 2006
Issued as Volume 135, Number 3

© 2006 by the American Academy
of Arts & Sciences
Incarnation: 9:30 am to 9:36 am
© 2006 by Jorie Graham
The Door in the Woods
© 2006 by Dorian Gossy
Leo Strauss in Chicago
© 2006 by Stanley Rosen

Editorial offices: *Dædalus*, Norton's Woods, 136 Irving Street, Cambridge MA 02138. Phone: 617 491 2600. Fax: 617 576 5088. Email: daedalus@amacad.org.

Library of Congress Catalog No. 12-30299

ISBN 0-87724-059-0

Dædalus publishes by invitation only and assumes no responsibility for unsolicited manuscripts. The views expressed are those of the author of each article, and not necessarily of the American Academy of Arts & Sciences.

Dædalus (ISSN 0011-5266; E-ISSN 1548-6192) is published quarterly (winter, spring, summer, fall) by The MIT Press, Cambridge MA 02142, for the American Academy of Arts & Sciences. An electronic full-text version of *Dædalus* is available from The MIT Press. Subscription and address changes should be addressed to MIT Press Journals, 238 Main Street, Suite 500, Cambridge MA 02142. Phone: 617 253 2889. Fax: 617 577 1545. Email: journals-orders@mit.edu.

Printed in the United States of America by Cadmus Professional Communications, Science Press Division, 300 West Chestnut Street, Ephrata PA 17522.

Postmaster: Send address changes to *Dædalus*, 238 Main Street, Suite 500, Cambridge MA 02142. Periodicals postage paid at Boston MA and at additional mailing offices.

Newsstand distribution by Ingram Periodicals Inc., 18 Ingram Blvd., La Vergne TN 37086. Phone: 800 627 6247.

Subscription rates: Electronic only for nonmember individuals – $38; institutions – $86. Canadians add 7% GST. Print and electronic for nonmember individuals – $42; institutions – $96. Canadians add 7% GST. Outside the United States and Canada add $20 for postage and handling. Prices subject to change without notice.

Institutional subscriptions are on a volume-year basis. All other subscriptions begin with the next available issue.

Single issues: current issues – $13; back issues for individuals – $13; back issues for institutions – $26. Outside the United States and Canada add $5 per issue for postage and handling. Prices subject to change without notice.

Claims for missing issues will be honored free of charge if made within three months of the publication date of the issue. Claims may be submitted to journals-claims@mit.edu. Members of the American Academy please direct all questions and claims to daedalus@amacad.org.

Advertising and mailing-list inquiries may be addressed to Marketing Department, MIT Press Journals, 238 Main Street, Suite 500, Cambridge MA 02142. Phone: 617 253 2866. Fax: 617 258 5028. Email: journals-info@mit.edu.

Permission to photocopy articles for internal or personal use is granted by the copyright owner for users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the per copy fee of $10 per article is paid directly to the CCC, 222 Rosewood Drive, Danvers MA 01923. The fee code for users of the Transactional Reporting Service is 0011-5266/06. Address all other inquiries to the Subsidiary Rights Manager, MIT Press Journals, 238 Main Street, Suite 500, Cambridge MA 02142. Phone: 617 253 2864. Fax: 617 258 5028. Email: journals-rights@mit.edu.

The typeface is Cycles, designed by Sumner Stone at the Stone Type Foundry of Guinda CA. Each size of Cycles has been separately designed in the tradition of metal types.

Comment by Marilyn A. Brown, Benjamin K. Sovacool & Richard F. Hirsh

Assessing U.S. energy policy

For decades, our political leaders have told us that we need to use energy more efficiently and derive more of it from domestic sources.[1] Since the energy crisis of 1973, U.S. presidents have declared the need to gain independence from unstable foreign energy suppliers and to do so with the same moral fortitude as if fighting a war. Some politicians have proposed massive government programs to achieve the goals of their energy policies; others have sought to unleash free-market forces that would encourage companies to develop novel sources of energy and motivate consumers to use energy more wisely.

Despite more than three decades of such efforts, the United States has not achieved the goal of energy independence. While progress in adopting more energy-efficient technologies has saved billions of dollars throughout the economy, most other indicators of energy autonomy – such as the percentage of imported fuel – demonstrate that the country has become less independent than ever. President Bush acknowledged this fact in his recent State of the Union address, telling Americans that the country has become "addicted to oil" and urging citizens to find alternative ways to satisfy their energy needs. For those with a sense of history, Bush's clarion call sounded eerily familiar.

Even though energy efficiency has taken root in some sectors of the economy,

Marilyn A. Brown is interim director of Oak Ridge National Laboratory's Engineering Science and Technology Division. An expert on the commercialization of new energy and environmental technologies and the evaluation of government programs and policies, she has authored more than 140 publications.

Benjamin K. Sovacool is a Eugene P. Wigner Postdoctoral Fellow at the Oak Ridge National Laboratory. Through the National Science Foundation's Electric Power Networks Efficiency and Security Program, he has investigated the social impediments to distributed and renewable energy systems.

Richard F. Hirsh is professor of History and Science & Technology Studies at Virginia Polytechnic Institute and State University. The author of "Technology and Transformation in the American Electric Utility Industry" (1989) and "Power Loss: The Origins of Deregulation and Restructuring in the American Electric Utility System" (1999), he currently manages Virginia Tech's Consortium on Energy Restructuring.

© 2006 by the American Academy of Arts & Sciences

1 The authors are grateful to Oak Ridge National Laboratory (operated by UT-Battelle, LLC, for the U.S. Department of Energy) and the National Science Foundation for supporting elements of the work reported here. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of Oak Ridge National Laboratory or the National Science Foundation.

it has not compensated for the growth in energy consumption that has occurred since 1973, nor will it (if current trends continue) accommodate the growth that forecasters anticipate in coming decades. Moreover, America's dependence on oil from insecure or politically unstable countries has required extensive diplomatic and military efforts that incur huge costs borne by energy users and taxpayers. Today's information economy also remains inextricably tied to reliable power and to just-in-time manufacturing and distribution processes that depend on fleets of petroleum-guzzling trucks and airplanes.

Disruptions in increasingly fragile energy systems can cause havoc to the nation's economy and to everyday life. We have already had a taste of such disruptions in the form of the California electricity crisis of 2000 to 2001, the 2003 Northeast blackout, and the fuel-supply interruptions resulting from the Gulf Coast hurricanes in 2005. These disruptions may be trivial preludes to what could be more substantial future catastrophes. Indeed, the country faces at least five immense and interconnected energy challenges due to (1) the risk of oil-supply disruptions; (2) increasing electricity usage; (3) a fragile electric-power (and overall energy) infrastructure; (4) the lack of sustained efforts to push energy-efficiency practices; and (5) the growing environmental impacts of increasing energy consumption.

First, the United States remains vulnerable to the risk of oil-supply disruptions, despite plenty of warnings over the past three decades. In 1973 the Arab members of the Organization of Petroleum Exporting Countries (OPEC) orchestrated an oil embargo, the first supply disruption to cause major price increases and a worldwide energy crisis. In unadjusted terms, the price of oil on world markets rose from $2.90 per barrel in September 1973 to $11.65 per barrel in December 1973. Further price hikes and economic repercussions accompanied the Iranian revolution in 1979. Eleven years later – in 1990 – when Iraqi forces invaded Kuwait, OPEC controlled roughly 5.5 million barrels per day (MBD) of spare capacity, enough to replace the oil from the combatant countries and to supply about 8 percent of global demand. Even so, the elimination of Iraqi and Kuwaiti shipments contributed to oil prices jumping from around $21.50 per barrel in January 1991 to $28.30 in February 1991.

In 2005, OPEC's spare production capacity stood at only 2 percent of world demand, with roughly 90 percent of this spare capacity located in Saudi Arabia. The rapidly growing demand for oil by China and India to fuel their expanding economies has placed unprecedented pressure on the world supply of oil, leading to recent prices of crude oil at $70 per barrel and higher. Because spare production capacity is both extremely limited and concentrated in one volatile region, world oil markets remain vulnerable to short-term disruptions. This situation will not likely improve since almost half of the world's proven reserves of conventional oil are in Saudi Arabia, Iraq, and Iran.

The United States remains more susceptible today to oil-supply disruptions and price spikes than at any time in the recent past. It has grown to become the world's largest oil consumer by a considerable margin while its domestic oil production has rapidly diminished. Oil imports have filled the expanding gap and accounted for 58 percent of total U.S. oil consumption in 2005 – up from 22 percent in 1970.

To obtain a sense of the consequences of a disruption in a constrained world

oil market, the National Commission on Energy Policy, a bipartisan group of sixteen leading energy experts, simulated an 'oil-supply shockwave' in 2005. Unrest in oil-producing Nigeria, an attack on an Alaskan oil facility, and the emergency evacuation of foreign nationals from Saudi Arabia precipitated the imagined shockwave, which removed three MBD from the world's market of oil. As result of these events, the price of gasoline in the United States rose to $5.75 per gallon, two million Americans lost their jobs, and the consumer price index jumped 13 percent. Worse, panelists who participated in the study concluded that we could do nothing to avoid these impacts after the hypothetical disruptions began.

The stagnating fuel economy of cars has contributed to America's vulnerability to oil disruptions. Corporate Average Fuel Economy (CAFE) standards for cars peaked in 1985 at 27.5 miles per gallon. For the past two decades, consumer (and manufacturer) preferences for larger and more powerful autos have negated technological advances in front-wheel drive transmissions, electronic fuel injection, enhanced power-train configurations, and computer-controlled engines, which would improve gas mileage even if nothing else were changed in cars. New-vehicle fuel economy therefore remains no higher today than in 1981, but automobile weight has increased by 24 percent and horsepower has almost doubled. In addition, more cars populate the roads, and are driven more miles each year. The net result of these trends has been growing demand for oil in the transportation sector and greater imports to meet that demand.

Second, the United States continues to see increasing demand for electricity in a way that threatens its ability to produce it. The country consumed about 167 percent more electricity in 2004 than it did in 1970, with power usage growing from 25 percent of the nation's total energy use in 1970 to 40 percent in 2004. And this demand for electricity will continue to grow: the Energy Information Administration forecasted in 2005 that electricity use will increase at a rate of 1.9 percent annually through 2025. Though much lower than the 7 percent annual growth rate experienced before the 1973 energy crisis, the current rate would still require a doubling of electricity production in about thirty-seven years.

Increased demand for power in the past decade has been met almost exclusively through the use of quickly built and increasingly efficient natural-gas combustion turbines or combined-cycle equipment. Indeed, more than 150 gigawatts (GW) of gas-fired power generation have been added to the power grid between 1999 and 2004, which totaled about 1,000 GW for the nation in 2004. Despite the high price of this clean-burning gas in the last few years, its use in new power plants seems likely to increase.

But energy analysts see problems with this trend. The National Petroleum Council predicts that current North American sources will be able to satisfy only 75 percent of domestic demand for natural gas. Questions of security will likely emerge as the trend of natural-gas imports begins to emulate the increasing trend of petroleum imports. Aggravating this concern is the possibility that today's nuclear-power plants could be retired over the next fifty years as current licenses expire, depriving the nation of one of its key noncarbon energy sources and pushing up demand for natural gas if that fuel replaces nuclear energy for electricity production.

What about other sources of power? Coal's high carbon content and added

cost of pollution abatement will continue to pose challenges for power providers. Clean coal technologies such as integrated gasification combined cycle and fluidized bed combustion offer policymakers a way to capture concentrated streams of carbon dioxide, but they still remain years away from commercial viability. Because of security problems related to fuel sources and waste disposal, as well as potential public opposition, new nuclear technology also cannot be counted on for widespread near-term use. And despite some impressive federal and state efforts to promote them, non-hydro renewables (such as biomass, geothermal, wind, and solar) have gained only a 2 percent share of electricity generation over the past thirty years. Reductions in the cost of power produced from renewables in this time have been impressive, making them look increasingly attractive for future use. Yet the intermittence of renewables – especially the most cost-effective wind turbines – coupled with high capital costs, a host of lingering utility-monopoly rules, and public opposition to local siting will likely prevent such technologies from taking over the bulk of the generation burden, at least in the next thirty years. Overall, it appears that meeting future demand for electricity will become an increasingly arduous undertaking.

Third, the electric-power-transmission infrastructure remains precarious and brittle, despite its increasing use. Data from the Edison Electric Institute and the Electric Power Research Institute note that utility investment in transmission peaked at almost $10 billion in 1970, but declined to an inadequate level of $2.2 billion in 1998 (in 2003 dollars). Spending grew to $3.8 billion in 2002 and $4.1 billion in 2003, though many analysts still feel more investment is necessary to transmit power to the growing wholesale and retail markets that have been created since utility-industry restructuring began in the 1990s.

But much higher spending may not be forthcoming, given that (as noted in a 2003 RAND Corporation study) incentives in the partially deregulated utility industry favor minimal investments in transmission facilities. Because federal regulators generally limit rates of return on transmission investments, companies often prefer to construct and operate new generation facilities, whose uncapped rates of return depend only on market conditions. To complicate matters more, local opposition to new power lines has grown over the years as the country has become more populated, resulting in delayed construction (or cancellation) of some transmission facilities. Taken together, these trends have resulted in a decreasingly reliable transmission network in many regions of the United States, with grid components being operated close to (or at) their technical limits.

The Energy Policy Act of 2005 includes provisions to respond to some infrastructural problems, such as incentives to increase investment in transmission lines and to simplify the planning and permitting process for building them. These measures may help, as thousands of miles of new transmission lines may be required if the electric-utility system expands along the same lines as it has for the past several decades. Increasing demand for other forms of energy in the future may also stress the country's infrastructure. Numerous new port terminals will be required to handle increased imports of liquefied natural gas and oil, for example. At the same time, new carbon-sequestration sites, bioenergy facilities, and hydrogen repositories and pipelines may be needed, espe-

cially as efforts increase to reduce environmental pollution. But these needs will not be easily met. Carbon sequestration, for example, may require use of depleted oil and gas fields, unmineable deep coal seams, or cavernous saline formations. The successful use of these geological formations will depend on techniques that resist operator and equipment failure, extreme weather, and malicious interference or attacks. Similar concerns over technical errors and assaults arise when considering the need for expansion of natural-gas and petroleum facilities. Opposition to construction of these new infrastructural elements has already become evident. Put simply, the future health of the country's energy infrastructure may be in peril.

Fourth, the country faces immense challenges in promoting more energy-efficient technologies. Before the 1973 OPEC oil embargo, U.S. energy consumption grew in lockstep with the nation's gross domestic product (GDP). Measured in terms of energy consumption per dollar of GDP, the energy intensity of the nation remained constant. Economic growth appeared to require consuming more energy.

This trend changed in the period after the 1973 energy crisis, when the economy (as measured by the inflation-adjusted GDP) grew by 148 percent (from 1973 to 2004). Total U.S. energy consumption, meanwhile, grew from about seventy-six quadrillion British Thermal Units of energy (quads) to almost one hundred quads in the same period, an increase of 32 percent. The energy intensity of the economy, in other words, dropped considerably.

What accounted for the change? Individuals purchased more fuel-efficient cars and appliances; they insulated and weatherproofed their homes; and they adjusted thermostats to reduce energy consumption. These measures led to a decrease in per capita residential energy use of 27 percent (and 37 percent per household) despite a 50 percent increase in new home size since 1970 and the growing use of air conditioning, electronic equipment, and a multitude of 'plug loads.' Businesses retrofitted their buildings with more efficient heating and cooling equipment and installed energy management and control systems, accounting for a 25 percent decline in energy use per square foot of commercial building space. Factories adopted more 'energy-stingy' manufacturing processes and employed more efficient motors for conveyors, pumps, fans, and compressors. These gains in energy productivity, prompted by high fuel costs and government policies, represent one of the great economic success stories of this century. If the nation's energy intensity remained the same today as it stood in 1970, the United States would be consuming twice as much energy, and its energy bill would be approximately $1 billion higher per day.

While such data suggest that energy-efficiency investments provide an economic and relatively rapid strategy for meeting the growing demand for energy services, many experts assert that efficiency can only play a limited policy role. For example, Hans Blix, the former director of the International Atomic Energy Agency, has argued, "The more efficient use of energy will only partially slow down the expanding use of energy. Although our light bulbs will save electricity, we shall have more lights." Similarly, Vice President Dick Cheney stated in 2001 that "conservation may be a sign of personal virtue, but it is not a sufficient basis for a sound, comprehensive, energy policy." And Spencer Abraham, President Bush's Secretary of Energy

from 2001 to 2005, reiterated this view when he told senators that "improved energy efficiency cannot do the whole job . . . . [T]he United States will need more energy supply." In short, efficiency may help the nation overcome some of its energy woes, but policymakers do not feel it will be the ultimate solution. As a result, the potential for improved energy efficiency is not being vigorously tapped.

Fifth and finally, the trend toward more energy consumption will exacerbate already prominent concerns about the environment. Since the 1960s, technically trained people, politicians, and the public have become aware of the health consequences of the exploration, extraction, transportation, and combustion of fuels used for making energy. They have also become alert to possible dangers of living near high-voltage power lines and radioactive-waste sites. More recently, people have pointed to the ecological damage created by hydroelectric dams and wind turbines, while also noting that the use of biomass from energy crops may promote agricultural monocultures that can pose severe risks to ecological diversity.

Efforts resulting from three decades of clean-air legislation have decreased sulfur-dioxide emissions from electric generators in the United States. Nevertheless, air pollution remains a serious threat to human and ecosystem health. Americans have experienced a rise in respiratory illnesses, and visibility continues to degrade in formerly pristine areas as a result of pollution from vehicles and coal-burning power plants. Rarely, for example, does visibility in the Great Smoky Mountains National Park achieve its 'natural' limit of ninety-three miles. Instead, average annual visibility has decreased to twenty-five miles in the winter and to twelve miles in the summer. Beyond air-pollution issues, current energy trends will lead to expanded emissions of greenhouse gases, which appear to be contributing to increased global temperatures, recession of glaciers, and more frequent and powerful weather events such as hurricanes.

The pollution associated with electric-power production was vividly documented by the August 14, 2003, Northeast blackout. Not only did the event shut off electricity for 50 million people in the United States and Canada, it also halted emissions from many fossil-fired power plants across the Ohio Valley and the Northeast. In effect, the power outage served as an inadvertent demonstration of the environmental consequences of electricity generation: twenty-four hours after the blackout, New York City's sulfur-dioxide concentrations dropped 90 percent; particulate matter fell by 70 percent; and ozone concentrations slipped to half.

Beyond federal clean-air initiatives, state-government policies have, in certain cases, made positive inroads to pollution abatement. Due to legislative and regulatory initiatives, California – which generates roughly one-fourth of its electricity from efficiently distributed and renewable energy technologies – emitted only 493 metric tons of carbon dioxide in 2002, a mere 12 percent increase from its emission levels in 1990, despite an increase in electricity demand of almost 25 percent.

Though making impressive inroads in pollution abatement efforts, California (and a few other states) remains the exception, not the rule. Few people dispute the fact that total U.S. emissions of carbon dioxide from energy consumption have increased significantly: from 4.3 billion metric tons in 1970 to 5.9 billion metric tons in 2004. Moreover, the Ener-

gy Information Administration forecasted in 2005 that carbon-dioxide emissions from energy use will grow an average 1.5 percent annually for the next twenty years, resulting in 8.1 billion metric tons of carbon-dioxide emissions in 2025. Clearly, the last thirty years have not seen the adoption of the low-carbon power and fuels needed to help stabilize atmospheric concentrations of greenhouse gases. Continued growth in energy usage will likely exacerbate environmental problems.

To conclude, despite three decades of 'progress' since the 1973 energy crisis, the United States faces a host of energy challenges that threaten the nation's economy, security, and lifestyle. Because of its huge dependence on imported oil to fuel a transportation sector that has seen little improvement in energy efficiency, the nation could be ravaged by disruptions to oil supplies due to weather, war, or terrorist attacks. At the same time, growing electricity consumption and reliance on power plants employing natural gas, along with a constrained transmission grid, make the electric-utility infrastructure increasingly vulnerable to service disruptions. And while efficiency efforts have successfully stemmed the growth rate of fuel consumption in the last few decades, population increases and economic expansion have forced up the nation's overall use of energy, exacerbating the country's environmental problems.

As a consequence of these trends, the goal of energy independence seems more distant in 2006 than it did in 1974, when President Nixon first proposed it as a way to deal with the oil embargo. While one can fruitfully debate whether complete reliance on domestic energy sources should be the objective of government policy, the fact remains that the United States cannot continue upon its present course. The country has become progressively vulnerable to economic, political, and military threats because of its growing fuel consumption and an increasingly challenged energy infrastructure. The nation's policymakers in business and government, as well as the citizenry, need to realize that the recent trends in energy consumption, production, and distribution reflected in this energy assessment cannot be sustained indefinitely. Americans must confront energy concerns as a top priority and learn to overcome the social, political, and technical obstacles that have hindered true progress for more than three decades.

Jorie Graham

Incarnation: 9:30 am to 9:36 am

She sits on the straightback chair in the room.
A ray of sun is calling across the slatwood floor.
I say *she* because my body is so still
in the folds of daylight
through which the one beam slants.
I say *calling* because it lays itself down
with a twang and a licking monosyllable

across the pine floor-boards –
making a meaning like a wide sharp thought –
an unrobed thing we can see the inside of –
less place than time –

less time than the shedding skin of time, the thought
of time,

the yellow swath it cuts
on the continuum –
now to the continuum
what she is to me,
a ceremonial form, an intransigent puissant corridor
nothing will intersect,

and yet nothing really
– dust, a little heat …
She waits.
Her leg extended, she waits for it –
foot, instep, calf –
the I, the beam
of sun –
the *now* and *now* –

it moving like a destiny across,
neither lured-on nor pushed-forward,
without architecture,
without
beginning,
over the book lying in the dust,

over the cracked plank – down into the crack – across –

not animal
nothing that can be deduced-from or built-upon,
aswarm with dust and yet
not entered by the dust,

not *touched* –
smearing everything with a small warm gaiety –

over the pillow-seam over the water glass –

cracking and bending but not cracking or bending –

over the instep now, holding the foot –

her waiting to feel the warmth then beginning
to feel it –
the motion of it and the warmth of it not identical –
the one-way-motion of it, the slow sweep,
approaching her as a fate approaches, inhuman but
resembling
feeling,

without deviation,
turning each instant a notch deeper towards
the only forwards,
but without beginning,
and never – not ever –
not moving
forwards…

Meanwhile the knowledge of things lies round,
over which the beam –
Meanwhile the transparent air
through or into which the beam –
over the virtual and the material –
over the world and over the world of the beholder –
glides:

it does not change, crawler, but things are
changed –
the mantle, the cotton-denim bunched at
the knees –

diamonds appearing on the tips of things then disappearing –
each edge voluble with the plushnesses of silence –

now up to her folded arms – warm under the elbow –
almost a sad smell in the honeyed yellow –
(the ridge of the collarbone) (the tuck of the neck)
till suddenly (as if by
accident)

she is inside – (ear, cheek) – the slice of time

now on the chin, now on
the lips, making her rise up into me,
forcing me to close my eyes,
the whole of the rest feeling broken off,

it all being my face, my being inside the beam of sun,

and the sensation of how it falls unevenly,

how the wholeness I felt in the shadow is lifted,
broken, this tip *lit*, this other *dark* – and stratified,
analysed, chosen-round, formed –

Jorie Graham, a Fellow of the American Academy since 1999, is Boylston Professor of Rhetoric and Oratory at Harvard University. She has published numerous collections of poetry, including "Hybrids of Plants and of Ghosts" (1980), "The End of Beauty" (1987), "The Dream of the Unified Field: Selected Poems 1974 – 1994" (1995), which won the 1996 Pulitzer Prize for Poetry, "Never" (2002), and, most recently, "Overlord" (2005). She has also edited two anthologies, "Earth Took of Earth: 100 Great Poems of the English Language" (1996) and "The Best American Poetry 1990." Graham served as a Chancellor of The Academy of American Poets from 1997 to 2003.

© 2006 by Jorie Graham

Antonio Damasio & Hanna Damasio

Minding the body

We spend a good part of our lives attending to the sights and sounds of the world outside of us, oblivious to the fact that we (mentally speaking) exist in our bodies, and that our bodies exist in our minds.[1] This neglect is both good and bad: good when it allows us to let our own physical suffering go undetected, bad when it screens us from the biological roots of our selves. Be that as it may, the body does come to mind, in no uncertain terms, when injury or disease breaks down its integrity and causes pain. The body also comes to mind, somewhat less demandingly, in moments of joy, when physical lightness and ease of function inevitably make us aware of the body.

What we would like to address in this essay, however, is not the obvious fact that the state of our bodies can be conveyed to our minds, but rather the neurological mechanisms that enable this spectacular phenomenon. How can the body, with its myriad physical compartments and complicated operations, show up in our minds and be felt?

The most common perspective on this question assumes that there is a 'mind-body problem,' which is best resolved using the tools of analytic philosophy. Our perspective here, however, is more restricted, focusing instead on the biological scaffolding without which the body certainly cannot be present in the mind. We are convinced, incidentally, that this perspective and its facts are relevant to the mind-body problem and go a long way toward solving it. Elsewhere we have addressed the connection,[2] but

Antonio Damasio, a Fellow of the American Academy since 1997, is David Dornsife Professor of Neuroscience and director of the Brain and Creativity Institute at the University of Southern California. He is the author of "Descartes' Error" (1994), "The Feeling of What Happens" (1999), and "Looking for Spinoza" (2003).

Hanna Damasio, a Fellow of the American Academy since 1997, is Dana Dornsife Professor of Neuroscience and director of the Dornsife Cognitive Neuroscience Imaging Center at the University of Southern California. She is the author of "Lesion Analysis in Neuropsychology" (1989) and "Human Brain Anatomy in Computerized Images" (1995), the new edition of which was published in 2005.

© 2006 by the American Academy of Arts & Sciences

1 The work discussed in this essay has been supported by a grant from The Mathers Foundation.

2 See Antonio Damasio, *The Feeling of What Happens* (New York: Harcourt, 1999), for an overview of the argument.

here we simply wish to discuss the latest in the nuts and bolts of how the body comes to mind.

In attempting to answer this question, we will use, as a stepping stone, the same assumptions William James made when he tried to explain how we perceive our emotions, a process he thought required a mental representation of the body: perceptions, any perception, occur in the brain; and perceptions of the body include, of necessity, brain processes that depend on a particular object – the body-proper.

James, along with a host of contemporary physiologists, already knew for a fact that there were nerve pathways conducting impulses from the body to the brain. He also had an inkling, given the then-emerging evidence for brain specialization, that the parts of the brain related to bodily perception would be distinct from those linked to visual or auditory perceptions. Today, we have no reason to doubt James's conjecture, and plenty of evidence to show that his account is fundamentally correct.

We have, however, many new developments to report on this score.[3] First, we now know that the details of James's basic arrangement are far more intricate than what might have been imagined a century ago. For example, the body uses chemical signals as well as neural signals to communicate with the brain; and the range of information conveyed to the brain is wider than expected, from the concentration of chemical molecules to the contractions of muscles anywhere in the body.

Second, while the brain does represent, with fidelity, body states that are actually occurring, it can do far more than that: it can also modify the representation of an ongoing state, and, most dramatically, it can simulate body states before they occur or body states that do not occur at all.

Third, the new knowledge has profound implications for our understanding of consciousness and of social behavior.

In what follows, we will address each of these developments in order.

Let us begin by clarifying that whenever we use the term 'body,' we mean 'body-proper,' so as to leave aside the brain. The brain is also a part of the body, of course, but it happens to have a particular status: it is the part of the body toward which every other body part is communicating, and that can communicate to every other part.

Misconceptions abound regarding how the body talks to the brain. For instance, those who are unacquainted with neuroscience (and, regrettably, some who are) assume incorrectly that the body operates as a single unit, a lump of flesh connected to the brain by live wires called nerves. Many also mistakenly believe that the main communication occurs between the body's organs – the viscera – and the brain, via the autonomic nervous system. Another erroneous belief is that the body-brain communication goes one way, from body to brain, but not in the reverse direction.

The reality is quite different.

First, rather than being one unit, the body has numerous separate compartments. To be sure, the viscera to which so much attention is paid – the heart, the lungs, the gut, the mouth, the tongue, the throat, and the equally vital but less recognized organ, the skin, which envelops our entire organism – are essential, but all of the body's compartments are

3 See Antonio Damasio, *Looking for Spinoza* (New York: Harcourt, 2003), for an overview.

indispensable for its normal operation. And the communication between viscera and brain is not confined to the autonomic nervous system. There are other neural channels. This communication is not even confined to nerves alone: there is a chemical channel as well.

The chemical bath in which all body cells live and of which the blood is an expression – the internal milieu – also ends up sending signals to the brain, not via nerves but via chemical molecules, which impinge directly on certain parts of the brain designed to receive their messages. The range of information conveyed to the brain in this manner is extremely wide. It includes, for example, the state of contraction or dilation of smooth muscles (the muscles that form the walls of the arteries, the gut, and the bronchi); the amount of oxygen and carbon dioxide concentrated locally in any region of the body; the temperature and the pH at various locations; the local presence of toxic chemical molecules; and so forth.

There is more to the body, however, than viscera and internal milieu. There are also striated muscles. When these muscles are connected to two bones articulated by a joint, the shortening of these muscles generates movement. Picking up an object, walking, talking, breathing, and eating are all actions that depend on muscular contraction. But note that whenever those contractions occur, the configuration of the body changes: except for moments of complete immobility, the image of the body in space is changing continuously.

The secret behind the changes is simple. In order to control movement with precision, the body must instantly convey information on the state of every striated muscle to the brain, using efficient nerve pathways, which are evolutionarily more modern than those that convey signals from the viscera and internal milieu. These pathways arrive in brain regions dedicated to sensing the state of these muscles. (The only partial exception to this scheme of things has to do with the heart, which is also made of striated muscles, and whose contractions serve to pump blood, but whose signals are segregated to brain sites dedicated to the viscera.)

However, just as important as the body-to-brain inputs described above is the fact that the brain also sends messages to the body. This fact is often forgotten. While body states are being continuously mapped in the brain, many aspects of those body states were caused by brain signals to the body in the first place. Just as with the communication from the body to the brain, the brain communicates with the body via neural channels – nerves whose messages lead to the contraction of muscles and the execution of actions – and via chemical channels. Examples of the latter include hormones, such as cortisol, testosterone, or estrogen. The release of hormones results in different modes of operation for the internal milieu and the viscera.

The idea here is that the body and brain are engaged in a continuous interaction that unfolds in time, within different regions of the body and within mental space as well. Mental states cause brain states, which cause body states; body states are then mapped in the brain and incorporated into the ongoing mental states. A small change on the brain side of the system can have major consequences for the body state (think of the release of a hormone); likewise, a small change on the body side (think of a knife cut, or a tooth infection, or the rupture of an ulcer) can have a major effect on the mind once the change is mapped as a nociceptive state and perceived as acute pain.

Curiously, German and British physiologists described the outlines of this basic arrangement over a century ago, but its importance for the understanding of both neurobiology and cognitive science was largely ignored until recently. And its intricate neuroanatomical and neurophysiological details have only been uncovered in the past few years.[4]

Those details reveal that the state of the body's interior is conveyed via neural channels dedicated to specific brain regions. Specific nerve-fiber types – C-fibers – bring signals from every nook and cranny of the body into specific parts of the central nervous system, such as the lamina-I section of the posterior horn of the spinal cord, at every level of its length, or the pars caudalis of the trigeminal nerve. The brain regions charged with handling the signals as they march toward the higher levels of the brain are also specific. Together with chemical information available in the bloodstream, these messages convey to the brain the state of a good part of the body's interior – all the visceral/chemical body components beneath the skin's outer perimeter. Complementing this interior sense, or interoception, is information regarding the state of striated muscles anywhere in the body. Messages from the striated muscles use different kinds of nerve fibers and different stations of the central nervous system all the way into the higher levels of the brain. The upshot of all this signaling is a multidimensional picture of the body in the brain and in the mind.

That the body, in most of its aspects, is continuously represented in the brain is thus a well-proven fact. The organism requires that sort of ongoing representation for rather transparent reasons: in order for the brain to coordinate physiological states in the body-proper, which it does without our being consciously aware of what is going on, the brain must be informed about the various physiological parameters at different regions of the body. The information must be faithful and current if it is to permit optimal controlling responses: call this information-processing network the 'body loop.'

But this is not the only network that links mind and body – there is another we call the 'as-if body loop.' Some fifteen years ago, Antonio proposed that in certain circumstances, as an emotion unfolds, the brain rapidly constructs maps of the body comparable to those that would result were the body actually changed by that emotion. The construction can occur well ahead of the emotional change, or even instead of the change. In other words, the brain can simulate a certain body state as if it were occurring; and because our perception of any body state is rooted in the body maps of the somatosensing regions, we perceive the body state as actually occurring even if it is not. (This functional arrangement can work for emotion because there is no need for fidelity of information concerning the body states that define an emotion provided the kind of emotion in question can be detected without ambiguity.)

At the time the 'as-if body loop' hypothesis was first advanced, the evidence we could muster in its favor was circumstantial. First, it made sense for the brain to know about the body state it was about to produce. The advantages of this sort of 'advance simulation' were obvious in the phenomenon of 'efferenz-copie.' Efferenz-copie is what allows

4 A. D. Craig, "How Do You Feel? Interoceptors: The Sense of the Physiological Condition of the Body," *Nature Reviews* 3 (2002): 655 – 666; Damasio, *Looking for Spinoza*.

motor structures that are about to command the execution of a certain movement, to inform visual structures of the likely result of that forthcoming movement. For example, when our eyes are about to move toward an object at the periphery of our vision, the visual region of the brain is forewarned of the impending movement and is ready to smooth the transition to the new object without creating a blur.

Second, it seemed obvious that simulating a body state without actually producing it reduces processing time and saves energy.

Third, the neuroanatomical structures needed for the as-if body loop to work were known to exist. The hypothesis required that the brain structures in charge of triggering a particular emotion be able to connect to the structures in which the body state corresponding to the emotion would be mapped. For example, the amygdala (a triggering site for fear) and the ventromedial prefrontal cortex (a triggering site for compassion) would have to connect to somatosensing regions such as the insular cortex, SII, and SI, where ongoing body states are continuously represented. We knew that such connections existed, thereby rendering possible the implementation of the hypothetical mechanism.

In recent years, more support for this hypothesis has come from several quarters, for example, from a series of remarkable experiments by Giacomo Rizzolatti and his colleagues. These experiments identified a class of nerve cells, known as 'mirror neurons,' by means of implanted electrodes in monkeys. In these experiments, a monkey would watch an investigator perform a variety of actions.[5] When a monkey saw the investigator move his hand or mouth, neurons in the monkey's brain regions related to hand or mouth movements became active, 'as if' the monkey were performing the action. But in reality, the monkey remained immobile.

Mirror neurons are, in effect, the ultimate 'as-if body' device. The mirror-neuron system achieves conceptually what we hypothesized as the 'as-if body loop' system: the simulation, in the brain's body maps, of a body state that is not actually taking place in the organism. The fact that the body state the mirror neurons are simulating is not the subject's does not minimize the power of this functional resemblance. On the contrary, it stands to reason that if a complex brain can simulate someone else's body state, it can simulate one of its own body states. Take, for example, a state that has already occurred in the organism: it should be easier to simulate since it has already been mapped by precisely the same somatosensing structures that are now responsible for simulating it. In fact, we suggest that the as-if system applied to others would not have developed had there not been an as-if system applied to the brain's own organism.

The nature of the brain structures involved in the process reinforces the suggestive functional resemblance between the as-if body loop and mirror neurons. For the as-if body loop, we hypothesized that neurons in areas engaging emotion – the premotor/prefrontal cortex (in the case of compassion), the anterior insular cortex (in the case of disgust), and the amygdala (in the case of fear) – would activate regions that normally map the state of the body and, in turn, move it

5 G. Rizzolatti, L. Fadiga, L. Fogassi, and V. Gallese, "Resonance Behaviors and Mirror Neurons," *Archives Italiennes de Biologie* 137 (1999): 85 – 100; V. Gallese, "The Shared Manifold Hypothesis," *Journal of Consciousness Studies* 8 (2001): 33 – 50.

to action. In humans such regions include the somatomotor complex in the rolandic and parietal operculum as well as the insular cortex. All of these regions have a dual somatomotor role, the insular cortex providing an especially good example of this functional duality. That is, these regions can hold a map of the body state, a sensory role, and they can participate in an action as well. By and large, this is what Rizzolatti's neurophysiologic experiments with monkeys uncovered. This discovery is quite consonant as well with human studies using magnetoencephalography (the studies of Rita Haari) and functional neuroimaging (again by the Rizzolatti group and by Tania Singer and her colleagues). Our own studies based on neurologic lesions point in the same direction.[6]

Explanations of the existence of mirror neurons have emphasized, quite appropriately, the role that mirror neurons can play in allowing us to understand the actions of others by placing us in a comparable body state. As we witness an action in another, our body-sensing brain adopts the body state we would have were we ourselves moving, and it does so, in all probability, not by passive sensory patterns, but rather by a preactivation of motor structures – ready for action yet not quite allowed to act – and in some cases by actual motor activation.

6 R. Haari, N. Forss, S. Avikainen, E. Kirveskari, S. Salenius, and G. Rizzolatti, "Activation of Human Primary Motor Cortex During Action Observation: A Neuromagnetic Study," *Proceedings of the National Academy of Sciences* 95 (1998): 15061 – 15065; T. Singer et al., "Empathy for Pain Involves the Affective but not Sensory Components of Pain," *Science* 303 (2004): 1157 – 1162; R. Adolphs, H. Damasio, D. Tranel, G. Cooper, and A. Damasio, "A Role for Somatosensory Cortices in the Visual Recognition of Emotion as Revealed by Three-Dimensional Lesion Mapping," *Journal of Neuroscience* 20 (2000): 2683 – 2690.

We must wonder, however, how such a complex physiologic system evolved. We doubt it arose *de novo*, simply because of the manifest advantage of better knowing the body state of others and, through it, the mind state of others. Instead, we suspect that the mirror system developed from an earlier 'as-if body loop' system, which complex brains used to simulate their own body states for a clear and immediate advantage: rapid and energy-saving activation of the maps of certain body states, which were, in turn, associated with relevant past knowledge and cognitive strategies. Eventually, the as-if system was applied to others and prevailed because of the equally obvious social advantages one could derive from knowing the body states of others, which are connected, of course, to their mental states.

In brief, we regard the remarkable evidence for mirror neurons as support for the 'as-if body loop' mechanism of placing the body in mind. In turn, we consider the 'as-if body loop' system within each organism as the precursor to the mirror-neuron system.

The as-if body loop, the body loop, and mirror neurons all point to a few remarkable features regarding the perception of the body during the experience of an emotion: The emotion ends up felt in our flesh. The process unfolds in time and is both sensory and motor. The sensing of body changes leads to motor activations that, in turn, can be sensed. All of these steps have the power to evoke related knowledge held in memory.

There is a peculiar relationship between the object perceived, our body, and the brain. The body and brain inhabit the same organism, and the relationship between the two can be entire-

ly circular. Feelings of emotions can help illustrate the situation.

Feelings of emotions are perceptions. They are, in the general scheme of things, comparable to other perceptions. For example, actual visual perceptions correspond to external objects whose physical characteristics impinge on our retinas and modify transiently the sensory maps in the visual system. Feelings of emotions also have an object at the origin of the process, and the physical characteristics of the object also prompt a chain of signals that impinges on maps inside the brain.

In other words, just as in the case of visual perception, a part of feelings of emotion is due to the object, and a part is due to the internal construction the brain makes of it. But something is quite different in the case of feelings, and the difference is not trivial. In feelings, the object at the origin of the process is inside the body rather than outside. Feelings of emotion are just as mental as any other perception, but a sizable part of the objects being mapped are states of the living organism in which the feelings arise.

As if this difference does not complicate things enough, there is another wrinkle in the process: feelings of emotion are linked to an object called the body, but they are also linked to the emotionally competent object that initiated the emotion-feeling cycle. A spectacular seascape is an object, but so is the body state that results from beholding that seascape, and it is the latter object at the origin of the emotional process that we perceive in the resulting feeling state.

In feeling, the brain can act directly on the very object that it is perceiving, because the object at the origin is inside the body. It can do so by modifying the state of the body, or by altering the transmission signals from it. Thus, the object at the origin, on the one hand, and the brain map of that object, on the other, can exert mutual influences in a sort of reverberative process that is not to be found in the perception of an external object. We can look at a painting we admire as intensely as we wish, for as long as we wish, and react emotionally to it. But nothing will happen to the painting itself. Our thoughts about it will change, but the object remains intact.

By contrast, in the feeling of emotion, the object itself – the body – can be changed radically. In other words, the feeling of emotion is not merely a passive perception or a flash in time. For a period of seconds or even minutes after a feeling begins, a dynamic engagement of the body, conducted almost certainly in a reiterative fashion, leads to a dynamic variation of the perception. Part of the variation may even be due to the homeostatic necessity of controlling the motor upheaval caused by the emotive process.[7] In brief, the body in the mind undergoes continuous transitions.

The fact that the body of a given organism can be fully represented in the brain of that organism opens important possibilities. The first relates to consciousness, specifically with the part of the process called the self. Elsewhere we have argued that the construction of the self would simply not be possible if the brain did not have available a dynamic representation of its body.[8] Con-

7 D. Rudrauf and A. Damasio, "A Conjecture Regarding the Biological Mechanism of Subjectivity and Feeling," *Journal of Consciousness Studies* 12 (8 – 10) (2005): 236 – 262.

8 Damasio, *The Feeling of What Happens*; J. Parvizi and A. Damasio, "Consciousness and the Brainstem," *Cognition* 79 (2001): 135 – 160.

sciousness is about the relation between a given organism and the objects perceived in its mind. In the mental process depicting the self, the integrated body representation serves as a stand-in for the organism. There is an invariant aspect to the body representation – its components and the schema according to which they are interconnected – and a variable aspect – the dynamic changes the components constantly undergo. Eventually the body representation behaves as an anchor for the construction of the self – a mental stand-in for the individual, for his or her personhood and identity.

These body representations have another major implication: after allowing us to represent our own actions and emotional states, actual or simulated, they allow us to simulate the equivalent states of others. And because we have established a prior connection between our own body states and their significance, we can subsequently attribute the same significance to the states of others that we come to simulate. The body in mind helps us construct our selves and then allows us to understand others, which is nothing short of astounding.

Gerald M. Edelman

The embodiment of mind

The word 'mind' is a loose one with many applications in use. As I use it here, I am restricting it to one definition in *Webster's Third International Dictionary*: "Mind – the sum total of the conscious states of an individual." I want to suggest a way of looking at consciousness in tune with, and responsive to, a statement on the subject by the American philosopher Willard van Orman Quine.[1] With his usual ironic candor, Quine said,

> I have been accused of denying consciousness, but I am not conscious of having done so. Consciousness is to me a mystery, and not one to be dismissed. We know what it is like to be conscious, but not how to put it into satisfactory scientific terms. Whatever it precisely may be, consciousness is a state of the body, a state of nerves.
>
> The line I am urging as today's conventional wisdom is not a denial of consciousness. It is often called, with more reason, a repudiation of mind. It is called a repudiation of mind as a second substance, over and above body. It can be described less harshly as an identification of mind with some of the faculties, states, and activities of the body. Mental states and events are a special subclass of the states and events of the human or animal body.

Philosophers have wrestled with the so-called mind-body problem for millennia. Their efforts to explore how consciousness arises were intensified following René Descartes' espousal of dualism. The notion that there are two substances – extended substances (*res extensa*), which are susceptible to physics, and thinking substances (*res cogitans*), which are unavailable to physics – still haunts us. This substance dualism forced confrontation with a key question: how could the mind arise in the material order? Attempts to answer this question have ranged widely. In addition to the various forms of dualism, a few proposals we might mention are panpsychism (consciousness inheres in all mat-

Gerald M. Edelman, a Fellow of the American Academy since 1968, is director of the Neurosciences Institute as well as chair of, and professor in, the Department of Neurobiology at the Scripps Research Institute. His publications include "The Mindful Brain" (1978), "The Remembered Present" (1989), and "Wider Than the Sky" (2004). He received the Nobel Prize in Physiology or Medicine in 1972.

© 2006 by the American Academy of Arts & Sciences

1 W. V. Quine, *Quiddities: An Intermittently Philosophical Dictionary* (Cambridge, Mass.: Belknap Press, Harvard University Press, 1987), 132 – 133.

ter in varying degrees), mind-body identity (the mind is nothing but the operation of neurons in the brain), and, more recently, the proposal that the understanding of quantum gravity will ultimately reveal the nature of consciousness.[2] There are many more proposals, but aside from the extremes of idealism espoused by Bishop Berkeley and Georg Hegel, they all wrestle with one question: how can we explain consciousness in bodily terms?

Attempts to answer this question often begin by examining the features of consciousness to generate a number of more pointed questions. I shall follow that path here. But I don't wish to consider the subject from a philosophical point of view. Rather, I will describe a theory of consciousness based on some significant advances in neuroscience.

FEATURES OF CONSCIOUSNESS: Consciousness is a process, not a thing. We experience it as an ongoing series of myriad states, each different but at the same time each unitary. In other words, we do not experience 'just this pencil' or 'just the color red.' Instead, within a period I have called the remembered present,[3] consciousness consists of combinations of external perceptions and various feelings that may include vision, hearing, smell, and other senses such as proprioception, as well as imagery, memory, mood, and emotion. The combinations in which these may participate are usually not fragmented, but instead form a whole 'scene.' Consciousness has the property of intentionality or 'aboutness' – it usually refers to objects, events, images, or ideas, but it doesn't exhaust the characteristics of the objects toward which it is directed. Furthermore, consciousness is qualitative, subjective, and therefore, to a large degree, private. Its details and actual feel are not obviously accessible to others as they are to the conscious individual who has wide-ranging first-person access to ongoing phenomenal experience.

This brief summary prompts me to single out three challenging questions: 1) How can the qualitative features of consciousness be reconciled with the activity of the material body and brain (the qualia question)? 2) Does the conscious process itself have effects? In other words, is the process of consciousness causal (the question of mental causation)? 3) How can conscious activity refer to, or be about, objects, even those that have no existence, such as unicorns (the intentionality question)?

BODY, BRAIN, AND ENVIRONMENT – THE SCIENTIFIC APPROACH: There is a voluminous body of philosophical thought that attempts to answer these questions. The efforts of nineteenth-century scientists in this regard were relatively sketchy. But a new turn dating from the 1950s has invigorated the scientific approach to consciousness.[4] Neuroscientific investigation has uncovered a rich store of anatomical, physiological, chemical, and behavioral information about our brains. It has become possible to lay the groundwork for a biologically based theory of consciousness, and I believe we are now in a position to reduce Quine's mystery. In this brief essay,

2 R. Penrose, *The Emperor's New Mind* (Oxford: Oxford University Press, 1989).

3 G. M. Edelman, *The Remembered Present: A Biological Theory of Consciousness* (New York: Basic Books, 1989).

4 T. C. Dalton and B. J. Baars, "Consciousness Regained: The Scientific Restoration of Mind and Brain" in *The Life Cycle of Psychological Ideas*, ed. T. C. Dalton and R. B. Evans (New York: Kluwer Academic/Plenum Publishers, 2004), 203 – 247.

I want to lay out some thoughts that bear directly on the nature of consciousness, as well as on how we know, how we discover and create, and how we search for truth. There is nature, and there is human nature. How do they intersect?

In the first place, we must recognize that consciousness is experienced in terms of a triadic relationship among the brain, the body, and the environment. Of course, the brain is the organ we wish to examine. But the brain is embodied, and the body and brain are embedded in the world. They act in the world and are acted upon by it.

We know that in vertebrate species, and specifically in humans, the development of the brain (for example, the organization of its sensory maps) depends on how our eyes, ears, and limbs receive sensory input from the environment. Change the sequence of actions and inputs to the brain, and the boundaries and response properties of brain maps change, even in adult life. Moreover, we sense our whole body (proprioception) and our limbs (kinesthesia), as well as our balance (vestibular function), and this tells us *how* we are interacting, consciously or not. We also know that damage to the brain – for example, from strokes involving the cerebral cortex – can radically change how we consciously 'sense' the world and interpret our bodies. Finally, through memory acting in certain sleep states, the brain can give rise to dreams in which our body seems to carry out actions of an unusual kind. The dreams of REM sleep, however fantastic, are in fact conscious states.[5]

5 Rapid eye movement (REM) sleep is characterized by fast jerky eye movements, dreams, and the absence in the electroencephalogram of so-called delta waves that are seen in so-called slow-wave or non-REM sleep.

The embodiment of mind

NEUROLOGY ESSENTIAL FOR CONSCIOUSNESS: What can we say about the brain structures whose interactions are responsible for such states? One such interactive structure is the cerebral cortex.[6] Most people are familiar with the cerebral cortex as the wrinkled mantle seen in pictures of the human brain. It is a thin six-layered structure, which, if unfolded, would be about the size of a large table napkin and about as thick. It contains approximately 30 billion neurons or nerve cells, and one million billion synapses connecting them. Moreover, its regions receive inputs from other parts of the brain and send outputs to other portions of the central nervous system such as the spinal cord. There are cortical regions receiving signals from sensory receptors that are functionally segregated for vision, hearing, touch, and smell, for example. There are other cortical regions, more frontally located, which interact mainly with each other and with more posterior regions. There are also regions concerned with movement, for example, the so-called motor cortex.

A key feature of the cortex is that it has many massively parallel nerve fibers connecting its various regions to each other. These cortico-cortical tracts mediate the interactions that are critical for binding and coordinating different cortical activities.

Another structure that is critical for consciousness is the thalamus. This is a relatively small, centrally located collection of so-called nuclei that mediate inputs to, and outputs from, various regions of the cortex. For example, the thalamus processes inputs coming from the eyes via the optic nerves and sends fibers called axons to a posterior corti-

6 G. M. Edelman, *Bright Air, Brilliant Fire: On the Matter of the Mind* (New York: Basic Books, 1992).

cal region called V1. V1, in turn, sends reciprocal fibers back to the thalamus. Similar thalamo-cortical and cortico-thalamic connections exist for all other senses except for smell; each sense is mediated by a specific thalamic nucleus.

It is known that strokes damaging a cortical area such as V1 lead to blindness. Similar losses of function in other regions can lead to paralysis, loss of speech function (aphasia), and even more bizarre syndromes in which, for example, a patient pays attention only to the right half of his perceptual world (hemineglect). Damage to particular portions of the cortex can thus lead to changes in the contents of consciousness.

The thalamus projects fibers from certain of its nuclei in a diffuse fashion to widespread cortical areas. Damage to these nuclei of the thalamus can have even more devastating effects than cortical strokes, including the complete and permanent loss of consciousness, in what has been called a persistent vegetative state. These thalamic nuclei thus appear to be necessary to set the threshold for the activity of the cortical neurons underlying conscious responses.

The thalamocortical system is essential for the integration of brain action across a widely distributed set of brain regions. It is a highly active and dynamic system – and its complex activity, in stimulating and coordinating dispersed populations of neuronal groups, has led to its designation as a dynamic core. The dynamic core is essential for consciousness and for conscious learning.[7] Interactions mainly within the core itself lead to integration of signals, but it also has connections to subcortical regions that are critical for nonconscious activities. It is these regions that enable you, for example, to ride a bicycle without conscious attention after having consciously learned how.

The structures I have mentioned thus far function dynamically by strengthening or weakening the synapses that interconnect them. These changes result in the activation of particular pathways after signals are received from the body, the world, and the brain itself. These dynamics allow the development of perceptual categories in the short term and memory in the long term.

In addition to changes that result from and accompany an individual's behavior, the brain also has inherited value systems selected for and shaped during evolution that constrain particular behaviors. These systems consist of variously located groups of neurons that send ascending axons diffusely into various brain areas. For example, the locus coeruleus consists of several thousand neurons on each side of the brain stem, sending fibers up to the higher brain. Like a leaky garden hose, the fibers release noradrenaline when a salient signal, such as a loud noise, is received. This substance modulates or changes the responses of neurons by changing their thresholds of activity.

Another important value system is known as the dopaminergic system. In situations of reward learning, neurons in this system release dopamine. This compound modulates the response threshold of large numbers of target neurons – for example, those in the cerebral cortex. Without such a value system, the brain would not function efficiently to relate behavior to the need for survival, i.e., to assure adaptive bodily behavior. Notice that 'value' as I discuss it here is not 'category.' While value systems constrain rewards or punishments, an individual's

7 G. M. Edelman, *Wider Than the Sky: The Phenomenal Gift of Consciousness* (New Haven and London: Yale University Press, 2004).

behavior, learning, perception of objects and events, and memory all derive from actions that occur during that individual's lifetime by means of ongoing selection from the brain's vast neuronal repertoires.

A word about the vastness of these repertoires may be in order. Taken together with the intricacy of brain anatomy, the dynamics of synaptic change can give rise to a huge number of possible functional circuits. For example, synaptic change acting on the million billion synapses of the cerebral cortex can provide hyperastronomical numbers of circuits subject to selection during behavior.

THE NEED FOR A BRAIN THEORY: The background for a theory of consciousness that I have presented so far puts a strong emphasis not just on the action of brain regions but also on their interaction. Some scientists have been tempted to speculate in the opposite direction, claiming that there are 'consciousness neurons' or 'consciousness areas' in the brain. It seems to me more fruitful to ask about the interactions among brain regions that are essential for consciousness.

To explain consciousness in biological terms requires a theory of brain action and a linked theory of consciousness, and both must be framed within an evolutionary perspective. To put these theories in such a perspective, it is useful to distinguish between primary consciousness and higher-order consciousness.[8] Primary consciousness (as seen, for example, in monkeys and dogs) is awareness of the present scene. It has no explicit conscious awareness of being conscious, little or no conscious narrative concept of the past and future, and no explicit awareness of a socially constructed self. Higher-order consciousness, which yields these concepts, depends on primary consciousness, but includes semantic capabilities that are possessed by apes, such as chimpanzees, and, in their highest reaches, by humans who have true language.

To simplify matters, let us focus on the evolutionary emergence of primary consciousness. Why do I insist that we base our explanation on an underlying brain theory? One reason stems from the idea that the neural structures underlying consciousness must integrate an enormous variety of inputs and actions. A parsimonious hypothesis assumes that the mechanism of integration of this great diversity of inputs and outputs is central and not multifarious. A contrasting hypothesis would require separate mechanisms for each conscious state – perception, image, feeling, emotion, etc.

What kind of theory can account for the unity in diversity of these states? I have suggested elsewhere that such a theory must rest on Darwin's idea of population thinking applied to individual vertebrate brains. The resultant theory, Neural Darwinism, or the theory of neuronal group selection (TNGS), states that the brain is a selectional system, unlike an instructional system such as a computer.[9] In a selectional system, a repertoire of diverse elements preexists, and inputs then choose the elements that match those inputs. The enormous diversity in the microscopic anatomy of the brain is created by a selectional rule during the brain's development: neurons that fire together wire together. This rule acts epigenetically, i.e., it does not depend primarily on genes. Overlapping this developmental selection is experiential selection: even after brain

8 Edelman, *The Remembered Present*.

9 Edelman, *Bright Air, Brilliant Fire*.

anatomy is developed, the connection strengths at the so-called synapses change as a result of an individual's experience. This alters the dynamic signaling across neuronal pathways. By these means, vast – indeed, hyperastronomical – repertoires of circuits, consisting of neuronal groups or populations, are created, upon which further selection can occur and upon which memory is based. As a result, no two brains are identical in their fine details.

The existence of these repertoires is essential as a basis for the selection of circuits leading to behavior. However, their existence cannot in itself account for the integration of the brain's responses in space and time. For this, a specific anatomically based dynamic feature of higher brains had to evolve. This critical feature is reentry: the recursive signaling between brain regions and maps across massively parallel arrangements of neural fibers called axons. Reentrant activity synchronizes and coordinates the activity of the brain regions linked by these axonal fibers. An outstanding example of such parallel connections is the so-called corpus callosum. This tract consists of millions of axons going in both directions to connect the right and left cerebral cortices. Reentrant activity across such a structure will change with behavior and also act to integrate and synchronize the dynamic activity of firing neurons. This integrative synchronization allows various brain maps to coordinate their activity by selection. No superordinate or executive area is required. This means that different maps of the brain can be functionally segregated – e.g., for sight, audition, touch, etc. – but, nonetheless, can become integrated, as reflected in the unitary scene of primary consciousness.

What might be useful at this point is an image or metaphor to capture how the reentrant thalamocortical system – the dynamic core – binds or integrates the complex activities of the various functionally segregated areas of the cortex in a manner consistent with the unitary scenes of primary consciousness. One such image is that of a densely coupled mass of numerous springs. Disturbance within one region of such a structure will be propagated through the whole structure, but certain of its distributed vibrational states will be integrated and favored over others. Less dense and looser coupling to other springs would correspond to interactions of the core with subcortical brain structures. The main point here is that the myriad interactions in such a densely connected mass will yield certain favored states, integrating various local changes in a more coherent fashion. This is, of course, only a gross mechanical analogy, but I hope it will help provide a grasp of the subtle electrochemical interactions of core neurons mediated by reentry that can yield such a great variety of distinct states.

Reentry is the central organizing principle in selectionistic vertebrate brains. It is of some interest that the underlying structures necessary for dynamic reentry appear to be missing from insect brains. For our purposes, reentry will turn out to provide an essential basis for the evolutionary emergence of consciousness. The implication is clear: animals lacking wide-scale reentrant activity are not expected to be conscious as we are.

A BIOLOGICAL THEORY OF CONSCIOUSNESS: We are now in a position to relate these observations of anatomy and neural dynamics to an analysis of consciousness. As I have suggested, a theory of consciousness based on interactions of the brain, body, and environment

must be grounded in an evolutionary framework.[10] According to the extended TNGS, primary consciousness first appeared several hundred million years ago at the time of the emergence of birds and mammals from their therapsid reptile ancestors. At these junctures, there appears to have been a large increase in the number and types of thalamic nuclei. Even more to the point, new and massive reentrant connectivity appeared among cortical regions responsible for perceptual categorization, and more anterior brain regions mediating value-category memory. This is the memory enabled by selective synaptic plasticity, which is constrained overall by value-system responses to reward or to a lack of reward. The integration achieved by this reentrant system, including the widely distributed thalamic connections, gave rise to unitary conscious or phenomenal experience.

Now we must confront an issue labored over by students of the mind-body problem. How can one relate the integrated firing of the dynamic core to the subjective experience of qualia? The term 'qualia' has been applied narrowly to the warmness of warmth, the greenness of green, etc. In view of the present theory, all conscious experiences – especially the various integrated unitary experiences accompanying core states – are qualia. How can they be explained in neural terms?

The answer harks back to evolution. According to the theory, animals possessing a dynamic core are able to discriminate and distinguish among the myriad interactions of different perceptions, memories, and emotional states.[11] This enormous enhancement of discriminatory capability is of obvious adaptive advantage. Animals lacking a dynamic core can make relatively few discriminations. In contrast, animals possessing primary consciousness can rehearse, plan, and generally increase their chances of survival through their ability to make the vast numbers of discriminations necessary for the planning of behavior.

This provides a key answer to our question concerning the relationship of neural states to qualia. Qualia *are* the discriminations afforded by the various core states. Thus, although each core state is unitary, reflecting integration of its activity, it changes or differentiates to a new state over fractions of a second, depending on outer and inner circumstances and signals. Still, you might ask: how can we connect neural activity to qualitative experience? The answer is that particular dynamic core states faithfully *entail* particular combinations of discriminations or qualia. Core states do not cause qualia any more than the structure of hemoglobin in your blood causes its characteristic spectrum – the quantum mechanical structure *entails* this spectrum. In this view, conscious states are not causal. The underlying brain and core activity is both causal and faithful. This reconciles the theory with physics – no readjustments for spooky forces need to be made to the laws of thermodynamics to account for consciousness.

What I have not emphasized is the relationship of this model of consciousness to the subjective self. Briefly, this relationship depends on the value systems – the agencies of the brain controlling endocrine and movement responses as well as emotions.[12] In the reentrant interactions of the core, the earliest and

10 Edelman, *Wider Than the Sky*.

11 A. R. Damasio, *The Feeling of What Happens* (New York: Harcourt Brace, 1999).

12 Ibid.

most inherent activities of these systems often supersede other inputs. There is, in fetuses as well as in babies and adults, constant proprioceptive and kinesthetic input to the core from the body and limbs. It is inevitable that elements of self-reference arise under these circumstances.

This account provides a background for certain features of higher-order consciousness present in humans. With the emergence of higher-order consciousness, through the evolution of larger brains with a new set of reentrant connections allowing semantic exchange, a socially defined self could appear. Narration of the past and extensive planning of future scenarios became possible. So arose the consciousness of being conscious.

Some find it a retreat to an abhorrent epiphenomenalism to assume that consciousness is not itself causal. But upon reflection, one sees that core processes are faithful ones – so much so that we can speak *as if* our discriminations or qualia are causal. Besides the fidelity of the proposed mechanism, we may point out its universality: all discriminations – whether sensory, abstract, emotional, or fantasy-ridden – are integrated by the same reentrant mechanisms operating in the thalamocortical core. This lays the burden of differences among qualia on their prior neural origins in regions sending inputs to the core. Qualia are different because the neural receptors and circuits for each differ. Touch receptors and circuits differ from visual receptors and circuits, as do neural circuits governing hormonal and movement responses. Each quale is distinguished by its position within the universe of other qualia, and there is, in general, no place for isolated qualia, except perhaps in the linguistic references of philosophers.

We may now encapsulate the picture put forth here.

According to Neural Darwinism, the brain is a selectional system, not an instructional one. As such, it contains vast repertoires of neurons and their connections, giving rise to enormous numbers of dynamic states. Behavior is the result of selection from these diverse states. While the brain responds epigenetically to signals from the body and the world, both in development and in behavior, it also has inherited constraints. These include not only morphological and functional aspects of the body, but also the operation of the brain's value systems. Such structures and systems were selected during evolutionary time. It is the interplay between evolutionary selection and somatic selection that leads to adaptive behavior.

To provide for this behavior, the combinatorial richness and uniqueness of each human brain are coordinated and integrated by the dynamic process of reentry. Indeed, it was the evolution of new reentrant circuitry in the dynamic thalamocortical core that allowed the emergence of the myriad discriminations among successive integrated states, which comprise the process of primary consciousness. The rich combinations of qualia constituting phenomenal experience are precisely these discriminations, which are faithfully entailed by core activity. The possession of primary consciousness allows for the planning of behavior, conferring adaptive advantages on the vertebrate species having this capability.

It is the activity of neuronal groups in the reentrant dynamic core that is causal, for it provides the means for planning adaptive responses. Consciousness as a phenomenal process cannot be causal in the physical world, which is causally closed to anything but the interactions

of matter-energy. Nonetheless, speaking as if conscious states are causal usually mirrors the truly causal core states.

Inasmuch as the set of historic selective events accompanying each individual's development is a function of the unique triadic interactions of body, brain, and world, no two selves or sets of brain states are identical. The privacy and subjectivity of conscious states and selves are an obligate outcome of body-brain interactions. In hominine evolution, a more sophisticated self emerged as a result of social interactions facilitated by the appearance of new reentrant core circuits that permitted the emergence of higher-order consciousness and, ultimately, language. As powerful as this system of higher-order consciousness is, it still depends critically on the operation of primary consciousness. In any event, the proposed reentrant core mechanism is universal, i.e., it applies to all mental states, whether they concern emotions or abstract thoughts.

As a result of higher-order consciousness enhanced by language, humans have concepts of the past, the future, and social identity. These enormously important capabilities derive from the activity of the reentrant dynamic core responding to a multiplicity of inputs from the body and the world, as well as the brain's use of linguistic tokens. The embodiment of mind that results is certainly one of the most remarkable consequences of natural selection.

These considerations provide provisional answers to both the qualia question and the question of mental causation. In this brief compass, I cannot delve deeply into the intentionality question.[13] But the framework I have described posits that consciousness requires reentry between systems of perceptual categorization and systems of memory. Perceptual systems, by their nature, depend upon interactions between the brain and signals from the body and the world. In one sense they are systems of referral. Moreover, memory systems allow the brain to speak to itself, providing a means for referral to what have been called 'inexistent objects,' such as unicorns or zombies. With the emergence of higher-order consciousness and language, intentionality achieves a range that is, for all intents and purposes, limitless.

SIGNIFICANCE: I have described a theory, the testing of which will depend on two factors. The first is the self-consistency of its underlying concepts. The second is the provision of support by experimental means. Clearly, it is important to search for neural correlates of conscious processes. There is already evidence that reentry plays a role in a person's becoming aware of an object.[14] What is required additionally is evidence of how the reentrant activity of the dynamic core changes when a person goes from an unconscious state to a conscious one. And, of course, we should welcome a variety of experiments exploring neural correlates of consciousness in the hope that some unforeseen correlation will either support or change our theoretical views.

For the present, it is useful to ask what consequences this theory would have,

13 John Searle considers intentionality extensively in J. R. Searle, *Consciousness and Language* (Cambridge: Cambridge University Press, 2002).

14 R. Srinivasan, D. P. Russell, G. M. Edelman, and G. Tononi, "Increased Synchronization of Magnetic Responses During Conscious Perception," *Journal of Neuroscience* 19 (1999): 5435–5448.

if we assume it is correct. If the theory holds up, we would no longer have to consider dualism, panpsychism, mysterianism, or spooky forces as explanations of our phenomenal experience. We would have a better view of our place in the world order. Indeed, we would finally be able to corroborate Darwin's view that the brain and mind of man are the outcome of natural selection.

Clearly such a theory, linking body, brain, and environment in terms of conscious responses, would, if correct, be of great use in gaining an understanding of psychiatric and neuropsychological syndromes and diseases. Even in the normal sphere, such a theory might give us a better picture of the bases of human illusions, useful and otherwise.

Tangent to these matters, such a brain-based theory might allow us to obtain a clearer understanding of the connection between the objective descriptions of hard science and the subjective, normative issues that arise in ethics and aesthetics. Theory pursued in this fashion might avoid silly reductionism while helping to undo the divorce between science and the humanities.

Quine, with whose quote this essay began, suggested that epistemology, the theory of knowledge, be naturalized by linking it to empirical science, particularly psychology.[15] His proposal encompassed physics, but restricted itself to sensory receptors, a position he justified by claiming that one could, by this restriction, maintain the extensionality of physics. His position, unfortunately, was allied to philosophical behaviorism, and to that extent it skirted the important issue of consciousness. The present excursions, if validated, are more expansive – they would allow the formulation of a biologically based epistemology, which would include the analysis of intentionality. While remaining consistent with physics, this would represent an accounting of knowledge in terms that relate truth to opinion and belief, as well as thought to emotion. Such an accounting would include aspects of brain-based subjectivity in its analysis of human knowledge. Intrinsic to such a study would be the understanding that knowledge, conscious or unconscious, depends on action in the world.

Finally, one must seriously consider the future possibility of an artificial embodiment of mind: we may someday be able to construct a conscious artifact. Brain-based devices capable of acting in the environment and able to develop conditioned responses and autonomously locate targets already exist.[16] Nonetheless, we are still very far from realizing a conscious artifact. To be sure that we had achieved this would require, I believe, that such a device have the ability to report its phenomenal states while we measured its neural and bodily performance. Would such a device sense the world in ways we cannot imagine? Only the receipt of extraterrestrial messages would exceed this enterprise in excitement.

In the meantime, we can take comfort in the fact that such a device, which will not have our body, will neither destroy nor challenge the uniqueness of our phenomenal experience.

15 W. V. Quine, *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969), chap. 3.

16 J. L. Krichmar and G. M. Edelman, "Machine Psychology: Autonomous Behavior, Perceptual Categorization and Conditioning in a Brain-based Device," *Cerebral Cortex* 12 (2002): 818 – 830.

Arne Öhman

Making sense of emotion: evolution, reason & the brain

We often define the basic goals of human striving in terms of emotion: we yearn for happiness and do our utmost to avoid misery.[1] But making the distinction between positive and negative emotions is not as simple as saying that we seek the former and shun the latter. Emotions often have a will of their own and may resist attempts to be disciplined. Victims of wartime atrocities and natural disasters, for example, may unwillingly suffer from involuntary flashbacks in which they re-experience the trauma, eliciting intense fright that threatens or undermines adjustment. But some individuals – such as journalists, photographers, and Peace Corps workers – are willfully drawn to those very fear-ridden circumstances, not to mention people who find their (sometimes compulsive) joy in activities most of us fear – parachute jumping, mountain climbing, or extreme skiing. Likewise, our lives may become devastated by the prototypical emotion we all desire, passionate love, and we may ruin our health with the delights of food and drink. Still, for most of us, life without emotion would not be worth living. But at the same time, others have regarded emotion as a dark, alien force to which we helplessly succumb, to our own detriment.

Clearly, emotions resist simple interpretation. The purpose of this essay is to discuss the conflicting nature of emotion in light of modern research in psychology and neuroscience. I start with some philosophical considerations that lead to a conceptualization of emotion that ties emotion to the body via evolutionary biology and neuroscience. I then review how contemporary science has addressed some of the classic questions of emotion research.

The conflict-ridden nature of emotion has been evident throughout recorded intellectual history. Almost 2,500 years

Arne Öhman is professor of psychology in the Department of Clinical Neuroscience at Karolinska Institutet. He has published close to two hundred journal articles and book chapters and coedited three books, including "Psychopathology: An Interactional Perspective" (with David Magnusson, 1987) and "The Structure of Emotion: Psychophysiological, Cognitive, and Clinical Aspects" (with Niels Birbaumer, 1993).

© 2006 by the American Academy of Arts & Sciences

1 This essay was completed while the author was a Fellow at the Center for Advanced Study in the Behavioral Sciences in Stanford, California.

ago, at the birth of Greek philosophy, Demokritos said that we need wisdom to cure the mind of emotion the way we need medicine to cure bodily ailments.[2] This idea was central to the Epicurean and Stoic philosophical movements, which predicated their notions of the good life on the insight that we are disturbed not by things themselves but by what we make of them. Reason tells us that we need not fear death because we shall not be there to experience it. We should enjoy food, drink, and intellectual exchange in the context of cultivating friendship. But we should not let emotions associated with insatiable desires for ephemeral things – such as wealth, fame, and power – seduce us. In contrast, the early Christians did not trust the power of reason to control emotion, but made a handful of problematic emotions central to the deadly sins (the committing of which did make death something to fear): avarice, lust, envy, gluttony, indifference, pride, and wrath.

The Stoics made an interesting distinction – between the first and second 'movements' of an emotion. The first movement is reflexive, such as when we instinctively duck for a swooping bird or stop dead when confronted by a snake. The second movement is what we make of this instinctive response: How dangerous is the situation? Will the bird attack again? Is the snake poisonous? This process of evaluation depends on voluntary mental activity. For example, after the initial surge of erotic excitement upon encountering an overwhelmingly attractive potential partner, one might then rationally analyze the situation, which may result in emotional deactivation by shifting one's attention to something less evocative. By making the second movement the essence of emotion, the Stoics changed the meaning of emotion from an automatic and involuntary response to something individuals could consciously control and take responsibility for.

The enigmatic nature of emotion may be one reason science has long neglected it. But there are other reasons as well. The way we normally know emotions is through feelings, which are elusive, capricious, and probably changed by the very act of observing them. Above all, they are observable only in the mind's eye of the emoter. Feelings, therefore, elude science, which aspires for an objective database in which observers can agree on raw data accessible to many observers. Accordingly, some have argued that the subjective nature of feelings excludes them from the realm of science.

However, few deny that they have feelings, and therefore a science of emotion remains incomplete without them. As Jeffrey Gray pointed out, feelings are the raw data of emotion for each of us, which we can use to test theories of emotion in our own mind.[3] Of course, such an exercise does not constitute a science, but it may help achieve one of the goals of science – helping people to understand the world in which they live.

Indeed, emotions *are* observable by an outsider, but only if we reject the notion of feelings as *the* raw data of emotion. The uniquely human ability of language provides a means for people to make their feelings known to the outside world, even though putting words to emotional experience poses challenges

2 This part on the history of emotion is very much inspired by Keith Oatley, *Emotions: A Brief History* (Malden, Mass.: Blackwell Publishing, 2005).

3 Jeffrey Gray, "The Content of Consciousness: A Neuropsychological Conjecture," *Behavioral and Brain Sciences* 18 (1995): 659 – 722.

for the verbal community. As behaviorist pioneers Edward C. Tolman and B. F. Skinner pointed out, language describing emotion is necessarily less precise than language depicting the outside world. Since an object or event in the world is available both to the language learner and the supervising verbal community, the community can reinforce the correct naming of objects and their characteristics. On the other hand, when trying to teach children to talk about their emotions, the verbal community can only interpret a child's body language as indicating fear rather than anger, for instance. Nevertheless, in the end, adults are reasonably good at labeling the emotion they feel, sometimes to the point of providing meaningful quantitative estimates of its intensity.

Evolutionary theory, in contending that humans have specifically evolved the capacity to sense the emotions of others, points to an even stronger argument for the possibility of objectively observing emotion. Evolutionary scientists commonly assume that the pressure of complex social organization catalyzed the rapid enlargement of the human brain during the last million years. Robinson Crusoe, as Nicholas Humphrey once remarked, illustrates this model of human evolution: the real challenge for Crusoe was not to survive alone on the island, but came with man Friday.[4] (Had Monday, Tuesday, Wednesday, and Thursday made their presence known as well, Crusoe really would have been put to the test.) Successful social navigation demanded not only that individuals could recognize many group members but also accurately decode their emotional states, in order to understand, predict, and exploit their actions. One could even claim that we have a special organ that allows (and sometimes impedes) emotion recognition: the face. Darwin himself provided a compelling argument that a primary function of the face is to communicate emotion. Indeed, a substantial research body attests to the fact that humans from diverse cultures are quite adept at distinguishing a set of apparently universal emotions from facial expressions.[5]

We have more than the movement of facial muscles to help us discern the emotional state of a fellow human. In a very real sense, emotions reside in the body, since they mobilize the body's metabolic resources for potentially vigorous action. Many of us have noticed the racing heart, the dry mouth, the cold sweat, and the 'butterflies in the stomach' in anticipation of a fearful encounter. These bodily changes are controlled by the autonomic nervous system, which is primarily responsible for matching metabolic resources to the muscular – and, to some extent, the mental – needs of the body.

Subtler physiological changes indicative of emotion, of which even the emoter may not be consciously aware, may also be readily apparent to observers. A blushing face reveals embarrassment; an opponent's pupils widened in fear can inspire confidence in a combatant; a date's pupil size can also help a person gauge the progress of his seductive efforts. Whereas some physiological changes, such as blushing, are specific to a particular emotion, the majority of them, like pupil size, indicates some unspecific emotional activation. In any case, physiological changes of this type

4 Nicholas Humphrey, *Consciousness Regained: Chapters in the Development of Mind* (Oxford: Oxford University Press, 1983).

5 Paul Ekman, *Emotions Revealed* (New York: Times Books, Henry Holt & Co., 2004).

are relatively easy to measure. In fact, an interdisciplinary field called psychophysiology[6] has developed a body of knowledge recording and interpreting peripheral bodily changes to psychologically meaningful stimuli, including emotional events.[7] Thus, psychophysiology provides one avenue for making emotions objects for scientific scrutiny.

Actions are another good indicator of emotion, since an important function of emotion is to prime and add urgency to action.[8] Thus, we can infer emotion from different aspects of action, both expressive and instrumental. As we have seen, we can detect and interpret emotion from facial responses, the primary example of expressive behavior. But emotion can also charge instrumental action by giving value to stimuli: what we like we will approach, what we dislike we will avoid.[9] We can observe this approach-avoidance dimension at many levels. For example, gaze direction is informative because we tend to look at things we like and avert our eyes from things that we dislike. General posture also gives clues to emotion; fear creates a tense posture, revealing an obvious readiness to escape. Then there is gross locomotion, which modulates the distance between ourselves and surrounding objects (including people). Some approach-avoidance is subtle, such as when we read new e-mails instead of answering the disturbing ones in our folders denoted 'urgent.'

Finally, we are often also aware of the stimulus situation eliciting an emotion, which provides abundant cues for likely emotional reactions and thus places useful constraints on the interpretation of bodily and behavioral responses as well.

In concert, all of these different domains of observation help supply an objective delineation of emotion accessible to scientific study. Furthermore, we can correlate these domains with neural events in brain imaging studies, thus advancing our understanding of the brain mechanisms of emotion.

Connecting emotion to different outputs – verbal reports as well as physiological and behavioral changes – does more than merely provide an operational definition of emotion. It provides a conceptual perspective on emotion that is easy to integrate with psychobiological considerations, which incorporate both evolutionary theory and neuroscience.

First, this scheme stands in opposition to the common notion of an emotion as a unified entity that is isomorphic with the felt emotion. Rather, felt emotion is one of several ways in which an emotion may manifest itself; an emotion is actually a complex reaction composed of several loosely coupled response components, none of which is necessary or sufficient to infer the emotion.[10] This

6 John Cacioppo, Lou Tassinary, and Gary G. Berntson, eds., *Handbook of Psychophysiology* (New York: Cambridge University Press, 2000).

7 Arne Öhman and Stefan Wiens, "On the Automaticity of Autonomic Responses in Emotion: An Evolutionary Perspective," in *Handbook of Affective Sciences*, ed. R. Davidson, K. Scherer, and H. Hill (New York: Oxford University Press, 2003), 256 – 275.

8 Nico Frijda, *The Emotions* (Cambridge: Cambridge University Press, 1986).

9 P. J. Lang, M. M. Bradley, and B. N. Cuthbert, "Emotion, Attention, and the Startle Reflex," *Psychological Review* 97 (1990): 377 – 398.

10 P. J. Lang, "Anxiety: Toward a Psychophysiological Definition," in *Psychiatric Diagnosis: Exploration of Biological Predictors*, ed. H. S. Akiskal and W. L. Webb (New York: Spectrum, 1978).

approach sees emotions as fuzzy concepts, best defined in terms of the degree of overlap with a prototype of a full-blown emotion, which includes an emotional stimulus, a reported feeling, a facial expression, psychophysiological activation, and emotional behavior.[11]

Second, this approach establishes behavioral and psychophysiological links between human and animal emotion, paving the way for an evolutionary analysis of emotion. Because evolutionary analyses center on adaptive function, they offer an interesting perspective on the long-standing belief that emotion undermines wisdom. If we think that cultivating wisdom is the uniquely human approach to bettering our position in the world, then it follows that natural selection must have favored human reasoning ability, and if so, emotion must have assisted, rather than undermined, wisdom. Indeed, a phenomenon as ubiquitous in mammalian life as emotion simply must have an important function; otherwise it would not have survived the natural selection process.

Yet, historically, psychology has been skeptical about emotion not only because of its subjective nature, but because of its questionable functional status as well. In fact, some investigators surmised that the primary effect of emotion was to disorganize behavior. But while we have all been pressed by overwhelming emotion to act stupidly, emotion would not have evolved had disorganizing behavior been its primary function. It would have been unlikely in animals as well, remaining a curious human ability with the obscure purpose of undermining higher cognition.

Using neuropsychological data on the effect of frontal-lobe lesion, Antonio Damasio built a strong case that emotions are critical to humans and human cognition.[12] Persons with lesions in the ventromedial prefrontal cortex (at the bottom of the frontal brain, just above the nasal cavity) show few obvious deficits (as revealed by psychological tests) in functions like perception, attention, memory, and language. Nonetheless, their lives fall apart. Even individuals who functioned at a high level before the lesion destroy their circumstances through a series of ill-advised economic and social decisions.

Damasio reasoned that their decision making had become dissociated from their emotions, which normally serve as "biasing devices" that assist in making decisions. When we are faced with a decision, positive associations (conscious or unconscious) surrounding some choices make them seem more appealing, while negative emotions surrounding others make them more or less impossible to choose. The ventromedial prefrontal cortex provides the interface between the cognitive and the emotional brain by evaluating "somatic markers" that convey information about emotion-related bodily changes. In the absence of the emotional backdrop provided by these somatic markers, the person with a lesion in the ventromedial prefrontal cortex is likely to get stuck pondering a multitude of alternatives, eventually making a bad choice.

This proposal fits into a broader perspective that views emotions as helping establish priorities for action.[13] Significant events in our world elicit different

11 P. Shaver, J. Schwartz, D. Kirson, and C. O'Connor, "Emotion Knowledge: Further Exploration of a Prototype Approach," *Journal of Personality and Social Psychology* 52 (1987): 1061–1086.

12 A. R. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: G. P. Putnam, 1994).

13 Frijda, *The Emotions*.

emotions, and these emotions guide action by highlighting important goals. An approach-avoidance dimension pervades the neural organization controlling goals and their emotional valence.[14] Some goals are negatively defined – that is, they are avoided because they are related to pain, fear, and loss – and others are positively defined because they activate powerful appetitive motivational states.

Evolution has equipped our brains with a system that is activated when we reach valued goals. It extends from the midbrain through the central parts of the brain, which control basic life functions, to the frontal cortex; it is served by dopamine neurons; and it modulates neural activity in large parts of the brain. It is activated by food and water, sexual activity and orgasm, defeating a rival, collecting resources, and so on.[15] Importantly, the neurons of this system are easily conditioned to fire to stimuli that signal reward.[16] It is this system that produces the kick we feel when reaching a goal ("Yeah, I did it, didn't I!"). It is also a system that can be co-opted by chemicals to produce addiction. In fact, all known addictive substances act at various receptor sites of the reward system.[17]

From an evolutionary point of view, this system is a clever device to make organisms honor goals vital to survival and procreation. Thus, we can see emotions as grounded in evolutionarily defined systems – i.e., reward, defense – which make us want to do what our forefathers had to do in order to make sure that their genes were represented in the next generation.[18]

This evolutionary analysis provides an explanation for one aspect of the inherently conflictual nature of emotions alluded to in the opening of this essay: at least the basic emotions operate according to an evolutionary agenda that may differ from our culturally defined agenda. The evolutionary agenda wants, above all, for genes to be propagated; therefore, the central task for humans (and other animals) is to mate, procreate, and take care of offspring. Hence, even though his Victorian imprinting precluded an explicit statement, Darwin would have agreed with Freud that sex is the most powerful source of motivation, bound to generate conflicts within and between the sexes, and within any group of people. Other emotionally charged basic motives that can produce conflict include dominance as well as competition for, and the hoarding of, resources.

From this perspective, an important priority for any culture is to domesticate emotions that are likely to generate conflicts and threaten group cohesion. An essential component of socialization, therefore, is to acquire the ability to regulate emotion. As a result, the success-

14 Lang et al., "Emotion, Attention, and the Startle Reflex," 377 – 398.

15 Trevor Robbins and Barry Everitt, "Motivation and Reward," in L. Squire et al., eds., *Fundamental Neuroscience*, 2nd ed. (San Diego: Academic Press, 2003).

16 W. Schultz, "Multiple Reward Signals in the Brain," *Nature Reviews Neuroscience* 1 (2000): 199 – 207.

17 George Koob, "Drug Reward and Addiction," in Squire et al., eds., *Fundamental Neuroscience*.

18 Arne Öhman, Anders Flykt, and Daniel Lundqvist, "Unconscious Emotion: Evolutionary Perspectives, Psychophysiological Data, and Neuropsychological Mechanisms," in *The Cognitive Neuroscience of Emotion*, ed. R. Lane and L. Nadel (New York: Oxford University Press, 2000), 296 – 327.

fully socialized individual has a cultural self with goals that may differ from (but are likely to be less evocative than) the goals of the evolutionary agenda. Thus, as philosophers realized long ago, we are more or less designed to get stuck on the horn of the dilemma between (culturally defined) reason and emotion.

The modern science of emotion has revived the Stoic distinction between a first and a second movement of emotion. The research literature has extensively documented the first, reflexive motion in particular. Here, Robert Zajonc's pioneering work demonstrated the "mere exposure effect."[19] In this critical experiment, participants were exposed to a set of Chinese ideograms very briefly – so briefly that they could not distinguish the ideograms. They were then shown and asked to rate the stimuli they had already viewed as well as stimuli to which they had not been exposed. Participants rated the previously exposed Chinese ideograms as more likable than similar nonexposed control ideograms.

Joseph LeDoux provided a potential neural underpinning for this effect by demonstrating that stimuli could reach the amygdala, a collection of nuclei in the temporal lobe and the hub of the brain's emotional network, by a direct and fast "low road" that did not pass through the cortex.[20] Through his work with animals, LeDoux argued that this direct route to the amygdala enabled rapid activation of defense in threatening situations.

LeDoux's work showed that emotions could be activated without fully processing the stimulus in the sensory cortices. Consequently, because evolution has been likely to retain this functional organization, it should be possible to evoke emotions in a person even if he or she is unaware of the stimulus. Studies using masking techniques to conceal an emotional stimulus have confirmed this supposition. In these studies, an emotionally evocative stimulus, such as a picture of a snake for individuals who are intensely afraid of snakes, is presented for hundredths of a second and then immediately masked by another stimulus. The experimental subject perceives only the masking stimulus consciously, remaining unaware of the preceding target stimulus. Psychophysiological recordings demonstrate that individuals specifically afraid of snakes show larger responses to the stimuli masking snakes than to the stiumuli masking spiders, and vice versa for individuals specifically afraid of spiders.[21]

Furthermore, humans spontaneously imitate emotional facial expressions (as assessed by electrophysiological measurements). Experimenters using masking stimuli have observed an unconscious imitation response to both angry and happy faces.[22] Without knowing it, we respond to miniscule facial gestures,

19 W. Kunst-Wilson and R. Zajonc, "Affective Discrimination of Stimuli That Cannot Be Recognized," *Science* 207 (1980): 557 – 558; R. Zajonc, "Exposure Effects," in *Feelings and Emotions: The Amsterdam Symposium*, ed. A. Manstead, N. Frijda, and A. Fischer (Cambridge: Cambridge University Press, 2004), 194 – 203.

20 Joseph LeDoux, *The Emotional Brain* (New York: Simon and Schuster, 1996).

21 A. Öhman and J. Soares, "'Unconscious Anxiety': Phobic Responses to Masked Stimuli," *Journal of Abnormal Psychology* 103 (1994): 231 – 240.

22 U. Dimberg, M. Thunberg, and K. Elmehed, "Unconscious Facial Reactions to Emotional Facial Expressions," *Psychological Science* 11 (2000): 86 – 89.

which add emotional color to our social interactions. Whether we feel relaxed or uncomfortable with some people may thus depend on nonconscious emotional cues. This implicit level of emotional give-and-take – which lies hidden behind explicit intentions, interpretations, and verbal statements – plays an important role in determining the outcome of human encounters. In addition, it is open to conditioning: when a mild electric shock to the fingers followed the presentation of a masked angry face, participants subsequently exhibited an elevated physiological response to the angry face when it was presented without the masking stimulus.[23] These findings imply that we not only respond emotionally to stimuli of which we are unaware, but we may also come to imbue such stimuli with a negative emotional valence through nonconscious conditioning.

These studies using peripheral physiological indices of emotion are supplemented by brain-imaging studies.[24] Consistent with the "low road" notion, these brain-imaging studies show that the nonconscious activation of the amygdala by emotional stimuli takes a subcortical route. For example, the amygdala can be activated by visual stimuli presented in a blind cortical field, that is, a lesioned part of the visual cortex that gives rise to blindness in the corresponding visual field.[25]

Essentially, the research reviewed here gives substance to the Stoic conception of a 'first movement' of emotion, which is automatic and reflexive in nature but can still respond to quite complex stimuli. This automatic emotional activation sets the stage for further emotional processing. In a sense, these data demonstrate that emotions are in the body before they are in the mind.

Damasio's somatic-marker hypothesis revives a central idea in the history of emotion, which was independently formulated more than a hundred years ago by Carl Lange, a Danish physiologist, and William James, the famous American philosopher and psychologist.[26] It suggests that feedback from the body's response to emotional circumstances is a central determinant of felt emotion. While Lange emphasized the role of the cardiovascular system as a stimulus source, James included both autonomic and motor responses in his formulation. But they agreed on reverting intuition: we do not cry because we feel sorry; we feel sorry because we cry.

23 F. Esteves, U. Dimberg, C. Parra, and A. Öhman, "Nonconscious Associative Learning: Pavlovian Conditioning of Skin Conductance Responses to Masked Fear-Relevant Facial Stimuli," *Psychophysiology* 31 (1994): 375 – 385; A. Öhman and S. Mineka, "Fear, Phobias and Preparedness: Toward an Evolved Module of Fear and Fear Learning," *Psychological Review* 108 (2001): 483 – 522.

24 K. Carlsson, K. M. Petersson, D. Lundqvist, A. Karlsson, M. Ingvar, and A. Öhman, "Fear and the Amygdala: Manipulation of Awareness Generates Differential Cerebral Responses to Phobic and Fear-Relevant (But Non-Feared) Stimuli," *Emotion* 4 (2004): 340 – 353; J. Morris, A. Öhman, and R. J. Dolan, "Conscious and Unconscious Emotional Learning in the Human Amygdala," *Nature* 393 (1998): 467 – 470.

25 J. S. Morris, B. DeGelder, L. Weiskrantz, and R. J. Dolan, "Differential Extrageniculostriate and Amygdala Responses to Presentation of Emotional Faces in a Cortically Blind Field," *Brain* 124 (2001): 1241 – 1252; A. Pegna, A. Khateb, F. Lazeyras, and M. Seqhier, "Discriminating Emotional Faces Without Primary Visual Cortices Involves the Right Amygdala," *Nature Neuroscience* 8 (2005): 24 – 25.

26 W. James, "What Is An Emotion?" *Mind* 9 (1885): 188 – 205.

Of course, an idea as radically breaking with common sense as this one did not go uncontested. The famous physiologist Walter Cannon launched what was taken as a devastating critique of the James-Lange position.[27] His basic argument was that the physiological activation seen in intense emotion is too crude and too slow to account for the richness and nuance of emotional experience. Indeed, Cannon himself demonstrated that the patterns of physiological responses observed in anger and fear are indistinguishable and that it takes several seconds (sometimes even tens of seconds) for the autonomic response to reach its maximum after an emotional provocation.

However, given what we know today, this critique is not as damaging as commonly thought. Emotional information may reach the amygdala and start activating the bodily response within some ten milliseconds after reaching sensory receptors, and before reaching the adequate cortical areas for identification. The feeling may then take several hundreds of milliseconds to develop. Meanwhile, it is amenable to changes in the stimulus situation. For example, feedback from facial responses is highly patterned and available within a few hundred milliseconds. Furthermore, this feedback remains available even after surgery that blocks information from the body from reaching the brain, which may help explain why animals with such surgery (or humans with spinal cord damage that block feedback from the body) still appear to have emotion – another of Cannon's critiques of the James-Lange theory.

Feedback from the slow autonomic responses may not have to await the full-blown peripheral responses but may start coming in as soon as the relevant brain nuclei are activated. As Damasio pointed out, "as-if body loops" provide simulations of previously experienced 'real' emotional body loops in a compressed time.[28] Thus, the brain may have quite specific information from the body early enough to make it a factor in shaping emotional experience. Indeed, Damasio and his colleagues have shown that simply recalling certain emotional episodes sets off distinct patterns of activity in brain structures that regulate and represent bodily states – patterns that differ between emotions.[29]

Feelings are mental images arising from changes in "neural maps" that represent bodily activations.[30] Experimental data show that the anterior insula, located in the convoluted cortex between the temporal and frontal lobes, was activated when participants "listened for their heart beats," and that this activation correlated with the participants' emotional characteristics.[31] Furthermore, masking studies suggest that the insula is one of the brain areas exclusively correlated with conscious

27 W. Cannon, "The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory," *American Journal of Psychology* 39 (1927): 106 – 124.

28 Antonio Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness* (New York: Harcourt Brace, 1999).

29 Antonio Damasio et al., "Subcortical and Cortical Brain Activity During the Feeling of Self-Generated Emotions," *Nature Neuroscience* 3 (2000): 1049 – 1056.

30 Damasio, *The Feeling of What Happens*.

31 H. D. Critchley et al., "Neural Systems Supporting Interoceptive Awareness: Evidence from Functional and Structural Magnetic Resonance Imaging," *Nature Neuroscience* 7 (2004): 189 – 195.

recognition of emotional stimuli.[32] Thus, there is good reason to associate the insula both with the registration of bodily responses in emotion and with emotional experience itself.[33]

So much for reflexive emotion – emotion is also a matter for reflection. It is obvious that people to a considerable extent construct their emotions. Depending on who we are, we may respond in vastly different ways to the same emotional stimulus. Many investigators have suggested that the perceived meaning of the situation is the central determinant of the emotional response. And emotional meaning, they claim, results from an appraisal process.

Appraisal theory is one of the dominant schools in the psychology of emotion.[34] An influential attempt to insulate the James-Lange idea about necessary bodily input in emotion from Cannon's critique proposed that cognition (i.e., appraisal processes) gives the specific emotional quality to an experience, while physiological activation determines its intensity.[35] For example, running up stairs produces an unquestionable activation of the cardiovascular system. The emotional ramification of this activation, however, is very different if we do it for exercise, to meet a lover waiting at the top of the stairs, or to escape from a maniac chasing us with an axe. These appraisal processes correspond to what the Stoics called the second movement of emotion.

Appraisal theory is too extensive a topic to go into detail here. The general idea is that a series of appraisal processes evaluates emotional stimuli and that the emotion is the outcome of the appraisal. The most basic evaluation is relevance. 'Relevant' in this context typically refers to whether the stimulus has any consequences for one's current goal scenario. If it has no goal relevance, there is no emotion. But if the stimulus has the potential to enhance or impede one's prospects of reaching a valued goal, it will evoke positive or negative emotions, respectively. Goal-congruent stimuli basically induce happiness: if one can attribute the enhanced goal prospects to one's own effort the likely emotion is pride; if they are attributable to another person the likely emotion is mutual affection or gratitude.

To further differentiate negative emotion, the involvement of one's self is crucial. If the event impeding goal prospects damages one's self esteem, the result is anger, particularly if another agent is involved. A threat to one's self results in fear; a loss to self, sadness. In this way, different emotions can be explained in terms of different appraisals. Indeed, the strong assumption is that a unique appraisal lies behind each emotional episode.

Another important dimension of appraisal concerns potential actions: "What can be done about the situation?" Here, controllability and its prerequisite, the stimuli's predictability, are critical: predictable and controllable

32 H. Critchley, C. Mathias, and R. Dolan, "Fear Conditioning in Humans: The Influence of Awareness and Autonomic Arousal on Functional Neuroanatomy," *Neuron* 33 (2002): 653–663; Carlsson et al., "Fear and the Amygdala."

33 A. (Bud) Craig, "Human Feelings: Why Are Some More Aware Than Others?" *Trends in Cognitive Sciences* 8 (2004): 239–241.

34 Klaus Scherer, Angela Schorr, and Tom Johnstone, eds., *Appraisal Processes in Emotion* (New York: Oxford University Press, 2001).

35 S. Schachter and J. Singer, "Cognitive, Social, and Physiological Determinants of Emotional State," *Psychological Review* 69 (1962): 379–399.

adverse stimuli generate less fear, anxiety, and pain than unpredictable and uncontrollable stimuli.

Although we discuss appraisal processes in terms of explicit mental activity, they need not be conscious. While originally conscious, appraisals, particularly immediate ones, may eventually become automatic.

Appraisal processes are also of different importance for different classes of emotion – primary versus complex or secondary emotions.[36] Basic emotions are hardwired and include a handful of universal emotions such as happiness, sadness, fear, and anger – and are triggered more or less automatically by biologically given sign stimuli. Indeed, as we have seen, we need not even consciously perceive these sign stimuli for them to elicit an emotion.

There are two classes of complex emotions. One class of complex emotions – which include attachment, caregiving, sexual desire, jealousy, and social rejection – is 'object oriented.' That is, one cannot consciously experience these emotions unless one is aware of the object. Another class of complex emotions builds on basic or even object-oriented ones, but is cognitively elaborated to reflect cultural and social influences and is predicated on a self-concept. For example, the culturally cultivated fear of nuclear holocaust or terrorism involves the basic emotion of fear, but woven into it is a network of objects – the military, 'the bomb,' 'evil others' – as well as socially determined beliefs about how the world is organized, characteristics of other nations and ethnic groups, and the perceived vulnerability of oneself and the group to which one belongs.

In terms of neural mechanisms, basic emotions primarily depend on the low road to the amygdala, whereas complex emotions require cortical processes. The link between the amygdala and the ventromedial prefrontal cortex is essential for transforming primary emotions into secondary ones.[37] In fact, the frontal cortex is central for the regulation of emotion. In order for study participants to succeed in inhibiting amygdala responses to gory pictures, the upper lateral areas of the frontal cortex, which are associated with executive cognitive control, were activated.[38] Similarly, these areas appeared to inhibit the enhanced amygdala response of 'nonprejudiced' white participants exposed to masked black faces (indicating an implicit racial bias) when the masking interval was extended to allow conscious recognition.[39]

So far the assumption has been that appraisal is a central determinant of emotion. However, the causal chain may be reversed. Phobias are intense, crippling fears of specific objects or situations. Most sufferers agree that their fear is out of proportion to the real danger involved. Nevertheless, when asked to rate the danger conveyed by a set of objects that includes the object of the phobia,

36 P. Johnson-Laird and K. Oatley, "Cognitive and Social Construction in Emotions," in M. Lewis and J. M. Haviland-Jones, eds., *Handbook of Emotions*, 2nd ed. (New York: Guilford Press, 2000), 458 – 475; Damasio, *Descartes' Error*.

37 Damasio, *Descartes' Error*.

38 K. Ochsner, S. Bunge, J. Gross, and J. Gabrieli, "Rethinking Feelings: An fMRI Study of the Cognitive Regulation of Emotion," *Journal of Cognitive Neuroscience* 14 (2002): 1215 – 1229.

39 W. Cunningham, M. Johnson, C. Raye, C. Gatenby, J. Gore, and M. Banaji, "Separable Neural Components in the Processing of Black and White Faces," *Psychological Science* 15 (2004): 806 – 813.

they rate the objects as more dangerous than do nonphobic individuals. This can be taken as evidence that the phobic fear reflects faulty appraisals. Alternatively, however, it may be an effect, rather than a determinant, of the phobic response, an attempt to make sense of, or justify, the irrational fear.[40]

Humans are prone to retrospective justifications. As dramatically stated by V. S. Ramachandran: "Your conscious life, in short, is nothing but an elaborate post-hoc rationalization of things you really do for other reasons."[41] Famous examples of this process were inspired by Leon Festinger's theory of cognitive dissonance, which stated that humans seek balance and consistency in their belief systems. As a consequence, we are motivated to restore balance when there are conflicts between beliefs or between belief and action. For example, when persuaded by shrewd social psychologists to publicly express a view that was inconsistent with their beliefs, research participants were more likely to actually change their beliefs if paid a small rather than a large sum of money. With a big reward, participants could explain away the dissonant action as 'I only did it for the money,' whereas, with a trivial reward, justifying the action was more likely to require a change in conviction.[42]

Similar processes may be at work in emotion. Even though specific stimuli automatically activate emotions, this automatic response often merely sets the constructive mind to work. We feel pressed to understand and to justify our emotions ('the man was so scary, so what could I do but try to escape?' or 'as adorable as she was, I just fell helplessly in love'), and we retrospectively manipulate emotion to justify our action ('I hit him because he made me so mad' or 'I certainly must be madly in love to act this stupidly'). Indeed, one attractive aspect of emotion may be that it provides a sanctuary from the social pressure on humans to make sense.

Largely based on his research on split-brain patients, who have had their two cerebral hemispheres surgically disconnected from each other as a treatment for epilepsy, Michael Gazzaniga argued that the pressure to justify one's actions reflects the operation of "an interpreter system" housed in the left frontal cerebral hemisphere.[43] According to this view, the brain automatically takes care of most of the exigencies raised by the interaction of person and environment. The fundamental interpretive component of the human mind comes in late to make sense of the unfolding scenario mindlessly managed by the brain, to fit it into one's worldview and self-image, and to keep constructing the narrative that we take to be our lives. Unlike all other creatures, humans can, by their access to language, keep a running commentary on their lives. As a consequence, we are prone to mixing up the commentary and the commented-on

40 A. Öhman and S. Wiens, "The Concept of an Evolved Fear Module and Cognitive Theories of Anxiety," in *Feelings and Emotions. The Amsterdam Symposium*, ed. A. Manstead, N. Frijda, and A. Fischer (Cambridge: Cambridge University Press, 2004).

41 V. S. Ramachandran, *A Brief Tour of Human Consciousness* (New York: Pi Press, 2004), 1.

42 L. Festinger and C. Carlsmith, "Cognitive Consequences of Forced Compliance," *Journal of Abnormal and Social Psychology* 58 (1959): 203 – 210.

43 Michael Gazzaniga, *The Mind's Past* (Berkeley: The University of California Press, 1998); Michael Gazzaniga, "Cerebral Specialization and Interhemispheric Communication: Does the Corpus Callosum Enable the Human Condition?" *Brain* 123 (2000): 1293 – 1326.

events in our memories, which may explain the unreliability of our memories.[44] But the commentary is not merely epiphenomenal activity. Rather, it gives consistency to the world and to our actions in it, and it helps us to cope with new situations by time-proven (and largely culturally and socially determined) formulas. In doing its work, the interpreter tries hard to be rational. Indeed, Gazzaniga claims that the interpreter is behind the human adoration of reason.

It appears that science is about to outline the neural geography of the eternal human struggle between emotion and reason inside our brains. Reason appears to reside in the left frontal cortex; its primary opponents, the basic emotions, are located in subcortical nuclei, including the amygdala; and the field of the struggle may be housed in the medial frontal cortex. However, as military analysts know, the geography in which battles take place is an important factor in their outcome, but it is far from a sufficient one. We must know the availability and quality of supporting forces and allies, the weaponry, the morale of the troops, and last but not least, the connectivity among the involved units in order to provide informed guesses about the battle's outcome. Similarly, knowing the location of certain functions in the brain is merely a first step in understanding as complex a phenomenon as emotion. But judging from the progress made during the last decade, there is reason to hope that the collective efforts of philosophers, anthropologists, psychologists, and neuroscientists may eventually pay off in a considerably improved scientific grasp of emotion's enigmas.

44 Elizabeth Loftus, "Planting Misinformation in the Human Mind: A 30-Year Investigation of the Malleability of Memory," *Learning & Memory* 12 (2005): 361–366.

Mark Johnson

Mind incarnate: from Dewey to Damasio

To be a human being requires a functioning human brain, in a living human body, interacting with complex physical, social, and cultural environments, in an ongoing flow of experience. What could be more self-evident than the fact that the human mind is intrinsically incarnate?

And yet, most people do not believe this. Traditional Western philosophical and religious traditions routinely assume the transcendence of mind over body. They assume that our inmost essence is mental and spiritual, which they regard as distinct from the bodily. To live in our culture is to unwittingly soak up the metaphysical mind-body dualism that pervades our commonsense views of cognition, knowledge, language, and values.

Until quite recently, only a handful of intellectually courageous philosophers have outspokenly embraced a nondualistic view of mind and pursued the radical implications of such a view. Baruch Spinoza stands out in this regard, followed much later by Friederich Nietzsche and then the pragmatic naturalists Charles Sanders Peirce, William James, and John Dewey in America, and also the phenomenologist Maurice Merleau-Ponty in France.

Over the past twenty years, the situation in philosophy has begun to change. The terms 'embodied mind' and 'embodied cognition' have become buzzwords in psychology and the other cognitive sciences – and also, increasingly, in philosophy itself. Taking this change seriously is no small matter. If we give up the notion of a transcendent soul and a disembodied mind, then we must give up as well some of our most commonly cherished assumptions about what it means to be human.

Whenever philosophers want to challenge mind-body dualism, they nearly always criticize René Descartes (1596 – 1650) – with good reason. Descartes claimed that reflection on our inner experience demonstrates that bodies are physical substances, extended in space and time, whereas minds are mental substances, having no spatial extension.

Mark Johnson is Knight Professor of Liberal Arts and Sciences in the Department of Philosophy at the University of Oregon. He is the author of "The Body in the Mind" (1987), "Moral Imagination: Implications of Cognitive Science for Ethics" (1993), and, with George Lakoff, "Metaphors We Live By" (1980) and "Philosophy in the Flesh" (1999).

© 2006 by the American Academy of Arts & Sciences

Bodily substance exhibits and supports one set of 'attributes' (e.g., digestion, perception, body movement, locomotion), whereas mental substance supports a quite different set of characteristics (e.g., thinking, willing, reasoning).

The appeal of the idea of disembodied mind – to Descartes and to many people today – appears to be based on three considerations.

First, if the mind exists apart from the body, then life after death would be metaphysically plausible because a 'mind-soul' might be able to survive the death of our fragile human bodies.

Second, mind-body dualism seems to explain how human freedom and moral responsibility might be possible in a physical world governed by cause and effect. If the seat of our moral reasoning and willing lies in nonphysical substance, then, indeed, a part of us (i.e., our moral personality) may not be causally determined and could be the source of free choice and action. This idea underlies the great appeal of Kant's assumption of a transcendent ego – the locus of rational willing that is not subject to the laws of nature governing all phenomenal beings and things. Kant eschewed Cartesian substance dualism, but his notion of the transcendent ego (as a "transcendental unity of apperception") is his substitute for Cartesian mental substance.

Third, our everyday experience appears to confirm the disembodied character of our thinking. We often seem to experience our minds as different from, and even independent of, our bodies. For example, at this very moment, as I write these words, I am going to will myself not to reach over to pick up my cup of tea that calls out to me to take a drink. 'I' must control 'myself,' so it would seem that the 'I' that does the controlling must be different from and independent of the 'self' that is controlled. Our conceptual system and therefore our language incorporate this ostensible dualism.

Merleau-Ponty attributed this apparent experience of disembodied mind partly to the fact that in perception we are not aware of our bodily organs doing the perceiving: "The moment perception comes my body effaces itself before it and never does the perception grasp the body in the act of perceiving."[1] More recently, the American philosopher Drew Leder, in his intriguing book, *The Absent Body*, has catalogued the many ways in which the very nature of our bodily capacities causes us to experience perception and thinking as disembodied. In a chapter on what he calls the "ecstatic body," for example, Leder shows how the structure of bodily perception hides the activity of the organs and processes of perception, as we attend only to what is being perceived and not to the conditions of that perception.

Scientists on the other hand do attend to the conditions of perception – and the growth of cognitive neuroscience over the past twenty years has provoked a revolution in our thinking about mind. Philosophers who have been following the remarkable recent work in neuroscience find the notion of disembodied thought increasingly implausible. For them as for most cognitive scientists, the new mantra is '*No body, never mind*.'

For a dualist like Descartes, a fundamental problem was how mental substance could hook up or interact with mere bodily substance. Descartes was scientifically sophisticated enough to realize that such a connection would

1 Maurice Merleau-Ponty, *The Visible and the Invisible*, trans. Alphonso Lingus (Evanston, Ill.: Northwestern University Press, 1968), 9.

somehow have to occur somewhere in the brain, and he speculated, quite mistakenly, that the pineal gland was the locus of this mental-physical interaction.

For a nondualist, this very 'mind-body problem' is a mistake because it presupposes that there are two distinct entities – body and mind – that must get yoked together. Consequently, the nondualist needs to reframe the problem entirely, asking not how two different metaphysical substances can interact, but rather how characteristics traditionally attributed to mind – the capacity to conceptualize, to understand, to reason, to know, and to will – emerge from physical processes.

The most popular nondualistic approach today is naturalism. To be a naturalist is to explain everything in nature – from the movements and changes of physical objects, to the emergence of living things, to the operations of mind – in terms of natural processes, that is, without reference to anything supernatural that might allegedly enter into and affect nature from beyond nature itself.

I regard American pragmatist philosophy, which came to prominence early in the twentieth century, as the most scientifically and philosophically sophisticated naturalistic, nondualistic approach to mind available to us even today. The pragmatists (especially Peirce, James, and Dewey) appreciated the critical importance of modern evolutionary theory for our understanding of human nature, and they realized that philosophy must grow hand in hand with the best science available. Consequently, the pragmatists gave us a model for how to develop an empirically responsible philosophy of mind.

Pragmatic naturalism starts with the assumption that human beings are natural organisms in ongoing interaction with their environments.[2] In other words, everything we attribute to 'mind' – perceiving, conceptualizing, imagining, reasoning, desiring, willing, dreaming – has emerged (and continues to develop) as part of an ongoing evolutionary process in which organisms seek to survive, grow, and flourish within various environments. As James remarks:

> Mental facts cannot be properly studied apart from the physical environment of which they take cognizance. The great fault of the older rational psychology was to set up the soul as an absolute spiritual being with certain faculties of its own by which the several activities of remembering, imagining, reasoning, and willing, etc. were explained, almost without reference to the peculiarities of the world with which these activities deal. But the richer insight of modern days perceives that our inner faculties are adapted in advance to the features of the world in which we dwell, adapted, I mean, so as to secure our safety and prosperity in its midst.[3]

In James's account, we do not have two entities or substances – body *and* mind – that somehow have to come into relation to each other for a human being to exist. Instead, 'mind' is an emergent process, never separate from body. Thus, experience is a series of purposive bodily activities immersed in the ongoing flow of organism-environment interactions.

Another way of expressing this rootedness of thinking in bodily experience is to say that there is no rupture in experi-

2 Part of the account of pragmatic naturalism that follows is taken, with minor changes, from Mark Johnson and Tim Rohrer, "We are Live Creatures," in *Body, Language, and Mind* (forthcoming).

3 William James, *Psychology (Briefer Course)* (New York: Holt, 1892), 3.

ence between such processes as perceiving, feeling, moving, and thinking. More complex levels of organic functioning are just that – levels – and nothing more, although within each level there arise emergent properties of 'higher' levels of functioning. John Dewey names this connectedness of all cognition the *principle of continuity*, a principle that denies any ontological gaps between various levels of functional complexity. According to Dewey:

> There is no breach of continuity between operations of inquiry and biological operations and physical operations. "Continuity"…means that rational operations *grow out of* organic activities, without being identical with that from which they emerge.[4]

The continuity thesis implies that any explanation of the nature and workings of mind, even of abstract conceptualization and reasoning, must have its basis in an organism's capacities for perception, feeling, object manipulation, and bodily movement. Dewey described at least three primary levels of organization that are relevant to an account of mind. First, there are inanimate material processes (the *physical* level). Second, there are living things that have needs, interests, and satisfactions (the *psycho-physical* level). Third, there are organisms that possess mind (the *mental* level). From this perspective, the problem for the naturalist is to explain how changes in organization and complexity give rise to ever more impressive functional processes, without introducing new ontological entities, structures, or forces. Dewey explains,

> The distinction between physical, psycho-physical, and mental is thus one of levels of increasing complexity and intimacy of interaction among natural events. The idea that matter, life and mind represent separate kinds of Being is a doctrine that springs, as so many philosophic errors have sprung, from a substantiation of eventual functions.[5]

In other words, the error of splitting off 'mind' from 'body' (or the animate from the inanimate, or the mental from the merely living) is a result of treating functional events and processes (Dewey's "eventual functions") as if they were different kinds of beings or entities.

For a naturalist like Dewey then, new organization is responsible for the fact that living organisms (the psycho-physical) have properties and can do things that are not possible for inanimate physical entities and structures:

> In the compound word [psycho-physical], the prefix "psycho" denotes that physical activity has acquired additional properties, those of ability to procure a peculiar kind of interactive support of needs from surrounding media. Psycho-physical does not denote an abrogation of the physico-chemical; nor a peculiar mixture of something physical with something psychical (as a centaur is half man and half horse); it denotes the possession of certain qualities and efficacies not displayed by the inanimate.[6]

Many people who might accept this continuous development from the inanimate to the animate will resist the idea that a similar continuity applies equally

4 John Dewey, *John Dewey, The Later Works, 1925 – 1953*, vol. 12, *Logic: The Theory of Inquiry (1938)* (Carbondale: Southern Illinois University Press, 1991), 26.

5 John Dewey, *John Dewey, The Later Works, 1925 – 1953*, vol. 1, *Experience and Nature (1925)* (Carbondale: Southern Illinois University Press, 1981), 200.

6 Ibid., 195 – 196.

to the emergence of mind. However, his principle of continuity demands that we treat mind not as a *thing*, but as another emerging *process* of interactions. Some organisms develop what we call mind when they achieve levels of functional organization that make communication and shared meaning possible for them, thereby opening up a host of unprecedented possibilities for dealing with the life problems they encounter.

> As life is a character of events in a peculiar condition of organization, and "feeling" is a quality of life-forms marked by complexly mobile and discriminating responses, so "mind" is an added property assumed by a feeling creature, when it reaches that organized interaction with other living creatures which is language, communication.[7]

To say that I have a 'mind' is to say that I am an organism whose potential for very complex interactions has risen to the level where I can share meanings, engage in various modes of inquiry and reasoning, and coordinate activities with other creatures who have minds, using symbols that have meaning for us.

Once we understand that mind is a functional achievement, it ceases to be surprising that mind is always continuous with body and could not exist without body. That is why Dewey always speaks of the "body-mind," and not of body *and* mind. Other philosophers have famously offered their own nondualistic accounts of the interfusion of mind and body. Spinoza avoided Cartesian mind-body substance dualism by arguing that there was but one substance, which he called Nature or God, and that 'body' and 'mind' are simply 'attributes' of that substance. Antonio Damasio's fondness for Spinoza's nondualistic metaphysics stems especially from Spinoza's view of mind as the idea of the human body ("The object of the idea constituting the human Mind is the Body," *Ethics* II, Prop. 13). Damasio shows how this conception of mind is compatible with recent empirical research in the neuroscience of emotions, consciousness, and thought.[8]

I began this essay by boldly proclaiming that acknowledging the embodiment of mind requires us to rethink some of our most cherished assumptions about human nature. Let us consider briefly some of the most significant implications that follow from the conception of the 'body-mind' that I have sketched above.

No mind without a body: Nobody can *prove* indisputably that a disembodied mind or soul cannot exist. However, cognitive neuroscience teaches us that, without certain bodily conditions, functions such as breathing, moving, perceiving, reasoning, feeling, and talking are not possible. So, if there is a 'bodyless' soul that survives after death, we can make no sense of how it could feel, experience, think, or value like we humans do. If you had a disembodied soul, that soul would not be *you*, for it would lack your body, and thus your thoughts, your memories, your feelings, and your emotions. Consequently, the doctrine of embodied cognition has very much a 'this-worldly' orientation – a philosophical perspective grounded in the experiences, thoughts, values, and actions of an intrinsically embodied consciousness that appears to be a tiny part of a sweeping and continual (if somewhat slow) evolutionary process.

7 Ibid., 198.

8 Antonio Damasio, *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain* (New York: Harcourt, Inc., 2003).

Mind is not a thing: Although we are born with many cognitive capacities necessary for human experiencing and thinking, it is a bit misleading to say that we are born 'with a mind,' as though that were some entity or given structure. To 'have a mind' is to rise to the level of being able to sustain a complex ensemble of functions that characteristically involve thinking, deciding, feeling, and communicating with others. When a person ceases to be able to execute these functions, it is fair to say that he has 'lost his mind,' which is not the loss of a *thing*, but rather a failure to sustain a certain dynamic process of higher-level functioning. (This is precisely what happens in cases of dementia.)

Neither is consciousness a fixed thing or a simple property. According to cognitive neuroscientists Gerald Edelman and Giulio Tononi, consciousness is an emergent dynamic unity that results from "a special kind of morphology – the reentrant meshwork of the thalamocortical system – as it interacts with the environment."[9] Consciousness is the temporary achievement of a "dynamic core," in which emerges an integration, within a certain narrow window of time, of various functional neuronal clusters that are highly differentiated functionally.

Body in mind/mind in body: The body is not just the seat of the mind, a mere resting place for a disembodied mind. 'Body' and 'mind' are just different aspects of an ongoing interactional process of experience. Thus, the nature of our human bodies determines both *what* we can experience and think and also *how* we think, that is, how we conceptualize and reason. The body is *in* (that is, working *in*) the mind, just as much as the mind is *in* the body. Damasio states this grounding hypothesis as follows:

> …the body, as represented in the brain, may constitute the indispensable form of reference for the neural processes that we experience as the mind.[10]

> [T]he apparatus of rationality, traditionally presumed to be *neo*cortical, does not seem to work without that of biological regulation, traditionally presumed to be *sub*cortical. Nature appears to have built the apparatus of rationality not just on top of the apparatus of biological regulation, but also *from* it and *with* it.[11]

> The lower levels in the neural edifice or reason are the same ones that regulate the processing of emotions and feelings, along with the body functions necessary for an organism's survival. In turn, these lower levels maintain direct and mutual relationships with virtually every bodily organ, thus placing the body directly within the chain of operations that generate the highest reaches of reasoning, decision making, and, by extension, social behavior and creativity.[12]

The recruitment of sensory-motor capacities to perform concrete and abstract conceptualizing and reasoning, and the crucial role of emotion in reasoning are foundational hypotheses of many contemporary naturalistic theories of mind, thought, and language. The challenge for 'embodied cognition' theories is thus to explain how all of our most marvelous acts of language, communication, abstract conceptualization and reasoning,

9 Gerald Edelman and Giulio Tononi, *A Universe of Consciousness: How Matter Becomes Imagination* (New York: Basic Books, 2000), 216.

10 Antonio Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: G. P. Putnam's Sons, 1994), xvi.

11 Ibid., 128.

12 Ibid., xiii.

and creativity involve the recruiting of sensory-motor functions for 'higher' cognitive functions.

Logic is a matter of body: In commonsense models and in philosophical and mathematical theories alike, logic has virtually always been thought of as the essence of rational thought, thus transcending the body. Like mathematics, it is supposed to be pure (disembodied), universal, and absolute. But if the ways of the body are actually constitutive of what and how we think, then logics (plural) have only as much validity as the shared patterns of bodily experience upon which they rest. Logic doesn't drop down from the heavens of pure reason; rather, it rises up from recurring patterns of embodied inquiry. Already in 1890, James in his *Principles of Psychology* argued that logic is tied to felt relations within bodily experience:

> If there be such things as feelings at all, *then so surely as relations between objects exist in rerum natura, so surely, and more surely, do feelings exist to which these relations are known* ....We ought to say a feeling of *and*, a feeling of *if*, a feeling of *but*, and a feeling of *by* quite as readily as we say a feeling of *blue* or a feeling of *cold*.[13]

A hundred years later, Damasio marshalled clinical and experimental neuroscientific evidence to argue for the role of emotion in certain types of reasoning.[14] Damasio's work has opened the door to a serious reconsideration of James's then seemingly preposterous claim that what we call logic requires an intact and functioning emotional system, and that our bodies play a crucial role in what makes sense to us and how we reason about it.

Language and symbolic interactions are also matters of body: What has come to be known as cognitive linguistics seeks to explain language as a result of many general cognitive capacities acting in consort, rather than as the result of 'autonomous' language modules. Furthermore, embodied approaches to cognitive linguistics present empirical evidence that patterns and processes of sensory-motor experience underlie linguistic meaning and other forms of symbolic interaction. Such evidence includes detailed analyses of how the words we use to talk about mind, and the mental activities of feeling, perceiving, thinking, deciding, and willing, are defined relative to cognitive models that are based either directly on structures of sensory-motor experience or else on systematic conceptual metaphors that are themselves indirectly based on aspects of sensory-motor experience. Within this embodied-meaning framework, George Lakoff and I have presented empirical research from psychology, linguistics, and other cognitive sciences, showing how patterns of sensory-motor experience (e.g., containment, balance, forced motion, iteration, motion along a path, increase/decrease in intensity, and verticality) structure both our concrete and abstract concepts.[15] These image-like patterns of body-based meaning (called *image* schemas) are then metaphorically elaborated to define abstract concepts.

Take, for example, the conventional metaphor, 'understanding is seeing.' Here, we conceptualize the abstract notion of understanding or knowing in

13 William James, *Principles of Psychology*, vol. I (New York: Dover Publications, 1950), 245–246.

14 Damasio, *Descartes' Error*.

15 George Lakoff and Mark Johnson, *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought* (New York: Basic Books, 1999).

terms of our sensory-motor conception of seeing. This metaphorical concept, operating mostly beneath the level of conscious awareness, gives rise to expressions such as 'I *see* what you mean,' 'What you said was quite *illuminating*,' 'She's *blind* to everything I say,' and 'From which *point of view* are you speaking?' In this way, most cognitive linguists seek to explain how patterns of organism-environment coupling and interaction – including perception, manipulation of objects, emotional responses, and body movements – can be the basis for patterns of abstract thought and language.

George Lakoff and Jerome Feldman's Neural Theory of Language (NTL) project carries this embodiment explanation further by trying to construct realistic models of the neural processes that make thought and language possible. They are developing "constrained" or "structured" connectionist neurocomputational models – models that utilize known neural architectures – of the workings of various body-based schemas, images, and concepts.[16] Both cognitive linguistics and the NTL paradigm typically argue that abstract conceptualization is based on metaphorical extensions of body-based concrete concepts and sensory-motor capacities. All of this work on the bodily basis of meaning, imagination, and reasoning is admittedly speculative at the neural level, but the neural models show how it is at least plausible that the mind could work in such a bodily fashion.

The body is more than flesh: In all of these developing accounts, it should be clear that 'the mind' cannot be reduced to 'the brain.' Likewise, 'the body' is never merely a material lump of skin and bones. The *bodily* aspect of 'body-mind' shows itself in many ways. First, there is the body as a physiological organism made of flesh, bones, blood, muscles, nerves, and the many organs of perception and life-maintenance, all organized into complex interactive functional systems. Second, there is the body that the brain and central nervous system permit us to experience – and to monitor and modify as we interact with our environment. Third, and quite importantly, there is 'the body' that does not terminate merely with the fleshy boundary of our skin but rather extends out into its environment, such that organism and environment are not independent, but rather *inter*dependent aspects of the basic flow of (bodily) experience. That is why no account of the body can exclude an explanation of the recurring affordances of the environment – the physical settings, cultural artifacts, institutions, rituals, and shared practices – that give the body its medium for action and determine its meaning for members of that culture. Consequently, scientific and philosophical notions of the body must encompass all of these aspects of embodiment; they cannot limit themselves to our narrow commonsense idea of the body as merely a *thing* consisting of flesh, blood, and bones.

Embodied values: One of the most underdeveloped areas within the embodied-cognition paradigm is the origin and nature of values. In his latest book, *Looking for Spinoza*, Damasio has speculated on where embodied creatures like us get our values. Naturalistic views of mind typically see values as emerging from the needs of organisms to survive, grow, flourish, and, for humans, find meaning within the types of environments they inhabit. Those human environments are at once physical, social, cultural, moral,

16 Terry Regier, *The Human Semantic Potential: Spatial Language and Constrained Connectionism* (Cambridge, Mass.: MIT Press, 1996).

economic, political, gendered, racialized, and spiritual. So, while many of our values are a fairly direct result of our bodies' instincts for survival and growth – we need air, water, food, shelter, warmth, and a host of biological conditions – other values will form as a consequence of our nature as social and political creatures, as gendered animals, and as purpose-seeking beings. It has become evident to those who look carefully at the range and variety of values found throughout cultures and across history that no one set of values can be certified as absolute, universal, and eternal. Although cultures will share many values because of the commonalities of our bodies and the recurring features of the environments we inhabit, *value pluralism* is an inescapable fact of the human condition.

The multidimensionality of the body-mind also explains why no single method or approach could ever capture the workings of mind. We need what Patricia Churchland has called the "co-evolution of theories"[17] – the dialectical collaboration of multiple strategies and methods from many disciplines. We need cognitive neuroscience to study the neurochemical bases of experience, thought, feeling, consciousness, and valuation. We need physiology to explore the whole-body perceptual and motor processes that underlie thought. We need phenomenology to describe the structures and qualities of experience. We need cognitive linguistics, psychology, and anthropology to investigate the bodily schemas and sensory-motor operations that underlie all aspects of cognition and symbolic interaction. We need developmental psychology to provide an account of the emergence of the self, of thought, and of language. We need all the humanistic disciplines to study the meaning humans make through literature, music, dance, and the plastic arts. And we even need philosophers of embodied cognition who try to see how all of these various accounts of embodied mind hang together – and what they tell us about who we are and how we should live.

17 Patricia Churchland, *Neurophilosophy: Toward a Unified Science of the Mind/Brain* (Cambridge, Mass.: MIT Press, 1986).

Carol Gilligan

When the mind leaves the body . . . and returns

I am sitting with a colleague on a platform at the front of a large university lecture hall. We are psychologists teaching in the same department, brought together on this occasion by students who want to hear how we converse. It is a Monday evening in the middle of the term, and the lecture hall is filled. We each speak briefly about our work and then begin the conversation. I notice that when I say "voice," my colleague, who studies cognition and intelligence, responds by saying "the notion of voice" or "the metaphor of voice." I move my chair away from his to signal the gap that has opened between us. The next morning, in class, my students want to talk about what happened. I write the word 'voice' on the blackboard, the sound sibilant in the still, morning air. One after another the students respond: "The notion of voice, the metaphor of voice." We talk about what happens when the body drops out of the conversation.

I am sitting with Sundi at a small table in an empty classroom of her public school. She is eleven, in the sixth grade, and a member of the writing and theater club that meets on Tuesday afternoons, part of a three-year project designed to strengthen healthy resistance and courage in girls. It is spring in the second year of the project, and I am interviewing Sundi. I place a photograph on the table in front of her and ask her to tell a story about what is happening. She stares into the face of the girl in the picture and says the girl has just had a fight with her friend – she is angry and sad. "Where is the anger?" I ask. Sundi replies: "In the pit of her stomach and in her throat." And the sadness? "The sadness is in her heart."

At age nine, Judy says that she knows how her friend will feel because "I just feel it in my mind." When she sees

Carol Gilligan is University Professor at New York University and currently a visiting professor at the University of Cambridge. Best known for her landmark book "In a Different Voice" (1982), Gilligan has since published numerous other books, including "Meeting at the Crossroads" (with Lyn Mikel Brown, 1992), named a New York Times Notable Book of the Year, and "The Birth of Pleasure" (2002). Her adaptation of "The Scarlet Letter" was presented at the 2005 Women Center Stage Festival in New York and will be produced by The Culture Project next year.

© 2006 by the American Academy of Arts & Sciences

someone walking away from her best friend, leaving her alone "just talking into space," she does not infer how her friend will feel or put herself in her friend's place. Instead, she says, "You can just kind of see them walking away or getting sad or something, but you can't tell right then and there she's going to get hurt or anything – but you just feel it. It's hard to explain." There is little language for this emotional connectedness and the knowing to which it gives rise.

By the age of thirteen, however, Judy has learned that knowing and feeling are "two different things." Striving to reconcile this distinction with her experience of knowing through feeling, she divides her mind, which she locates in her gut, from her brain, which is in her head:

> The knowing sort of comes from the brain, like your intelligence part. Like your smartness, your brightness, your education part. And your feeling is something that it doesn't matter if you have an education or not. It's just like something that you can't put into words, that you can't really explain, but it's not, I don't know, it's just like a deeper sort of knowing than intelligence knowing.

In following her disclaimer ("I don't know") by speaking of "a deeper sort of knowing," Judy elaborates a split, not between mind and body but between an embodied mind and a disembodied brain:

> The mind sort of has your real thoughts and a brain sort of has the intelligence . . . what you learn in school . . . but your mind is sort of associated with your heart and your soul and your internal feeling and your real feelings.

Separating her mind – her real thoughts and feelings – from her intelligence and her education, she offers an observation about development: "Children," she says, "have the most mind, but they are starting to lose it actually."

I begin with Judy to illustrate the findings of a five-year study of development involving girls between the ages of seven and eighteen. Prior to this research, adolescent girls, in the words of the 1980 *Handbook of Adolescent Psychology*, had "simply not been much studied." By listening to girls narrate their experiences in coming of age – first in yearly interviews and then in the more intensive writing and theater clubs that met weekly or in week-long sessions over a period of three years – my colleagues and I came to see girls as messengers, like canaries in a mine.[1] They alerted us to a process of initiation that required them to separate their minds from their bodies, their thoughts from their emotions, themselves from their relationships. The initiation entailed a paradoxical sacrifice of relationship for the sake of having 'relationships,' a sacrifice that was at once culturally sanctioned and psychologically incoherent. The resistance of Judy and other girls to making this sacrifice led me to zero in on the question: what happens when the mind leaves the body and returns?

In *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, Antonio Damasio describes core consciousness, or a core sense of self, as our ability to register our experience from moment to moment, like a film running continually inside us, as well as our awareness of watching the film, which extends the sense of self through time and history, leading to memory

1 The Harvard Project on Women's Psychology and Girls' Development began in 1981 and continued for twenty years, expanding in the 1990s to include studies of boys' development and the culture of manhood.

and identity. He contrasts the core self, grounded in the body and in emotion, with what he calls "the autobiographical self," the self that is wedded to a story about itself. I have found this distinction helpful in thinking about dissociation: how we can know and also not know what we know; how it is possible for our experience not to become part of our story.

For example, in the years between nine and thirteen, Judy begins to tell a story about herself that is at odds with what she knows within herself to be true. At the age of ten, she says with pride, "I hardly ever get into fights with my friends because usually we like the exact same things and we do the exact same things." Yet she knows that disagreement is a natural if upsetting part of relationships, integral to the process of rupture and repair. In the absence of the ability to address the inevitable breaks in connection, relationships lose their resilience and become fragile, reduced to sameness, a matter of liking and doing "the exact same things." Judy at ten is aware of a change in her relationships. Heeding an injunction to be nice, she backs off from conflict, fearful that if she says what she is feeling and thinking, people will leave her or "move out."

The impetus to rein herself in also comes from a growing awareness of danger in the world at large and a fear that acting on impulse will lead her to get hurt. Signs of dissociation accompany this appraisal of reality, as Judy begins to have difficulty remembering her experience. The presence of an impediment to accessing what has happened becomes evident when she tells her interviewer about something that sounded fun and exciting, something she recalls but can't quite remember:

> [My friend and I] were deciding whether or not to do something, and, I don't know, it might have been – I guess it was – kind of dangerous because both of us were not sure whether to do it or not. ["*You can't remember what it was?*"] No, I have a short memory. It was recently, too.

The phrase 'I don't know,' spoken four times by Judy at age nine, once after she implied a connection between her brother's anger and her hamster's death, occurs twenty-four times a year later when she is ten, in an interview of comparable length. Rather than an admission of ignorance or an expression of uncertainty about something she has said, it seems more literally to indicate a barrier to speaking – the injunction 'don't' standing between 'I' and 'know.' Associating knowing now only with her intellect, what goes on in her head, Judy begins to talk about, rather than to speak, her feelings – and the ground of experience slips away. As the interview draws to a close, she confesses her sense of a problem: "I don't know what's wrong here; I keep stuttering here. It was tough . . . . I know what the question was, but as soon as you asked me, my mind went blank."

When the mind leaves the body, thought becomes divided from emotion, and we lose an inner compass for navigating the world. Judy registers this loss, but staying in connection with her body now means owning desires that are, at once, exciting and dangerous. Holding thoughts and feelings together, she can read beneath the surface and pick up what is going on around her. She knows that the fight at the dinner table over eating the carrots is not really about the carrots. Yet she fears that saying what she knows would only heighten conflict and lead to trouble.

Except in dreams and flights of fantasy, the mind leaves the body when it be-

comes, for whatever reason, unbearable or untenable to know what in our bodies and our emotions we know. The return of the mind to the body then undoes a dissociation that however adaptive or culturally valued is psychologically problematic. With the approach of adolescence, girls often hover between knowing and not knowing, as if testing the climate they are now entering. Is it possible for them to say what they know without losing relationships and jeopardizing their future? The phrases 'I don't know' and 'you know' rose exponentially in our interview transcripts, implicitly asking: Can I know? Do you know? Do you need me not to know?

When Judy is interviewed at eleven, the lines of her dilemma sharpen. A new framework is evident as she responds to the opening question, "Looking back over the past year, what stands out for you?" with "Well, this was the first year that I started meeting boys, just recently, because someone had a boy/girl party, and I started meeting boys." The parameters of this shift into a different kind of encounter with boys become clear when she is asked about a time she had to make a decision but wasn't sure what she should do. She begins by saying, "Lots of decisions are really simple things," yet the decision she chooses to talk about is anything but simple: "My parents are divorced, and like, next year, I have a choice of which one to live with, and I have just been thinking over that a lot recently. I haven't really decided. I have kind of made up my mind to stay here [with my mom]."

Exploring the choice to stay with her mom, Judy thinks, "I'm getting a really good education, and education means a lot to me now; and I like where I live, where my mom lives . . . and I like the way I am living right now." But she also measures her life against the standard of what she calls "the typical life of a child," by which she means, "just growing up with a regular family, and like, I think I would get a regular family at my dad's because there are two parents; they are a two-parent family with already two kids." Asked how she feels about this issue, she says, "I don't know. I feel like either my mom or my dad will feel bad, whichever decision I make."

She moves to resolve the dilemma by privileging her father's feelings:

> My dad . . . would feel bad, because he would feel like I really didn't want to live with him, but it wouldn't be that big a thing if I left my mom instead of staying with my mom; just the feelings, I think, would be different toward the parent, my parents, and it's a hard decision to make. It's like . . . whatever I do, it depends on my future.

In dismissing her own and her mother's feelings, she minimizes the loss ("It wouldn't be that big a thing") and distances herself by speaking of "the feelings" and "the parent." Her grammatical incoherence ("it depends on my future") mirrors her psychological confusion.

Yet, in fact, she is planning to do what she wants and stay with her mom.

> I am thinking I am pretty much going to stay here . . . I think I'd be just as happy there, but it's hard to explain – I just think like . . . this is what I want to do, so I think just me wanting to do this makes it right, because there is no really wrong answer unless I make it wrong.

The issue then centers on judgment: Does her wanting to live with her mom make it right? Is there a wrong choice in this situation? Or more pointedly, can she avoid making what feels right to her wrong?

Thus, Judy guides us through an initiation into ways of seeing and speaking

about herself and the world that would require her, in the name of morality and for the sake of her future, to dissociate herself from what she wants and knows. In resisting this initiation, she combats internalized voices that call her 'a troublemaker' and 'childish,' that enjoin her to be nice and to have 'a good attitude,' that encourage her to leave the life she wants and values in order to have 'the typical life of a child.' Sent to her room and grounded for getting "really mad" at her mom, she thinks, "If I had just kept my mouth shut and didn't say anything . . . that would have been the end of it." Yet, she says, "in my mind I was still angry."

In fitting herself into a framework that leads her to silence her expression of anger, her honest voice, and her sexuality, Judy internalizes an honor code that divides girls and women into the good and the bad. She worries that in concealing the "bad" parts of herself, she is creating a false impression. She wonders if this is lying, but she decides it is not. She knows that falsity has entered her relationships and wrestles with the question of integrity. She does not want to be "selfish" or "rude," like girls who do not "think about anyone but themselves." Holding herself apart from her relationships, she strives to put herself in the place of others, attending to their feelings by asking herself, "How would I feel?" She wants to be a good rather than a bad girl, to have a good rather than a bad attitude. Yet this entails losing relationship – her connectedness with others and also with vital parts of herself.

The dilemma she faces is one of relationship, and it appears insoluble: either way she anticipates losing relationship, whether by withholding herself from others or being left by them. Thus an initiation, mandating dissociation and enforced through codes and scripts of gender, leads Judy to what psychoanalysts have called a compromise formation. By splitting her mind, which is connected to her heart and her soul, from her brain, which she associates with her intelligence and her education, Judy, in coming of age, is resisting losing her embodied mind.[2]

Tracy, her classmate, reveals how this loss can come to seem inconsequential. "When we were nine, we were stupid," she says. The five-year study has ended, and I have come to ask the girls how they want to be involved now that my colleagues and I are presenting our findings and preparing to publish them in a book. The thirteen-year-olds respond without hesitation: "We want you to tell them everything we said, and we want our names in the book." Tracy then voices her concern that their nine-year-old selves will sound stupid. I say it would never have occurred to me to use the word 'stupid' because what struck me most about them when they were nine was how much they knew. "I mean," Tracy says, "when we were nine, we were honest."

When the mind is forced to leave the body in the name of intelligence and for the sake of education, when thought becomes divorced from emotion as a way of avoiding conflict and trouble, when the self moves out of relationship in order to have 'relationships,' an honest voice – the voice of the core self that registers experience – comes to sound stupid. Thus we become wedded to what within ourselves we know is a false story.

2 A more extensive discussion of Judy as well as additional excerpts from her interview transcripts can be found in Lyn Mikel Brown and Carol Gilligan, *Meeting at the Crossroads: Women's Psychology and Girls' Development* (Cambridge, Mass.: Harvard University Press, 1992).

At a time when research in neurobiology has exposed Descartes' error, when attention to girls and women has revealed a systematic bias in psychological theory, the question arises: how do we come to tell a story about ourselves that is at odds with our human nature, neurologically unfounded, and psychologically untrue?

Damasio has shown that the severing of thought from emotion is a result of brain injury or trauma.[3] Psychologists filming infants with their mothers have discovered that the baby's world is an interpersonal world.[4] As the psychoanalyst Donald Winnicott observed, "There is no such thing as a human baby"; the human infant is a member of a couple. We are, male and female alike, inherently relational, responsive beings, born with a voice and into relationship. The separation of the self from relationships, once considered a milestone of development, is a sign of dissociation. It signals a rift in the psyche, a split in consciousness, a need to shelter parts of ourselves.

Studying girls first led me to recognize that separations long associated with development bear some of the hallmarks of trauma: a loss of voice, gaps in memory, the inability to tell one's story. The privileging of mind over body, thought over emotion, self over relationships reflects a culture that elevates qualities associated with masculinity over those gendered as feminine. What seemed at one time a problematic resistance on the part of girls and women to taking what were considered crucial steps on the road to maturity, leading to rationality and autonomy, now appears as a healthy resistance to psychologically and politically costly losses of voice and relationship – losses that would compromise their ability to love and to function as citizens in a democratic society. Yet this healthy resistance is often met with surprising opposition or force.

What is at stake? Seventeen-year-old Iris, the valedictorian of her high school class, observes, "If I were to say what I was feeling and thinking, no one would want to be with me, my voice would be too loud," and then adds, by way of explanation, "but you have to have relationships." I say, "But if you are not saying what you are feeling and thinking, then where are you in these relationships?" Iris sees the paradox in what she is saying: she has given up relationship in order to have 'relationships,' muting her voice so that 'she' could be with other people. The rewards of this adaptation are clear; the costs, for the moment, less apparent. In Shakespeare's play *The Tempest*, when Miranda asks her father why he is raising a sea-storm, he responds by urging her to sleep: "Here cease more questions," he tells her, "'Tis a good dullness." Later in the play, goddesses arrive to bestow "Honor, riches, marriage, blessing," the gains for entering her father's order. Miranda will preserve that order. When Ferdinand, her husband-to-be, says he would not "for the world" play her false, she tells him, "For a score of kingdoms you should wrangle, and I would call it fair play."[5]

To resist the structures of patriarchy is to challenge long-standing adaptations on the part of both women and men. It means reopening a wound, revivifying a loss, and questioning a sacrifice made for the best of reasons. Yet children's

3 Antonio R. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: G. P. Putnam, 1994).

4 See, for example, Daniel N. Stern, *The Interpersonal World of the Infant* (New York: Basic Books, 1985).

5 William Shakespeare, *The Tempest* (New York: Washington Square Press, 1961).

reluctance to incorporate these structures into their psyches exposes their psychological costs. The initiation into and enforcement of gender splits and hierarchies typically begin at an earlier time in development for boys than for girls, in early childhood rather than at adolescence, and are consequently written more deeply into the stories they tell about themselves. For this reason, girls become informants. At adolescence, they are mature enough to recognize and reflect on what is happening and also more aware of the gap between their sense of themselves and their stories.

But children are children, and as Iris says, you have to have relationships and live in the world. Their resistance to an initiation that is socially rewarded, culturally driven, and yoked to gender – which affects feelings about one's body, one's desires, oneself, and one's future – inevitably becomes embattled, leading to inner conflict, open struggle, and signs of psychological distress. The difference in the timing of boys' and girls' initiation can explain what otherwise appears as a series of psychological puzzles: why does a heightened risk to resiliency set into boys' lives around the ages of five, six, and seven; why do boys often begin at this time to show signs of depression as well as learning and speech disorders and various forms of out-of-control and out-of-touch behavior; why are boys throughout childhood more subject to depression than girls; what protects girls' resiliency until adolescence; why does this hardiness in girls tend to falter at adolescence, when the incidence of depression, eating disorders, and destructive behavior sharply rises? The symptoms themselves – attention disorders, learning disorders, eating disorders, and depression – reflect a disturbance in mind and body; and the loss of voice and relationship, signaling dissociation, leads to behavior problems. To the usual explanations that vacillate between nature and nurture, evolution and socialization, I add a psychological factor. The heightened risk to resiliency reveals a threat to psychological integrity.

In the years before adolescence, girls, in the absence of severe trauma, tend to hold self and relationship together. The preadolescent years are a time of honest voices and shrewd perceptions, recorded across history and culture by artists ranging from Euripides to Toni Morrison. The frank and fearless girls of this age, with their candid voices and open faces, belie stereotypes of femininity. "We have our voices," they tell me. It is this directness, this willingness to say what they see and speak their experience – that older girls and women often come to call stupid or bad or wrong or crazy.

Preschool boys show a similar ability to read the human emotional world, including emotions that are being withheld. Four-year-old Sam asks his mother one day, "Mommy, why are you sad?" When she, wanting to shield him from her sadness, says, "I'm not sad," he tells her, "Mommy, I know you. I was inside you." Five-year-old Nick responds to his father's expression of remorse for having "lost it" and hit Nick the previous day: "You are afraid that if you hit me, when I grow up, I'll hit my children." Alex, Nick's father, had been hit by his father and had vowed to break the cycle. Nick picks up his fear that the pattern now will continue into the next generation.

Yet when manhood is established through a gender binary and through hierarchy, when being a man means not being a woman and also being on top, boys separate their sense of themselves from feelings associated with women –

sadness and fear – and sacrifice love for hierarchy. Covering their vulnerability and calling an emotionally open voice 'babyish,' they become less attentive, less direct, less articulate, and less authentic. They are turning themselves into 'one of the boys' and entering a competitive male hierarchy.[6]

Violent rituals of shaming enforce this initiation, marked by the internalization of and identification with a father's voice or law, and like the initiation of girls at adolescence with its vicious games of inclusion and exclusion, it registers internally in the body as a loss of relationship and of pleasure, a loss quickly covered by a voice that labels what has been lost 'babyish' or 'stupid.' It is a voice that follows dissociation – the voice of a mind split off from the body, of a self divorced from relationship, telling a story about separation that has become linked not with betrayal and trauma but with development and civilization. It is a history written after the fact.

It was her lawless passion that released her from the iron framework of reasoning and enabled her to see the frame.[7] "Is the world then so narrow?" she asks the anguished minister who had been her lover, a man who loved the truth but was living a lie. The Puritan settlement, she observes, was once a forest floor. Built up in one way, it could be torn down and built up anew. She encourages him to leave "these iron men, and their opinions."

6 For additional examples of boys' emotional astuteness and an extended discussion of the responses of their fathers and mothers, see Part II, "Regions of Light," in Carol Gilligan, *The Birth of Pleasure* (New York: Knopf, 2002).

7 Nathaniel Hawthorne, *The Scarlet Letter* (New York: The Modern Library, 2000).

Meeting in the forest, alone for the first time in seven years, their minds return to their bodies. "Do I feel joy again?" the minister asks, amazed at this resurrection of himself. "I seem to have flung myself – sick, sin-stained, and sorrow-blackened – down upon these forest-leaves, and to have risen up all made anew, and with new powers to glorify Him that hath been merciful. This is already the better life!" Hester undoes the clasp that fastened the scarlet letter and throws it among the withered leaves. "She had not known the weight, until she felt the freedom."

> Her sex, her youth, and the whole richness of her beauty, came back from what men call the irrevocable past, and clustered themselves, with her maiden hope, and a happiness before unknown, within the magic circle of this hour . . . . Such was the sympathy of Nature – that wild, heathen Nature of the forest, never subjugated by human law, nor illuminated by higher truth – with the bliss of these two spirits! Love, whether newly born, or aroused from a deathlike slumber, must always create a sunshine, filling the heart so full of radiance, that it overflows upon the outward world.

It was an age, the narrator observes, when "men of the sword had overthrown nobles and kings. Men bolder than these had overthrown and rearranged – not actually but within the sphere of theory, which was their most real abode – the whole system of ancient prejudice, wherewith was linked much of ancient principle." Imbibing this spirit, Hester Prynne, charged with raising a daughter amid a host of difficulties and seeking to cherish and develop "the germ and blossom of womanhood," envisions what seems a hopeless task:

> As a first step, the whole system of society is to be torn down, and built-up anew. Then, the very nature of the opposite sex, or its long hereditary habit, which has become like nature, is to be essentially modified, before woman can be allowed to assume what seems a fair and suitable position [in the new society]. Finally, all other difficulties being obviated, woman cannot take advantage of these preliminary reforms, until she herself shall have undergone a still mightier change; in which, perhaps, the ethereal essence, wherein she has her truest life, will be found to have evaporated.

It is a prospect more daunting than overthrowing nobles and kings because it involves a psychological as well as a political transformation, a change in what has come to seem like human nature.

At the end of the story, Hester, having freed her daughter, returns to Boston to take up, as a radical ministry, her lover's failed mission. She assures the people who come to her for counsel and comfort that

> at some brighter period, when the world should have grown ripe for it, in Heaven's own time, a new truth would be revealed in order to establish the whole relation between man and woman on a surer ground of mutual happiness.

The word 'patriarchy' runs through Nathaniel Hawthorne's astonishing novel *The Scarlet Letter*, which he wrote in a heightened state of emotional openness and turmoil during the six months following his mother's death. As a boy of four, he had seen his mother scorned by his father's family after his father died at sea. Hawthorne places the romance of Hester Prynne and Arthur Dimmesdale in seventeenth-century Boston, and with the exception of Hester, Dimmesdale, Chillingworth, and Pearl, his characters are historical personages. Most of the action takes place in 1649, the year the English king was beheaded.

As Hawthorne reminds us, we are in the vicinity of Anne Hutchinson, in a world riddled with contradiction: the radical Protestant vision of an unmediated relationship with God, who is everywhere and thus able to be worshipped at home or in the forest as well as in church, clashes with the institutionalized power of an all-male, clerical hierarchy; the vision of a democratic society conflicts with the continuation of patriarchal power and privilege. Anne Hutchinson, assuming a direct relationship with God, criticized the ministers' sermons. They convicted her of heresy and insubordination and banished her from Massachusetts.

Summoning moonlight to illuminate an inner landscape, to show the familiar at once "so distinctly . . . yet so unlike a morning or noontide visibility," Hawthorne explores the psychological tensions that reflect and magnify these contradictions. The antagonism between democracy and patriarchy plays out in the lives of women and men as a strain between passion and Puritanism, love and hierarchy. It registers as unhappiness.

Writing in 1850, at the height of abolitionist feminism, Hawthorne, neither an abolitionist nor a feminist, saw into the heart of a problem. A woman must bring the new truth, be "the angel and apostle of the coming revelation." Yet the very passion that releases a woman from the "iron framework" of Puritanism and enables her to envision a new order of living also disables her by leading others to view her as an impure woman. Midway through the novel, however, the framework shifts. As Hester's "nature showed itself warm and rich, a well-spring of human tenderness,

unfailing to every real demand, and inexhaustible by the largest," many people "refused to interpret the scarlet A by its original signification. They said it meant Able; so strong was Hester Prynne, with a woman's strength."

This ability of the mind to reframe the world on the basis of experience poses a far greater threat to orthodoxy than sexual transgression because it reveals that the very terms of the orthodoxy are a human construction, a way of thinking rather than reality. Hawthorne's narrator observes that once Hester's mind, released from "an iron chain," had "assumed a freedom of speculation . . . which our forefathers, had they known of it, would have held [it] to be a deadlier crime than that stigmatized by the scarlet letter." It is not the sexual transgression itself as much as the freeing of a sexual voice, the joining of mind and body, that releases Hester from the psychic imprisonment signified in the novel by the term "goodwife."

Having embroidered her scarlet A with gold thread, Hester recognizes in her unruly daughter the seeds of a noble woman: "The stedfast principles of an unflinching courage, – an uncontrollable will, – a sturdy pride, which might be disciplined into self-respect, – and a bitter scorn of many things which, when examined, have the taint of falsehood in them." Hawthorne was the father of a six-year-old daughter when he wrote *The Scarlet Letter*, and Pearl has always seemed to me more observed than invented. It is seven-year-old Pearl who sees what the Puritans cannot discern: the connection between her mother who wears the scarlet letter and the minister who keeps his hand over his heart.

In exploring the connection between mind and body and the conversion of psychic suffering into physical pain, Hawthorne anticipates Freud by almost a half century. He also illuminates dissociation: in the split names of his two male characters, Dimmesdale and Chillingworth, we see what happens when concerns about privilege and honor lead sensitive and intelligent men to conceal their nature and their worth – they render one dim and the other chilling. Chillingworth, moving stealthily to uncover the truth in the minister's heart, tells him: "You, Sir, of all men whom I have known, are he whose body is the closest conjoined, and imbued, and identified, so to speak, with the spirit whereof it is the instrument." When Dimmesdale interrupts to insist on the separation of his soul from his body, Chillingworth overrides the interruption:

> Thus, a sickness – a sore place, if we may so call it, in your spirit, hath immediately its appropriate manifestation in your bodily frame. Would you, therefore, that your physician heal the bodily evil? How may this be, unless you first lay open to him the wound or trouble in your soul?

In their 1895 *Studies on Hysteria*, Josef Breuer and Freud reported discoveries they had come to in a relatively short time by listening to hysterical women: the intimate connection between our minds and our bodies; the symbolic nature of symptoms; the phenomenon of dissociation and its relation to trauma; and the power of association, the stream of consciousness, and the touch of relationship, to undo dissociation – the power of the talking and listening cure. For Hawthorne as for Breuer and Freud, myself and others studying psychological development, the voices of women and girls have been key to seeing into the psychological and political structures of patriarchy that seem like nature, because they are so closely aligned with manhood and incorporated so early into boys' psyches.

Maybe it is simpler than we have imagined – to recover what we know and free ourselves from a false story. If we are indeed neurologically hardwired to register our experience in our body, then the body holds the clue. And maybe this is another reason why the body – associated with vulnerability, mortality, sexuality, and women – is so readily suspect; why truth becomes aligned with a disembodied intelligence and education; and why associative methods, long relied on by artists, lack credibility, although their power is repeatedly proven. The separation of self from relationships, which seems at once objective and protective, leads us to overlook what otherwise would be self-evident: that one cannot exist without the other. But when a healthy resistance, the psyche's defense of its integrity like the body's immunity against disease, takes on some of the characteristics of a political resistance in its refusal of false authority, dissociation offers a kind of sanctuary: a way of preserving what we know and sheltering ourselves from harm until such time as we are able to confront the problem.

I have taken the 'strengthening healthy resistance and courage' project I began with nine-, ten-, and eleven-year-old girls into a law school classroom, teaching a seminar on sexuality, voice, and resistance with David Richards, a philosopher and constitutional law scholar. The subject of our seminar is the persistence of ethical contradictions in the history of Western democracies along with a tradition of ethical resistance. Over the years of teaching, we have become increasingly aware of the role of the body as touchstone or wellspring, both empirical and democratic, and also of art and the way artists across time and culture have gone into the problem of loss of voice, showing how culture can crush voice and revealing dissociation from bodily experience as a central, crucial problem. The repression of sexual voice thus becomes a key to the structure of dissociation.

Freud saw this in his 1908 essay, "Civilized Sexual Morality and Modern Nervous Illness," when he traced the so-called intellectual inferiority of women to their sexual suppression. Restricted from knowing their sexuality, women are forced to constrain their intelligence, to keep their minds out of their bodies, not to know what in their bodies they know. Sexuality, as Freud discerned in the early days of psychoanalysis, is a nexus of the psychological and the political, a site of repression and a source of resistance. The traumatizing of sexuality, interpreted broadly to mean the dissociation of mind and self from body and relationship, is the *caput Nili*, the source of neurotic suffering in the sense of condemning the psyche to live, as it were, east of desire and knowledge.

Within a political context, the divorcing of sexuality from knowledge leads to what David Richards has called "moral slavery," the imprisonment of an ethical voice that would contest injustice and harm.[8] In the book of Genesis, the word '*da-at*,' referring both to the tree of knowledge and to Adam's knowing of Eve, signifies embodied knowledge – the knowledge that comes through experience; what you know in your bones, in your gut, by heart. This knowledge is forbidden or hidden because it poses a threat to the establishment of hierarchy: God over Adam, Adam over Eve, Eve over the serpent, sorrow over joy. Thus we bind ourselves to a tragic story.

8 David A. J. Richards, *Women, Gays, and the Constitution: The Grounds of Feminism and Gay Rights in Culture and Law* (Chicago: University of Chicago Press, 1998).

The work of our students led me to reflect on what I first saw as an unintended consequence of our teaching: the freeing of a creative voice. The encouragement to overcome dissociations commonly assumed and even valued in the academy – the separation of mind from body, thought from emotion, oneself from one's intellectual work – and the teaching of associative methods, primarily through reading literature but also through writing and theater work, led to papers of exceptional quality. As students breached the restrictions on voice and the inner divisions that had limited their intelligence, as they gained access to experiences they recalled but had forgotten, the range of their knowledge expanded and their insights became surprising.

As a psychologist, I asked how a two-hour seminar that met once a week in a law school could overcome what often were long-standing patterns that inhibited creative work. Reflecting on the pedagogy of the seminar, I have arrived at a provisional explanation: it is the combination of providing a psychological map that directs attention to times in development when an embodied voice becomes muted or silenced, and a political map, dating back to fifth-century Athens, that illuminates the historical costs of dissociation. A psychological puzzle – why do we wed ourselves to a false story? – joins with a political puzzle – how do we come to overlook the obvious? A critical part of our pedagogy then consists of legitimizing associative methods as a means to undo dissociation, and making our classroom a resonant space for a voice connected with the body.

In a forward to a recent edition of *Beloved*, Toni Morrison writes about her experience in leaving her job. While working as an editor at a publishing house, she had written four novels, and writing had become her central work. Yet she was surprised when "a few days after my last day at work, sitting in front of my house on the pier jutting out into the Hudson River, I began to feel an edginess instead of the calm I had expected." There was nothing new or unexpected in any of the problem areas of her life, nothing to explain what was "so unexpectedly troubling on a day that perfect, watching a river that serene." Yet she heard her heart "stomping away in my chest like a colt," felt "this apprehension, even panic" that she knew was different from fear.

> Then it slapped me: I was happy, free in a way I had never been, ever. It was the oddest sensation. Not ecstasy, not satisfaction, not a surfeit of pleasure or accomplishment. It was a purer delight, a rogue anticipation with certainty. Enter *Beloved*.[9]

Morrison reflects, "I think now it was the shock of liberation that drew my thoughts to what 'free' could possibly mean to women." Inevitably, this thought led her to the different history of black women in this country and thus to the writing of her novel.

Perhaps collectively we have responded to the shock of liberation with a similar apprehension, even panic. Perhaps the Puritanism ingrained in American consciousness leads us to be suspicious of joy. But picking up Morrison's question in light of the new truths revealed by neurobiology and developmental psychology, my thoughts lead me to imagine what it could mean to free ourselves, men and women, from a false story about human nature, to release ourselves from the prison of a rewritten history and break a cycle of tragedy and trauma.

9 Toni Morrison, *Beloved* (New York: Vintage International, 2004).

William E. Connolly

Experience & experiment

Social scientists and interpretive theorists of culture have struggled with the 'mind-body problem' since the inception of the human sciences. To emulate the natural sciences as they understand them, many social scientists pursue a predictive science ('in principle') that curtails attention to the creative dimension of culture. Cultural theorists, on the other hand, sometimes minimize the role of biology in human life in order to preserve a space for creativity in thought, emotion, and culture. Even culturalists who study bodily *representations* seldom examine the body as a site of biocultural dispositions and relay point for political mobilization. The anxiety is that to do so would be to play up the importance of genetic determination. In fact, cultural reductionism – that is, the minimization of how biology and culture are always mixed together in human life – threatens to generate the result its practitioners fear. It depreciates the layered character of the body/brain/culture network and thus ignores some aspects of that network implicated in cultural creativity.

The contemporary revolution in neuroscience offers the possibility of opening a new dialogue between advocates of a science of society and those of cultural interpretation. The most promising route, in my judgment, is to forge links between neuroscience – the observational and experimental study of body-brain processes – and phenomenology, understood as the explication of implicit structures of experience that infuse perception, desire, and culture. But what philosophy of mind and body can inform such inquiries without lapsing into either cultural or biological reductionism?

The approach that inspires me is a descendant of Baruch Spinoza's doctrine of parallelism.[1] His philosophy has gone through several modifications by those indebted to him. I will present some of them, trying to make my own position plausible as I proceed.

William E. Connolly is Krieger-Eisenhower Professor at Johns Hopkins University. He has published numerous books on political theory, including "Political Theory and Modernity" (1988), "Why I Am Not A Secularist" (1999), "Neuropolitics: Thinking, Culture, and Speed" (2002), and, most recently, "Pluralism" (2005).

© 2006 by the American Academy of Arts & Sciences

1 See Baruch Spinoza, *Ethics* (with the *Treatise on the Emendation of the Intellect* and *Selected Letters*), trans. Samuel Shirley (Indianapolis: Hackett Publishing Co., 1992).

William E. Connolly on body in mind

Spinoza projects a world of one substance without embracing mind-body reductionism. He asserts that each change of the body is matched by a parallel change of mind (and vice versa), even though neither body nor mind can be understood through the concepts appropriate to the other. There is, rather, one substance with two attributes: extension and ideas. A few formulations in Spinoza suggest that while God possesses the concepts to subsume ideas and extension under one rubric, human beings are capable of knowing that substance is univocal but incapable of understanding bodies and ideas through the same concepts.

In Spinoza's system, efficient causality gives way to immanent causality. The model of efficient causality, in which B is fully separate from A and follows from A in a predictable (in principle) pattern of succession, morphs into one in which new patterns of regularity come into being as 'expressions' of heretofore unrealized possibilities implicit in substance. His system – where for every change in thinking there is a corollary change in bodily state (and vice versa) – thus inspires several recent philosophies – best known through the work of Michel Foucault, Stuart Hampshire, and Gilles Deleuze. They emphasize the importance of techniques, "arts of the self," and "micropolitics" applied to bodies to alter established patterns of thought, judgment, and feeling. Spinoza thus sets the stage for modern encounters between experiment and experience.

Spinoza's theory faces several questions and challenges, however. One is the place of responsibility and freedom in the system. I will bracket that issue here.[2] Another is what it *means* to say that mind and body are parallel. If the process of thinking is incommensurate with the movement of body-brain processes, how could you know that the two attributes run on parallel tracks? Some theorists indebted to Spinoza bypass this issue. Others respond by modifying his claim.

The former response is found in the philosophy of "anomalous monism," a position advanced by Donald Davidson and taken by some to provide a useful updating of Spinoza. The latter response is what I will call "immanent naturalism," a position that emerges from the conjunction of the English philosopher Stuart Hampshire and the French philosopher Gilles Deleuze, two recently deceased thinkers inspired by Spinoza.

Davidson is a monist in one sense. He says that though body and mind belong to one world, the explanation of bodily processes and the interpretation of thought processes differ because they are governed by different aims, roughly to foster prediction in the first case and to clarify meaning and responsibility in the second. The difference in aim enables the two systems to be part of one world while differing in their mode of account. But Davidson's model of physical explanation is too narrow to speak to Spinoza, and it is also not closely attuned to developments in physics, neuroscience, and biology that call into question the sufficiency of the law-like model of explanation.

Davidson says that "if two events are related as cause and effect, there is a strict law under which they may be subsumed . . . . Thus laws must belong to a closed system: whatever can affect the system must be included within it."[3] He

2 I address that question in "Spinoza and Us," *Political Theory* 29 (4) (August 2001): 583 – 595.

3 Donald Davidson, "Donald Davidson," in Samuel Guttenplan, ed., *A Companion to the Philosophy of Mind* (Oxford: Blackwell, 1994), 231.

might have considered the possibility of moving closer to the idea of immanent causality, coming to terms with systems in which resonances and feedback loops roll back and forth between multiple nodes, in which some elements blend into others, and in which a degree of creativity and unpredictability emerges periodically from these patterns.[4] This, at any rate, would move his thinking closer to that of the scientists and philosophers to be engaged here.

Another concern is that while Davidson distinguishes between interpretation (applied to "mental events") and explanation (applied to "physical systems"), and while he is alert to how reflexive understanding by an agent of the causes of its own bodily state can introduce new considerations into future thought and action, he does not consider how techniques applied to the body can alter thought and feeling. Anomalous monism does not experiment with tactics of the body that then find *expression* in thought, feeling, and desire.

Stuart Hampshire surmounts these limitations in Davidson's version of monism. He articulates five themes, all recognizably connected to the Spinozist philosophy they modify.

First, Hampshire contends that while substance is one, we humans have two irreducible perspectives on it, what he sometimes calls the first-person and the third-person perspectives. Second, he says that a change in either body or thought is always correlated with *some* change in the other, even though we cannot specify the exact shape and extent of that change except through live experiments. So he drops the theme of a strict parallel while retaining that of an intrinsic connection. Third, he emphasizes how new findings in neuroscience can, once reviewed by the human objects of inquiry, be folded into their own thinking, informing future capacities of thought and action. Fourth, he asserts that coming to terms with such external knowledge can also prompt the invention of *techniques* and *technologies* to act upon the body/brain network, so as to alter, in turn, future patterns of thought, feeling, and action. (An example of the latter, unavailable when Hampshire wrote, is neurotherapy. Here the subject observes a screen that signals several of his own brain states. The subject then tries to move the signals this way or that according to the instructions of a therapist. If the body/brain patterns are altered in the specified direction over several sessions, a pattern of depression might be lifted or a new mood cultivated.)[5] Fifth, Hampshire presents his (neo-)Spinozism as a defensible yet contestable philosophy: neither it nor mind-body dualism has been demonstrated to date, though an accumulation of evidence and argument may gradually increase its plausibility.

4 See in biology, for example, Lynn Margulis and Dorion Sagan, *What is Life?* (Berkeley: University of California Press, 1995), and Brian Goodwin, *How the Leopard Changed Its Spots: The Evolution of Complexity* (Princeton, N.J.: Princeton University Press, 1994); in paleontology, Stephen Gould, *The Structure of Evolutionary Theory* (Cambridge, Mass.: Harvard University Press, 2002); in neuroscience, Walter J. Freeman, "Consciousness, Intentionality and Causality," in Freeman and Rafael Núñez, eds., *Reclaiming Cognition* (Bowling Green, Ky.: Imprint Academic Press, 1999); and in chemistry, the thinker who inspired several of the above studies, Ilya Prigogine, *The End of Certainty*, trans. Odile Jacob (New York: The Free Press, 1996).

5 For one account see Alondra Oubre, "EEG Neurofeedback for Treating Psychiatric Disorders," *Psychiatric Times*, February 2002, at http://www.neurofeedback-institute.com/cgi-bin/articles.pl.

Here is a formulation, from an essay written before the explosion of new work in neuroscience, in which Hampshire indicates how the plausibility of this philosophy may become enhanced:

> The confirmation, if it comes, will not be like the confirmation of an empirical hypothesis. Rather the confirmation would be that some notions closely resembling Spinoza's key notions become widely accepted as peculiarly appropriate in studying and evaluating human behavior. New psychological knowledge might fit better into this framework than into any other. Certainly anyone who altogether rejects Spinoza's naturalistic standpoint, and anyone who has some religious and transcendental grounds for doing so, would remain unpersuaded, and given his premises, justifiably so. But those of us who have no such transcendental grounds may at least pause and consider the possibility that our habitual moralizing about the ends of action is altogether mistaken. Certainly we should not deceive ourselves by dismissing Spinoza as the kind of determinist who allows no possibility of deliberative self-improvement, as if this were the dividing line between him and traditional moralists. It is not.[6]

Gilles Deleuze, the author of two books on Spinoza, augments Hampshire's contribution.[7] More than Hampshire, Deleuze anticipates modifications in the concept of causality – modifications that neuroscience and allied disciplines are currently exploring. He elaborates a philosophy of immanent naturalism. It is *naturalistic* in refusing to embrace dualism or supranatural forces. It is *immanent* in identifying protean forces – forces that can disturb the "actuality" of relatively stable things, beings, processes, systems, etc. These forces, when activated under the right conditions, periodically introduce, say, a new species, weather system, or human brain/body pattern into the universe. Deleuze thus radicalizes Spinoza's idea of immanent causality and breaks more sharply than Hampshire does with the law-like regularity Davidson attributes to physical systems. Deleuze and his collaborator, Felix Guattari, first, challenge faith in a transcendence (often associated with mind/body dualism) "lodged in the mind of a god, or in the unconscious of life, of the soul, or of language . . . , always inferred." Second, they affirm historically "shifting relations of movement and rest, speed and slowness between unformed elements, or at least between elements that are relatively unformed, molecules and particles of all kinds."[8]

Such a philosophy of energetic "movement and rest" does not reduce the world to chaos. It suggests, rather, that each system – when examined in the timescale appropriate to it – oscillates between periods of relative arrest and heightened imbalance and change, followed in turn by new stabilizations, some of which may assume a composition never fully manifest before. The Nobel Prize–winning chemist Ilya Prigogine summed up a similar thesis more

6 Stuart Hampshire, "Spinoza and the Idea of Freedom," in *Freedom of Mind and Other Essays* (Princeton, N.J.: Princeton University Press, 1971), 203 – 204. The five themes are presented in that essay and its companion piece, "A Kind of Materialism," ibid., 210 – 231.

7 Gilles Deleuze, *Expressionism in Philosophy: Spinoza*, trans. M. Joughin (New York: Zone Books, 1990); and Gilles Deleuze, *Spinoza: Practical Philosophy*, trans. Robert Hurley (San Francisco: City Lights Books, 1970).

8 Gilles Deleuze and Felix Guattari, *A Thousand Plateaus*, trans. Brian Massumi (Minneapolis: University of Minnesota Press, 1987), 266.

briefly, "Our universe is far from equilibrium, nonlinear and full of irreversible processes."[9]

The conjunction of Hampshire and Deleuze suggests the value of translating Hampshire's idea of parallelism and Deleuze's idea of immanent causality into that of emergent causality. According to such a conception, neuroscientists can deploy advanced technologies to observe and alter body/brain processes. Such technologies may well become much more sophisticated and refined in the future. But to date, and perhaps forever, we cannot observe how this complex pattern of entries and multiple feedback loops *blends* layers of past experience into current encounters, carrying both into future action.[10] Observation can thus isolate *body/brain processes* of entry and reentry, but it takes embodied, mobile beings to absorb and catalyze a body/brain/cultural network into specific patterns of thinking, feeling, and judgment. To replicate thinking, a new technology would have to *participate* in these emergent patterns. It would have to become a feeling and thinking agent.

To those who define the physical world as a closed system, terms such as 'blend,' 'absorb,' 'catalyze,' and 'emergence' will seem evasive. But they dramatize a key difference between neo-Spinozists in philosophy and neuroscience, on the one hand, and classical models of science, on the other. For, as we saw earlier, neo-Spinozists play up the role of volatile forces in *both* nonhuman and human worlds. And they reduce the distance between human culture and nature further by locating a capacity of 'self-organization' to varying degrees in several zones of nonhuman nature as well as in human life. Put another way, neo-Spinozists discern an element of real creativity in both nature and culture, inviting exploration of selective affinities between them in a universe in which the future is open to an uncertain degree.

Neo-Spinozists are thus encouraged to examine multiple ways in which culture becomes sedimented into different layers of the body/brain network, to incorporate that knowledge into future thought and action, and to experiment with techniques of body/brain intervention that might find expression in altered patterns of thought, feeling, and judgment. The term 'expression' here means a process of infusion irreducible to efficient causality, partly because of the multiple entries and reentries between different sectors of the body/brain network and partly because of novel capacities of self-organization

9 Ilya Prigogine, *Is Future Given?* (River Edge, N.J.: World Scientific Publishing Company, 2003), 65. Here, and in *The End of Certainty*, Prigogine criticizes the concepts of law and causality in classical science, including some recent interpretations of science. Prigogine and Deleuze are connected by at least two streams: the debt each owes to the work of Henri Bergson on time as alteration, and the debt of Prigogine's collaborator, Isabelle Stengers, to the work of Deleuze. For the latter, see Isabelle Stengers, *Power and Invention: Situating Science*, trans. Paul Bains (Minneapolis: University of Minnesota Press, 1997).

10 In his posthumously published book, *Spinoza and Spinozism* (Oxford: Clarendon Press, 2005), Hampshire seems to take another step toward the position I am attributing to Deleuze. The book consists of his classic 1951 book on Spinoza, the essay on freedom listed above, and a previously unpublished essay, "Spinoza and Spinozism." In the latter, Hampshire expresses a debt to Antonio Damasio, whose book *Looking for Spinoza* (New York: Harcourt, 2003) is the first close engagement by a neuroscientist with the work of Spinoza. I explore this stimulating study in "The Radical Enlightenment: Faith, Power, Theory," *Theory & Event* 7 (3) (2004).

periodically activated as these processes are underway. When the neuroscientist V. S. Ramachandran speaks of "reverberations" rolling back and forth between sectors of the body/brain network, he points to the first phenomenon.[11] When the chemist Ilya Prigogine identifies a capacity for self-organization in relatively simple physical systems, he suggests that this human endowment is shared, though unevenly, with other systems in the world.[12]

Mind and body are intrinsically connected, though the experimental knowledge and experiential capacities of human beings are not fully commensurable. It is through creative movement back and forth among experience, reflection upon it, experimental observation, reflexive awareness of such experiments, and the cautious application of specific techniques to individuals and groups that the most promising and dangerous possibilities emerge. We here note a couple of examples.

Consider lucid dreaming, the process by which people participate in their own dreams, steering them this way or that. How can nonlucid dreamers verify whether such dreams occur in others? What are the effects of lucid dreaming, if and when it occurs? Are there ways to amplify lucidity?

One study, drawing upon high-tech observations and the subtle experience of Buddhist monks, gleaned insight into these questions. It turns out that a skilled dreamer can signal to an observer, by blinking, that he has entered the state of lucidity. EEG measurements, combined with muscle and skin monitors, then make it possible to correlate that signal with the blinker's specific body/brain states. In this experiment, REM movement indicated that dreaming was taking place when the blinking signal was given; and the muscle and skin tone of the dreamers indicated that the dreaming was of a particularly intense type.

The interaction between monks and neuroscientists in this experiment has already generated a new technology to prompt and record such experiences,

> a compact device to help people develop lucid dreaming and remember their dreams better . . . ; a mask worn on the face while sleeping, with a small signaling light so the machine can communicate with the sleeper. The mask is attached to a small computer. Sensors distinguish when the user is in REM sleep, and the computer gives them a gentle signal. The user can then make a conscious effort to be aware of the dream and remember it.[13]

Why study lucid dreaming? And why invent a technology to prompt it? The dreamers, its practitioners claim, tap a latent reserve of compassion in them-

11 V. S. Ramachandran, *Phantoms in the Brain: Probing the Mysteries of the Human Mind* (New York: William Morrow, 1968). He says, "Brain connections are extremely labile and dynamic. Perceptions emerge as a result of reverberations of signals between different levels of the sensory hierarchy, indeed even across different senses." Ibid., 56. A similar approach appears in Gerald Edelman and Giulio Tonino, *A Universe of Consciousness* (New York: Basic Books, 2000). They theorize patterns of entry and reentry in the body/brain network, finding consciousness to be the cumulative result of them.

12 I discuss this dimension of Prigogine's thought in chapter 3 of *Neuropolitics: Thinking, Culture, Speed* (Minneapolis: University of Minnesota Press, 2002). Chapter 4 of that book lists numerous 'techniques of the self' that can be applied to body/brain processes.

13 Francisco J. Varela, ed., *Sleeping, Dreaming and Dying: An Exploration of Consciousness with the Dalai Lama* (Boston: Wisdom Publications, 1997), 106 – 107.

selves, a reserve that then finds expression in future conduct.[14] This effect is also pertinent to the quality of ethical life as understood by both Spinozists and neo-Spinozists. For, as the earlier quotation from Hampshire suggested, we deny that goodness takes the form of obedience to a universal law, as claimed in the dualist traditions of Augustine and Kant. We also contend that command-and-obedience models of morality too often contain within them a drive to revenge against the human condition, finding expression in punitive and accusatory orientations toward the diversity of life. Goodness, to Spinozists and neo-Spinozists, grows out of cultivation of positive attachment to this world in conjunction with reflection into the complexity of specific situations. So, in suggesting techniques through which to amplify care for the future of this world, the engagement between monks and neuroscientists in the study of lucid dreaming speaks to Spinozists.

The second example comes from the lab of Antonio Damasio, a neuroscientist whose work has been influenced by Spinoza. A female patient he names S is highly intelligent, and she is excellent at learning new facts. The hippocampus, the brain nodule that launches the subsystem to lay down new memories, is thus in fine shape. But her positive attitude toward life is never ruffled by past experiences of danger, betrayal, or abuse. Tests to gauge her ability to distinguish between dangerous and benign situations verified this disposition: "It was as if negative emotions such as fear and anger had been removed from her affective vocabulary, allowing the positive emotions to dominate her life." Scanning tests then revealed that "both amygdalae . . . were almost entirely calcified."[15] The amygdala is a little brain nodule that both generates rapid, affective responses of fear on its own and sends signals to more refined brain zones for slower, more complex processing. Using a multidimensional scaling technique, Damasio's colleague showed that "S cannot consistently tell the expression of fear in another person's face . . . , she has no problem with the recognition of other facial expressions."[16]

S's case suggests how well-developed the neuroscientist's ability is to establish correlations between observed body/brain states and the quality of lived experience, even though the observer cannot report actual experience without help from the client. It also reveals how much of perception and judgment is prior to consciousness. In another study, for instance, skin-conductance tests of the subjects revealed the ability to distinguish between favorable and unfavorable situations before a conscious judgment registered.[17] Those of us with fluent relays between the amygdala, skin, and other brain regions thus make preliminary affect-imbued discriminations

14 The instruments are limited in what they detect. "If you look at a person's EEG, you cannot tell if he or she is full of compassion or completely oblivious." Ibid., 104. This book provides a model of how the interplay between experience and experiment can work. As a researcher says, "In statistical analyses, we found that there is more body movement in lucid dreams, and more sound. Body balance seems to be very important. All this leads us to ask whether there are psychological, cognitive predispositions to lucid dreaming. It turns out that there are, notably in the realm of spatial skills such as body balance." Ibid., 105.

15 Antonio Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness* (New York: Harcourt Brace, 1999), 65, 62.

16 Ibid., 65.

17 Ibid., 301–302.

before a *feeling* of fear or annoyance floods over us.

Mothers, Spinozists, Freudians, Buddhist monks, Christian preachers, phenomenologists, novelists, filmmakers, advertisers, and charismatic leaders have intervened in nonconscious processing for centuries. But today more systematic knowledge of body/brain processes attracts corporate advertisers and political consultants as well. Robert Heath, a highly successful advertiser working in England, draws upon recent work in neuroscience to improve the effectiveness of TV ads. The most successful ads, he says, are "low involvement ads," in which the higher reflective capacities of the viewers are either placed on hold or diverted to a side issue, allowing nonconscious processing to take over. The advantage of nonconscious processing is that "it is on all the time"; it is also automatic and inexhaustible in its capacity and more durable in retentive power.[18] Such ads can trigger intensive charges of thought-imbued affect that flow into a consumer's perception and judgment even before they become explicit.

Political leaders, talk show hosts, and product advertisers seek to mobilize such nonconscious patterns of resonance across large constituencies and to encourage the results to flow into consciousness. Some of those patterns demean particular groups and instill consumption demands ill suited to the health of consumers or the collective future.[19] A major contemporary challenge is to devise ways to expose and respond to such technologies of collective mobilization.

What insights can neo-Spinozism offer? It is important, certainly, to publicize how such strategies work, drawing upon studies in neuroscience, advertising, political campaigns, and phenomenology to do so. But it is also critical *to devise countertechniques of cultural-corporeal infusion*, tactics that work upon individuals and constituencies at the visceral level as they also engage the higher intellectual registers. This is dangerous territory. But it is also unavoidable territory in a media-rich world, in which there is never a vacuum in the micropolitics of corporeal-cultural infusion.

How can you participate in such strategies without becoming an envoy of cultural manipulation? I support a three-tiered strategy: you expose the tactics of those who do not themselves call attention to them; you introduce counterstrategies of cultural-corporeal infusion attached to a more generous vision of public life; and you publicize, as you proceed, how these counterstrategies themselves impinge upon the affectively rich, nonconscious layers of life.

The way in which Stephen Colbert and Jon Stewart mimic and exaggerate the orchestration of image, voice, music, sound, and rhythm by media stars such as Bill O'Reilly provides one starting point. They do not simply expose factual misstatements – an inadequate response to influences exerted in part upon affective states situated below the refined intellect. Instead, they fight fire with fire, reenacting media strategies of inculcation by parodying them. Clearly, however, much more thought and experiment is needed in order to both expose and respond to the media tactics that attempt to code the visceral register of affect-imbued judgment.

18 Robert Heath, *The Hidden Power of Advertising* (Henley-on-Thames: Admap Publications, 2001), 67.

19 I examine a contemporary instance in "The Evangelical-Capitalist Resonance Machine," *Political Theory* 33 (6) (December 2005): 869–886.

Neither the arguments nor examples provided here prove the truth of neo-Spinozism. I join Hampshire in doubting that such a definitive proof is apt to emerge, either for the neo-Spinozist vision, or for its dualist and reductionist competitors. Nonetheless, as new experiments in neuroscience are linked to reflection on cultural experience, the plausibility of neo-Spinozism may be enhanced.

Two avenues seem particularly promising to pursue. The first is to place neuroscience and phenomenology into closer communication. In his late work, *Nature*, Maurice Merleau-Ponty moved close to neo-Spinozism. If and as we absorb his experiential accounts of the interinvolvement of the senses, the essential role of the body's implicit self-image in perception, the intersubjective dimension of experience, and the layering of bodily dispositions, our awareness of the imbrications between body/brain observation and lived experience may become more supple.[20]

The second avenue of inquiry is related to the first. We should experiment cautiously with bodily techniques that then find expression in thought and feeling. This is particularly pertinent to intellectuals who embrace a neo-Spinozist image of ethics. Such strategies might include visualization, priming dreams by reviewing a perplexing issue before going to sleep, lucid dreaming, meditation, and neurotherapy. As we move back and forth among experiential awareness, media studies, knowledge of body/brain processes, and subtle technologies of body/brain intervention, we may also gain more insight into how to confront and counteract the politics of cultural revenge that exerts so much power in the media and elsewhere today.

20 See Maurice Merleau-Ponty, *Nature: Course Notes from the College de France*, trans. Robert Vallier (Evanston, Ill.: Northwestern University Press, 2003). His classic *Phenomenology of Perception*, trans. Colin Smith (New York: Routledge, 1989), should be read in conjunction with that text.

Jacques d'Amboise

The mind in dance

At eighteen years of age, I had already been a member of the New York City Ballet for three years and had just been made a principal dancer. While performing at the ballet company's home, the City Center of Music and Drama on 56th Street in New York City, I developed a system that I believe has maximized and improved the quality of my performances.

Early in the morning or on my days off, I sit in the empty auditorium, gazing at the stage. I am envisioning a variation from my repertoire, imagining, in detail, first how I will look in costume, then how I will enter the stage and from which wing. As if watching a movie, I then dance the variation in my mind the very best that I can, or even better – the leaps a foot higher, the space covered double what I have done in the past. I picture the expression on my face, the use of my arms and hands, and the speed at which I move. A dream of the possible, glorified, runs on an imaginary loop through my mind, sometimes in slow motion, sometimes accelerated.

At first, I run this imaginary film to rhythmic counting alone (without music, melody, theme, harmony, etc.) – creating a blueprint of mathematical time. For example, I launch into a leap on the first count (or beat), float through the second and third counts, and land noiselessly on the fourth. Next, I rerun these movements, adding, in my head, the melody of the music in place of the counts. Each of these processes I repeat multiple times.

Now I am ready to make the imagined concrete. Up on the stage, I rehearse what I have envisioned – step by step, count by count, without music, over and over again. Sometimes I spend as much as two hours on a dance sequence that is perhaps one-and-a-half minutes long. During these repetitions, I count the

---

Jacques d'Amboise is a well-known dancer and choreographer. He became a soloist with the New York City Ballet in 1953. Best known for his roles in works such as "Filling Station," "Stars and Stripes," and "Apollo," d'Amboise also danced in several movies, including "Seven Brides for Seven Brothers" (1954) and "Carousel" (1956). His own ballets include "The Chase" (1963), "Quatuor" (1964), and "Irish Fantasy" (1964). He also taught at the School of American Ballet and the State University of New York. In 1976, d'Amboise founded the National Dance Institute, a nonprofit organization that introduces the arts into public schools, using dance as a catalyst.

© 2006 by the American Academy of Arts & Sciences

beats out loud as I dance, even rehearsing how I will breathe. I also practice the dance movements in three different tempos: slow motion, ideal, and accelerated (in case the orchestra conductor has an adrenaline rush during the performance). I am now prepared to handle any tempo that may emanate from the orchestra pit.

To end my practice session, I dance the entire variation, singing the melody as though it were an aria. Sometimes as I dance, I speak out loud to an imaginary audience. I comment on what I am doing and sell it to them: "Watch this! Did you like that? Here comes the biggest leap!"

Many years after creating this process, I read about athletes who demonstrably enhanced their performances using visualization techniques. Scientific articles on the mind-body connection confirmed my own experiences. Similar anecdotes from fellow artists further corroborated my belief in the connection between the body and the mind. Several times, the ballerina Suzanne Farrell described how she would lie in a warm bath and mentally rehearse a ballet. Conrad Ludlow, a principal dancer with the New York City Ballet, would put on his makeup while wearing a wool skullcap. "Why, Conrad, are you putting your makeup on with a wool hat?" His reply: "I'm warming up my brain. While I'm using the time putting on my makeup, I imagine my body doing my warm-up. That way, when I actually have to do them, I don't have to do as many – or maybe, if I'm lucky, I won't have to do them at all!" I laughed and thought him eccentric. But he was not alone: Rudolph Nureyev would also wear a wool knit cap as he led barre exercises. "It all comes from the brain to the feet," he told me. "If the brain is warm, it gets to the feet faster."

Kay Gayner, a National Dance Institute teacher and a performing artist in the fields of dance, theater, and music, once described the exercises she does in drama class: "In order to create a character, we program pictures in the mind and then trust that the body will respond. You don't have to do anything; just trust your mental pictures." Her teacher, John Osbourne Hughes, also believes that physical actions are always a direct result of mental images or impressions. If you can create, prepare, and store mental images from, say, Lady Macbeth's life, you will begin to walk, sound, and look like her, and even think her thoughts.

After thirty-five years and many thousands of performances, I can recall a handful of moments where I experienced a distortion of time and space. Time slowed down markedly – it seemed that I could create a freeze-frame of the moment – and I became an observer, detached, watching and enjoying every moment unfold as it was happening. The dancer on stage was my Siamese twin, parted momentarily – one of us watching from somewhere above the audience; the other, on stage, in perfect control of the timing of every movement and gesture in space. I believe this phenomenon activates the same brain center or centers as when one is under great stress – during a car accident, when you watch your every action in slow motion; battle; or even death, when some people have described leaving their bodies and watching themselves die. It is a function of the brain under stress – at times enjoyable, and at times not.

Oh, what a beautiful thing the human being is, and how extraordinary the mind! It is not divorced from the body. Rather, they are entwined, as a successful marriage, a union, and a bond.

Ray Dolan

The body in the brain

In J. M. Coetzee's novel *Elizabeth Costello*, the protagonist, fictional author Elizabeth Costello, delivers a prize lecture entitled "The Lives of Animals." Costello contrasts reason, exemplified by the Descartian credo *Cogito Ergo Sum*, with what she describes as

> fullness, embodiedness, the sensation of being – not a consciousness of yourself as a kind of ghostly reasoning machine thinking thoughts, but on the contrary the sensation – a heavily affective sensation – of being a body with limbs that have extension in space, of being alive to the world. This fullness contrasts starkly with Descartes' key state, which has an empty feel to it: the feel of a pea rattling around in a shell.

Here, Costello radically asserts that consciousness is multilayered – and that a fundamental layer is 'embodiedness,' the representation of our body in our brains. She later points out that embodiedness provides a sentient bridge between human consciousness and that of animals.

In making the seemingly paradoxical claim that bodily and mental states are intimately conjoined, Coetzee, via his fictional alter ego, raises a provocative question, one that a neuroscientific analysis, let alone a definition, of consciousness has yet to address. His account of consciousness echoes a similar challenge William Golding made in *The Inheritors*, which extensively documents the mental lives of Neanderthals – where thought and felt experience, engendered by perturbed bodily states, are indistinguishable.

I cannot claim to rise completely to the challenge of explaining the continuities between representations of the corporeal self and the conscious self. Nonetheless, I will approach it, starting by accepting the general premise, articulated most eloquently by Antonio Damasio, that the brain's representations of bodily states are fundamental to our ongoing mental experience.[1] As Damasio stated, "*We only know that we feel an emo-*

Ray Dolan is a professor at and director of the Wellcome Department of Imaging Neuroscience in London, England. The author of numerous articles in cognitive neuroscience, he is also the recipient of the Alexander von Humboldt international research award.

© 2006 by the American Academy of Arts & Sciences

1 Antonio Damasio, *Descartes' Error* (New York: G. P. Putnam, 1994).

tion when we sense that emotion is sensed as happening in our organism."[2] In this essay, I will address what we currently know about how our brains generate and, in turn, remap one component of bodily states, specifically those mediated by the autonomic nervous system. My choice of the autonomic system reflects not only my own scientific interest but also the fact that the autonomic nervous system provides the most dynamic mapping between the body and brain and vice versa.

Homeostatic regulation is fundamental to the physiological organization of complex organisms. Examples of homeostatic control include body temperature, acid (pH), and salt regulation. All physiological systems operate within narrow constraints that must be continuously maintained to ensure the viability of an organism. In turn, any perturbation in bodily state, such as altered cardiovascular status, invokes a need for homeostatic control. This homeostatic control is largely accomplished via multiple feedback loops that include contributions from specialized vestibular, proprioceptive, and pain afferent signals. Information from these diverse systems is integrated within higher brain centers to provide an ongoing, continually updated image of the state of the body in the brain.

The importance of maintaining the internal milieu has been a central organizing principle in physiology since the time of Claude Bernard (1813 – 1878), who noted,

> So far from the higher animal being indifferent to the external world, it is on the contrary in a precise and informed relation with it, in such a way that its equilibrium results from a continuous and delicate compensation, established as by the most sensitive of balances.

Among the body's homeostatic systems, the autonomic nervous system is the most flexible and dynamic. Output signals to the body, generated in the brain's central autonomic effector sites, give the body a continuously updated picture of the motivational state of the brain. In turn, afferent signals from the body provide the brain with an updated representation of the current state of the body. In concert, these regulatory signals integrate bodily states with short- and long-term motivational needs, which depend on the current and anticipated future environments. My prime focus is to account for how such autonomic bodily states, which reflect a dynamic signaling between body and brain, are generated and, in turn, remapped within the brain.

Anatomically and functionally there are two segregated components to the autonomic nervous system: the sympathetic and parasympathetic subdivisions. Among its key functions sympathetic activity facilitates motor action, increasing cardiac output and reducing blood supply to the gut. In contrast, parasympathetic activity promotes more recuperative functions such as reducing heart rate, lowering blood pressure, and slowing gut motility. Consequently, bodily states associated with survival behaviors (e.g., the fight-or-flight response) are characterized by increased sympathetic activity and, in most instances, decreased parasympathetic activity. It is also notable that many autonomic responses, which tend to be outside one's conscious control (for example, sweating, piloerection, and vasomotor change such as blushing), contribute to the emotional expression and subtle

2 Antonio Damasio, *The Feeling of What Happens* (New York: Harcourt Brace, 1999).

social signaling that guide human interaction.

Autonomic output is continuously adjusted to reflect our physical, cognitive, and motivational needs with respect to our environment. This adjustment entails homeostatic control, which, in turn, requires feedback from the body to the brain's central effector-control mechanisms. Central homeostatic autonomic control is supported primarily by a functional organization of deep brain nuclei within structures such as the hypothalamus, pons, and medulla. Posterior and lateral hypothalamic nuclei influence sympathetic function via brain-stem centers such as the tegmentum; raphé nuclei; periaqueductal grey; and paraventricular, parabrachial, and medial reticular nuclei. Meanwhile, the anterior hypothalamus influences parasympathetic efferent responses via medullary nuclei such as the nucleus ambiguus and dorsal motor nucleus of the vagus, Edinger-Westphal nucleus, and salivatory nucleus. Regions such as the central nucleus of the amygdala, as well as the bed nucleus of the stria terminalis and locus coeruleus, also contribute to autonomic control, as they project directly to brain-stem autonomic nuclei and sympathetic cell bodies in the spinal cord. At the level of the brain stem, afferent feedback of visceral information is represented in the nucleus of the solitary tract and in hypothalamic nuclei that lie in close proximity to these efferent autonomic centers.

The default mode for autonomic control of bodily states is autoregulatory. But higher brain centers supporting cognitive functions, such as memory and mental imagery, also modulate lower autoregulatory autonomic control centers. One can gather this from introspective observation alone. For example, recollecting a threatening event from the past can induce changes in respiration, skin conductance, and heart rate despite the absence of any overt change in the demands upon our bodies. The influence of higher centers is also evident in anatomical and physiological observations, whereby stimulation of distinct cortical and subcortical regions can evoke autonomic responses. These areas of stimulation include the insula (implicated in somatic and visceral representations), the motor cortex, the neostriatum and cerebellum (involved in initiation and control of limb movements), the amygdala and hippocampus (involved in emotion perception, threat responses, and episodic memory), and the anterior cingulate and ventromedial prefrontal cortices (implicated in attention, motivation, and decision making). Thus, there appears to be an obligatory yoking of autonomic regulatory control in lower brain regions to functions mediated by higher brain centers.

A contemporary approach to understanding the relationship between bodily and brain states involves using brain-scanning techniques, such as functional magnetic resonance imaging (fMRI), to measure brain activity noninvasively, close to real time. By scanning the brains of subjects performing specific mental or physical tasks that alter cardiovascular autonomic responses (e.g., increases in heart rate and blood pressure), it is now possible to correlate indices of altered bodily states with corresponding states of the brain.

In one of the first investigations of higher cortical control of autonomic function, my colleagues and I adopted this approach. First, we tested for brain activity common to performance of effortful mental and physical tasks. Our principal observation was the correlation of activity in the anterior cingulate

cortex and the pons with increased cardiovascular output.[3] This finding indicated that in addition to classical autoregulatory sites, such as the pons, higher cortical regions implicated in cognitive control, such as the anterior cingulate cortex, are important for integrating volitional behavior (the sustained effort of doing the tasks) with peripheral states of cardiovascular arousal.

Extending this approach, we then determined whether specific cortical regions contribute to influences on bodily states mediated by the actions of either the sympathetic or parasympathetic system. Again, we scanned subjects while they performed motor and cognitive tasks to induce variability in heart rate. Measures of heart-rate variability (HRV) were then derived from R-wave intervals of the recorded electrocardiogram (ECG), utilizing a power-spectral analysis of R-R interbeat intervals. This type of analysis provides a means for distinguishing parasympathetic from sympathetic nervous control: 'high-frequency' spectral power (0.15 – 0.50 Hz) reflects parasympathetic nervous control, whereas 'low-frequency' spectral power (0.05 – 0.15 Hz) reflects sympathetic control.

From this study, we inferred that there are distinct cortical inputs into sympathetic and parasympathetic control of bodily arousal. An increase in the low-frequency sympathetic component of the ECG was associated with increased anterior cingulate, somatomotor, and insula cortex activity. Of particular interest was the connection between anterior cingulate cortex activity and influences expressed through the autonomic nervous system's sympathetic axis.[4] Increases in the parasympathetic component, on the other hand, were expressed in the anterior temporal cortex and deep brain structures, in particular, the basal ganglia.

Altered cardiovascular state is one index of bodily arousal resulting from increased autonomic drive; electrodermal activity (EDA), or enhanced electrical conductance of the skin with increased sweat secretion, is another. EDA is generated via the sympathetic system, and, unlike cardiovascular reactivity, its expression is not confounded by concurrent parasympathetic influences.

To determine how central brain states contribute to the generation of EDA, we designed a gambling task in which we presented subjects with a playing card (with a face value between one and ten) and then had them guess if the value of the successive card would be higher or lower.[5] We then rewarded a correct decision by giving the subject money and punished a wrong decision by taking the subject's money. Each time, we waited eight seconds to reveal the second card; during this delay subjects exhibited an anticipatory state of autonomic arousal that reflected the degree of uncertainty in the outcome. If the first card had a face value of five, there was maximal uncertainty; if its value was one or ten, there was no uncertainty.

3 H. D. Critchley et al., "Cerebral Correlates of Autonomic Cardiovascular Arousal: A Functional Neuroimaging Investigation in Humans," *Journal of Physiology (Lond)* 523 (1) (2000): 259 – 270.

4 H. D. Critchley et al., "Human Cingulate Cortex and Autonomic Control: Converging Neuroimaging and Clinical Evidence," *Brain* 126 (2003): 2139 – 2152.

5 H. D. Critchley, R. Elliott, C. J. Mathias, and R. J. Dolan, "Neural Activity Relating to Generation and Representation of Galvanic Skin Conductance Responses: A Functional Magnetic Resonance Imaging Study," *Journal of Neuroscience* 20 (2000): 3033 – 3040.

Using this game, we could explore which brain regions corresponded to a sympathetic state of arousal (as indexed by enhanced EDA) during the process of appraising each decision's risk value. Activity in the anterior cingulate and dorsolateral prefrontal cortices correlated with the degree of anticipatory EDA response, while both risk and arousal modulated activity in the anterior cingulate and insula cortex. These results led us to infer that these regions mediate between a cognitive state of risk and uncertainty and an associated bodily state of increased sympathetic arousal.

Although EDA changes automatically in response to the environment, subjects can be taught to exert control over their autonomic responses with biofeedback techniques. Using EDA as an index of sympathetic arousal, we trained subjects to perform a biofeedback relaxation task. In this study, we presented subjects with a 'thermometer' measuring their level of EDA arousal. By scanning these subjects as they reduced their EDA, we showed that this decrease in EDA correlated with increased activity in the anterior cingulate cortex, inferior parietal cortex, and globus pallidus. We then concluded that these regions contributed to intentional influences on sympathetically mediated bodily states.

In this study, exteroceptive information, as provided by the 'thermometer,' enabled biofeedback control of autonomic states. But what happens if this form of feedback is inaccurate or noisy? In such situations subjects can still accomplish the task by ignoring the exteroceptive signal and focusing instead on their own interoceptive state (for example, their state of cardiovascular arousal). In a follow-up experiment we used a similar biofeedback relaxation task but scrambled the accuracy (by adding random 'noise') or the sensitivity (by scalar adjustments of feedback) of the visual index of electrodermal arousal (EDA). These manipulations, as expected, only enhanced the subjects' reliance on their own interoceptive states.

Performance of biofeedback relaxation tasks activated the anterior cingulate, insula, thalamus, hypothalamus, and brain stem, as well as the somatosensory cortex and the dorsolateral prefrontal cortex. Both accuracy and sensitivity of feedback influenced activity within the anterior insula in particular. Thus, the perceptual qualities (sensitivity) of the feedback could increase the already enhanced insula response to noise in the feedback signal. These findings led us to surmise that activity in the anterior cingulate cortex mediates an intentional drive to decrease sympathetic activity, whereas the insula supports *sensory* integration of interoceptive and exteroceptive information, reflecting the current bodily state of autonomic arousal.[6]

The experimental approach described so far relies exclusively on studies conducted with healthy subjects. Homeostasis is crucially dependent on feedback signals. Once something perturbs the body, afferent-feedback circuits provide the brain with a representation of the body's altered state (or a metarepresentation, meaning a representation that indexes a disturbance in homeostasis). But where are these states mapped in the brain? This question should garner wide interest in view of Damasio's suggestion that feedback representations of visceral and somatomotor activity give not only emotional color to ongoing experience

6 H. D. Critchley et al., "Volitional Control of Autonomic Arousal: A Functional Magnetic Resonance Study," *Neuroimage* 16 (2002): 909–919.

but also support a bedrock representation of 'the self.'

Patients with discrete lesions to efferent or afferent limbs of these control loops can, in theory, provide a means to investigate the dynamics of autonomic regulation and, indeed, the wider influence of bodily responses on emotion and cognition. But the neurological literature lacks any clear lesion that effectively compromises feedback for autonomic states. Even a patient with a high spinal cord lesion that impairs most forms of somatic sensory feedback still has intact autonomic feedback through the vagus nerve.

So, instead, my colleagues and I have studied patients with pure autonomic failure (PAF), an acquired syndrome that results in degeneration of the peripheral autonomic system.[7] This syndrome affects postganglionic sympathetic and parasympathetic neurons in the absence of a central neurological pathology. Patients with this disorder no longer generate peripheral autonomic responses to stress, supplying a unique glimpse into how autonomically generated bodily states are remapped in the brain.

We reasoned that the primary difference in brain activity between PAF and control subjects performing identical stress tasks would be an absence of afferent feedback from the body. Initially, we observed increased activity in the pons in PAF subjects, across effortful and effortless cognitive and physical tasks, which is consistent with the idea that this region is responsible for continuous autoregulatory, autonomic control. The absence of afferent regulatory feedback in PAF subjects accounts for their enhanced level of activity in this region. In addition, PAF patients demonstrated reduced activity in the insula and primary somatosensory cortices across all tasks, indicating the involvement of these areas in a metarepresentation of altered bodily states.

Of particular interest was identifying the brain areas in which PAF patients and controls exhibited differences during physically and mentally taxing tasks (accompanied by autonomic arousal in controls), but not during effortless tasks. Significantly greater anterior cingulate activity was evident in PAF patients, who generated no cardiovascular arousal, compared to matched healthy controls. This finding is consistent with the idea that the anterior cingulate is responsible for context-specific autonomic modulation of bodily states to meet ongoing behavioral demands. In healthy subjects, context-specific autonomic responses provide a negative feedback signal to regulate the efferent sympathetic drive. Without such feedback, PAF subjects do not experience a decrease in anterior cingulate cortex activity.

In a similar experiment, we exposed both PAF patients and control subjects to a threat stimulus (a face paired with a shock), inducing bodily changes associated with fear. Again, we could determine how the absence of afferent information from the body alters brain activity. While PAF patients can perceive a threat and generate a central output signal from the brain, they lacked activity in the insula cortex, pointing again to the insula as providing the map of interoceptive states of autonomic arousal.[8]

7 H. D. Critchley, C. J. Mathias, and R. J. Dolan, "Neuroanatomical Basis for First- and Second-Order Representations of Bodily States," *Nature Neuroscience* 4 (2001): 207 – 212.

8 H. D. Critchley, C. J. Mathias, and R. J. Dolan, "Fear Conditioning in Humans: The Influence of Awareness and Autonomic Arousal on Functional Neuroanatomy," *Neuron* 33 (2002): 653 – 663.

Ray Dolan on body in mind

Many of our neuroimaging investigations implicated regions such as the anterior cingulate and insula cortices in autonomic control, particularly in the contextual generation and representation of states of bodily arousal, respectively. Thus, lesions affecting these regions should generate abnormal autonomic responses during volitional behavior. First, we studied three patients who had acquired damage to the anterior cingulate cortex. During the performance of effortful tasks all three patients demonstrated an absence of adaptive cardiovascular responses.[9] The patients' heart-rate variability (HRV) supplied additional evidence of an association between anterior cingulate function and control of autonomic-induced change in bodily states. Power-spectral analysis of their HRV revealed abnormalities primarily in sympathetic power, consistent with our prior conjecture that the anterior cingulate cortex provides an interface between cognitive effort and generation of sympathetic bodily arousal. Indeed, these observations strikingly support our theory that the anterior cingulate cortex is involved in integrating higher states of cognition, as required during volitional behaviors, and autonomic states of bodily arousal.

The possibility that the insula cortex is the primary substrate for a second-order mapping of bodily states led us to believe that this region mediates conscious awareness of bodily states – or, more precisely, feeling states – an idea also central to Damasio's account of feeling states.[10] Feeling is the subjective, private, or experiential component of emotion. In the present context, feeling refers to *the responses of our sense organs and internal milieu, including visceral states, to the environment, and the consequential responses within central mechanisms that monitor induced internal change*. To address the question of whether visceral awareness is associated with enhanced activity in this cortical region, we studied a group of interoceptively aware subjects.

The critical experimental hurdle here was creating a reliable index of visceral awareness. It turns out that a heartbeat-detection task can test sensitivity to visceral states. In this task, subjects must indicate the timing of their own heartbeat. The subjects are played back a signal, visual or auditory, triggered by their own heartbeats – with or without an experimentally manipulated lag. The subjects must then determine whether or not the feedback signal is synchronous with their interoceptively monitored heartbeat. Approximately half the population is good at detecting their heartbeats; interestingly, these subjects are also more emotionally aware and expressive.

Our study showed that enhanced attention to one's visceral state was associated with increased activity in a number of brain regions, in particular, the somatosensory and insula cortices.[11] Furthermore, detecting the desynchronization between the feedback signal and one's own heartbeat, a manipulation that places greater demands on interoceptive awareness, involved increased activity in the right insula. This finding converged with recent neuroanatomical discoveries indicating this region's importance to awareness of one's feelings. In particular, results showing that this region receives a dedicated lam-

9 Critchley et al., "Human Cingulate Cortex and Autonomic Control."

10 Damasio, *The Feeling of What Happens*.

11 H. D. Critchley et al., "Neural Systems Supporting Interoceptive Awareness," *Nature Neuroscience* 7 (2004): 189 – 195.

ina-1 spinothalamic input that converges with vagal inputs (a major source of visceral feedback) in the right anterior insula, and neuroanatomical evidence revealing that this region has a distinct laminar architecture found only in humans and higher primates, are in keeping with this interpretation of our neuroimaging findings.[12]

A question arising from these findings, though, is the degree to which neural mechanisms that mediate awareness of one's own bodily state contribute to awareness of the bodily states of others. Such awareness is at the core of the psychological attribute of empathy, and, arguably, this attribute underpins a human disposition to altruism and compassion. The suggestion here is that our ability to empathize with others relies on neuronal systems that underlie a higher-order representation of our own bodily and emotional states.

We addressed this very question in a study in which subjects received either a highly painful or not painful stimulus, and then viewed a loved one receive a painful stimulus. The intriguing result was that the emotional component of pain evoked when a subject actually received a painful stimulus was also triggered when the subject witnessed a loved one receive the same pain.[13] The regions of shared activation, for pain to self and to others, involved anterior cingulate and insula cortices – areas that I have suggested mediate the generation of sympathetic autonomic output and the afferent mapping of autonomically generated perturbation in bodily states.

Human neuropsychology has traditionally embraced a dualist model of brain and body. The fact that virtually all cognitive states are, to a greater or lesser degree, yoked to bodily states means that our conventional ideas of a fundamental division between the mental and the corporeal are in need of revision. What we apprehend in the realm of the senses is represented by images and concepts that facilitate representation in memory and awareness. What we sense from our bodies is, for the most part, only recognized as a vague background state, a current of feeling within our ongoing mental life, that is largely without symbolic mediation and conceptual form. However, this absence of conceptual form arguably provides the basis for the felt immediacy of experience, and a subjective consciousness, that emerges out of a dynamic mapping between brain and body.

12 J. Allman, A. Hakeem, and K. Watson, "Two Phylogenetic Specializations in the Human Brain," *Neuroscientist* 8 (2002): 335 – 346; A. D. Craig, "Human Feelings: Why Are Some More Aware Than Others?" *Trends in Cognitive Science* 8 (2004): 239 – 241.

13 T. Singer et al., "Empathy for Pain Involves the Affective But Not Sensory Components of Pain," *Science* 303 (2004): 1157 – 1162.

Jerry Fodor

How the mind works: what we still don't know

One could make a case that the history of cognitive science, insofar as it's been any sort of success, has consisted largely of finding more and more things about cognition that we didn't know and didn't know that we didn't. 'Throwing some light on how much dark there is,' as I've put it elsewhere. The professional cognitive scientist has a lot of perplexity to endure, but he can be pretty sure that he's gotten in on the ground floor.

For example, we don't know what makes some cognitive states conscious. (Indeed, we don't know what makes *any* mental state, cognitive or otherwise, conscious, or why any mental state, cognitive or otherwise, bothers with being conscious.) Also, we don't know much about how cognitive states and processes are implemented by neural states and processes. We don't even know *whether* they are (though many of us are prepared to assume so *faut de mieux*). And we don't know how cognition develops (if it does) or how it evolved (if it did), and so forth, very extensively.

In fact, we have every reason to expect that there are many things about cognition that we don't even know that we don't know, such is our benighted condition.

In what follows, I will describe briefly how the notions of mental process and mental representation have developed over the last fifty years or so in cognitive science (or 'cogsci' for short): where we started, where we are now, and what aspects of our current views are most likely to be in need of serious alteration. My opinions sometimes differ from the mainstream, and where they do, I will stress that fact; those are, no doubt, the parts of my sketch that are least likely to be true.

The 1950s 'paradigm shift' in theories of the cognitive mind, initiated largely by Noam Chomsky's famous review of B. F. Skinner's book *Verbal Behavior*, is usually described in terms of a conflict between 'behaviorism' and 'mentalism,' from which the latter emerged victorious. Behaviorists thought something

Jerry Fodor is State of New Jersey Professor of Philosophy at Rutgers University. He is the author of several publications in philosophy and cognitive science, including "Modularity of Mind" (1983), "A Theory of Content and Other Essays" (1990), and "The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology" (2000).

© 2006 by the American Academy of Arts & Sciences

was methodologically or ontologically controversial about the claim that we (and, presumably, other advanced kinds of primates) often do the things we do because we believe and desire the things we do. Chomsky's reply was, in essence, 'Don't be silly. Our behavior is characteristically caused by our mental states; therefore, a serious psychology must be a theory about what mental states exist and what roles they play in causing our behavior. You put gas in the tank because you believe that, if you don't, the car will grind to a stop, and you don't want the car to do so. How could anyone sane believe otherwise?'

That was, to put it mildly, all to the good. Behaviorism never was a plausible view of the methodology of psychology, any more than instrumentalism was a plausible view of the methodology of physics. Unsurprisingly, the two died of much the same causes. Many of the arguments Chomsky brought against the proposed reduction of the mind to behavior recall arguments that Carl Hemple and Hilary Putnam brought against the proposed reduction of electrons (to say nothing of tables and chairs) to 'fictions' or 'logical constructions' out of sensory experience. 'Don't be silly,' they said. 'Sensations and the like are mind-dependent; tables and chairs are not. You can sit on chairs but not on sensations; a fortiori, chairs can't *be* sensations.' Chomsky's realism about the mental was thus part of a wider realist agenda in the philosophy of science. But it's important to distinguish (as many of us did not back in those days) Chomsky's objections to Skinner's *behaviorism* from the ones he raised against Skinner's *associationism*. In retrospect, the latter seem the more important.

Behaviorism was and remains an aberration in the history of psychology. In fact, the mainstream of theorizing about the mind (including both philosophical empiricists and philosophical rationalists, and the 'sensationist' tradition of psychologists like Wilhelm Wundt and Edward Tichner) wasn't behavioristic. Rather, it was a mentalistic form of associationism that took the existence of mental representations (what were then often called 'Ideas') and their causal powers entirely for granted. What associationism mainly cared about was discovering the psychological laws that Ideas fall under. And the central thesis – which, Hume said, was to psychology what gravitation was to Newtonian physics – was that Ideas succeed one another in cognitive processes according to the laws of association.

For nearly three hundred years, associationism was the consensus theory of cognition among Anglophone philosophers and psychologists. (It's still the view assumed by advocates of 'connectionism,' a movement in cognitive science that hopes to explain human intellectual abilities by reference to associations among 'nodes' in 'neural networks,' the latter corresponding, more or less, to Ideas and the former corresponding, more or less, to minds that contain them. If, in fact, you take away the loose talk about 'neurological plausibility,' the connectionist's account of cognition is practically indistinguishable from Hume's.) Associationism was widely believed to hold, not just for thought but for language and brain processes as well: thoughts are chains of associated concepts, sentences are chains of associated words, and brain processes are chains of associated neuron firings. In all three cases, transitions from one link in such a chain to the next were supposed to be probabilistic, with past experience determining the probabilities according to whatever Laws of Association happened to be in fashion.

These Anglophone theorists notwithstanding, it's been clear, at least since Kant, that the associationist picture can't be right. Thoughts aren't *mere* sequences of ideas; at a minimum, they are *structured* sequences of ideas. To think 'there's a red door' isn't to think first about red and then about a door; rather, it's to think *about* a door *that* it is red. This is, I suppose, a truism, though perhaps not one that the cogsci community has fully assimilated. Likewise, sentences aren't just lists of words. Instead, they have a kind of internal structure such that some of their parts are grouped together in ways that others of their parts are not. Intuitively, the grouping of the 'parts' of 'the red door opened' is: [(the) (red door)] (opened), not (the red) (door opened).

So sentences have not just lexical contents but also constituent structures, consisting of their semantically interpretable parts ('the red' doesn't mean anything in 'the red door opened,' but 'the red door' does). One of the main things wrong with associationism was thus its failure to distinguish between two quite different (in fact, orthogonal) relations that Ideas can enter into: *association* (a kind of causal relation) and *constituency* (a hierarchical kind of geometrical relation). Ironically, as far as anybody knows, the first isn't of much theoretical interest. But the second, the constituency relation, does a lot of the heaviest lifting in our current accounts of cognition.

For example, as Chomsky famously pointed out, sentences are 'productive.' The processes that construct sentences out of their parts must be recursive: they must be able to apply to their own outputs, thereby generating infinite sets. It turns out that these recursions are defined over the constituent structures of the expressions they apply to. Typically, they work by embedding a constituent of a certain type in another constituent of the same type, like a sentence within a sentence. (For example, the sentence 'John met the guy from Chicago' is some sort of construction out of the sentences 'John met the guy' and 'the guy is from Chicago,' with the second sentence embedded in the first.) The same sort of story goes for mental representations, since they are also productive: if there weren't boundlessly many thoughts to express, we wouldn't need boundlessly many sentences to express them.

Their potential for productivity isn't, of course, the only thing that distinguishes constituent structures from associative structures. The strength of the association between Ideas is traditionally supposed to depend largely on the frequency and spatiotemporal contiguity of their tokenings. In contrast, as Chomsky also pointed out, the structural relations among the constituents of a complex representation hold for novel representations as well as for previously tokened ones, and are typically independent of the propinquity of the relata.

In sum, by far the most important difference between the traditional theories of mind and the ones those of us who *aren't* connectionists endorse is the shift from an associationist to a constituent, or *computational*, view of cognition.

Here, then, are the two basic hypotheses on which the current computational theory of cognition rests:

First, mental representations are sentence-like rather than picture-like. This stands in sharp contrast to the traditional view in which Ideas are some kind of images. In sentences, there's a distinction between *mere* parts and constituents, of which the latter are the *semantically interpretable* parts. By contrast, every part of a picture has an interpretation: it shows part of what the picture shows.

Second, whereas associations are operations on parts of mental representations, computations are operations defined on their constituent structures.

So much for a brief (but, I think, reasonably accurate) summary of what most cognitive scientists now hold as a working hypothesis about mental representations and mental processes (except, to repeat, connectionists, who somehow never got beyond Hume). Now, the question of interest is, for how much of cognition is this hypothesis likely to be true? The available options are *none of it*, *some of it*, and *all of it*.

My guess is: at best, not very much of it. This brings me to the heart of this essay.

Here's what I'm worried about. As we've been seeing, constituent structure is a species of the part/whole relation: all constituents are parts, though not vice versa. It follows that constituency is a *local* relation: to specify the parts of a thing you don't need to mention anything that's *outside* of the thing. (To specify the part/whole relation between a cow and its left leg, you don't have to talk about anything outside the cow. Even if this cow were the only thing in the whole world, it would bear the same relation to its left leg that it bears to its left leg in this world. Only, in that world, the cow would be lonelier.) Likewise for representations – mental representations included. Since constituents are parts of representations, operations defined on constituents apply solely in virtue of the internal structure of the representations.

The question thus arises: are there mental structures, with mental processes defined on them, that aren't local in this sense? If there are, then we are in trouble because, association having perished, computation is the only notion of a mental process that we have; and, as we've just seen, computations are defined over local properties of the representations that they apply to.

Well, I think there's pretty good reason to suppose that many of the mental processes crucial to cognition are indeed *not* local. So I guess we're in trouble. It wouldn't be the first time.

There are at least two pervasive characteristics of cognitive processes that strongly suggest their nonlocality. One is their sensitivity to considerations of relevance; the other is their sensitivity to the 'global' properties of one's cognitive commitments. It is very easy to run the two together, and it's a common practice in cogsci literature to do so. For polemical purposes, perhaps nothing much is lost by that. But some differences between them are worth exploring, so I'll take them one at a time.

Consider the kind of thinking that goes on in deciding what one ought to believe or what one ought to do (the same considerations apply both to 'pure' and to 'practical' reason). In both cases, reasoning is typically *isotropic*. In other words, *any* of one's cognitive commitments (including, of course, currently available experiential data) is relevant, in principle, to accepting or rejecting the options – there is no way to determine, just by inspecting an empirical hypothesis, what will be germane to accepting or rejecting it. Relevance isn't like constituency – it's not a local property of thoughts.

So how *does* one figure out what's relevant to deciding on a new belief or plan? That question turns out to be very hard to answer. There is an infinite corpus of prior cognitive commitments that might prove germane, but one can actually visit only some relatively small, finite subset of them in the 'real time' during which problems get solved. Relevance is long, but life is short. Something, somehow,

must 'filter' what one actually thinks about when one considers what next to believe or what next to do.

Hence the infamous 'frame problem' in theories of artificial intelligence: how do I decide what I should take to be relevant when I compute the level of confidence I should invest in a hypothesis or a plan? Any substantive criterion of relevance I employ will inevitably risk omitting something that is, in fact, germane; and one of the things I want my estimate to do (all else equal) is minimize this risk. How on earth am I to arrange that?

I think the frame problem arises because we have to use intrinsically local operations (computations, as cogsci currently understands that notion) to calculate an intrinsically nonlocal relation (relevance). If that's right, the frame problem is a symptom of something deeply inadequate about our current theory of mind.

By contrast, it's a widely prevalent view among cognitive scientists that the frame problem can be circumvented by resorting to 'heuristic' cognitive strategies. This suggestion sounds interesting, but it is, in a certain sense, empty because the notion of a heuristic procedure is negatively defined – a heuristic is just a procedure that only works from time to time. Therefore, everything depends on *which* heuristic procedure is alleged to circumvent the frame problem, and about this the canonical literature tends to be, to put it mildly, pretty causal.

Here, for example, is Steven Pinker, in a recent article, explaining what heuristics investors use when they play the stock market: "Real people tend to base investment decisions on, among other things, what they hear that everyone else is doing, what their brother-in-law advises, what a cold-calling stranger with a confident tone of voice tells them, and what the slick brochures from large investing firms recommend. People, in other words, use heuristics."[1]

Pinker provides no evidence that this is, in fact, the way that investors work; it's a story he's made up out of whole cloth. At best, it's hard to see why, if it's true, some investors make lots more money than others. But never mind; what's really striking about Pinker's list is that he never considers that *thinking about the stock market* (or paying somebody else to think about it for you, if you're lazy like me) might be one of the 'heuristics' that investors employ when they try to figure out whether to buy or sell. Ironically, thinking seems largely to have dropped out of heuristic accounts of how the mind works. Skinner would have been greatly amused.

There have been, to be sure, cases when cognitive scientists have tried to tell a story about the use of heuristic strategies in cognition that amounts to more than the mere waving of hands. To my knowledge, the heuristic most often said to guide decisions about what action to perform or belief to adopt is some version of 'if things went all right with what you did last time, do the same again this time.' We owe a rather opaque formulation to the philosopher Eric Lormand: "A system should assume *by default* that a fact persists, unless there is an axiom specifying that it is changed by an occurring event .... [G]iven that an event E occurs in situation S, the system can use axioms to infer new facts existing in S+1, and then simply 'copy' the remainder of its beliefs about S over to S+1."[2] Likewise, Peter Carruthers says

1 Steven Pinker, "So How Does the Mind Work?" *Mind and Language* 20 (1) (February 2005): 1 – 24.

2 Zenon W. Pylyshyn, ed., *Robot's Dilemma: The Frame Problem in Artificial Intelligence* (Norwood, N.J.: Ablex, 1987), 66.

that "there's no reason why the choices [about what to do next] couldn't be made by higher-order heuristics, such as 'use the one which worked last time.'"[3] The idea, then, is to adopt whichever plan was successful when this situation last arose. Cogsci literature refers to this heuristic as the 'sleeping dog' strategy. Last time, I tiptoed past the sleeping dog, and I didn't get bitten. So if I tiptoe past the sleeping dog again now, I probably won't get bitten this time either. So the plan I'll adopt is *tiptoe past the sleeping dog*. What could be more reasonable? What could be less problematic?

But, on second thought, this suggestion is no help since it depends crucially on how one individuates situations, and how one individuates situations depends on what one takes as *relevant* to deciding when situations are of the same kind. Consider: What was it, precisely, that *did* happen last time? Was it that I tiptoed past a sleeping dog? Or was it that I tiptoed past a sleeping *brown* dog? Or that I tiptoed past a sleeping pet of Farmer Jones? Or that I tiptoed past *that* sleeping pet of Farmer Jones? Or that I tiptoed past a creature that Farmer Jones had thoughtfully sedated so that I could safely tiptoe past it? It could well be that these are *all* true of what I did last time. Nor, in the general case, is there any reason to suppose that I know, or have ever known, what it is about what I did last time that accounts for my success. So, when I try to apply the sleeping dog heuristic, I'm faced with figuring out which of the true descriptions of the situation *last time* is relevant to deciding what I ought to do *this* time. Keeping that in mind is crucial. If the dog I tiptoed past last time was sedated, I've got no grounds at all for thinking that tipping my toe will get me past it now. Philosophers have gotten bitten that way from time to time.[4]

So I'm back where I started: I want to figure out what action my previous experience recommends. What I need, in order to do so, is to discern what about my previous action was relevant to its success. But relevance is a nonlocal relation, and I have only local operations at hand with which to compute it. So the 'sleeping dog' strategy doesn't *solve* my relevance problem; it only begs it. You might as well say: 'Well, you decided on an action that was successful last time; so just decide on a successful action this time too.' My stockbroker tells me he has a surefire investment heuristic: 'Buy low and sell high.' It sounds all right, but somehow it keeps not making me rich. It's well-nigh useless to propose that heuristic processing is the solution to the problem of inductive relevance because deciding how to choose, and how

3 Peter Carruthers, "Keep Taking the Modules Out," *Times Literary Supplement* 5140 (October 5, 2001): 30.

4 Another way to put the same point: What counts as the last time *this* situation arose depends on how I describe this situation. Was the last time *this* happened the last time that I tiptoed past a sleeping dog? Or was it the last time that I tiptoed past a brown sleeping dog? Or was it the last time that I tiptoed past a sedated brown sleeping dog? What determines which heuristic I should use in this situation thus depends on what kind of situation I take it to be. And what kind of situation I take it to be depends on what about my successful attempts at dog-passing was *relevant* to their succeeding. We're very close here to Nelson Goodman's famous point that, in inductive inference, how you generalize your data depends crucially on what you take them to have in common – what their relevant similarity is. The frame problem is thus an instance of a perfectly general problem about the role of relevant similarity in empirical inferences. Lacking an account of that, the advice to do the same as you did last time is, quite simply, *empty*.

to apply, a heuristic itself typically involves estimating degrees of inductive relevance.

It's remarkable, and more than a bit depressing, how regularly what is taken to be a solution of the frame problem proves to be simply one of its formulations. The rule of thumb for reading the literature is: if someone thinks that he has solved the frame problem, he does not understand it; and if someone even thinks that he understands the frame problem, he doesn't understand it. But it does seem clear that, whatever the solution of the frame problem turns out to be, it isn't going to be computationally local. Its constituent structure is all of a mental representation that a mental process can 'see.' But you can't tell from just the constituent structure of a thought what tends to (dis)confirm it. Clearly, you have to look at a lot else as well. The frame problem is how you tell *what* else you have to look at. I wish I knew. If you know, I wish you'd tell me.

The frame problem concerns the *size* of a field of cognitive commitments that one has to search in order to make a successful decision. But there are also cases where the *shape* of the field is the problem. Many systems of beliefs that are germane to estimating confirmation levels have 'global' parameters; that is, they are defined over the whole system of prior cognitive commitments, so computations that are sensitive to such parameters are nonlocal on the face of them.

Suppose I have a set of beliefs that I'm considering altering in one way or another under the pressure of experience. Clearly, I would prefer that, all else equal, the alteration I settle on is the simplest of the available ways to accommodate the recalcitrant data. The globality problem, however, is that I can't evaluate the overall simplicity of a belief system by summing the intrinsic simplicities of each of the beliefs that belong to it. There is, on the face of it, *no such thing* as the 'intrinsic' simplicity of a belief (just as there is no such thing as the intrinsic relevance of a datum). Nothing *local* about a representation – nothing about the relations between the representation and its constituent parts, for example – determines how much it would complicate my current cognitive commitments if I were to endorse it.

Notice that, unlike the problems about relevance, this sort of worry about locality holds even for *very small* systems of belief. It holds even for *punctate* systems of belief (if, indeed, there can be such things). Suppose that *all* that I believe is P, but that I am now considering also adopting either belief Q or belief R. What I therefore want to evaluate, if I'm to maximize overall simplicity, is whether the belief P&Q is simpler than the belief P&R. But I can't do that by considering P, Q, and R severally – the complexity of P&Q isn't a function of the complexity of P and the complexity of Q taken separately. So it appears that the operations whereby I compute the simplicity of P&Q can't be local.

The same goes for other parameters that anyone rational would like to maximize, all else being equal. Take, for example, the relative *conservatism* of such commitments. Nobody wants to change his mind unless he has to; and if one has to, one prefers to opt for the bare minimum of change. The trouble is, once again, that conservatism is a *global* property of belief systems. On the face of it, you can't estimate how much adding P would alter the set of commitments C by considering P and C separately; on the face of it, conservatism (unlike, for example, consistency) isn't a property that beliefs have taken severally.

In short, it appears that many of the principles that control (what philoso-

phers call) the nondemonstrative fixation of beliefs have to be sensitive to parameters of whole systems of cognitive commitments.[5] Computational applications of these principles have to be nonlocal. As a result, they can't literally be computations in the sense of that term that our current cognitive science has in mind.

If you suppose (as I'm inclined to) that nondemonstrative inference is always a species of argument to the best available explanation, this sort of consideration will be seen to apply very broadly indeed: what's the best available explanation always depends on what alternative explanations are available; and, by definition, the presence or absence of alternatives to a hypothesis isn't a local property of that hypothesis.

I should, however, enter a caveat. Suppose that something you want to measure is a property of complex beliefs but not of their parts; for example, suppose that you want to assess the simplicity of P&Q relative to that of P&R. My point has been that, *prima facie*, the computations you have to perform aren't local; they must be sensitive to properties that the belief that P&Q has *as such*. One could, however, *make* the computations local by brute force. So while the complexity of P&Q isn't determined by local properties of P together with the local properties of Q, it is determined, trivially, by local properties of the representation P&Q. In effect, you can always preserve the *locality* of computations by inflating the size of the *units* of computation. The distance between Washington and Texas is a *nonlocal* property of these states, but it's a *local* property of the Northern Hemisphere.

That sort of forced reduction of global problems to local problems is, however, cheating since it offers no clue about how to solve the local problems that the global ones are reduced to. In fact, you would think that nobody sensible would even consider it. To the contrary: recent discussions of confirmation (from, say, Duhem forward) have increasingly emphasized the *holism* of nondemonstrative inferences by claiming that, in the limiting case, *whole theories* are the proper units for their evaluation. This saves the locality of the required computations by fiat, but only at the cost of making them wildly intractable. More to the point: even if we *could* somehow take whole theories as the units in computing the confirmation of our hypotheses, the patent fact is that we don't. Though we

5 It's very striking how regularly the problems cognitive psychologists have when they try to provide an explicit account of the nondemonstrative fixation of belief exactly parallel the ones that inductive logicians have when they try to understand the (dis)confirmation of empirical theories. Pinker, among many others, objects to conceptualizing individual cognition as, in effect, scientific theorizing writ small. "Granted that several millennia of Western science have given us nonobvious truths involving circuitous connections among ideas; why should theories of a single human mind be held to the same standard?" Pinker, "So How Does the Mind Work?" In fact, however, it's increasingly apparent that the philosophy of science and the psychology of cognition are beating their heads against the same wall. It is, after all, a truism that, by and large, scientists think much the same way that we do.

The similarity between the two literatures can be quite creepy given that they seem largely unaware of one another's existence. Thus, Arthur Fine raises the question: how can someone who is not a realist about the ontological commitments of scientific theories explain the convergence of the scientific community on quite a small number of explanatory options? He says it's in part because we all follow "the instrumentally justified rule 'if it worked well last time, try it again.'" David Papineau, ed., *The Philosophy of Science, Oxford Readings in Philosophy* (New York: Oxford University Press, 1996), 78. He doesn't, however, even try to explain the consensus about what 'it' is.

don't alter our cognitive commitments one by one, it's also not true that everything we believe is up for grabs all of the time, which is what the idea that the units of confirmation are whole theories claims if it is taken literally. The long and short is: in science and elsewhere, it appears that the processes by which we evaluate nondemonstrative inferences for simplicity, coherence, conservatism, and the like are *both* sensitive to global properties of our cognitive commitments *and* tractable. What cognitive science would like to understand, but doesn't, is how on earth that could be so.

There is quite possibly something deeply wrong with the cognitive psychology that we currently have available, just as there was something deeply wrong with the associative cognitive psychology that couldn't acknowledge recursion or the constituent structure of linguistic and mental representations. What, then, are we to do?

Actually, I don't know. One possibility is to continue to try for an unvacuous heuristic account of how we might compute relevance and globality. I haven't heard of any, but it's perfectly possible that there are some out there somewhere – or that there will be tomorrow, or the day after. I don't believe it, but hope is notorious for springing eternal.

Alternatively, what our cognitive psychology needs may be a new notion of computation, one that doesn't have locality built into it. That is, of course, a lot easier to say than to provide. The current computational account of mental processes is at the core of our cognitive science. The notion of a computation is what connects our theory of mental representations to our theory of mental processes; it does for our cognitive science what the laws of association promised (but failed) to do for our empiricist forebears. As things stand, we have no idea at all how to do without it. But at least we may be starting to understand its intrinsic limitations. In the long run, that could lead to revising it, or rejecting it, or, best of all, replacing it with some theory that transcends its limitations – a consummation devoutly to be wished.

So the good news is that our notions of mental representation and mental process are much better than Hume's.

The bad news is that they aren't nearly good enough.

Steven Pinker recently wrote a book called *How the Mind Works*. It is a *long* book. In fact, it is a *very* long book. For all that, my view is that he doesn't actually know how the mind works. Nor do I. Nor does anybody else. And I suspect, such is the state of the art that, if God were to tell us how it works, none of us would understand Him.

Fiction by Dorian Gossy

The Door in the Woods

The cancer did not so much kill Frieda's mother as engulf her like rising water. Within a week of her death, Frieda's father had locked himself in the cabin at the edge of Frieda's property. He had the clothes on his back and the few amenities already in the cabin: a tuberculosis cure cot with raising back, a quilt, a door skin on cinder blocks for a desk. A chair, a functional woodstove, and a spring-fed spigot outside.

John Prade was seventy-eight. Despite the Prades' fractious marriage, their neighbors in Pittsburgh stood ready with casseroles and good cheer after the funeral. He fended off all generosities and phoned Frieda to come gather her mother's things. She drove down from her house on retired farmland in the Adirondacks.

"I want to get the hell out of here," he said when she arrived.

Dorian Gossy, author of "Household Lies" (2005), was the recipient of the 2004 First Book Award in Fiction from Winnow Press. Her fiction has appeared in a dozen literary journals, including "North American Review," "Other Voices," "Denver Quarterly," and "The Sun." She teaches creative writing at Indiana University each spring.

© 2006 by Dorian Gossy

The house had convulsed into unprecedented clutter, as though a huge hand had shaken everything off its shelves and out of drawers.

"Come live with me." Frieda had not planned to say it. "My house is big enough."

"Hell, no. I've seen it. Odd little box of a place. Too much land. Too cold up there. Here there's a furnished apartment across town."

"Let's pack you up today," Frieda said. Where had she acquired such calm? Neither parent had had any to spare. "I'll hire someone to clean this up and ship us the good stuff. Ralph and Cathy can help."

There would be a fight, of course. Days of cajoling, colluding with her brother, Ralph, on strategy. John Prade had the furious visage of a demonic Chinese mask. "You always got your way, didn't you?"

Frieda suddenly thought of her mother, pinched and dying, but strong enough in the last eight days of her life to banish her husband from her sickroom, pointing the way out with a waxy finger. At the time, Frieda imagined this act a kindness, though, in retrospect, there was nothing kind about her mother's fevered eyes.

Fiction by Dorian Gossy

"Well, I got a blue dress one time," she said. "I remember promising everything for it."

Frieda's house was ten years old, modest but solid. Large windows brought her to a startling intimacy with firs, white pines, poplar, beech, maple. Her forty acres of browning autumnal grasses opened west toward the tower of Whiteface Mountain like an invitation.

Frieda had moved up from Pittsburgh after Lane, her husband of twenty years, fell in love with a student in one of the Spanish night classes he taught – a woman older than Frieda by ten years, with patrician cheekbones and gleaming silver blonde hair. They now lived in Utah. The whole of it had crept up on Frieda, who had felt the waning of Lane's attention; fielded the calls from Stephanie, always so very cordial; and finally, marked the passion for stargazing Lane had evolved out of nowhere, requiring late nights out and the purchase of a new telescope whose case never seemed to scuff. The stars above her house on the farmland, so far away from where she and Lane had lived, shone in virulent profusion on cloudless nights.

"I understand wanting to get away," Frieda ventured the second night, serving her father pea soup with lamb. He ate steadily, pausing for a swallow of his beer. "Coming up here was good for me. I can edit medical textbooks anywhere."

She regretted this last admission. Her father seldom passed up an opportunity to berate her for dropping out of medical school in 1977, when she was just twenty-four.

John Prade finished his soup. "I was wondering about that cabin out there." He jerked his head toward the south window, beyond which the rough one-room cabin stood a quarter of a mile away. Frieda guessed that it had been thrown together as a hunting cabin in the late nineteenth century. "Show it to me."

"Dad, it's night. How about tomorrow?" Frieda said. "It's pretty primitive. Rusted farm junk out there next to it. Broken old bottles."

"Miss Frieda, you have been good to bring me here." John Prade put his elbows on the table and rested his chin on his folded hands. He'd not called her Miss Frieda since she was small enough to be wrestled into the itchy crinoline dresses thought cute for little girls in 1958. The rush of pleasure that swept her felt like shame. "But I don't think it too much skin off your nose to take me out there right now," he continued. "Or I'll find my way by myself."

Frieda stood and busied herself with collecting their dishes and tipping them into the sink.

"All right, I'll take you out there," Frieda said. She tried out a note of exasperation.

He snorted. "Don't do me any favors, girl." But after a long time in the bathroom, John Prade appeared at the back door with his coat on, shod in an abused pair of Wellingtons. Frieda readied her biggest flashlight and led them through the field.

Inside the cabin, he squinted at the sooty ceiling joists and walls as if they caused him pain. When he said he wanted a minute to look things over, Frieda left him the flashlight and returned to the house with only the smoke of the Milky Way for guidance. It wasn't until midnight, while dozing in an overstuffed chair waiting for him, that Frieda began to worry.

She sweated inside the heavy coat she'd worn as she walked back down the thin path in the grass with a small Maglite. "Dad?" she called as she ap-

proached the cabin. All the drapes were drawn, but the door window had no cover. The door was locked, and she had no key. She hadn't known the door could lock. "Dad – are you in there?"

She shone the flashlight inside. At first the window threw the beam back into her own face. Then she could make out the recliner, the dead cluster flies and ladybugs piled in the corners, the edge of the cure cot. She shifted the light toward the cot and leaned her cheek against a windowpane. John Prade sat on the end of the cot, eyes on her. He blinked when the light struck his face. Frieda knocked on the glass. "Dad? It's after midnight. Are you all right?"

Without answering, John Prade bent down and pulled off his battered Wellingtons. He drew back the bedclothes and arranged his lean frame on the cure cot, pulling the comforter up to his neck. Once he stilled, only the contours of his body under the comforter marked anything different about the cold disuse of the room.

"Dad?" Frieda knocked once more. "I'm leaving my coat on the doorknob. I'll be back in the morning." A sensible person would have smashed open the door window, broken the lock. Frieda turned away coatless toward her house with inexplicable contentment.

Frieda's elder brother was a lawyer in Philadelphia. With two reasonable teenaged children and his nice wife, Cathy, Ralph's life had the well-crafted appearance of a Christmas crèche. He was happy, as he always told Frieda, because he had decided to be happy. Frieda found this willed happiness a human miracle, like a great gift for athletics or music.

Ralph approved of their father's move as a temporary measure. "He's still upset about Mom's death, of course," he told Frieda over the phone. "It's so new, even though we knew it was coming."

John Prade had been in the cabin a week. Each night when Ralph asked after him, Frieda said, I think he's okay, which was the truth. "Ralph, have you ever wanted to get away from home? Just up and leave it all behind?"

"Of course not," Ralph said. "But I'm lucky, you see. And I work at keeping it that way."

"Well, what if you did? How would you get away?"

Ralph pushed out a breath. "What do you mean? I just said I never have."

The bare birches and white pines that grew along her land's rivulets shouldered each other irritably in the wind. Frieda turned away from the window, took a fresh tight grip on the phone. "Dad's locked himself in my old hunter's cabin."

"He's done what?" Ralph bellowed. "How is he surviving out there? It's December now, for Pete's sake! Won't that spigot freeze up? What's he supposed to do for water then?"

"I've been taking food out and putting it on the doorstep." Frieda said. "I take out blankets, and batteries for the flashlight. I got some clothes from the Catholic thrift store in Ausable Forks that should fit him. At least they're warm. He has a saw and a shovel and an axe. He's figured out the woodstove. I think he's thawing snow on the stove for wash water."

She could practically hear Ralph shaking his head slowly back and forth, as if at an opposing lawyer's dim client on the witness stand. Frieda had seen him do this – he had cajoled her into watching him argue a case before a jury once when she and Lane were visiting. "Sis, you said you were reeling him back into civilization."

Fiction by Dorian Gossy

"He's keeping the place neat, I think." Frieda tugged hard on the permanent braid she'd recently made of her long tangled hair. "He writes me thank-you notes for the food. He was never a big thanker, you know. You remember how he barked at Mom all the time."

"Well," Ralph's voice rose, "are we talking about the same man? My father was gracious and sociable. Happily married. How do you think I got this way?"

"You? The self-made happy man?"

"Okay. Listen. I'll come up this weekend, knock on the door, and if he won't come out, we'll get someone to get him out. You have a local sheriff, right?"

"Ralph, no. Don't do that – not yet. Let me have some more time with him. I know it sounds idiotic, but I think – " she scrambled for sentences – "I think I can get him to actually go home if I give him more time – "

"Really? Why?"

"He's making a sculpture," Frieda lied completely. "The sort of lawn ornament thing Mom used to like in her garden. I think it's a memorial for her."

"Well, let me know," Ralph said. "Sounds good, actually."

It was Frieda who began the sculpture, using things from the old household dump outside the cabin door. She explored the dump in the afternoons, when John Prade was inside the cabin, the tang of his woodstove fire the only sign of life. He had fixed a sheet of newspaper to the door window. Out of the cold ground came bottles: milk of magnesia bottles in cobalt blue; amber motor oil and Clorox bottles; and Cra-Rock seltzer bottles, in glass as aqua as river ice. Frieda coaxed these lost things and others from the pit with lightly gloved hands. A truck license plate from 1940. A jar of Lustre-Creme Hair Dressing. Warped and rusted gears, severed from the machines they had served. She rinsed everything in the tap from the spring and arranged the items in pyramids on the porch, by which, she imagined, someone's mother might have been amused. Her own mother had disliked clutter, however, so her grave had only a low, polished marble marker. After her death, John Prade had vetoed the suggestion that the stone also include his name, or space for it.

Frieda let one editing deadline go by a few days while she dug in the dirt. And then another, a week. The medical textbook company liked her – it was all right. She even had benefits. At times she forgot the deadlines, forgot about her father inside the cabin while she dug. She had not knocked after the first night. His thank-you notes, scrawled on scrap paper Frieda brought out from her printer, were left pinched between the doorjamb and the door like a thumb.

"Dad, Ralph called. He wants to come get you out of there."

Frieda had brought a chair out from the house and placed it next to the door, where she sat while scraping her finds from the dump. Below the doorknob a crack in the wood ran parallel to the door's length, and it was near this crack that Frieda put her lips when she spoke. "Ralph always was kind of a meddler, wasn't he? He used to come into my room without knocking."

The snows held off for the first week in December, but the television news promised the first storm by week's end. Still the spigot had not frozen. Frieda had bought a new warm coat for herself from the Catholic resale shop. She bought extra-heavy gloves from a mountaineering outfitter in Keene Valley. "I've found a Noxema jar out here, Dad, in that navy blue glass. With Deco letter-

ing – isn't that from the '30s?" John Prade had been an architect in his working life. His language had been finely rendered angles, his stories the blueprints that bore his name. "You always liked that style," Frieda said. "I remember not liking it when I was a kid – it spooked me. It reminded me of the Wizard of Oz."

John Prade didn't answer, but the next day when she went back, the arrangement of tool and bottle she'd been constructing had changed. A rooster weathervane dangled a ringed doorknocker from its cockscomb, both balanced on a feed bucket. A dozen of the bottles snaked nose to tail, no two of the same color touching.

"Dad, one time you asked what was wrong with me that made Lane leave. You said I might be – frigid, and maybe that was my problem." There seemed the tiniest breath coming from the fissure in the door. Flickering stove light – or was she imagining it? "Do you remember how I cried when you said it?"

He had commenced his own excavation in the dump. Fresh dirt piled up beside the pit. He'd hit a mother lode of pest-control items. The "Dead Easy" rattrap; a medieval-looking "Nash Mole Trap," all spikes and collars; and a choker mousetrap, wire guillotines set in a circle atop a square block of wood. These he arranged in a row across the pitted wood of a broken harvester.

A presence manifested on the other side of the crack in the door. Nothing visible, but when she inhaled the space through the door, Frieda thought of how one intuits an object or a silent person nearby in jet dark. She rubbed at a rust stain on a cornflower-blue Bromo-Selzer bottle.

"I used to – sort of – be attracted to other women, Dad." Frieda put down the jar. She scrubbed at a flake of paint on the doorjamb with a chapped fingertip. "Not that I ever did anything about it. But there was a girl in junior high. Her name was Marla. She'd wear dresses out of Qiana, that silky material I've not seen since the '70s. And she was small and slim, like a caterpillar. I used to want to take hold of her around the waist. That was it."

A breeze whistled low in the door crack. "I thought about women later, though. When Lane found Stephanie. You know, he had his Spanish class come to the house for a party at the end of the semester that she was in his class. They stayed away from each other the whole party, but I saw him hand her a glass of wine. Then I knew. And then I started thinking about women and their breasts, when I was alone. The more he disappeared, the more I had those thoughts."

A sound like scurrying. A floorboard creak. Then a wedge of folded paper squeezed out from under the door. Frieda retrieved it. Like the other notes John Prade pinched in between the door and its jamb, the triangle of paper unfurled bore only the words, THANK YOU.

"You're talking to him." Ralph had that ruffled tone again.

"Yeah."

"And he's not talking back."

"Well, no – except with the notes I told you about." Frieda dug her fingers into her braid. She knew she should undo the braid, wash her hair all the way down. It had been two weeks.

"And this is conversation?"

"Well, he listens," Frieda said. "And I tell him things."

Ralph cleared his throat noisily. All his life, those neglected allergies. Frieda could remember him snoring when she was still sleeping in a crib. "What things?"

"Things I've never told anyone. He doesn't respond, but I know he's listening."

"So." Ralph's voice grew muted. "What have you told him, exactly?"

Occasionally Frieda had gotten the upper hand with her brother during their childhood. Her ability to do so had been like predicting the weather – she could generally tell which situations would turn to her advantage, but nothing was guaranteed. She hadn't sought it much. Ralph seemed to be a good brother. But she felt the same change in the timbre of his question that she'd come to recognize long ago as the shift of power. "I've told him about – my marriage. About Lane. About a girl from school. About you, some."

"What did you tell him?" Ralph's big baritone condensed to a whisper. "About me?"

Under the bridge a mile away from Frieda's land the Ausable River furled itself into menacing rapids. Their churn took hold of her and Ralph, as though the craft of their conversation had lost its rudder. "I said you were pushy when we were young. I said you came into my room without knocking."

"Did you tell him what – happened one time? You know, that time?"

Frieda searched the dark city of her memory. Something had happened. But then, something had always happened. What you did was knuckle your forehead and try to forget. She suddenly felt so, so tired, the way she'd worn out from her own grief when Lane left. "I don't know what you're talking about," she said to the pinprick holes of the phone's mouthpiece.

"You did tell him. About when – I came in. I just wanted to see – a girl. Curiosity used to be healthy. The hairbrush was a really bad idea. You know I said I was sorry."

"Ralph, I didn't – "

"I was only thirteen years old!"

The recollection as ordinary as that of breakfast, or a fishing trip in the summer. Ralph pinning her to the bed, bribing her with promises of candy and good behavior, so he could look between her legs and put things up inside her. She'd been nine. "Look, Ralph, how do you think I could tell Dad anything like that? I wouldn't tell anyone that. I never even told Lane."

"I'll bet you're lying. I'll bet you told Dad."

"Oh, go make yourself happy over this one," she said, so weary. "For Christ's sake."

Ralph cleared his throat again, sharp and loud. "I'll come, that's all. I'll come up and explain things to him. I'll get in the car tomorrow. You'll be there, right? Though I guess it doesn't matter if you are or not. I know the way. You've got the key to your house under the stairs on a nail."

"Go to hell, Ralph."

"No joke. I'll be there by dinner tomorrow. Don't say anything to him."

But Frieda had not been joking. "Okay," she said.

Back in 1962, Frieda's mother had discovered Ralph at his investigations of Frieda and punished them both. That was the end of it. "I don't know if you heard about it," Frieda continued to the crack in the cabin door. "We were grounded the same as if we'd been caught stealing, or gotten bad grades at school."

Nothing stirred. The first storm had left five inches of snow. The spigot worked briefly at midday and then froze up again in the afternoon. John Prade had been building fires to warm it; charred logs spiraled out from the spigot's entry point into the ground.

"Dad, he's coming tomorrow to explain things to you. He thinks I told you before now. You – and me, too, I guess – are standing in the way of his happiness. You know how he feels about happiness."

The evening was so still that the tiniest thing moving in the woods resounded like the snap of a leg breaking. It might have been a grouse, or a deer. A red squirrel. "Won't you explain something to me, Dad?" Frieda said. "It doesn't have to be why you're – here in my cabin. It's your cabin now. I mean, how about the theory of stresses in a skyscraper? Is there a formula you can rely on? How do architects know one building will stand and another won't? I never paid attention. I was – looking into the body. Did you know the body is like a building? That's what my anatomy teacher said."

Frieda was about to rest her forehead against the doorjamb when the door opened with a ripping sound. She jerked back and stood up. Inside, the orange lights of a fire pulsed through the slits in the woodstove door. She waited, frozen on her feet. It had begun to snow. John Prade appeared and motioned her to come in.

The smell of the room was chaotic with extremes: unwashed body and balsam fir. Food beginning to turn, and the thick, sweet odor of hot wax from a few struggling candles. Crushed old newspaper, clothing from the last century, pine needles, mouse and squirrel shit spilled out of a fresh opening in the south wall. Charred drips streaked the flanks of the woodstove, and the floorboards blurred under a new layer of grime.

Her father stood like a stake in the center of the room. His new beard was stone white, his eyes glassy, and his bearing absolutely erect. When he tugged off his black knit cap, as an antique gesture of respect, Frieda supposed, his hair whorled around his head as if windwhipped. She had not seen him face to face in a month.

"I have in mind my last design," John Prade said. His terrible glance swept the room. "Engineering the end of futurity."

There was no precedent for this moment. The most Frieda had ever directly addressed with her father were calculus problems when she was in high school, and that had not gone well. "That's – well, that's just crazy sounding, Dad."

When he swung his gaze to her, Frieda had to squint, as if at sun thrown off bright metal.

"You are not the only one who wants to get away," he said. "I'm just not coming back."

Their eyes were the same color: blue gray. She stared back. "Why?"

"Look at you, sequestered up here like a nun. And why is that? Because you can't abide the smell of your own life. You and that boy-husband. Just try to tell me you'll ever get over it."

Frieda couldn't draw a full breath. "You – you and Mom. What a lie – "

A wretched, dry smile. "Exactly correct."

"And all those years I thought your disdain a kind of love."

"So you see."

Frieda thought her whole body might fly apart. She felt she could kill and find it good. "So I see WHAT? That you're giving up on life because you made a wreck of it? That you can't do without Mom because she gave you someone to blame for your misery?"

"One side of the arch hates the other and pushes on it like a bull. That keeps the roof up."

"What – you and Mom were a roof?"

"What do you think, Miss Frieda? Did we keep the rain away and your brother out of your drawers?"

Fiction by Dorian Gossy

Frieda felt the death-chill in odd places: the palms of her hands, inside her elbows, deep inside the curving walls of her hipbones, like cramps from the menstrual periods she'd stopped having over a year before. "I should never have told you that."

A green log in the woodstove hissed a long time as its sap boiled. John Prade considered his black knit cap as he crushed it in his hands. "Only a corroborating detail."

He was on a plane about to crash, alive but doomed. "I don't have a mother anymore," she said. "I'm not ready to give up a father, too."

"I'm sorry for that."

"Don't leave me." It cut Frieda's throat to say it. She'd said it to Lane.

"But I will. Sooner or later. It can't be helped."

Frieda went to her father and put her arms around him. He caught her in his arms like a lover. Never had they hugged so, not at graduation nor wedding day nor funeral. His body was as bony as a tree and smelled of rank, wild things. She loathed tears, but they poured down her face. She released him, and he stepped away.

"I still don't understand," she said, scrubbing her cheeks with a wool glove.

"I think you do." John Prade fetched and drew on another coat – the oversized one Frieda had given him his first night in the cabin – over the three he was already wearing. "You could have turned the dogs loose on me long ago."

"I still can, you know," Frieda said. "I only wanted you to be mine."

John Prade went to the door and opened it. The frigid air poured its leaden weight into the room. "I will be, from out there," he said, pointing at the thickening snowfall and to the blurring trees beyond. "If you'll let it."

Frieda dove into her parka and found her footing on the snowy porch. "I'll get Ralph. I'll get searchers."

But her father had already shut the door behind her.

By the time Ralph arrived three hours later, the snow was falling fast. He stooped to kiss his sister but caught himself and drew away after an awkward shoulder squeeze. Without removing his coat, he retightened his scarf and turned up his collar, turning to Frieda. "You know, I'm sorry about that – fooling around back then. Did I ever apologize?"

"No. But never mind about that now," Frieda said. She'd finally washed her hair, which cascaded damp and loose down her back. "It doesn't matter anymore."

Ralph returned in an hour. John Prade was gone. "Could he be somewhere else?" Ralph asked. "Could he be in the house? The garage?" When Frieda shook her head, Ralph threw his arms wide. "No footprints, of course – not with snow falling this hard. You got snowshoes?"

They searched until thwarted by a muddled, gray twilight boiling with snowflakes. Inside the house, they smacked snow off their clothing and kicked it out of their boots. Frieda could see Ralph trembling once he shed his parka. "That place he lived in – did you see it?"

"Yes."

"How did the walls get all charred like that? Somebody must have tried to burn it down. The bed scorched. No furniture. All the walls gutted? The insulation torn out? Did he do that?"

Frieda had avoided a direct glance at the cabin while they were searching, as though it were a former friend encountered on the street whom she'd badly

failed. "It was never in the greatest shape," she said.

Ralph found a beer in the refrigerator, one of a six-pack their father had never finished. He downed it quickly. "We need to call rescuers, maybe the sheriff. You got their number?"

The snow had fast laid down a six-inch blanket. In her mind's eye, Frieda watched her father moving through the furred quiet of the snowy woods. He'd be dressed in his tinker's layers, but his topcoat, the one Frieda had left for him his first night, would flap around unbuttoned. He'd have his shovel and his axe. He'd go up into the wildest part of her land, beset with blackberry brambles and willow and infant poplar all struggling for daylight. Nothing particular would mark the place where he would dig himself the grave that a trespassing dog will discover and bark at for hours the following spring. After taking what must have been days opening a hole in snow and hard ground only a week from freezing, John Prade will lie down in the hole, and cover himself well with dirt. In the spring, defying her attempts to shoo him away, the dog will hover with unmistakable sorrow as Frieda brushes aside the leaves and dirt enough to behold her father's corpse in its hole, his flesh and clothes merging with loam, before she closes the hole above him again.

This strange dog, black with white question marks over each eye, will always appear when she approaches the grave with a weekly offering of river stone that she will bring in by wheelbarrow through the muddy forest. The dog will mark the swelling cairn with his urine each time, gazing at her regretfully as he holds his leg back like a salute. Covered by duff and stone, his scent masked by dog pee, John Prade will belong to Frieda at last.

All I want to know is where Dad is, Ralph will say one day to the wide, lush field during a visit that summer, drinking gin on her porch. He will have had a full menu of law enforcement search and fail to find from Lake Champlain to Montreal. As Ralph's wife and teenagers play badminton on the coarse grass, he will look heavenward. Can't I just know where he is?

No, Frieda will think. No, you may not. And something like bliss will fill her.

Memoir by Stanley Rosen

Leo Strauss in Chicago

I first met Leo Strauss when I was nineteen years old and a student in the College of the University of Chicago. It was the spring of 1949 – this was during the epoch of the presidency of Robert Maynard Hutchins, when the University was at the height of its glory. At that time, the College was famous for the eccentricity and precociousness of many of its students, and also for its highly unusual custom of allowing entering students to take examinations on the basis of which they were assigned course requirements. The intention of this program was to extend the time we spent in graduate school, provided that we already possessed the necessary foundation. It was therefore possible to graduate with a B.A. from the College in less than a month of residence. Apparently, the graduate of a Swiss private lycée accomplished this some years after my departure. In 1949, though, the record was one year, which was matched by eighteen members of my class, including myself and my classmate and friend Seth Benardete.

Another peculiarity of the College was that one could enter it at any age, and among my classmates were a number of virtual children. I still remember a party given by some of the older students. There, I entered into conversation with a man who seemed to be in his mid-thirties, a guess that his thick glasses and advanced baldness only strengthened. He informed me that he had broken with Catholicism and, thanks to a recent visit to Europe, with existentialism as well. I first inquired whether he was an instructor at the University, and then a graduate student. He replied in the negative to both queries and informed me that he was an undergraduate. "How old are you?" I asked. "Thirteen," he replied. I should add that when I arrived in 1948, I was, relatively speaking, an old man. I had been admitted to the College following graduation from high school in 1947, but I had chosen to live in New York for a semester, under the mistaken impression that I was a burgeoning novelist.

Stanley Rosen is University Professor and Borden Parker Bowne Professor of Philosophy at Boston University. He is the author of numerous books, including "Nihilism: A Philosophical Essay" (1969), "The Limits of Analysis" (1985), "The Ancients and the Moderns: Rethinking Modernity" (1989), "The Elusiveness of the Ordinary" (2002), and, most recently, "Plato's Republic: A Study" (2005). He is currently at work on a memoir of Alexandre Kojève and Leo Strauss.

© 2006 by Stanley Rosen

By the time I arrived in Chicago, my vocation had shifted from fiction to poetry. If I am not mistaken, I was the only one of Leo Strauss's long-term students who came to him from poetry. I was also virtually uninterested at the time in politics, unlike the majority of Strauss's students. Instead, I was an avowed metaphysician, who had elaborated a philosophical position partly influenced by T. S. Eliot, one of whose main tenets was that philosophy and poetry are two different languages about the same world. In addition to these intellectual propensities, which most of Strauss's students regarded as deficiencies, I was undisciplined in the academic sense and spent most of my time writing poetry, with some professional success and with reasonable hopes for future progress. These hopes were sustained by Hayden Carruth, who was then the editor of *Poetry Magazine*, and Henry Rago, who was about to assume that position, but also by Allen Tate, who taught in the College for a year.

High on my list of things that I had no intention of doing was to become a professor of philosophy. To my adolescent vision, being a philosopher and a professor were incompatible, and besides, I regarded myself as already a philosopher. Needless to say, this was about to be changed by my encounter with Leo Strauss.

I had a number of unusual classmates during my year in the College. Perhaps the most interesting was Seth Benardete, who went on to become Strauss's favorite student. Benardete stands out in my memory as a spirit of genuine distinction and, even at that early age, of rare scholarship. At the time my friends and I assumed that Benardete would go on to a distinguished career as a classical philologist, as in a sense he did. But he wrote his books in so oblique a style that he was widely ignored by the orthodox classical establishment, with some important exceptions, including Pierre Vidal-Naquet.

In 1949 he was for me a formidable exotic. I remember vividly to this day a long conversation we had one night in his dormitory room during which Benardete informed me that he regarded it immoral to love a human being. As a youth with a certain proclivity to this form of immorality, I was incredulous and asked him what we should love. He replied in a magisterial tone: "Greek vases." This struck me as the most sophisticated view I had ever heard, but a view with one flaw: it was nonsensical.

My friendship with Benardete, whom I saw virtually every day during that first year in Chicago, was my only real preparation for my first meeting with Leo Strauss. I was a poet, a romantic, and a metaphysician, who had somehow wandered into the lair of the philosopher, the classicist, and the historian. There was for me no quarrel between philosophy and poetry, as there apparently was (albeit in a subtle form) for Benardete and for Strauss, who both followed Plato. The atmosphere around Benardete was redolent of Socratic irony and continental sophistication, whereas I represented something quite different. One of the entering examinations in the College required us to write an essay describing our philosophical position. Afterwards, a member of the philosophy department told me that my views were Fichtean, something of which I had never heard. A poet of Fichtean leanings was not in the best position to meet either the young Benardete or the middle-aged Strauss, to say nothing of my distinct deviation from the paradigm of the Aristotelian gentleman.

I should say at once that Strauss was not at all a snob, and that his conception

of decorum was quite reasonable. He was quite right to note in the margin of a first draft of my doctoral dissertation that I liked to *épater le bourgeois*. His astringent follow-up – "I could wish that the entire dissertation had been written in the style of paragraph 2 on page 153" – taught me more about scholarly writing than a dozen texts on hermeneutics. Strauss's own style, at its best, comes very close to the appropriate blend of the daring of thought veiled by the web of prudence. He was nevertheless capable of flexibility in selecting his students. Many years after I had left Chicago, I encountered an old professor and former colleague of Strauss's who was noted for his elegance and aristocratic tastes. This colleague, a minor Latvian baron, told me that he used to complain to Strauss about my youthful uncouthness, to which Strauss would reply each time: "He's getting better." I owe my education to this willingness to overlook baronial standards.

Strauss had recently arrived in Chicago from New York and was, at the time, unknown to most of the Chicago student community. This may help to explain his charitable reception of so unpromising a potential student. Strauss's stepson, who was also a student in the College, arranged the meeting. I had been preparing an honors paper on a Yiddish writer named Achad Ha'am. Strauss was a professor of political science, but his son told me that his father was also an expert on Yiddish writers, and asked if I would like to consult him on my paper. I agreed and set out, having been warned that Strauss would probably give me twenty or thirty minutes of his time.

It was a warm spring evening, and mosquitoes filled the humid air. Strauss received me in shirtsleeves, gesturing with a cigarette holder as if it were a baton. He was a rather short man with a thin, high-pitched voice. His initial demeanor was polite but understandably reserved. He opened the conversation by asking me what I did. I replied, "I am a poet." Strauss immediately inquired whether I knew what Plato says about poets. To this I answered something like, "I don't care what Plato says about poetry. I am a poet, and I understand it better than he does."

This drove Strauss like an uncoiled spring from the easy chair in which he had been sitting, and he paced up and down the room, gesticulating with his cigarette holder, as if trying desperately to bring an unruly orchestra back to orderly response. I will not attempt to reproduce the entire conversation, which lasted for at least two hours. At the end of it, he invited me to become his student.

I respectfully declined, as I was planning to return to New York for another go at the literary life. When I told Strauss that I intended to study at the New School for Social Research, he suggested that I mention his name. I did and was promptly awarded a scholarship, so great was Strauss's reputation at the New School. My experience in New York, however, proved unsatisfactory, and I decided to return to Chicago in 1950 in order to study with Strauss.

In the intervening year, Strauss had attracted the attention of a number of very gifted students, among them Benardete, Hilail Gildin, Victor Gourevitch, Muhsin Mahdi, and Allan Bloom. At or shortly after this time, Richard Rorty began to attend Strauss's lectures, but left to take his doctorate at Yale. Despite his subsequent adventures with analytical philosophy and postmodernism, Rorty had great respect for Strauss and the best members of his circle.

I will not attempt to give a complete list of my contemporaries who studied with Strauss. Let me say only that the students were divided into two main groups: the political scientists and the members of the Committee on Social Thought. There were only a few members of the philosophy department, including myself (until 1952, when I transferred to the Committee), as well as a steadily increasing selection of visitors from a variety of faculties, both at Chicago and elsewhere, who had been attracted by news of the pied piper of the Midway. Eventually Strauss's audience included several priests, of whom perhaps the most interesting was Ernest Fortin.

One could also divide the students in a very general way into those who were primarily interested in American politics and those who were students of classics or one of the major periods in the history of political philosophy. This is, of course, not a rigid classification, but it is not entirely nebulous. Whereas all of us, I suspect, regarded ourselves as engaged somehow in the pursuit of wisdom, there is a difference between constitutional law, the *Federalist Papers*, or the fact-value distinction in contemporary political science, on the one hand, and Plato's analysis of the soul, Ibn Khaldun's philosophy of history, or Rousseau's anthropology, on the other.

In view of the subsequent explosion of interest in Strauss's politics, it is worth mentioning that several of his closest students were originally Communists. One of them told me that as a boy scarcely into his teens, he would walk the streets of Manhattan saying, "Someday this will all be mine!" Others, however, were, and often remained, Socialists or New Deal Democrats.

Strauss made no attempt to alter their political views. I doubt he was even aware of most of our political orientations. What he did was teach us how to read, and how to think about what we had read. Very far from producing extremist reactionaries, of which he is often accused today, he presented us with the path of moderation and practical wisdom. It is a tribute to Strauss's tolerance and brilliance as a scholar and a teacher that he could serve to unify a band of investigations as diverse as those just mentioned, undertaken by students of the most diverse convictions.

Strauss also led private reading groups devoted to topics that he could not conveniently teach in the political science department, such as Maimonides' *Guide of the Perplexed* or Hegel's *Logic*. For obvious bureaucratic reasons, most of his students were not writing a dissertation on 'pure' philosophy, in the academic sense of that term. Most were officially political scientists, not specialists in causality, ontology, or German idealism, and, of course, not in the philosophy of science, epistemology, or the foundations of mathematics, which lay outside Strauss's (and most of our) competence. Instead, Strauss, the great enemy of historicism, was, at first sight, training historians of political thought and political scientists who were prepared to use that history as a foundation for rethinking the cardinal tenets of their discipline. From this standpoint, he was engaged in a radically reconceived version of the Heideggerian 'destruction' of Western philosophy – that is, a radically new close reading of canonic philosophical texts – with two massive qualifications: First, Strauss was primarily concerned with politics rather than ontology; and second, his archaeological excavations were designed to return us to the thoughts of the heroes of the Western tradition, that is, to the thoughts as these heroes had thought them and not, as in Heidegger's case, to the ostensibly deep-

er and unthought thoughts that constituted the authentic *Seinsgeschichte* of Western metaphysics.

Despite these important differences, both Strauss and Heidegger were engaged in the enterprise of 'uncovering' a concealed truth. It is not much of an exaggeration to say that for both thinkers the crucial figure in this excavation was Plato. For Strauss as for Heidegger, moreover, the appropriation of Plato for the reconstitution of late modernity entailed a type of strong interpretation that Strauss did not sufficiently emphasize, perhaps because doing so might have tied him too closely to Heidegger's even more extreme form of critical interpretation. One can surmise that Strauss's reticence on questions of this sort was part of his program to inoculate his students against Heidegger. Nevertheless, it is true that Strauss was deeply impressed by Heidegger, especially by the Heidegger of the early and middle period. I am certain that Strauss learned much from Heidegger about how to read a Greek text, not to mention that he also assimilated through Heidegger Nietzsche's critique of modernity and nihilism.

Heidegger had radicalized Nietzsche's critique of Enlightenment by extending it to Plato; his intention was to go behind or above Platonism to another way, a way entirely free of the presumably reified and subjectivist thinking of Western Platonist metaphysics. From this exalted standpoint, the Enlightenment was itself a version of Platonism and something to be overcome. Strauss, on the contrary, took us back through the history of philosophy to Plato, not in order to overcome the modern Enlightenment but to find its mistakes and to correct these in the name of a genuine liberalism and freedom of thought.

Heidegger has sometimes been called a liberator while Strauss is often denounced for his conservatism. Both judgments are largely nonsense unless they are given careful qualification. Heidegger's conception of radical freedom has nothing to do with Western European liberalism. Strauss was, no doubt, a social conservative, but his political views were, in my opinion, designed to compensate for the exaggerated decline of modern liberalism into nihilism. His fundamental goal was to preserve philosophy, and it is from this point that serious arguments about his politics must begin. Otherwise stated, he undertook to preserve philosophy with the political tools of classical liberalism.

Strauss's form of liberalism is perhaps most obvious in his popular writings, which contain an eloquent defense of modern political freedom and a critique of Marxism. To the extent that this form of liberalism had deteriorated into what we came to call 'postmodernism,' Strauss of course responded as a conservative. Perhaps one could say that Strauss would have preferred to argue with Georg Lukács than Jacques Derrida. But he was altogether more flexible and more moderate than his close friend, the radical Alexandre Kojève, who accepted the bankruptcy of the West up to the Napoleonic counterrevolution, but who also accepted its purification by his own post-Hegelian fusion of Hegel, Feuerbach, Marx, and Heidegger.

When I arrived in Chicago for my 'second sailing' in the fall of 1950, I entered as a graduate student in the department of philosophy. Officially my mentor was Richard McKeon. But my serious education took place in Strauss's seminars or during conversations in his office or home. As time went by, I began to detach myself from the philosophy department, spending time in the Chicago art museum instead of attending

philosophy seminars regularly. But the serious part of the day was devoted to listening to or speaking with Strauss, and immersing myself in the texts that were being analyzed in his graduate courses. Interestingly enough, my status in the philosophy department improved under this regime, and just when I had decided to shift to the Committee on Social Thought, I was offered financial assistance to begin doctoral study in philosophy.

The story of my shift from the philosophy department to the Committee deserves mention because it illustrates the difference in admission procedures between bureaucrats and what would be called in France the *grands seigneurs*. I was a waiter at the University faculty club, where most of my customers were scientists like Enrico Fermi, Harold Urey, Edward Teller, and Leo Szilard. More immediately accessible to me was the eminent sociologist Edward Shils, a member of the Committee. I had a number of conversations with Shils, mainly about topics like how to cook oxtail soup in a pressure cooker. As a result, he asked me if I would like to join the Committee. Acting on the assumption that an organization with an interest in oxtail soup could not be all bad, I accepted. The second stage was an interview with the classicist David Grene, which I began by spilling tea over his orange tweed suit. The net result was a fellowship. At no point was I asked to supply references or to fill out forms and provide a statement of purpose, nor did Strauss raise these formalities. Direct contact supervened over conventional bureaucracy.

During the first two years of my study at Chicago, then, as a member of the philosophy department, I was, in a real sense, a man without a country, moving back and forth across the frontier with forged papers. One could call this a practical application of Strauss's understanding of esotericism.

Recounting my first impressions of Leo Strauss inevitably raises a question that has played a puzzling and, to be frank, irritating role in my professional life. It is also one of those aspects of the relation between teacher and student that is most difficult to explain. But it is important to get right for the sake of the general validity of my portrait of Strauss. Stated crudely, the question is whether I was then, or am now, a 'Straussian.'

My first inclination upon hearing this question is to reply, "Do you refer to Johann or to David Friedrich Strauss?"

In a less frivolous vein, one can say that the expression 'Straussian' has many counterparts among academics; one finds the students of all charismatic teachers being described as 'Hegelians,' 'Marxists,' 'Heideggerians,' 'Wittgensteinians,' even 'Quineans.'

Strauss himself liked to quote Nietzsche to the effect that the best thing a student can do for his teacher is to cut the umbilical cord that connects them. I will never forget a conversation with Strauss during the last year of my stay at Chicago, when we discussed the question of my finding a teaching position. "Disown me!" Strauss said, smiling hugely and straightening up in his swivel chair to facilitate the punctuation of his remark with the inevitable flourish of his cigarette holder. By this advice, Strauss did not mean to suggest that I forget or dishonor everything that I had learned from him, but that I do what is necessary to carry out his deepest teaching: live the life of a philosopher.

Let me therefore say at once that there were no 'Straussians' among the inner circle of Strauss's best students. Those who could legitimately be called such

were, to put it bluntly, second-raters. There was, of course, a certain area of agreement about the solidity of Strauss's scholarship; if this had not been so, it would have been folly to study under his supervision. We were all convinced that he was right about the tradition of esoteric writing – that is, the fact that a great many early modern philosophers, writing under conditions of censorship and persecution, felt compelled to write in a style that had to be carefully decoded in order to separate the true meaning from the superficial message of their texts. As a corollary, we accepted the need to read serious philosophical works written at least before the French Revolution with a kind of Talmudic eye. Even more important, we felt as a direct force the erotic strength of Strauss's spirit, and we were ourselves 'turned around' (to use Socrates' metaphor of the *periagōgē tēs psuchēs*) by that strength in a way that goes beyond inspiration to a reattunement of the soul and an opening of the eye of the intellect. This is something that does not come from reading books, but only from direct contact with a great teacher. And a great teacher is one who encourages critical analysis of his own views.

I am myself an example of the fact that Strauss did not demand unquestioning obedience from his students. I admired him enormously, and in due course I came to revere him as a rare blessing, without whose training and guidance my life would have been seriously diminished. But I was not from the outset, nor did I ever become, a 'Straussian.' I put this word in scare-quotes to indicate that it is a pseudo-term employed by ideologues as an excuse to avoid serious thought, or as a mask for their own ignorance. No one could deny that Strauss had his own views, and he was not quick to assume that his students knew more or had thought more deeply than he. He was correct in both respects, of course. But he always accepted modifications or addenda to his interpretations when the evidence warranted it, and he welcomed competent, or at least sincere and well-argued, disagreement.

In my own case, I developed reservations about Strauss's general orientation on three main points: his position on the famous quarrels between philosophy and poetry; between the ancients and the moderns; and, finally, between Athens and Jerusalem (in other words, between reason and revelation).

On the first point, paradoxically enough, I might cite Strauss himself as my authority: he often interpreted Plato to imply that poetry is, in fact, one of two necessary components of the philosophical nature, the other being mathematics. If we think this through, it leads us directly toward the role of strong interpretation, and also of Kantian constructivism, in modern philosophy; this, in turn, reopens the problem of historicism, which we cannot resolve simply by recourse to common sense or the popular affirmation of natural right.

On the second point, I have always leaned toward the moderns in their famous quarrel with the ancients. Stated as simply as possible, the ancients, and Plato in particular, protect humankind from nature whereas the moderns take a progressively more aggressive stand toward freedom. I regret to say that Nietzsche's remark, first called to my attention by Strauss, is correct: "Advice whispered into the ear of a conservative. Man is not a crab." In other words, we must march forward through the semidarkness of nihilism. This forward march may fail, but it is inevitable. The serious question is how to keep up the morale of the marchers.

At this juncture, I want to interject a relevant observation about Strauss's critique of modernity. Although it is similar to that of Heidegger, it differs from his in a significant respect. There is virtually no emphasis upon the problem of technology, certainly none in the style of Heideggerian ontology. Instead, we are given numerous criticisms of *methodology*, particularly of the methodology of modern social science, with its mathematical model of rationality. A critique of methodology is useful, but it does not go to the heart of the matter. Whatever we may think of it, Heidegger's analysis of *technē* does go to the heart of the matter. I am certainly no Heideggerian, but I have studied him closely for fifty years, and I have been sometimes surprised and always struck by how much Strauss learned from Heidegger, but also by how much Heideggerian depth he sacrificed. It was, of course, part of Strauss's deconstruction of Heidegger to move back to the surface as a preparation for the descent into the depths. But one must in fact descend, and this Strauss did not do, whether from conviction or lack of theoretical strength.

With respect to the third quarrel, I have always found questionable Strauss's statement on the mutual irrefutability of reason and revelation. To pose the question of irrefutability already gives an edge to philosophy: the evaluation of arguments for and against religion is not so much religious as philosophical. At the same time, to repudiate reason is not to refute it. (I am reminded of Strauss's regular reference to an observation of Plutarch, that whoever asks, "What is a god?" is already an atheist.)

Perhaps Strauss's greatest accomplishment, at least for me, is that he inoculates the young against Heidegger's speculative excesses and hermeneutical brutality. His major shortcoming is that he is neither a metaphysician nor a poet. Many will regard this as a compliment. I shall not debate the point, but simply record it as a difference between Strauss and me that did not at all interfere with my great admiration for him nor diminish all that I have learned from him. He certainly understood this difference. Also, there is no doubt that he tried to mitigate my own metaphysical and poetical excesses, just as there is no doubt that he was right to do so. What needs to be emphasized is that, whereas my contemporaries and I would not have chosen to study with Strauss had there not been unmistakable evidence of his superior gifts, this choice was never in the best cases a passport to discipleship.

And yet, in the 1950s, there was already something like a Strauss 'school.' I have already noted that this phenomenon is not at all uncommon among philosophers, and it is simply bad faith to criticize Strauss because he had many students who admired him enormously. The animus addressed toward Strauss, and, in a general sense, 'Straussians,' had to do with the substance of his teaching, not with his success in attracting students. Strauss was disliked for his critique of the sacred cows of modern social and political science, in particular, Max Weber (whom he highly admired), and because of his rediscovery of the tradition of esotericism, a tradition known to every well-educated scholar until the late nineteenth century, and one that is referred to by thinkers of the rank of Descartes, Leibniz, Condorcet, Hume, Kant, Lessing, Renan, and Nietzsche, to mention only a few of the most prominent examples from modern times. Too many professors, in a radical distortion of genuine liberalism, condemned Strauss's views on esotericism because they were ignorant of the evidence.

In speaking of Strauss's bad reputation, I have to come back to the question of conservatism. Strauss was widely condemned as a conservative, whereas someone like W. V. Quine, who was at least as conservative as Strauss, was not. I still remember a full-page ad in the *New York Times* shortly before the decisive outbreak of the Watergate scandal, defending then-president Nixon, and signed by a large array of well-known academics, including Strauss – and Quine. It was not self-evident then that support for Nixon was the equivalent of a declaration of fascism, as people so widely assume today. I say this as one who was at the time a New Deal Democrat who despised Nixon. Be that as it may, I never heard Quine being accused of fascism, nor, indeed, was his political position widely discussed in academic circles, although it was certainly known. The key, of course, is that Quine was a 'technical' philosopher. One could separate his logical views from his politics, just as, to move to the other end of the political spectrum, one could do in the case of Noam Chomsky's linguistics. In other words, Strauss's reserve with respect to the application of scientific models to the study and interpretation of human life struck his critics as a reactionary repudiation of modernity in a way that was not associated with conservative logicians or, say, physicists.

As a result, Strauss became an immediate target of opprobrium for the liberal academic establishment. He was challenging the theoretical and methodological soundness of modern social science, which claimed to represent the scientific Enlightenment and the progress of the human race. This challenge, incidentally, was the attenuated or surface version of Heidegger's critique of technicism. In Strauss's version, the challenge was manifestly political, whereas in Heidegger's case, one could claim to speak as an ontologist or seeker after Being, and thereby be permitted to detach the ontological question from contingent political appropriations. This is very much like saying that the attempt to reduce the study of human nature to a branch of mathematics and physics – or, in today's idiom, neurophysiology and electrical engineering – has no political implications and is leading us to a radically extended form of Enlightenment – as if that were not itself a political position of utmost force and seriousness. By analogous reasoning, Heidegger's French followers associated his attack on modernity in general, and the Enlightenment in particular, with a doctrine of liberation and creativity, two of the favorite principles of the majority of late-modern thinkers. Thus, Heidegger was assimilated into the left-wing interpretation of Nietzsche. This camouflage could never be sustained in Strauss's case.

One more remark on this topic. Just as Quine was forgiven his conservatism because of his technical orientation in logic, so Heidegger was often forgiven his obnoxious political views on the grounds that he was an ontologist, or 'thinker,' and so naive in practical affairs. Some of his students went so far as to blame Heidegger's politics on the influence of his wife, who was the daughter of a Nazi general. Those who dispense such nonsense are nevertheless touching the surface of a serious problem, that of the relation between theory and practice. Strauss attributes to Lessing the 'concealed' view that "all practical or political life is essentially inferior to contemplative life, or that all works, and therefore also all good works, are 'superfluous' insofar as the level of theoretical life, which is self-sufficient, is reached." Speaking about Thucydides, Strauss says, apparently

in his own voice, that "wise men will always be inclined to see in political life an element of childishness."[1] One can reply that evil works are also superfluous for theoretical life, but this does not answer the question of the ground of good works. The decency of the philosopher seems to be a contingent character trait. If morality is exoteric, could one not argue that philosophy is the most dangerous of all gifts?

Rather than develop further the tangled motives for Strauss's great unpopularity as the rediscoverer of esotericism, let me close by raising the question of whether he used the right rhetoric for his own time and place. On balance, I prefer Hegel's treatment of the same question in his lectures on the history of philosophy. Hegel, we recall, denies the claims, current in his own time, that Plato practiced esotericism, on the ground that the philosopher cannot philosophize with his ideas in his pockets. Immediately after, however, Hegel says that philosophy is, by its nature, esoteric. In order to show the truth of this judgment, however, one must actually philosophize, and this means with one's ideas on the table, not concealed in one's pockets.

This is especially true in times during which one need not employ esotericism in order to preserve one's life or liberty. In such a time, it is incumbent upon the philosopher to present as forceful and detailed a statement of his or her doctrines as is humanly possible. Only this will attract the best young minds to a genuine philosophical encounter, assuming, as did Strauss, that philosophy was in radical decline. Perhaps one may suggest that Strauss, very far from practicing esotericism himself, as is widely believed today, was on all crucial points extraordinarily frank. Straussian frankness, of course, is not the same as Nietzsche's. In Strauss's case, frankness can itself be a form of esotericism.

1 The first quotation is from Leo Strauss, "Exoteric Teaching," in *The Rebirth of Classical Political Rationalism*, selected and introduced by Thomas L. Pangle (Chicago and London: University of Chicago Press, 1989), 64. The second is from Leo Strauss, "Thucydides: The Meaning of Political History," in ibid., 74.

Joel F. Handler

on welfare reform's hollow victory

Almost a decade has passed since President Clinton fulfilled his pledge to "end welfare as we know it" by signing the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA). PRWORA, commonly known as 'welfare reform,' replaced Aid to Families with Dependent Children (AFDC), a government cash-aid program that primarily supported single mothers and their children, with Temporary Aid for Needy Families (TANF), a block-grant program that increased state discretion substantially.

The goal of the reform was to promote work in the paid labor force among welfare recipients. Within five years, 80 percent of the states' caseloads had to be working or searching for jobs. And for the first time in welfare history, strict time limits were imposed on individual aid: a two-year limit on continuous aid and a cumulative lifetime limit of five years.

Most states adopted a 'work first' strategy – recipients had to take any job, even a low-wage, entry-level one – rather than providing recipients opportunities for education and training first. The assumption was that working would improve the economic and social outcomes of poor families: even workers who started at low-wage jobs would be able to work their way up the economic ladder. Policymakers believed that these mandatory requirements were necessary to move recipients off the rolls. Furthermore, they thought that the children of welfare recipients would gain a sense of pride and direction from seeing their parents working.

Everyone has declared the reform a success. Caseloads have declined from 12.2 million to 5.3 million recipients. Labor-force participation has increased 15 to 20 percentage points among single mothers with children under the age of five – the population most likely to be welfare recipients. Welfare reform, a burning political issue since the 1970s, has disappeared from the radar screen for almost a decade.

But this reform has actually resulted in a hollow victory. While caseloads have declined and work participation has increased, most families are still living in poverty and enduring significant hardship.

Joel F. Handler, a Fellow of the American Academy since 2004, is Richard C. Maxwell Professor of Law and professor of policy studies at the University of California, Los Angeles. An authority on social-welfare law and poverty, he has conducted numerous studies on poverty, political participation, and administration of justice. He is the author of "The Poverty of Welfare Reform" (1995), "Down From Bureaucracy" (1996), "Social Citizenship and Workfare in the United States and Western Europe" (2004), and, most recently, "Blame Welfare, Ignore Poverty and Inequality" (with Yeheskel Hasenfeld, 2006).

© 2006 by the American Academy of Arts & Sciences

Studies of 'leavers,' recipients who have left welfare, have shown that between half and two-thirds of leavers find jobs shortly after leaving the rolls, primarily in sales, food preparation, clerical support, and other service-sector jobs. Their wages range from $5.67 to $8.42 per hour, though, leaving many full-time working families impoverished. Benefit loss is also a significant problem. Although leavers still qualify for other government programs, including food stamps and Medicaid, they have to go to separate offices and make separate appointments (often available only during working hours) to obtain them. Thus, the majority who are eligible are not enrolled.

As a result, leavers, despite earnings, are not better off than when they were on welfare. Evaluations of eleven state welfare-to-work programs have demonstrated that the net income of recipients remained the same or decreased despite the fact that employment increased. Few of the eleven programs generated substantial gains in incomes or declines in poverty.

Besides the loss in benefits, working also creates costs such as childcare and transportation. Perhaps the most significant is childcare. A survey of childcare centers in forty-seven urban areas revealed that in over half of the areas surveyed, childcare cost an average of $6,000 annually. Thus, a two-parent family in which both parents earn the minimum wage spends "one-quarter or more of their income to afford average-priced care for an infant in 37 out of 47 urban areas surveyed," according to a 2003 report of the Children's Defense Fund. During the 1990s, state budget surpluses ensured that a substantial amount was spent on child-care subsidy programs and other supports for the working poor. But even now, more than one-third of states have long waiting lists for these programs – over two hundred thousand in California, forty-eight thousand in Florida, twenty-two thousand in Georgia, and thirty-eight thousand in New York City. Subsequently, most low-wage working families rely on informal childcare – relatives and friends who often also have jobs – creating an unending scramble to find childcare in a world of shifting schedules and flexible, temporary jobs. Many children are even left alone or in the care of older, but still young, children.

While the government has helped alleviate poverty for low-wage workers by considerably expanding the Earned Income Tax Credit (EITC) in the 1990s (federal spending on EITC is almost three times that of TANF and more than that of food stamps), the working poor still face tremendous challenges. With an average wage of about $8 per hour, even working thirty hours per week for a full year grosses only $12,000, and federal and state income payroll taxes offset a sizable amount. Although families may earn more income by working rather than relying on cash aid, their reduced receipt of other government benefits and increased work-related expenses mean that these families are not necessarily better off.

This discussion has focused thus far on employed leavers. However, more than one-third of leavers are unemployed. Sarah Brauner and Pamela Loprest have found that approximately 6 percent of unemployed leavers are disabled, sick, or otherwise unable to work. Approximately 69 percent of those without disabilities are looking for work. And 14 percent of unemployed leavers rely on the earnings of a spouse or partner as income. Like employed leavers, very few of them utilize other government benefits, including unemployment, food

Note by Joel F. Handler

stamps, and Medicaid. In any case, a high proportion of leavers, regardless of employment status, reported suffering: 33 percent cut down on or skip meals; 39 percent are unable to meet rent, mortgage, or utility payments; and 7 percent end up moving in with others.

Unfortunately, the pressure to meet the new requirements has also exacerbated welfare agencies' use of sanctions and diversions to discourage existing and potential welfare recipients. Sanction rates in various programs now range from 20 percent to over 60 percent of the caseload. Reports reveal that clients have been sanctioned for minor errors: a missed appointment, a failure to file a form, or a lapse in administrative procedures. In a California study, recipients either didn't know that they had been sanctioned (since grants often fluctuate) or why. Moreover, many states have failed to accept applicants. Cook County, Illinois, for example, only accepted 19 percent of applicants. In several states, caseworkers used informal tactics to discourage applications, such as requiring applicants to engage in multiple job searches or to go on multiple interviews before even considering their applications.

The addition of work requirements has undoubtedly complicated an already overworked, undertrained bureaucracy. PRWORA assumed that with the new work requirements, increased state flexibility, and block grants, welfare agencies would turn into employment agencies offering individualized services. This has not happened. Instead, the work programs have been a mostly unwelcome addition. For most agencies, whether public or private, their main task is still accuracy and timeliness in determining eligibility and distributing payments. Though the titles, rules, and programs of the offices and its workers have changed, most employees say that their jobs have remained basically the same. They have received little training in the new rules and programs.

Further, contracting with private agencies (nonprofit and for profit), touted as a cure for the problems of welfare administration, has not helped. Advocates claimed that, with competition, services would be delivered more efficiently, quality improved, and government made more flexible. Thus, in 2001, over $1.5 billion of TANF and state funds went to private contractors, according to 2002 GAO estimates. But contracting has given birth to its own problems: contractual incompleteness, asymmetry of information, difficulty of government monitoring, lack of competition, dependence of government on contractors, and opportunistic behavior, among others.

In fact, because of the difficulty of calculating the cost of services, various pay-for-performance, cost-reimbursement, and fixed-price contracts are used. Thus, contractors have more incentive to reduce caseloads, rapidly move clients into the job market, and prevent their return. Outcome measures, particularly of the quality of service, are especially problematic. Evidence now shows that contractors are not more efficient than public agencies. Indeed, evaluations have revealed the limited qualifications and poor training of staff in these agencies; the poor quality of services; a high use of sanctions; a diversion of funds for corporate promotion; and a lack of improvement in client earnings. Yet public officials have had to justify the existing contracts because contractors have become a powerful interest group.

Instead of declaring victory, it's time to recognize the severity of poverty and hardship in the United States. Despite having one of the highest average in-

comes, the United States has the astonishing distinction of having one of the highest rates of poverty in the industrialized world. More ominously, it also has the highest child-poverty rate compared to fourteen Western European countries and Canada. Our problem is the assumption that once a family crosses the 'poverty line,' everything is all right.

However, the poverty line is seriously out of date. In 2000, Jared Bernstein and his colleagues constructed a budget just to meet "basic needs and achieve a safe and decent standard of living." This budget included food, housing, health care, transportation, childcare, and other necessary expenses (e.g., telephone, clothing, personal care, household items, school supplies, television, etc.). It did not include restaurant meals (including fast food); vacations; movies; and saving for education, retirement, and emergencies. They found that in nineteen different locations, this budget ran well above the poverty line. In Baltimore, for example, a family of four would have to have an annual income of $34,732 just for basic needs – twice the amount necessary to cross the poverty threshold. A single-parent family, with the parent working full-time, a seven-year-old, and a three-year-old, would need $30,000 to live – again twice the poverty line. And the National Low Income Housing Coalition recently reported that a one-bedroom apartment at average market rates anywhere in the country was not affordable for a full-time minimum-wage worker.

The poor are 'playing by the rules,' but not enough full-time jobs paying decent wages and benefits are available. To raise a family of four even above just the poverty line, one would have to earn $9.36 per hour, working full-time (thirty-five hours per week for fifty-two weeks). Yet nearly one-third of heads of households earn less than $10 per hour. The state of poverty in low-wage America – with millions of children lacking quality childcare, health care, and education – is not only unjust but national suicide. Welfare reform is not the answer. Adequate income support for all families and decent jobs for the vast majority of parents who want to work are the preferred solutions. Instead, TANF has been reauthorized, increasing the mandatory work requirements for two-parent families.

William A. Galston

on the reemergence of political pluralism

As recently as a century ago, the forms of government known in classical antiquity – empire, monarchy, aristocracy, democracy, and mixed regimes – still existed. Moreover, people regarded them as legitimate. However, in the aftermath of the political instability that World War I unleashed across Europe, new forms of government representing radical challenges to the status quo emerged. Leading thinkers lent their support to fascism and communism, in part because these new forms promised to solve economic and social problems that parliamentary governments could not.

But with World War II the range of political legitimacy began to narrow dramatically. Fascism was defeated on the battlefield and also discredited as a morally respectable form of politics. The subsequent founding of the United Nations initiated the development of universal norms of governance. Article 21 of the Universal Declaration of Human Rights specified that "the will of the people shall be the basis of the authority of government" and that "this will shall be expressed in periodic and genuine elections which shall be by universal and equal suffrage." For the first time in history, the nations of the world officially endorsed republicanism as the sole source of political legitimacy.

To be sure, this norm has often served to cloak authoritarian regimes – some oxymoronic (e.g., Indonesian President Sukarno's "guided democracy"), others almost laughably hypocritical (e.g., the "Democratic People's Republic of Korea"). Nonetheless, these examples show that governments, however repressive, now feel compelled to justify themselves in republican terms.

It is not surprising then that this development changed political theory as well. At one time, theorists had weighed the relative merits of government by the One, Few, or Many – or, in more contemporary terms, the rule of the Leader or the vanguard party versus democratic elections. But after World War II, debates about political theory focused increasingly instead on differences among forms of democracy – elitist versus populist, representative versus direct, consensus driven versus agonistic, deliberative versus bargaining based, substantive versus procedural.

The Universal Declaration of Human Rights also made it clear that legitimate

William A. Galston was most recently Saul I Stern Professor of Civic Engagement and director of the Institute for Philosophy and Public Policy at the University of Maryland School of Public Policy. He is now a senior fellow in governance studies at The Brookings Institution. Among his many publications are "Liberal Pluralism" (2002), "The Practice of Liberal Pluralism" (2005), and "Public Matters" (2005). He has been a Fellow of the American Academy since 2004.

© 2006 by the American Academy of Arts & Sciences

governments are not only democratic in form but also limited in scope. Not even democratic majorities can deny individuals freedoms of thought, conscience, religion, and association, among many others. Notably, the document made no effort to specify the basis on which human beings are endowed with rights that constrain the legitimate exercise of public power. The U.S. Declaration of Independence famously characterized this matter as self-evident, a stance that leaves theorists unsatisfied. Since sincere democrats disagree among themselves about the range of legitimate public authority, appeals to unargued consensus are unlikely to get very far.

Many contemporary democratic theorists endorse the view that only the norms of democracy itself circumscribe the scope of properly constituted democratic authority. Others espouse an alternative – a pluralist account of limited government. According to this view, social life comprises multiple forms of activity and association. Political association is but one of these forms. Others include families, civil associations, religious faith and faith communities, the pursuit of knowledge and knowledge-seeking organizations, and the activities of solitary reflection and conscience. Each of these forms has an identity partly independent of the others, and each generates claims to authority, no one of which is dominant in all spheres, for all purposes, on all occasions.

This is not a new idea. It takes as its point of departure an account of politics found in the writings of British political pluralists such as John Neville Figgis, G. D. H. Cole, Harold Laski, and thinkers working in the Reformed Calvinist tradition. Nor is the argument between pluralists and their adversaries a recent development. The pluralist movement began to take shape in the nineteenth century as a reaction to the growing tendency to see state institutions as all-powerful or total, a tendency that took various practical forms in different countries: French anticlerical republicanism, British parliamentary supremacy, and the drive for national unification in Germany and Italy against subordinate political and social powers.

The reemergence of political pluralism

Historically, one can discern at least three distinct secular-theoretical arguments for political monism rather than pluralism. (Theological arguments, which raise a different set of issues, are beyond the scope of this note.) The first is the idea, traced back to Aristotle, that politics enjoys general authority over subordinate activities and institutions because it aims at the highest, most comprehensive good for humans. The *Politics* virtually begins with the proposition that "all partnerships aim at some good, and that the partnership that is most authoritative of all and embraces all the others does so particularly, and aims at the most authoritative good of all. That is what is called . . . the political partnership."

Thomas Hobbes offered a second justification: Any less robust form of politics would, in practice, countenance divided sovereignty, an open invitation to conflict and civil war. Thus, sovereignty cannot be divided, even between civil and spiritual authorities. In Hobbes's view, this undivided sovereign authority must also have unlimited power to decide whether, and under what conditions, individuals and associations would enjoy liberty of action. No entity, individual or collective, can assert rights against the public authority. Indeed, civil law may rightfully prohibit even the teaching of truth, if it is contrary to the requirements of civil peace.

Jean-Jacques Rousseau inspired a third argument for political monism: a society

Note by William A. Galston

cannot achieve civic health and morality without its citizens' wholehearted devotion to the common good. Loyalties divided between the republic and other ties, whether to civil associations or to revealed religious truth, dilute civic spirit. Likewise, the liberal appeal to private life as against public life only legitimates selfishness at the expense of the spirit of contribution and sacrifice, without which the polity cannot endure. Representing this tradition, Emile Combes, a turn-of-the-century premier in the French Third Republic, declared, "There are, there can be no rights except the right of the State, and there [is], and there can be no other authority than the authority of the Republic."

It is in the context of these claims that political pluralism emerges as an alternative to monism. Political pluralism rejects efforts to understand individuals, families, and associations simply as parts of a political whole. Nor does it accept the argument that they are merely instrumental, existing 'for the sake of' some political purpose. For example, religion is valuable, not only for the contribution it may make to politics and society, but in its own right, and there is no guarantee that religion faithfully practiced will always support the existing political or social order.

Instead, political pluralism regards human life as consisting of a multiplicity of spheres, some overlapping, but each with distinct inner norms and a limited but real autonomy. Moreover, it rejects any account of political community that creates a hierarchy of these spheres. Rather, different forms of association and activity interrelate in a complex manner.

Furthermore, political pluralism does not seek to overcome, but rather endorses, the division of human loyalty and political authority created by the rise of revealed religion, summarized in Jesus' maxim "Render unto Caesar what is Caesar's and unto God what is God's." Obviously, this division creates problems of practical governance. But pluralists refuse to resolve these problems by allowing public authorities to determine the substance and scope of allowable belief, or by elevating devotion to the common civic good as the highest human value. They recognize instead that fundamental tensions rooted in the structure of human existence cannot be abolished in a stroke but must rather be acknowledged and adjudicated with regard to specific circumstances and controversies.

In short, pluralist politics is a politics of recognition rather than of construction. It respects the diverse spheres of human activity; it does not understand itself as creating or constituting those activities *ex nihilo*.

However, a pluralist polity is responsible for coordinating other spheres of activity, especially in adjudicating the inevitable overlaps and disputes among them. This form of politics requires the mutual limitation of some freedoms, individual and associational. For example, it monopolizes the legitimate use of force, except in cases of self-defense, when the polity cannot or does not protect its members. It also foresees the possibility of group tyranny and therefore protects the individual against some associational abuses. But a pluralist polity also presumes that the enforcement of basic citizenship and exit rights, suitably understood, will usually suffice. Indeed, it respects that groups have a broad, though not unlimited, right to define their own membership, to exclude as well as include.

While political pluralism undoubtedly has moral implications, it rests on empirical claims about the diverse forms of

human activity and association. For example, in most cultures, family ties limit the scope of political authority. Sophocles' *Antigone* revolves around the primordial imperatives of kinship that stand opposed to the imperatives of patriotic loyalty. In the story, the king, Creon, attempts to keep Antigone from burying one of her brothers. The fact that her brother was slain in battle fighting against his own city does not justify per se Creon's effort to prevent her from burying him. To be sure, Antigone is as deaf to Creon's legitimate concerns for his city as he is to her family obligations. One could make a case for Creon's stance. Still, Sophocles clearly implies that the disaster that befalls Creon is the result of his assertion of political authority beyond its rightful bounds. Whether we ultimately agree with Sophocles or not, *Antigone* contains this basic message: political leaders should not disregard nonpolitical values completely.

American constitutional law also endorses the proposition that family ties restrict political authority; parenthood is its own sphere of authority. For example, in the famous case *Pierce v. Society of Sisters*, the Supreme Court denied the right of a public authority – in this case, the state of Oregon – to require all parents within its jurisdiction to send their children to public schools. In justifying its stance, the Court declared, "The fundamental theory of liberty upon which all governments in this Union repose excludes any general power of the State to standardize its children by forcing them to accept instruction from public school teachers only. The child is not the mere creature of the State; those who nurture him and direct his destiny have the right, coupled with the high duty, to recognize and prepare him for additional obligations."

Of course, this does not mean that family activities are immune from political regulation. *Pierce* explicitly recognized a substantial degree of legitimate governmental regulation of the family, including its educational decisions. The deep theory of liberty on which the Court relied allows the state to require "that all children of proper age attend *some* school." It is the task of so-called pluralist casuistry to distinguish legitimate from illegitimate assertions of political authority over families.

Consider another example. U.S. law and jurisprudence restrict the sway of public authority over religious associations. It is a commonplace that these associations may establish their own criteria for their religious offices, notwithstanding general public norms of nondiscrimination to the contrary. For example, public law cannot invoke these norms to compel the Catholic Church to admit women into its priesthood. But the limits are even stricter. As Laurence Tribe observes, courts and other agencies of the U.S. government "may not inquire into pervasively religious issues." The rationale for this restriction goes beyond prudential fears of entanglement and political divisiveness. It reflects, as well, the belief that doctrinal and scriptural interpretation is beyond the competence and rightful authority of political power.

At first glance, this position, and the pluralist thesis it buttresses, might seem to be a classic example of cultural particularism masquerading as universalism. Yes, Jesus instructed his followers (and adversaries) to render unto Caesar what is Caesar's. But don't all the Abrahamic faiths actually tend in principle toward theocracy, to which each has succumbed at points in its history? And isn't this theocratic thrust alive today – among Protestant fundamentalists who call for

Note by William A. Galston

the official designation of America as a Christian nation, among messianic Jews in Israel who invoke God's will in support of their settlements and against the authority of their own government, among the Wahhabi clerics in Riyadh and the Shia clerics in Tehran?

Pluralists respond that each of these faiths contains resources that validate distinctions among different spheres of authority in social life. Both Christianity and Judaism developed doctrines that drew lines between civil and religious authority, a stance that eventually defeated the entrenched Constantinian tradition within the Catholic Church. And Islam, which traditionalists interpret to require the rule of God's law in every aspect of life, nonetheless endorses propositions that open a path toward pluralism. Article X of the Universal Islamic Declaration of Human Rights begins by quoting the Quranic principle that "there is no compulsion in religion," which it insists governs the religious rights of non-Muslims. Therefore, "in a Muslim country religious minorities shall have the choice to be governed in respect of their civil and personal matters by Islamic Law, or by their own laws."

Like every other kind of political theory, political pluralism is contestable. But its guiding insight, anchored in widely recognized features of human life, offers an appealing basis for liberty and diversity. The pluralist wager is that this approach will prevail, not only over civic totalism, but also over theories that find it easier to assert rights than to justify them.

Letters to the Editor of *Dædalus*

On history in the twentieth century

May 10, 2006

To the Editor:

Anthony Grafton's essay, "History's postmodern fates," in the Spring 2006 issue of *Dædalus* was interesting and informative, but I share neither his pessimism about the fate of history nor his limited view of important twentieth-century developments. In summarizing other trends among historians in the United States, I will also focus on the post–World War II period. More works deserve mention if readers are to avoid a restricted view of history in this period.

Grafton covers major works by microhistorians, especially Carlo Ginzburg, Natalie Davis, and Robert Darnton, and discusses some contributions by Lawrence Stone and Emmanuel Le Roy Ladurie. He mentions other giants of history like William McNeill and C. Van Woodward, but not their main contributions. I can suggest some additional accomplishments and controversies.

Postwar politics – from HUAC investigations and loyalty oaths to mass movements against war and for civil, women's, and gay rights – affected the personnel and content of historical studies. A leftward turn among students, with many choosing academe over more lucrative professions, changed history departments from their former largely conservative white male character.

The geographical expansion of history brought most of the world into history courses. It also created new views of world and comparative history. World historians like William McNeill, Leftan Stavrianos, and Marshall Hodgson brought differing novel approaches to many aspects of world history, a growing field that challenged the privileging of the West over the rest that characterized earlier writings and theories regarding global history. Elements of Marxism were important in post–World War II schools founded by social scientists but adopted by many historians – Dependency Theory, which originated in Latin America, and World Systems Theory, which was begun by Immanuel Wallerstein and divided the modern world into a changing core, a semi-periphery, and a periphery, with the former exploiting the latter. World-oriented historians have often focused on trade, conquest, and migration as transnational forces: Many study trade regions – the Mediterranean, the Indian Ocean, the Atlantic, and the so-called Silk Route. Migration and conquest sometimes create diasporas, another topic of research.

While European history remained a pioneer, the rapid development of non-Western history brought them a new range and sophistication that could have had more impact had more historians paid attention. Joseph R. Levenson brought a subtlety and sophistication to Chinese intellectual history matched only by Benjamin I. Schwartz. Another historical giant was Thomas C. Smith.

His early works analyzed the development of industry and agriculture in modern Japan, finding many parallels to Western Europe. Smith later applied demographic methods pioneered by the Cambridge Group in England and found significant parallels to Western trends.

Indian and South Asian history have flourished, with an emphasis on theory – often postmodern and/or postcolonial – which have been welcomed by some and contested by others like Richard Eaton. Partha Chatterjee's view of the nation and nationalism is popular among historians of the global South.

Studies of major historical empires have also become significant. Historians of the Middle East have revised many earlier views, especially of the Ottoman Empire. For example, there has been a reaction against the view of Ottoman decline from the late sixteenth century on, with many writing only of a relative decline as compared to that of the European Northwest. The imperial harem, often blamed for the alleged decline, has been rehabilitated in Leslie Peirce's *The Imperial Harem*, which also revises views of how the empire was governed.

History has expanded to cover countries without writing or with writing systems that did not cover as much as did those of much of Eurasia. Methods for using oral history were pioneered by Jan Vansina; historical linguistics to measure migration and material culture, by Christopher Ehret and others; and the use of living languages to supplement incomplete writings, by James Lockhart. Such methods are mainly used for Africa and the Americas.

U.S. history since the 1950s has expanded in many directions, notably revision of views of slavery, Reconstruction, and black history: beginning with the works of W. E. B. DuBois and John Hope Franklin, and then Kenneth Stampp's *The Peculiar Institution*, which overthrew the dominant view of benevolent slavery, and proceeding through the works of Eugene Genovese, which saw slavery as part of an interlocking social and cultural system. Recent examples include Ira Berlin's *Generations of Captivity* and Eric Foner's works on Reconstruction. David Brion Davis has pioneered in many, including cultural and comparative, aspects of the history of slavery.

Robert Fogel and Stanley Engerman brought a novel mathematical economic history approach to slavery in *Time on the Cross*. Some economic historians followed their combination of 'what if' history with mathematics into 'counterfactual history,' while others pursued the mathematics without the 'what if.' Economic historians who came from history departments rather than economics departments tended rather to stress social, geographic, and cultural factors.

The intersection of economic history with ecological and demographic history has flourished in recent decades. For instance, many works on disease have followed Woodrow Borah's work on Latin America's drastic population decline owing to the conquerors' diseases. Alfred Crosby's *Ecological Imperialism* is a leading work on ecological history.

Among subjects crossing geographical boundaries, women's history stands out for its contributions and for encouraging the development of other fields – gender, family, male, and gay history. Gerda Lerner, Natalie Davis, and Joan Scott are among its several pioneers. In gay history, George Chauncey's *Gay New York* goes beyond its title in discussing varieties of homosexual culture and practice. General histories can no longer ignore gender, though their mode of incorporating it is not always sophisticated or adequate. The study of women and gender has extended to all parts of the world:

three winners of the relevant AHA Joan Kelly prize since 1997 were books about China, Syria, and Iran.

Also popular have been histories of human relations to the nonhuman – often to commodities. There has been a rise in writing about consumption and a relative decline in writing about production, although both the history of technology and of modes of production have received some attention.

In various fields number crunching, aided by computers, has yielded information about the social and economic history of people for whom we have scant records. Other novel methodologies, such as the Freudian psychohistory of Peter Loewenberg, also have adherents. Many historians use new media, chiefly audio and visual records of individuals, events, and material culture.

The history of science and medicine has become a major field in many history departments. Thomas Kuhn's seminal *Structure of Scientific Revolutions* was written when he was in UC Berkeley's History Department. In addition to the usual intrafield controversies, there are disagreements between scientists, who often see their history as a progressive discovery of scientific laws and theories, and historians, who stress the social and cultural aspects of science.

On the theoretical side, a few books with different approaches have been influential, notably, E. J. Carr's *What is History?* which combined a quasi-Marxist approach with relativism toward what is important in different periods. For more relativist postmodernists, Hayden White's *Metahistory* is important, while Joyce Appleby, Lynn Hunt, and Margaret Jacobs' *Telling the Truth about History* combines a partially postmodern approach with others.

The students I have known have been interested in controversy and in diverse approaches to understanding the past, and the writers I have known have been enthusiastic, not bored. Surveying the varieties of twentieth-century history seems more a cause for optimism than the opposite.

– Nikki R. Keddie

Nikki R. Keddie, a Fellow of the American Academy since 1994, is professor emerita of history at the University of California, Los Angeles. She has written numerous publications, including "Modern Iran: Roots and Results of Revolution" (new ed., 2006) and "Women in the Middle East: Past and Present" (2006). Keddie was founding editor of "Contention: Debates in Society, Culture and Science" from 1991 to 1996. A longer version of this letter with citations is at http://nikkikeddie.blogspot.com/2006/05/history-writing-in-us-since-world-war.html.

June 1, 2006

Anthony Grafton responds:

One point of my piece was that argument is the lifeblood of history. Professor Keddie's response nicely bears this out. In fact, I didn't set out to survey the field, and my original draft was in any case much compressed to fit this issue of *Dædalus*. So I don't propose to argue with Professor Keddie, who offers a different point of view and much supplementary information. I only wish I could share Professor Keddie's optimism about the condition of our discipline – especially as I too work with many gifted and eager students at all levels. But there too, it's salutary to have two points of view expounded.

On peace in the Middle East

May 8, 2006

To the Editor:

If William B. Quandt is to have any credibility as an expert "on the peace process in the Middle East" (published in the Spring 2006 issue of *Dædalus*), he should at least know that the first name of former Israeli Prime Minister Barak is Ehud (the same as current Prime Minister Olmert's) and not Aharon. He fails to mention that among those blaming Arafat for the failure of the Camp David talks were not only Clinton and Ross, but Saudi Arabia's Prince Bandar. While mentioning the unfounded speculation that Arafat's death was due to poison, he fails to mention the more plausible speculation that it was due to AIDS. He conveniently ignores the fact that the charter of Hamas declares that the elimination of Israel as a political entity is an unalterable principle that cannot be amended or moderated. He joins the far right in Israel when he blames rather than praises Ariel Sharon for turning over the entire Gaza Strip to the Palestinians. It is naïve in the extreme to believe that "peace" can be achieved when one of the parties refuses to recognize the other party's right to exist.

– Solomon W. Golomb

Solomon W. Golomb, a Fellow of the American Academy since 2003, is University Professor at the Communication Sciences Institute of the University of Southern California.

May 16, 2006

William B. Quandt responds:

Solomon Golomb is of course correct that former Israeli Prime Minister Barak's first name is Ehud. I stand corrected. The rest of his intemperate letter is either a matter of opinions on which we differ, or a significant distortion of what I actually wrote. I'm not sure that anyone knows what caused Yasir Arafat's death, but it is certainly accurate to say that many Palestinians believe that he was poisoned. I have no knowledge or opinion on the matter.

More seriously, I fail to see how he can say: "He joins the far right in Israel when he blames rather than praises Ariel Sharon for turning over the entire Gaza Strip to the Palestinians." Even a casual reader of my short article will see that I neither blame nor praise Sharon for evacuating Gaza. Instead, I merely state that his decision seemed to reflect his "vision of the future – one that does not involve a negotiated peace with the Palestinians." That also is, I believe, an accurate statement of how Sharon saw the future, even before the election of a Hamas-led government.

Mr. Golomb seems very intense in his views, but that does not absolve him of the obligation to state my views accurately when he chooses to criticize them.

On teaching in European universities

May 14, 2006

To the Editor:

In their essay, "American philosophy in the twentieth century," which appeared in the Spring 2006 issue of *Dædalus*,

Professors Dagfinn Føllesdal and Michael Friedman write: "Unlike their European counterparts, American [graduate] students hand in written work several times a term and receive detailed comments from their teachers. European universities are too understaffed to give similar attention and feedback to their students . . . ." Of my own university, Oxford, this is completely false, in respect to undergraduates as well as graduate students. How am I to tell of how many other European universities it is equally false?

– Michael Dummett

Michael Dummett, a Foreign Honorary Member of the American Academy since 1985, is Emeritus Professor of Logic at the University of Oxford.

May 24, 2006

Dagfinn Føllesdal and Michael Friedman respond:

Sir Michael Dummett is right. Oxford, and also Cambridge, gives attention and feedback to its students similar to what students receive in good American universities. Unfortunately, this is not generally the case in European universities. While American colleges and universities, all 3,500 of them, have an average student/teacher ratio of ten students per teacher, similar to that of Oxford and Cambridge, most European universities have two to four times as many students per teacher with far less feedback for each student. And the situation in Europe is getting worse rather than better: Many countries are introducing a three-year B.A. without increasing teaching resources; East European countries, and also France and Germany, draw the best researchers out of the universities and offer them good research conditions, while the university teachers become second-class citizens; Germany increases the teaching load for those who remain in the universities from eight to nine hours of teaching per week. With policies like these, good European students will continue to go to America for their graduate work, and America will retain its dominance in science and scholarship.

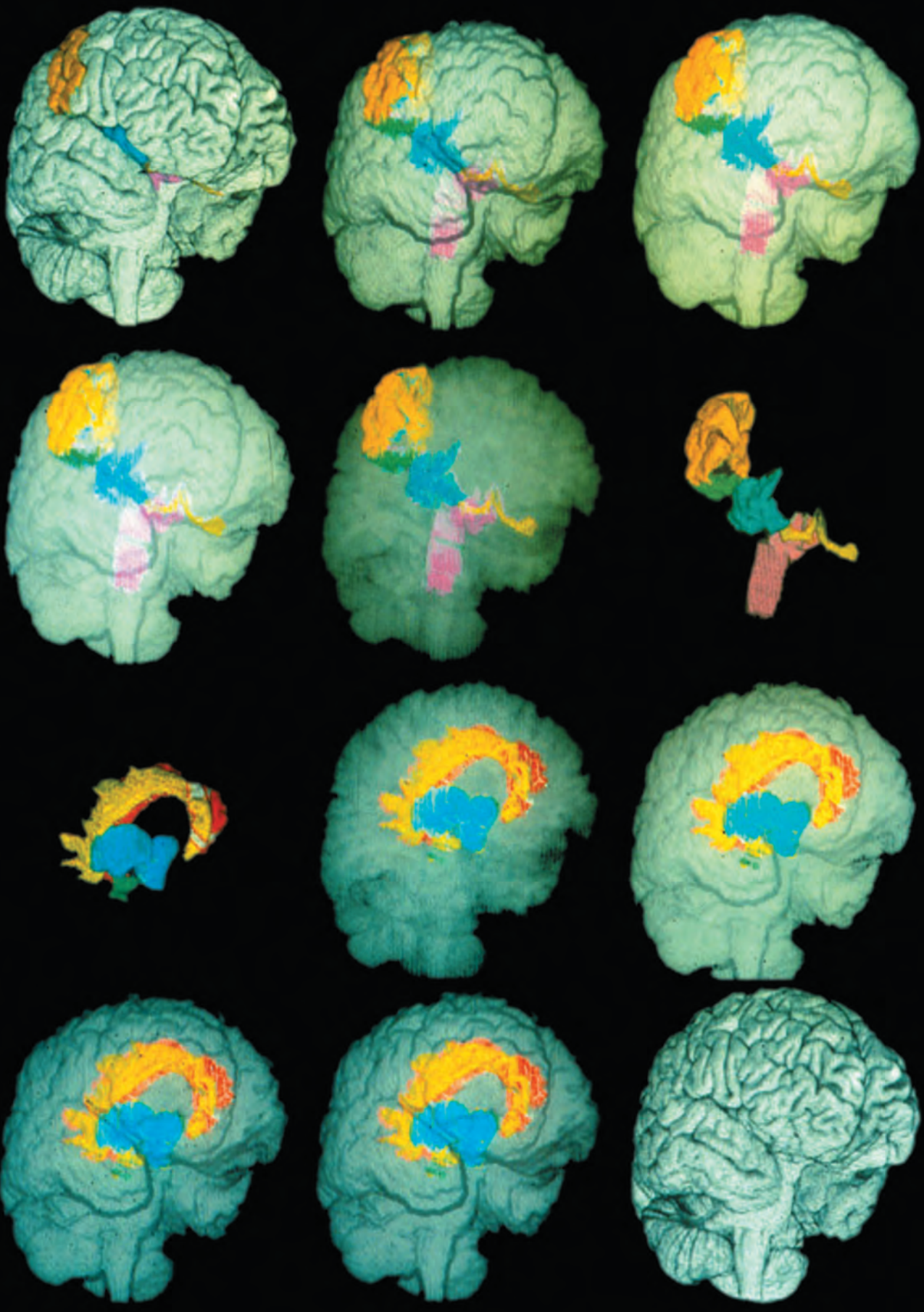
AMERICAN ACADEMY
OF ARTS & SCIENCES

President
Patricia Meyer Spacks
Chief Executive Officer
Leslie C. Berlowitz
Vice President and Chair of the Academy Trust
Louis W. Cabot
Treasurer
John S. Reed
Secretary
Jerrold Meinwald
Editor
Steven Marcus
Librarian
Robert C. Post
Vice President, Midwest
Geoffrey R. Stone
Vice President, West
Jesse H. Choper

Dædalus Board of Editors Committee
Steven Marcus, *Chair,* Joyce Appleby, Russell Banks, Stanley Hoffmann, Donald Kennedy, Martha C. Nussbaum, Neil J. Smelser, Rosanna Warren, Steven Weinberg; *ex officio*: Patricia Meyer Spacks, Leslie C. Berlowitz

Committee on Publications
Jerome Kagan, *Chair,* Emilio Bizzi, Jesse H. Choper, Denis Donoghue, Jerrold Meinwald; *ex officio*: Patricia Meyer Spacks, Steven Marcus, Leslie C. Berlowitz

Inside back cover: Many of the brain structures used to represent and manipulate the body and its workings perform double duty and also play a role in the construction of the self, a process indispensable for consciousness. Somatosensory cortices are shown in orange yellow; the cingulate gyri in yellow and red; the subcortical nuclei of the hypothalamus and brain-stem tegmentum in pink; and the combination of insular cortex and thalamus in dark blue. See Gerald M. Edelman on *The embodiment of mind,* pages 23–32: "As a result of higher-order consciousness enhanced by language, humans have concepts of the past, the future, and social identity." Photography courtesy of Hanna Damasio, MD, and the Dornsife Cognitive Neuroimaging Center, University of Southern California.



coming up in Dædalus:

on identity	Akeel Bilgrami, Wendy Doniger, Amartya Sen, Stephen Greenblatt, Kwame Anthony Appiah, Sydney Shoemaker, Joseph Koerner, Susan Greenfield, David A. Hollinger, Carol Rovane, Ian Hacking, Todd E. Feinberg, and Courtney Jung
on nonviolence & violence	William H. McNeill, Adam Michnik, Jonathan Schell, James Carroll, Breyten Breytenbach, Mark Juergensmeyer, Steven LeBlanc, James Blight, Cindy Ness, Neil L. Whitehead, Mia Bloom, and Tzvetan Todorov
on sex	Joan Roughgarden, Terry Castle, Steven Marcus, Elizabeth Benedict, Brian Charlesworth, Lawrence Cohen, Anne Fausto-Sterling, Tim Birkhead, Catharine MacKinnon, Margo Jefferson, and Stanley A. Corngold
on capitalism & democracy	Joyce Appleby, John C. Bogle, Lucian Bebchuk, Robert W. Fogel, Jerry Z. Muller, Richard Epstein, Benjamin M. Friedman, John Dunn, and Robin Blackburn
on life	Anthony Kenny, Thomas Laqueur, Shai Lavi, Lorraine Daston, Paul Rabinow, Michael S. Gazzaniga, Robert George, Robert J. Richards, Jeff McMahan, Nikolas Rose, John Broome, and Adrian Woolfson
on the public interest	William Galston, E. J. Dionne, Jr., Seyla Benhabib, Jagdish Bhagwati, Adam Wolfson, Lance Taylor, Gary Hart, Nathan Glazer, Robert N. Bellah, Michael Schudson, and Nancy Rosenblum

plus poetry by Kevin Carrizo di Camillo, John Kinsella, Charles Simic, Lawrence Dugan &c.; and notes by Michael Cook, Richard Morris, Susan T. Fiske, Daniel G. Nocera, Omer Bartov, Robert I. Rotberg, Samuel Weber, Harriet Ritvo &c.

U.S. $13; www.amacad.org