



# Dædalus

Journal of the American Academy of Arts & Sciences

Fall 2005

50 years

David Baltimore	Limiting science: a biologist's perspective (1978) 7
Lorraine Daston	Fear & loathing of the imagination in science (1998) 16
Steven Weinberg	Physics & history (1998) 31
Robert N. Bellah	Civil religion in America (1967) 40
Clifford Geertz	Deep play: notes on the Balinese cockfight (1972) 56
Robert M. Solow	How did economics get that way & what way did it get? (1997) 87
Thomas C. Schelling	Reciprocal measures for arms stabilization (1960) 101
John Hope Franklin	Two worlds of race (1965) 118
Jill K. Conway	Politics, pedagogy & gender (1987) 134
Tu Wei-ming	Cultural China (1991) 145
Amartya Sen	Indian traditions (1997) 168
Robert Frost	On Emerson (1959) 186
Umberto Eco	Modern & postmodern aesthetics (1985) 191
Susan Sontag	On beauty (2002) 208











*Inside front cover*: A late nineteenth-century postcard, showing a Fourth-of-July fête. See Robert N. Bellah on *Civil religion in America*, pages 40 – 55: “Holidays provide an annual ritual calendar for the civil religion.” Image © PoodlesRock/Corbis.

James Miller, Editor of *Dædalus*

Phyllis S. Bendell, Managing Editor  
and Director of Publications

Esther Yoo, Assistant Editor

Contributing Editors:

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

### *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 (Executive Officer)

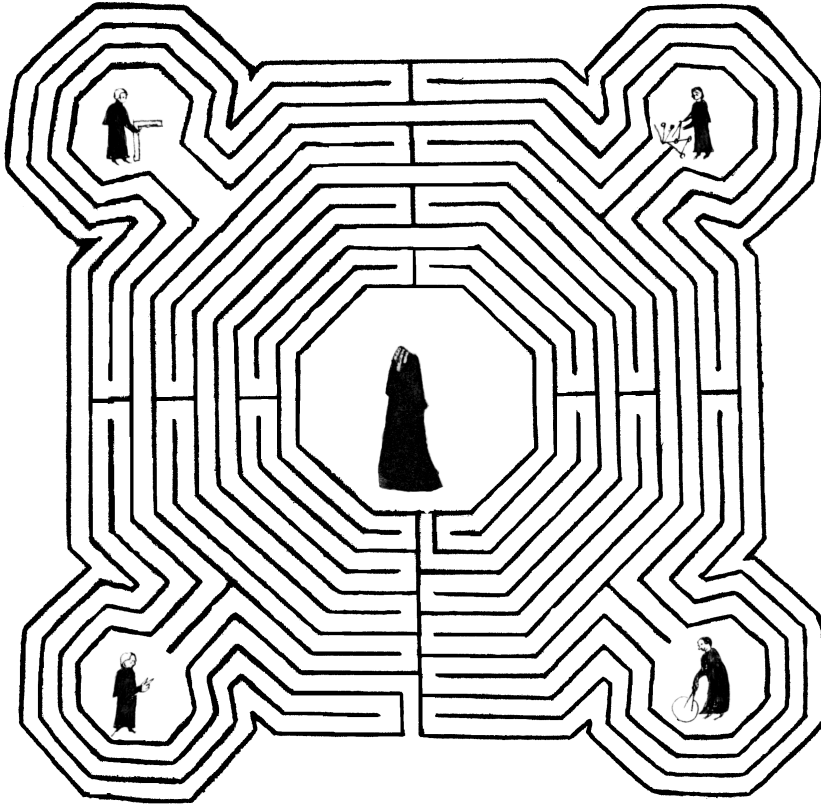
### *Editorial advisers*

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)

*Dædalus* is designed by Alvin Eisenman

# Dædalus

Journal of the American Academy of Arts & Sciences



The pavement labyrinth once in the nave of Reims Cathedral (1240), in a drawing, with figures of the architects, by Jacques Cellier (c. 1550 – 1620).

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* Fall 2005  
Issued as Volume 134, Number 4

© 2005 by the American Academy  
of Arts & Sciences  
*An argument about beauty*  
© Susan Sontag

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 0-87724-051-5

*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.

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

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.

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.

Subscription rates: Electronic only for non-member 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 0011-5266/05. 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.

This year marks the 50th anniversary of *Dædalus* and the 225th anniversary of the American Academy of Arts & Sciences. *Dædalus* was founded as the Journal of the American Academy of Arts & Sciences in 1955 and established as a quarterly in 1958.

*Dædalus* takes its place in a series of distinguished Academy journals and publications. The first and longest-lived was the *Memoirs of the American Academy of Arts & Sciences*, which consisted largely of learned papers submitted to the Academy for publication. Appearing first in 1785, the *Memoirs* were published, apart from a few interruptions, until 1946, when they were replaced by the *Bulletin of the American Academy of Arts & Sciences*. In the meantime, the Academy had begun to distribute a second publication, the *Proceedings of the American Academy of Arts & Sciences*, which reported on the intellectual activities of the Academy. Eighty-five volumes of the *Proceedings* appeared between 1846 and 1958, and like the *Memoirs* they were superseded by the *Bulletin*. The *Bulletin* continues today to report on the meetings and activities of the Academy.

*Dædalus* from the start was meant to reach the broadest possible reading pub-

lic. This was no easy task. By 1955, after all, specialization had beset every field of the arts and sciences. *Dædalus* was designed, in part, to remedy this situation. Over the years, the journal has evolved into a preeminent forum for interdisciplinary work, complementing the Academy's other activities by bringing a variety of specialists from every field of endeavor into ongoing contact with educated readers from all walks of life.

To celebrate these two anniversaries – the 225th of the Academy and the 50th of its journal – we have selected a few essays that reflect the quality and range of what *Dædalus* has published over the years. As a whole, these essays express the continuing aspiration of the Academy, in the words of its eighteenth-century founders, “to cultivate every art and science which may tend to advance the interest, honor, dignity, and happiness of a free, independent, and virtuous people.”

James Miller, Editor of *Dædalus*

Leslie Berlowitz, Executive Officer  
of the American Academy of Arts  
& Sciences

September 2005





# David Baltimore

## *Limiting science: a biologist's perspective*

Contemporary research in molecular biology has grown up in an era of almost complete permissiveness. Its practitioners have been allowed to decide their own priorities and have met with virtually no restraints on the types of work they can do. Viewed as a whole, the field has not even met with fiscal restraints because, relative to “big science,” molecular biology has been a relatively cheap enterprise.

Some of the funds that fueled the initial, seminal investigations in molecular biology were granted because of the medical implications of work in basic biology. Most continue to come from agencies concerned with health. Although basic research in biology has yet to have a major impact on the prevention and treatment of human disease, a

backlash already seems to be developing in which various groups in our society question whether the freedom that has characteristically been granted to research biologists by a permissive public requires modification. Among the numerous elements prompting this questioning are impatience with the lack of practical results, and fears that direct hazards might result from the experimentation, that basic research may not be an appropriate investment of significant funds, and that dangerous new technologies may flow from discoveries in basic biology. Lay critics as well as a few members of the profession have argued that molecular biologists should concentrate their efforts in certain areas of research, like fertility mechanisms, while other areas, like genetics or aging, are possibly dangerous and certainly not worthy of financial support from the public. These critics believe that they can channel contemporary biology to fit their own conception of appropriate research.

I wish to argue that the traditional pact between society and its scientists, in which the scientist is given the responsibility for determining the direction of his work, is a necessary relationship if basic science is to be an effective endeavor. This does not mean that society is at

---

*David Baltimore, president of the California Institute of Technology, has been a Fellow of the American Academy since 1974. This essay appeared in the Spring 1978 issue of “Dædalus.” At the time of its publication, Baltimore was American Cancer Society Professor of Microbiology at the Massachusetts Institute of Technology. In 1975 he received the Nobel Prize in Physiology or Medicine.*

---

© 2005 by the American Academy of Arts & Sciences

the mercy of science, but rather that society, while it must determine the pace of basic scientific innovation, should not attempt to prescribe its directions.

What we call molecular biology today had its origins in individual decisions of a small number of scientists during the period from the late 1930s through the early 1950s. These people were trained in diverse fields, among them physics, medicine, microbiology, and crystallography. They created molecular biology out of the realization that the problems posed by genetics were central to understanding the structures of living systems. No one channeled them towards this line of thinking, no one cajoled them to tackle these problems; rather, their own curiosity and sense of timing led them to try to elucidate these mysteries. This history provides a model of how the most effective science is done.

It is partly the successes of molecular biology that have brought on the questioning of whether scientists should be allowed their freedom of decision. It is therefore worthwhile tracing the development of concerns about whether certain areas of science should be closed to investigation. Molecular biology is the science that has revealed to us the nature of one of the most fundamental of all substances of life, the gene. There is a very simple underlying reality to the transmission of characteristics from parents to children: a code based on four different chemicals, denoted by the letters A, T, G, and C, is used to store the information of heredity. The order in which these four chemicals appear in a virtually endless polymer, called DNA, is the language of life.

Knowing that DNA was the physical storehouse of the genes, Watson and Crick in 1953 first solved the problem of how the DNA is organized to assure that information is transferred with almost

perfect reliability from parent cells to progeny cells. They showed that DNA is made of two strands intertwined together into a helix and also that the four subunits in DNA are physically related so that only A and T form a specific pair as do G and C. The two strands are held together by this specific pairing so that the two strands carry redundant information. Each gene is thus really a gene and a mirror-gene; this self-complementarity allows DNA to repair itself – and therefore maintain the fidelity of its information – and also to duplicate itself, sending identical copies into the two daughter cells resulting from a round of cell division.

Following the monumental discovery of the structure of DNA, many scientists have contributed to learning how to make DNA, how to read DNA, how to cut DNA, how to rejoin DNA, and in general, how to manipulate at will the genes of very simple organisms.

What good has all of this new knowledge brought to the average person? First and of great importance is the contribution that scientific advances make to human culture. The continued accumulation of knowledge about ourselves and our environment is a crucial cultural aspect of contemporary life. Science as well as art illuminates man's view of himself and his relation to others. Our knowledge of how we work, how one person differs from and is similar to another, what is health and what disease, what we need to support health, and so on, helps to set the ground rules for the debates of politics and the productions of art.

The more practical benefits of biology can be expected to come in the future in the form of medical advances, or increases in food production, or in other manipulations of life processes that will be able to provide positive contributions

to civilization. But molecular biology, for all of its power as a basic science, has not been easily translated into tangible benefits. This is a situation that could change very soon. New discoveries are rapidly bringing molecular biology closer to an ability to affect the lives of the general public.

Of the advances that have occurred, a critical one has been the development of a process called recombinant DNA research. This is a technique whereby different pieces of DNA are sewn together using enzymes; the chimeric DNA is then inserted into a bacterium where it can be multiplied indefinitely. Because the method allows genes from any species in the world to be put into a common type of bacterium, there is a theoretical possibility of hazard in this research. The potential for unforeseen occurrences led a number of scientists, including me, to issue in 1974 a call for restraint in the application of these new methods. We were addressing a limited problem, whether there could be a recognizable hazard in the performance of certain experiments. That limited question opened a floodgate; other questions came pouring out and are still coming. They have led to front-cover stories in weekly magazines, to serious attempts at federal legislation, to a demoralization of some of the community of basic research biologists, and, most significantly, to a deep questioning of whether further advances in biology are likely to be beneficial or harmful to our society.

Much of the discussion about recombinant DNA research has centered on whether the work is likely to create hazardous organisms. The mayor of Cambridge, Massachusetts, raised the specter of Frankenstein monsters emerging from MIT and Harvard laboratories, and speculations about the possibility

of inadvertent development of a destructive organism like the fictitious Andromeda Strain have been much in the news. I am personally satisfied that most of such talk is simply science fiction and that the research can be made as safe as any other research. The people who understand infectious diseases best make the arguments most strongly that recombinant DNA research is not going to create monsters. But rather than defend my judgment that the safety issue has been blown out of proportion, I want to consider some of the more general issues that have been raised by the controversy.

If safety were the most important consideration behind the debate about recombinant DNA, then we might expect the debate to focus on the hazards of doing recombinant DNA experiments. Instead, many of the discussions that start considering such questions soon turn to a consideration of genetic engineering.

Two techniques labeled genetic engineering exist. Both originated because not everybody's genes are perfectly designed for the job of being a functioning human being, so that many instances of blood disorders, mental problems, and a host of other disabilities are traceable to a malfunctioning gene. It would be a triumph of medicine if the effects of such genes could be countered, and two approaches for countering them have been considered, both of which are called genetic engineering. One approach involves altering some cells of the body so that they can carry out the needed function. A patient could, for instance, be treated in this way for a blood disease caused by an abnormal protein made by a mutant gene. A normal gene would be inserted into the precursor cells – immature bone marrow cells that ultimately develop into functioning blood cells.

*Limiting science: a biologist's perspective*

In this way, a normal protein could be made in place of, or along with, the aberrant protein. The genetically altered blood cell precursor could then cure the patient's disease. But the malfunctioning gene would still be transmitted to the patient's offspring. Because this form of genetic engineering would not change the gene pool of the species and because it may prove an effective medical treatment of disease, it does not present the same moral problem as the other form. It is likely to be the first type of genetic engineering tried on human beings, and might be tried within the next five years.

The second type of genetic engineering presents more of a dilemma because it could change the human gene pool. This would involve replacement of genes in the germ cells, cells that transmit their genes to our offspring. Such a change would represent a permanent alteration of the types of the genetic information that constitutes our species. Replacement of germ cell genes would be very difficult and is, I suspect, at least twenty years away. It presents no theoretical problems, only formidable logistic problems.

Both forms of genetic engineering, but especially the engineering of germ cells, present two very deep and perplexing problems: who is to decide, and how shall they decide what genes are malfunctioning? Decisions about which genes are good and which bad truly represent decisions of morality and are therefore highly subjective. Fear that dictators will decide which genes should be suppressed and which promoted, and that their criteria for decision will be how best to maintain their own power, has made the phrase "genetic engineering" symbolic of the moral problems that can be created by modern biology.

Genetic engineering of human beings is not the same as recombinant DNA research. Genetic engineering is a process carried out on human beings; recombinant DNA methods are ways to purify genes. Such genes might be used in genetic engineering procedures but are much more likely to be used as tools in studies of biological organization or as elements in a biological manufacturing process. Although genetic engineering is not the same as recombinant DNA research, the two are rightly linked because recombinant DNA work is hastening the day when genetic engineering will be a feasible process for use on certain human diseases. Since recombinant DNA work is also bringing closer the discovery of many other possible new medical treatments, and is likely to bring other new capabilities, why is there so much focus on genetic engineering? I believe that genetic engineering, because of its tabloid appeal, has become a symbol to many people of the frightening potential of modern-day technology. Rather than seeing in molecular biology the same complex mixtures of appropriate and inappropriate applications that characterize all powerful sciences, many people have allowed the single negative catch phrase "genetic engineering" to dominate discussions. People worry that if the possibility of curing a genetic defect by gene therapy should ever become a reality, the inevitable result would be "people made to order." It is argued that unless we block recombinant DNA research now, we will never have another chance to control our fate.

To see the form of the argument against recombinant DNA research most clearly and to highlight its danger to intellectual freedom and creativity, we should realize that similar arguments



have been put forward relative to other areas of basic biology. One of the most respected critics of recombinant DNA research, Robert Sinsheimer, has for instance, made comparable analyses of two other research topics, research on aging and attempts to contact extraterrestrial beings.<sup>1</sup> He argues that if research on aging were to be successful, people could live to much older ages and the changed age structure of the population would bring serious stresses to society. His fear about searching for extraterrestrial beings is, again, that we could be successful, and he considers that the discovery of civilizations much more advanced than ours would have effects on us like those the Europeans had on native Americans after the discovery of the New World. In the case of recombinant DNA research as well as the other topics, Dr. Sinsheimer says we should avoid studying these areas of science. Rather, he believes, we should put our resources into investigating areas of proven need, such as fertility control.

These examples – and I could choose many others, especially outside of biology – have the same general form: certain people believe that there are areas of research that should be taboo because their outcome might be, or in some scenarios will be, detrimental to the stable relationships that characterize contemporary society. I have heard the argument in a different and more pernicious form from members of a Boston area group called Science for the People. They argue that some research in genetics should not go on because its findings might be detrimental to the relationships they believe should characterize a just society. Such arguments are reminis-

cent of those surrounding the eugenics movements that developed in Germany and Russia in the 1920s. After a period of intense debate, these countries with opposing ideologies settled on opposing analyses of the role of genetics in determining human diversity.<sup>2</sup> German scientists and politicians espoused a theory of racial purification by selective breeding, while Russia accepted the Lamarckian principle of transmission of acquired characteristics. In both cases, science was forced into a mold created by political and social ideology, and in both cases the results were disastrous.

As I see it, we are being faced today with the following question: should limits be placed on biological research because of the danger that new knowledge can present to the established or desired order of our society? Having thus posed the question, I believe that there are two simple, and almost universally applicable, answers. First, the criteria determining what areas to restrain inevitably express certain sociopolitical attitudes that reflect a dominant ideology. Such criteria cannot be allowed to guide scientific choices. Second, attempts to restrain directions of scientific inquiry are more likely to be generally disruptive of science than to provide the desired specific restraints. These answers to the question of whether limits should be imposed can be stated in two arguments. One is that science should not be the servant of ideology, because ideology assumes answers, but science asks questions. The other is that attempts to make science serve ideology will merely make science

1 R. L. Sinsheimer, "Inquiry into Inquiry: Two Opposing Views," *The Hastings Center Report* 6 (4) (1976): 18.

2 L. R. Graham, "Political Ideology and Genetic Theory: Russia and Germany in the 1920's," *The Hastings Center Report* 7 (5) (1977): 30–39.

impotent without assuring that only desired questions are investigated. I am stating simply that we should not control the direction of science and, moreover, that we cannot do so with any precision.

Before trying to substantiate these arguments I must make a crucial distinction. The arguments pertain to basic scientific research, not to the technological applications of science. As we go from the fundamental to the applied, my arguments fall away. There is every reason why technology should and must serve specific needs. Conversely, there are many technological possibilities that ought to be restrained.

To return to basic research, let me first consider the danger of restricting types of investigation because their outcome could be disruptive to society. There are three aspects of danger. One is the fallacy that you can predict what society will be like even in the near future. To say, for instance, that it would be bad for Americans to live longer assumes that the birth rate will stay near where it is. But what happens if the birth rate falls even lower than it is now in the United States and also stabilizes elsewhere in the world? We might welcome a readjustment of the life span. In any case, we have built a world around a given human life span; we could certainly adjust to a longer span, and it would be hard to predict whether in the long run the results would be better or worse.

In a general form, I would call this argument for restricting research the Error of Futurism. The futurist believes that the present holds enough readable clues about the future to provide a good basis for prediction. I doubt this assumption; to think that the data of today can be analyzed well enough to predict the future with any accuracy seems nonsensical to me.

The second danger in restricting areas of scientific investigation is more crucial: although we often worry most about keeping society stable, in fact societies need certain kinds of upheaval and renewal to stay vital. The new ideas and insights of science, much as we may fight against them, provide an important part of the renewal process that maintains the fascination of life. Freedom is the range of opportunities available to an individual – the more he has to choose from, the freer his choice. Science creates freedom by widening our range of understanding and therefore the possibilities from which we can choose.

Finally, attempts to dictate scientific limits on political or social considerations have another disastrous implication. Scientific orthodoxy is usually dictated by the state when its leaders fear that truths could undermine their power. Their repressive dicta are interpreted by the citizens as an admission of the leaders' insecurity and may thus lead to unrest requiring further repression. A social system that leaves science free to explore, and encourages scientific discoveries rather than trying to make science serve it by producing the truths necessary for its stability, transmits to the members of that society strength, not fear, and can endure.

The other argument I mentioned in opposing imposition of orthodoxy on science is the practical impossibility of stopping selected areas of research. Take aging as a prime example. It is one of the mystery areas of modern biology. The questions are clear. Why do we get older? Why do organ systems slowly fail? Why does one species of animal live three years and another live for one hundred? These seemingly straightforward questions are unapproachably difficult for modern biology. Not only can we not

understand events that occur over years, but we even have difficulty understanding questions about events that require minutes to transpire. In fact, molecular biologists are really only experts in the millisecond range of time. Such times are those of chemical reactions; to understand events in longer time frames probably requires knowing how individual reactions are integrated to produce clocks that measure time in seconds to years. Clues to the great mysteries of biology – memory, aging, and differentiation – lie in understanding how biological systems tell time.

There are a few hints about where answers to the puzzle of aging might be found, but they are only the vaguest suggestions. In such an area of science, history tells us that successes are likely to come from unpredictable directions. A scientist working on vitamins or viruses or even plants is just as likely to find a clue to the problem of aging as is a scientist working on the problem directly. In fact, someone outside the field is more likely to make a revolutionary discovery than someone inside the field.

Another example will help to show the generality of this contention. Imagine that we were living at the turn of the last century and had wanted to help medical diagnosis by devising a method of seeing into the insides of people. We would probably have decided to fund medical scientists to learn how to use bright lights to see through patients' skin. Little would we have guessed that the solution would be revolutionary and would come, not from a medical research scientist, but from a physicist who would discover a new form of radiation, X-rays.<sup>3</sup>

<sup>3</sup> Dr. Mahlon Hoagland first suggested to me this example.

Major breakthroughs cannot be programmed. They come from people and areas of research that are not predictable. So if you wanted to cut off an area of fundamental research, how would you be able to devise the controls? I contend that it would be impossible. You could close the National Institute of Aging Research, but I doubt that any major advance in that field *could* be prevented. Only the shutdown of all scientific research can guarantee such an outcome.

Although they would fail to produce their desired result, attempts explicitly to control the directions of basic research would hardly be benign. Instead, disruption and demoralization would follow from attempts to determine when a scientist was doing work in approved directions and when he was not. Creative people would shun whole areas of science if they knew that in those areas their creativity would be channeled, judged, and limited. The net effect of constraining biologists to approved lines of investigation would be to degrade the effectiveness of the whole science of biology.

Put this way, the penalty for trying to control lines of investigation seems to me greater than any conceivable benefit. I conclude that society can choose to have either more science or less science, but choosing *which* science to have is not a feasible alternative.

I must repeat a qualification of this broad generalization: the less basic the research area in which controls are imposed, the less general disruption will be caused by the controls. The development of a specific sweetener, pesticide, or weapon could be prevented with little generalized effect.

While it seems necessary that scientific research be free of overt restraints, it would be naive to think that science is not directed at all. Many crucial deci-

sions are made about general directions of science, usually by the control of available resources. Again, the formula I used before is applicable: the less basic the area of research, the easier it is to target the problems. A fallacy behind some of the hopes for the "War on Cancer" was the assumption that the problems to be solved were sufficiently well defined to allow a targeted, applied approach to the disease. Some problems could be defined and those were appropriate targets for a war on the disease, but the deeper mysteries of cancer are so close to the frontiers of biology that targets are hard to discern.

I have painted a picture of inexorable, uncontrollable development of basic scientific knowledge. The response of many people to such a vision might well be "If we can't put any controls on research, maybe we shouldn't have any research." I see the rationale in this criticism because it is conceivable that the rate of accretion of knowledge could become so high that a brake might be needed. A way exists to produce a slow-down, that is, by controlling the overall availability of resources. A nonselective brake on fundamental biology would decrease productivity without the disruptive and dangerous effects of trying to halt one area and advance another.

Because such a brake was applied in the late 1960s, today we have less basic research, measured in constant dollars spent on it, than we had then. As a result of the slowdown, the danger seen by many is not that we have too much basic research, but rather that we are living on our intellectual capital and that an infusion of funds into basic areas of research is needed before these scientific resources are exhausted. It must also be realized that our commitment to the solution of problems like cancer re-

quires that we develop much more basic knowledge.

I may seem to have strayed from my topic of recombinant DNA research, but I think not. In 1977, there was a draft bill in the U.S. Senate to set up a National Commission on Recombinant DNA Research. The charge to this commission discussed mainly questions of safety, but at the end of the bill questions of ethics and morality entered. The bill was a clear invitation to begin the process of deciding what research shall be allowed and what research prevented. The battle to stop the creation of this commission was waged by scientists, university presidents, and other concerned individuals across the country. I believe that the long-term success or failure of these efforts will determine whether America continues to have a tradition of free inquiry into matters of science or falls under the fist of orthodoxy.

To finish this discussion, let me quote from the most eloquent analyst of contemporary biology, Lewis Thomas, the president of the Memorial Sloan-Kettering Cancer Center in New York. Writing in the *New England Journal of Medicine* about the recombinant DNA issue, Dr. Thomas ended with this analysis of the role of scientific research in the life of the mind:

Is there something fundamentally unnatural, or intrinsically wrong, or hazardous for the species, in the ambition that drives us all to reach a comprehensive understanding of nature, including ourselves?

I cannot believe it. It would seem to me a more unnatural thing, and more of an offense against nature, for us to come on the same scene endowed as we are with curiosity, filled to overbrimming as we are with questions, and naturally talented as we are for the asking of clear questions, and then for us to do nothing about it, or



worse, to try to suppress the questions. This is the greater danger for our species, to try to pretend that we are another kind of animal, that we do not need to satisfy our curiosity, exploration, and experimentation, and that the human mind can rise above its ignorance by simply asserting that there are things it has no need to know.<sup>4</sup>

*Limiting  
science:  
a biologist's  
perspective*

<sup>4</sup> Lewis Thomas in the *New England Journal of Medicine* 296 (1977): 328.

# Lorraine Daston

## *Fear & loathing of the imagination in science*

Recently a reader responded with dismay to a *New Yorker* article by historian Daniel J. Kevles about the charge of scientific fraud brought by Margot O'Toole against Thereza Imanishi-Kari. What distressed this reader was not so much the issue of fraud itself as Kevles's argument that the exercise of judgment and imagination in science was essential and should not be conflated with fraud:

... I am troubled by Kevles's acceptance of a need for scientists to be imaginative in analyzing research results. What might the public's realization that this practice exists do to its confidence in the hard sciences? Will we next be expected to believe that accountants require imagination in their work?<sup>1</sup>

Such expressions of uneasiness about the role of the imagination in science are not new. When the physicist John Tyndall delivered a "Discourse on the Scientific Use of the Imagination" to the British Association for the Advancement

---

*Lorraine Daston, a Fellow of the American Academy since 1993, wrote this essay for the Winter 1998 issue of "Daedalus." Daston has served as director of the Max Planck Institute for the History of Science in Berlin, Germany, since 1995.*

---

© 2005 by the American Academy of Arts & Sciences

of Science in 1870, he too drew shocked reactions from the press. The *London Times* was severe:

The glory of a Natural Philosopher appears to depend less on the power of his imagination to explore minute recesses or immeasurable space than on the skill and patience with which, by observation and experiment, he assures us of the certainty of these invisible operations.... [Tyndall] confesses that Mr. Darwin "has drawn heavily upon time and adventurously upon matter." We ask ourselves whether we are listening to one experimental philosopher describing the achievement of another experimental philosopher. We had been under the impression that Natural Philosophers drew no bills.<sup>2</sup>

The echo of fiscal analogies reverberates over the space of more than a century: scientists should be as methodical (and as plodding) as accountants ("Natural Philosophers draw no bills"). To permit the imagination to infiltrate science is to tamper with the books, to betray a public trust.

1 Aaron Fischbach, "In the Mail," *New Yorker*, July 22, 1996, 6. Kevles's article appeared in the May 27, 1996 issue of the *New Yorker*.

2 *London Times*, September 19, 1870; reprinted in John Tyndall, *Essays on the Use and Limit of the Imagination in Science* (London: Longmans, Green, & Co., 1870), 1–2.

My aim here is not to show that first-rate science requires imagination; others have already pleaded this point with vigor and eloquence.<sup>3</sup> Rather, I would like to explore how and why large portions of the educated public – and many working scientists – came to think otherwise, systematically opposing imagination to science. I shall argue that the critical period was the mid-nineteenth century, when new ideals and practices of scientific objectivity transformed the persona of the scientist and the sources of scientific authority. More specifically, I shall focus on the apparent paradox, also first framed in the early decades of the nineteenth century, that the more scientists insisted upon the obduracy and intransigence of facts, the more they feared the power of their own imaginations to subvert those facts. Why would scientists convinced of the power of ugly facts to murder beautiful theories, as Thomas Henry Huxley famously put it, nonetheless take heroic precautions to protect those burly facts from gossamer-spun imagination?

The key to this paradox lies buried within the histories of the scientific fact, on the one hand, and of the faculty of the imagination, on the other. In order to dramatize the novelty of the mid-nineteenth-century developments, I shall begin with a brief account of how eighteenth-century natural philosophers and natural historians understood the relationship between scientific facts and the scientific imagination. The pivot of my story is the polarization of the personae of artist and scientist, and the migration of imagination to the artistic pole. At roughly the same time that art-

ists working in a romanticist vein emphasized creativity over mimesis, scientists troubled by the overthrow of one time-honored theory after another in quick succession sought more durable achievements. This early nineteenth-century confrontation of individualistic, brashly subjective art with collective, staunchly objective science was not simply the collision of some timeless faith in the imagination with an equally timeless faith in facts. Rather, it signaled a mutation in the meanings both of imagination and of facts that still shapes the moral economy of science.

Experience we have always had with us, but facts as a way of parsing experience in natural history and natural philosophy are of seventeenth-century coinage. Aristotelian experience had been woven of smooth-textured universals about “what happens always or most of the time”; early modern facts were historical particulars about an observation or an experiment performed at a specific time and place by named persons.<sup>4</sup> What made the new-style facts granular was not only their specificity but also their alleged detachment from inference and conjecture. Ideally, at least, “matters of fact” were nuggets of pure experience, strictly segregated from any interpretation or hypothesis that might enlist them as evidence. Some seventeenth-century philosophers were as skeptical as their twentieth-century successors about the bare existence of what we now (redundantly) call theory-free facts. René Descartes, for example, trusted only those experiments per-

3 Gerald Holton, “Imagination in Science,” in his *Einstein, History, and Other Passions: The Rebellion against Science at the End of the Twentieth Century* (Reading, Mass.: Addison-Wesley, 1996), 78–102.

4 For an account of the transformation of scientific experience in the early modern period, see Peter Dear, *Discipline and Experience: The Mathematical Way in the Scientific Revolution* (Chicago: University of Chicago Press, 1995).

formed under his own supervision, because those reported by others distorted the results to “conform to their principles.”<sup>5</sup> Even the most vigorous promoters of “matters of fact” acknowledged that these nuggets of pure experience were hard won: Francis Bacon thought only the strict discipline of method could counteract the inborn tendency of the human understanding to infuse observation with theory.<sup>6</sup> The 1699 *Histoire* of the Paris Académie Royale des Sciences confessed that the “detached pieces” of experience the academicians offered in lieu of coherent theories or systems had been wrenched apart by a “kind of violence.”<sup>7</sup> Chiseling out “matters of fact” from the matrix of interpretation and conjecture was hard work.

But it was the hard work of smelting and purifying, not that of building and constructing. One of the most striking features of the new-style scientific facts of the seventeenth century is how swiftly and radically they broke with the etymology that connected them to words like “factory” and other sites of making and doing. In Latin and the major European vernaculars the word “fact” and its cognates derives from the verb “to do” or “to make,” and originally referred to a deed or action, especially one remarkable for either valor or malevolence: *facere/factum, faire/fait, fare/fatto, tun/Tatsache*.<sup>8</sup> English still bears traces of

this earlier usage in words like “feat” and, especially, in legal phrases like “after the fact.” When the word “fact” acquired something like its familiar sense in the early seventeenth century as “a particular truth known by actual observation or authentic testimony, as opposed to what is merely inferred, or to a conjecture or fiction,” to quote from its entry in the *Oxford English Dictionary*, it snapped the philological bonds that had tied it to words like “factitious” and “manufacture.” Conversely, by the mid-eighteenth century, once-neutral words like “fabricate” (originally, to form or construct anything requiring skill) or “fabulist” (teller of legends or fables) had acquired an evil odor of forgery and deception in addition to their root senses of construction. For most Enlightenment thinkers, facts *par excellence* were those given by nature, not made by human art. “Facts” and “artifacts” had become antonyms, in defiance of their common etymology.

In keeping with the opposition of natural facts to human artifacts, the errors that most terrified Enlightenment servants in theory and practice were the errors of construction, of a world not reflected in sensation but made up by the imagination. Sensory infirmities worried Enlightenment epistemologists of science relatively little, prejudices and misconceptions instilled by bad education rather more so, the distortions wrought by strong passions still more, and the unruly creations of the imagination most of all. These latter seemed so pervasive as to make the simplest factual narrative a triumph of vigilance, discipline, and civilization in the minds of some Enlightenment writers. Bernard de Fontenelle, Perpetual Secretary of the Académie Royale des Sciences in Paris, thought the inclination to embellish the facts of the matter in any re-

5 René Descartes, *Discours de la méthode* [1637], pt. VI. Unless otherwise specified, all translations are my own.

6 Francis Bacon, *Novum organum* [1620], pt. I.46.

7 *Histoire de l'Académie Royale des Sciences. Année 1699*, 2nd ed. (Paris, 1718), “Préface.”

8 See relevant entries in the *Oxford English Dictionary*, *Grimms Wörterbuch*, and the *Dictionnaire historique de la langue française*.



telling so irresistible that “one needs a particular kind of effort and attention in order to say only the exact truth.” It took centuries before society advanced to the point of being able to “preserve in memory the facts just as they happened,” before which time “the facts kept in [collective] memory were no more than visions and reveries.”<sup>9</sup>

The chronic inability to hold fast to fact, to keep the inventive imagination in check, was a midpoint along a continuum to madness. Scientists were as much at risk as poets from the diseases of the imagination. In Samuel Johnson’s allegorical novel *The History of Rasselas, Prince of Abissinia*, the philosopher Imlac meets a learned astronomer “who has spent forty years in unwearied attention to the motions and appearances of the celestial bodies, and has drawn out his soul in endless calculations.” Upon further acquaintance the astronomer proves as virtuous as he is learned, “sublime without haughtiness, courteous without formality, and communicative without ostentation.” Surely the astronomer is the long-sought-after happy man, content in his science and virtue? Alas, no; the astronomer is stark raving mad. He discloses to Imlac his delusion that he alone can control the world’s weather and that he therefore bears the crushing responsibility for the welfare of the world’s population on his shoulders. Imlac reflects that no one is immune from the deprecations of the imagination: “There is no man whose imagination does not sometimes predominate over his reason, . . . All power of fancy over reason is a degree of insanity; . . . By degrees the reign of fancy is confirmed; she

grows first imperious, and in time despotic. Then fictions begin to operate as realities, false opinions fasten upon the mind, and life passes in dreams of rapture or anguish.”<sup>10</sup>

It was not only novelists and philosophers who worried about “fictions [that] begin to operate as realities,” about the fragility of facts in the face of overweening imagination. Practicing naturalists also fretted openly. In his monumental *Mémoires pour servir à l’histoire des insectes* (1734 – 1742) the French naturalist and experimental physicist René Antoine Réaumur warned that “although facts were assuredly the solid and true foundations of all parts of physics,” including natural history, not all reported facts in science could be trusted. It was not simply a matter of weeding out hearsay or dubious sources; even sincere, well-trained naturalists could adulterate observations with imaginings. Citing the example of Godaert’s observation that some insects could spawn insects of a different species, Réaumur preached caution: “Too often the observer has the disposition to see objects quite otherwise than they [actually] are. The extravagant love of the marvelous, a too strong attachment to a system fascinates his eyes.”<sup>11</sup> An errant imagination was also Georges Cuvier’s diagnosis of how Jean-Baptiste Lamarck had gone astray in natural history: for all of his scientific gifts, Lamarck was one of those minds that “cannot prevent themselves from mixing [true discoveries, *découverts véritables*] with fantastic conceptions . . . [T]hey laboriously

9 Bernard de Fontenelle, *De l’origine des fables* [1724], ed. J.-R. Carré (Paris: Librairie Félix Alcan, 1932), 14, 33.

10 Samuel Johnson, *The History of Rasselas, Prince of Abissinia* [1759], ed. J. P. Hardy (Oxford: Oxford University Press, 1968), 98 – 99, 104 – 105.

11 René François Ferchault de Réaumur, *Mémoires pour servir à l’histoire des insectes*, vol. 2 (Paris: Imprimerie Royale, 1736), xxxiv – v.

construct vast edifices on imaginary bases, similar to the enchanted palaces in our old romances which disappear when the talisman upon which their existence depends is broken.”<sup>12</sup>

Cuvier’s opposition of “true discoveries” to “romances,” of fact to fiction, was at least as old as Bacon and was echoed countless times before, during, and after the Enlightenment. Equally banal and enduring was the parallel opposition of the faculties of reason and imagination. What was striking about eighteenth-century views of the imagination in light of later developments is their firm insistence that the imagination, despite its perils, was as essential to philosophy and science – the pursuits of reason – as to the arts. Moreover, both art and science drew on the same kind of healthy imagination – and both were at risk from the same pathologies of the imagination. Both science and art were, in the view of most of their eighteenth-century practitioners, dedicated to revealing the truths of nature; imagination enlisted to this aim was a sound, sane one, that is to say, an imagination subject to rules. Even the most inventive genius should, Enlightenment critics insisted, bow to the authority of nature and its rules. John Dryden, for example, wondered whether Shakespeare might not have gone too far in creating the monstrous character of Caliban in *The Tempest*, “a person which was not in Nature,” and Goya explained the famous epigram of his *Caprichos* – “The sleep of reason produces monsters” – as a call to the union of reason and imagination: “Imagination deserted by reason produces impossible monsters. United with reason, imagination is the mother

12 Georges Cuvier, *Recueil des éloges historiques lus dans les séances publiques de l’Institut de France* [1819 – 1827], vol. 3 (Paris: Firmin Didot Frères Fils, 1861), 180.

of the arts and the source of their wonders.”<sup>13</sup>

Images of the monstrous pervaded Enlightenment accounts of the diseased imagination in the arts and sciences. Voltaire distinguished between the “active imagination,” which inspired the finest works of mechanics, mathematics, poetry, and the fine arts, and the “passive imagination,” which caused violent passions, fanaticism, delusions, and monsters both figurative and literal. The passive imagination in the arts and sciences welded together “incompatible objects” into chimeras; in the womb of a pregnant woman it could impress the soft embryo with the form of some hideous perception – for example, of a convicted criminal broken on the wheel – received by the mother.<sup>14</sup> The French critic Jean-François Marmontel acknowledged that fiction was no servile imitation of nature, but even fiction that perfected nature still kept the imagination on a short leash. What was variously called the “marvelous,” “monstrous,” or “fantastic” imagination in the arts led to the “debauchery of genius.”<sup>15</sup> Poets and artists were instead directed to obey the cardinal rule of verisimilitude: “A verisimilar fact is a fact possible in the circumstances where one lays the scene.

13 Quoted in Rudolf Wittkower, “Genius: Individualism in Art and Artists,” in Philip P. Wiener, ed., *Dictionary of the History of Ideas*, vol. 2 (New York: Charles Scribner’s Sons, 1973), 307, 308.

14 [Voltaire], “Imagination,” in Jean d’Alembert and Denis Diderot, eds., *Encyclopédie, ou Dictionnaire raisonné des sciences, des arts et des métiers*, vol. 8 (Paris/Neuchâtel, 1751 – 1780), 560 – 563. See also Wendy Doniger and Gregory Spinner, “Misconceptions: Female Imaginations and Male Fantasies in Parental Imprinting,” in *Daedalus* 127 (1) (Winter 1998).

15 [Marmontel], “Fiction,” in *Encyclopédie*, vol. 6, 679 – 683.

Fictions without verisimilitude, and events prodigious to excess, disgust readers whose judgment is formed.”<sup>16</sup> Enlightenment good taste demanded that even fictions be decked out as possible facts and that art as well as science follow nature. Both art and science required imagination, but in neither should the imagination be allowed to invent at will.

Or rather, *against* will, for in the view of Enlightenment writers like Voltaire and Marmontel the pathological imagination overthrew the reasonable sovereignty of the will. Whereas the healthy, active imagination always partakes of judgment and “raises all of its edifices with order,” the diseased, passive imagination acts imperiously, so that its victims are no longer “master” of themselves.<sup>17</sup> Here the distinction between the healthy and the diseased imagination took on moral as well as epistemological (and aesthetic) undertones. The consequences of submitting weakly to the domineering imagination could be dramatic, as the members of the French scientific commission formed in 1784 to investigate alleged phenomena of animal magnetism emphasized. After observing the remarkable convulsions and cures displayed by mesmerized patients, the commission – which included the astronomer Jean-Sylvain Bailly, the chemist Antoine Lavoisier, and the electrician Benjamin Franklin – decided to undergo animal magnetism themselves. Seated in the great mesmeric tubs, under the magnetizer’s wand, the commissioners contrasted their own calm impassivity with the spectacular crises of the *convulsionnaires*:

16 [Chevalier de Jaucourt], “Vraisemblance,” in *ibid.*, vol. 17, 484.

17 Voltaire, “Imagination,” 561.

The Commissioners could not help but be struck by the difference between the public treatment and their own particular treatment in the tubs. The calm and silence of the one, the motion and agitation of the other; there, the multiple effects of violent crises, the habitual state of mind and body interrupted and troubled, nature exalted; here, the body without pain, the mind untroubled, nature preserving both its equilibrium and its ordinary course, in a word the absence of all effects.<sup>18</sup>

Tranquil and self-controlled savants versus shaking and shrieking patients: for the commissioners there could be no clearer contrast between the sound and the diseased imaginations. They concluded that the cures wrought by animal magnetism were often genuine, and the convulsions mostly sincere, but that all were the work of the imagination, “that active and terrible power that produces the great effects that one observes with astonishment at the public treatment.”<sup>19</sup> Although gender and class played some role in how the commissioners gauged degrees of susceptibility to the imagination, the ultimate defense against “that active and terrible power” was enlightenment (*lumières*), a combination of intelligence and self-mastery. Despite their palpable disapproval of such excesses of the imagination, the savants of the Royal Commission paid tribute to its extraordinary power over mind and body. No roman-

18 [J. S. Bailly], *Rapport des commissaires chargés par le Roi, de l'examen du magnétisme animal* (Paris: Imprimerie Royale, 1784), 18–19. On animal magnetism in late eighteenth-century France and the background to the Royal Commission, see Robert Darnton, *Mesmerism and the End of the Enlightenment in France* (Cambridge, Mass.: Harvard University Press, 1968).

19 Bailly, *Rapport*, 59.

tic poet was ever more firmly convinced of the force of the unfettered imagination than the Parisian savants.

In Enlightenment art and science, the imagination was Janus-faced: on the one hand, it was essential to creative work in both realms; but on the other, it could betray the natural and the verisimilar by breeding monsters. Its power verged on the supernatural. It could drive brilliant artists and scientists mad, it could trigger violent seizures, it could cure the hopelessly ill, it could distort and obliterate facts. So long as art and science shared a common goal of truth to nature, they also shared a code of aesthetic, epistemological, and moral values that praised one face of the imagination and deplored the other. Genius – be it in poetry, sculpture, or natural philosophy – was the expression of heightened imagination. Whether the genius in question was Milton or Leibniz, Michelangelo or Descartes, the natural endowment that made their achievements possible was in essence the same: a soaring imagination that “produces more than it discovers . . . [that] hatches brilliant systems or discovers great truths.”<sup>20</sup> Imagination was not yet immiscible with science, and it was arguably more robust than facts.

Between about 1780 and 1820 this configuration changed dramatically. Put in the briefest terms, facts hardened, the imagination ran riot, and art and science diverged in their aims and their collective personae. Within the narrow confines of this essay, it is only possible to offer emblematic episodes to illustrate the nature and extent of these major transformations in the self-images of artists and scientists. Immanuel

Kant’s account of genius in his *Kritik der Urteilskraft* heralded things to come. Kant took it for granted that originality was the *sine qua non* of genius and that “[e]veryone is agreed on the point of the complete opposition between genius and the *spirit of imitation*.” But that which can be learned, reasoned Kant, can also be in a sense imitated. Hence even the greatest triumphs of the natural sciences could no longer count as true works of genius:

So all that Newton has set forth in his immortal work on the *Principles of Natural Philosophy* may well be learned, however great a mind it took to find it all out, but we cannot learn to write in a true poetic vein, no matter how complete all the precepts of the poetic art may be, or however excellent its models. The reason is that all the steps that Newton had to take from the first elements of geometry to his greatest and most profound discoveries were such as he could make intuitively evident and plain to follow, not only for himself but for every one else.<sup>21</sup>

Kant was second to none in his admiration for Newton and the revelations of the natural sciences, but he nonetheless denied even Newton the title of genius. For Kant, the very transparency and communicability of mathematics and the natural sciences removed them from the realm of profound, ineffable originality inhabited by Homer or even Christoph Wieland. Kant’s emphasis on communicability linked the natural sciences to an emergent opposition between objectivity and subjectivity that Kant himself pioneered. Kant employed these terms in several distinct senses in his critical philosophy; I wish to draw attention here only to the sense that res-

20 Denis Diderot, “Génie,” in *Encyclopédie*, vol. 7, 583.

21 Immanuel Kant, *The Critique of Judgment* [1790], trans. James Creed Meredith (Oxford: Oxford University Press, 1952), 168–170.

onated most loudly for nineteenth-century scientists and that meshed most tightly with Kant's rejection of the bare possibility of scientific genius. In the closing pages of the *Kritik der reinen Vernunft*, Kant offered a rough-and-ready test for distinguishing objectively valid convictions from merely subjectively valid persuasions:

If the judgment is valid for everyone, provided only he is in possession of reason, its ground is objectively sufficient [*objektiv gültig*], and the holding of it to be true is entitled conviction. If it has its ground only in the special character of the subject, it is entitled persuasion . . . The touchstone whereby we decide whether our holding a thing to be true is conviction or mere persuasion is therefore external, namely, the possibility of communicating it and of finding it to be valid for all human reason.<sup>22</sup>

In the middle decades of the nineteenth century this ideal of objectivity as communicability, shorn of every idiosyncrasy and particular perspective, was realized in the emergence of international, long-term scientific collaborations like the Internationale Gradmessung or the Carte du Ciel, which committed participants around the globe and across generations to instruments, procedures, and research agendas standardized in the name of commensurability and solidarity. Charles Sanders Peirce, who himself participated in some of these far-flung collaborations as an experimental physicist, drew the philosophical moral that scientific objectivity depended on the existence of a vast scientific community, extended over time and space, "beyond this geological epoch, beyond all

22 Immanuel Kant, *Critique of Pure Reason* [1781, 1787], trans. Norman Kemp Smith (New York: St. Martin's Press, 1965), A820–821/B848–849, 645.

bounds."<sup>23</sup> Or as the experimental physiologist Claude Bernard put it with lapidary concision: "L'art c'est moi, la science c'est nous."<sup>24</sup>

But if science – and with it, objectivity – had come to be identified with the communal and the communicable, how did art wander to the pole of solitude and the individual? Within the Enlightenment framework, both savants and artists, especially those touched by genius, were often idealized as solitary seekers of deep truths on the model of hermetic saints, whatever the biographical realities might have been.<sup>25</sup> One might therefore argue that there is nothing to be explained on the side of art: artists, at least in their idealized personae, simply remained lonely geniuses while their scientific brethren became clubby, and thereby ungenial. So simple a conclusion would, however, overlook the impact of far-reaching changes in aesthetics and in views of the artistic imagination that occurred in the early decades of the nineteenth century. Again, I can offer only a small sampling over the many possible examples to make my point vivid.

The ramifications of post-Kantian theories of the imagination fan out into a broad and branching tree, from Johann Gottlieb Fichte to Friedrich Schelling to Samuel Taylor Coleridge to Jules Miche-

23 Charles Sanders Peirce, "Three Logical Sentiments," in Charles Hartshorne and Paul Weiss, eds., *Collected Papers of Charles Sanders Peirce*, vol. 2 (Cambridge, Mass.: Harvard University Press, 1932), 398.

24 Claude Bernard, *Introduction à l'étude de la médecine expérimentale* [1865], ed. François Dagognet (Paris: Garnier-Flammarion, 1966), 77.

25 Dorinda Outram, "The Language of Natural Power: The 'Eloges' of Georges Cuvier and the Public Language of Nineteenth-Century Science," *History of Science* 16 (1978): 153–178.



let and beyond.<sup>26</sup> There is probably no generalization that holds for all of these developments, but tendencies can be discerned. For my purposes, the most significant are, first, the heavy emphasis upon the almost mystical originality of the imagination, independent or even in defiance of reason and will; and, second, the allied cult of individual subjectivity, what the art historian Rudolf Wittkower once called the “egomania” of romanticism. Each element had distinguished antecedents – Plato’s poetic furor or the Renaissance master as *Deus artifex* – but the combination of the two was novel to the early nineteenth century. Quasi-divine inspiration overwhelming will and judgment had not been traditionally paired with towering individualism: for example, the pythian priestesses through whom the godhead spoke at the oracle of Delphi were inspired but interchangeable. The intertwining of these two elements – originality and subjectivity – effectively rehabilitated what Enlightenment theorists had regarded as the pathological imagination. For the romantics, it was the unbidden, darkling force of the so-called passive imagination that was the wellspring of genial creativity, not the well-regulated active imagination subservient to will and reason. As William Blake retorted to Sir Joshua Reynolds, “What has Reasoning to do with the Art of Painting? ... One power alone makes a poet; Imagination, the Divine Vision.”<sup>27</sup> Hence the strong association in nineteenth-century

26 For general overviews of these developments, see Karl Homann, “Zum Begriff Einbildungskraft nach Kant,” *Archiv für Begriffsgeschichte* 14 (1970): 266–302; Mary Warnock, *Imagination* (London: Faber and Faber, 1976); Eva T. H. Brann, *The World of the Imagination: Sum and Substance* (Savage, Md.: Rowman and Littlefield, 1991).

27 Quoted in Wittkower, “Genius,” 306.

psychological studies of genius – which restricted their subjects almost exclusively to artists and poets – between extraordinary creativity and the unconscious, or even insanity.<sup>28</sup>

In conjunction with the elevation of the passive imagination aesthetic norms shifted away from verisimilitude. A genuinely productive, as opposed to reproductive, imagination could be bound neither by the rules of decorum nor those of the natural order. Strict mimesis had never been the avowed ideal of Enlightenment critics, but they had subscribed to a standard of truth to nature, if not truth to fact. Romantic poets and artists attacked this aesthetic openly, under the twin banners of originality and individual subjectivity. Charles Baudelaire parodied what he called the credo of nature – “I believe in nature, and only in nature” – and called for art infused with imagination, for landscapes in which “human egotism replaces nature,” for works to which the artist or poet “adds his soul.” It was idolatry for art to prostrate itself before nature; any photograph could surpass the most faithful artistic replica in “absolute material exactitude.” Deploring the public infatuation with photographic landscapes and portraits, Baudelaire insisted that ideals of truth and beauty not only did not coincide, they were inalterably opposed to one another: “With us the natural painter, like the natural poet, is almost a monster. The exclusive taste for the True (noble though it may be when limited to its true applications) here oppresses and suffocates the taste for the Beautiful.” For Baudelaire, imitation of nature shaded imperceptibly into imitation of other artists: “The

28 Most famously in Cesare Lombroso, *Genio e follia. Prelezione ai corsi di antropologia e clinica psichiatrica* (Milan: G. Chiusi, 1864).

artist, the true artist, the true poet . . . must be really faithful to his own nature. He must avoid like death borrowing the eyes and the sentiments of another man, however great," just as he must avoid depicting "the universe without man," without the intervention of the imagination.<sup>29</sup>

It is customary to classify such views as "romantic," a term Baudelaire himself occasionally used. However, this label covers over fault lines that opened up within romanticism between subjective art and objective science, between the acolytes of beauty and those of truth. Although early nineteenth-century science had its own avowed romantics, such as Johann Wolfgang Goethe, Johann Ritter, Sir Humphrey Davy, or Alexander von Humboldt, they were notably wary of the exalted imagination and individualism of the new aesthetics. The experimental physicist Ritter, who discovered ultraviolet radiation in his search for polarities in nature and who was given to utterances such as "Light is the external intuition of gravity, love the internal," nevertheless balked at allowing the imagination free rein in science: "The most beautiful thoughts are often no more than soap bubbles: filled with the hydrogen of our fantasy they rise quickly, and one does not realize that all the delightful play of their colors is nothing more than the reflection of their deceptive interiors."<sup>30</sup> Goethe warned the experimentalist against "the imagina-

tion [*Einbildungskraft*], which raises him to heights on its wings while he still believes his feet to be firmly planted on the ground";<sup>31</sup> Alexander von Humboldt scrupulously divided his monumental survey of nature into a first part containing "the main results of observation, which, stripped of all the extraneous charms of fancy, belong to the purely objective domain of a scientific delineation of nature," and a second part on "impressions reflected by the external senses on the feelings, and on the poetic imagination of mankind."<sup>32</sup> The wild imagination and individualism now held to be the birthright of true artists frightened even romantic scientists.

The point is that the newly erected divide between the objective and the subjective – the very words first enter dictionaries as a pair in German, French, and English in the 1820s and 1830s<sup>33</sup> – ran deeper than any opposition between neoclassicism and romanticism. My claim is not that there ceased to be fastidious realists among artists or daring speculators among scientists. Baudelaire found plenty of nature-worshippers to criticize among the paintings on display at the Paris Salon of 1859; Tyndall did not want for examples of scientists guided by their sense of beauty. But the new polarity of the objective and subjective

29 Charles Baudelaire, "Salon de 1846," "Salon de 1859," in *Curiosités esthétiques. L'art romantique et autres oeuvres critiques*, ed. Henri Lemaître (Paris: Editions Garnier, 1962), 97–200, 305–396.

30 Johann Ritter, *Fragmente aus dem Nachlasse eines jungen Physikers*, ed. Steffen and Brigit Dietzsch (Leipzig/Weimar: Gustav Kiepenhauer, 1984), 96, 260.

31 Johann Wolfgang Goethe, "Der Versuch als Vermittler von Objekt und Subjekt" [1792, publ. 1823], in *Goethes Werke*, vol. 13, *Naturwissenschaftliche Schriften I*, ed. Dorothea Kuhn und Rike Wankmüller (München: C. H. Beck, 1994), 14–15.

32 Alexander von Humboldt, *Cosmos* [1844], vol. 2, trans. E. C. Otté and W. S. Dallas (New York: Harper and Brothers, 1850–1859), 19.

33 For a brief account of the history of the words, see Lorraine Daston, "How Probabilities Came to Be Objective and Subjective," *Historia Mathematica* 21 (1994): 330–344.

structured how such boundary-straddling was perceived. When the novelist Gustav Flaubert attempted in *Madame Bovary* (1856) to depict a provincial adultery with clinical, impartial accuracy, both he and his critics seized upon the word “objective” to describe a style in which “subjects are seen as God sees them, in their true essence.”<sup>34</sup> When embryologist Wilhelm His described the advantages of scientific drawings, he called the result “subjective.”<sup>35</sup> Successful art could and did emulate scientific standards of truth to nature, and successful science could emulate artistic standards of imaginative beauty. But whereas in the eighteenth century both artists and scientists had seen no conflict in embracing both standards simultaneously, the chasm that had opened up between the categories of objectivity and subjectivity in the middle decades of the nineteenth century – words that, as Thomas De Quincey wrote in 1856, had once sounded pedantic and yet had so quickly become “indispensable to accurate thinking and to *wide* thinking”<sup>36</sup> – forced an either/or choice.

Hence a figure like Goethe, who combined artistic and scientific interests, became an uncomfortable paradox, especially for German scientists who could hardly escape the long shadow cast by the official national genius. The obligatory addresses delivered by leading Ger-

man scientists on Goethe’s scientific work provide a sensitive indicator of how entrenched the divide between objective and subjective had become. The physicist Hermann von Helmholtz gave two such addresses, in 1853 and 1862, and both turned on what Helmholtz took to be the opposition between scientific and artistic ways of thinking. Goethe’s regrettable (in Helmholtz’s view) attack on Newtonian optics could be explained, if not excused, by the impossibility of mingling the ineffable, almost divinatory intuitions of the artist with the crystalline concepts of the scientist. As in Kant’s touchstone for distinguishing the objective from the subjective, communicability was central to Helmholtz’s analysis of the distinction between artistic and scientific thinking: “Since artistic intuitions are not found by way of conceptual thinking, they cannot be defined in words . . . .”<sup>37</sup>

At the crossroads of the choice between objective and subjective modes stood the imagination. Very few nineteenth-century writers went so far as to deny scientists any imagination. Baudelaire, for example, acknowledged that imagination was as essential to the great scientist – or for that matter, the great diplomat or soldier – as to the artist. But in the next breath he relegated photography, whose exact rendering of what is seen he took to be diametrically opposed to the artistic imagination, to the sphere of science, where it might serve without corrupting.<sup>38</sup> By the last quarter of the nineteenth century, psychologists who investigated creativity routinely distin-

34 Quoted in Erich Auerbach, *Mimesis. The Representation of Reality in Western Literature* [1946], trans. Willard R. Trask (Princeton, N.J.: Princeton University Press, 1953), 487.

35 Wilhelm His, *Anatomie menschlicher Embryonen* (Leipzig: F. C. W. Vogel, 1880), 6. I am grateful to Robert J. Richards for this reference.

36 Thomas De Quincey, *The Confessions of an English Opium Eater* [1821], in *The Works of Thomas De Quincey*, 2nd ed., vol. 2 (Edinburgh: Adam and Charles Black, 1863), 265.

37 Hermann von Helmholtz, “Über Goethes naturwissenschaftlichen Arbeiten” [1853], and “Goethes Vorahnungen kommender naturwissenschaftlichen Ideen,” in *Vorträge und Reden*, 4th ed., vol. 2 (Braunschweig: Friederich Viewig und Sohn, 1896), 344.

38 Baudelaire, “Salon de 1859,” 319 – 322.

guished between different species of imagination, including the artistic and the scientific. In what was perhaps the most exhaustive treatment of the subject, the French psychologist Théodore Ribot defended science against the charges that it “sometimes extinguished the imagination” but nonetheless insisted that the “plastic imagination” of artists and poets and the “scientific” imagination belonged to different species (and further distinguished varieties within each species). Whereas the plastic imagination was free to invent and to grant its inventions a degree of emotional reality, the scientific imagination was constrained by “rational necessities that regulate the development of the creative faculty; it cannot wander aimlessly; in each case its end is determined, and, in order to exist, that is to say, in order to be accepted, the invention must be subjected to predetermined conditions.”<sup>39</sup> For all his insistence on the existence and fecundity of the scientific imagination, Ribot could not free himself from a certain suspicion that imagination was linked to scientific error: the “false sciences” of astrology, alchemy, and magic represented for Ribot “the golden age of the creative imagination” in the history of science. In its 1902 survey of the psychology of creative mathematicians, the journal *Enseignement Mathématique* asked respondents, inter alia, whether “artistic, literary, musical, or, in particular, poetic occupations or relaxations seem to you of a nature to hinder mathematical invention, or to favor it, by the momentary rest they offer the mind?” It was apparently inconceivable that the exercise of the artistic imagination could promote the work of the mathematical imagination, except as a distraction in

the same category as “physical exercises” and “vacations.”<sup>40</sup>

It is against this historical background that we must read distrust of the imagination in science. The power of the imagination had long awakened fear among scientists – and theologians, poets, artists, and doctors, to boot – because it could make up a world of its own that was livelier, lovelier, or more logical than the real world. In extreme cases the imagination could conquer the body as well as the mind, leading not only to madness but also to violent somatic crises. But Enlightenment theorists of the imagination had been confident in the right and competence of reason to discipline the imagination. Geniuses of art and science exercised the same brand of controlled imagination, in contrast to the wild imaginations that tyrannized pregnant women, religious fanatics, or mesmerized *convulsionnaires*. Only in the early nineteenth century was fear of the imagination in science compounded with loathing. The causes lay in new views of the artistic imagination as freed from all constraints of reason and nature, and in a new polarity between objectivity and subjectivity. Wild, ineffable imagination became the driving force of creativity in art – and the bogey of objectivity in science. In their ideals, practices, and personae both art and science had mutated, and drifted apart.

What kind of objectivity bans the imagination from science? I have mentioned one moment of objectivity, the communitarian impulse that urges sci-

39 Théodore Ribot, *Essai sur l'imagination créatrice* (Paris: Félix Alcan, 1900), 198.

40 “Enquête sur la méthode de travail des mathématiciens,” *Enseignement Mathématique* 4 (1902): 208–211, Questions 19, 26, 28. Jacques Hadamard, *The Psychology of Invention in the Mathematical Field* [1945] (New York: Dover, 1949), reports on some results of the survey.

entists to standardize their instruments, clarify their concepts, and depersonalize their writing styles to achieve communicability and commensurability across continents and centuries, perhaps even across planets. Max Planck spoke in the name of this form of communitarian objectivity when he yearned for a physics that would be accessible “to physicists in all places, all times, all peoples, all cultures. Yes, the system of theoretical physics lays claim to validity not merely for the inhabitants of this earth, but also for the inhabitants of other heavenly bodies.”<sup>41</sup> Communitarian objectivity could not coexist with the artistic cultivation of individualism, which enshrined personal perspectives and identified the ineffable with originality.

There was, however, a second moment of scientific objectivity that emerged alongside communitarian objectivity in the mid-nineteenth century. In an earlier article, Peter Galison and I have called this second moment “mechanical objectivity”; it replaces judgment with data-reduction techniques, observers with self-registering instruments, hand-drawn illustrations with photographs.<sup>42</sup> Mechanical objectivity strives to eliminate human intervention in the phenomena, to “let nature speak for itself.” The free imagination celebrated by Baudelaire and other romantics threatened mechanical objectivity by projecting its own creations onto the facts of nature. Yet the facts envisioned by nineteenth-century scientists were not the fragile, pliable facts so carefully protected by their eighteenth-century prede-

41 Max Planck, *Acht Vorlesungen über theoretische Physik* (Leipzig: S. Hirzel, 1910), 6.

42 Lorraine Daston and Peter Galison, “The Image of Objectivity,” *Representations* 40 (Fall 1992): 81–128.

cessors from the distortions of system-builders. It was a byword that facts were angular, even truculent entities, sturdily resisting all attempts to ignore them or bend them to fit the Procrustean bed of theory. Huxley insisted that “a world of facts lies outside and beyond the world of words.”<sup>43</sup> In part, this change in scientific perception corresponded to a very concrete change in scientific practice: in the last quarter of the eighteenth century a new generation of instruments and measuring techniques made it possible to stabilize and replicate results with a success undreamed of fifty years earlier.<sup>44</sup> In some real sense, scientific facts had become more robust. Why, then, were the automated ideals and practices of mechanical objectivity necessary at all? Why couldn’t hard facts defend themselves against wild imagination?

The answer lies in a very different kind of fear that began to haunt scientists in the 1830s – the fear of vertiginous, open-ended progress. When Kant denied scientists genius, he had consoled them with progress: “The talent for science is formed for the continued advances of greater perfection in knowledge, with all its dependent practical advantages, as also for imparting the same to others. Hence scientists can boast a ground of considerable superiority over those who merit the honor of being called geniuses, since genius reaches a point at which art makes a halt, as there is a limit imposed upon it which it cannot transcend.”<sup>45</sup> In

43 Thomas Henry Huxley, “Scientific Education: Notes of an After-Dinner Speech” [1869], *Science and Education. Essays* (New York: Appleton, 1894), 115.

44 Christian Licoppe, *La formation de la pratique scientifique: Le discours de l’expérience en France et en Angleterre (1630–1820)* (Paris: Editions de la Découverte, 1996), 243–317.

45 Kant, *Critique of Judgment*, 170.



the late eighteenth century, the sciences did indeed seem destined for smooth, steady, unlimited progress. Between 1750 and 1840, a stream of histories of various sciences poured from the press, all purporting to demonstrate the existence and extent of progress in those disciplines.<sup>46</sup> But the progress envisioned in these optimistic histories was of change without transformation. Once the foundations for the new science had been laid in the seventeenth century, as the standard story went, the edifice could be expanded but not remodeled. In the 1830s this placid view of scientific progress received a rude shock when the wave theory unseated the Newtonian emission theory of light, most notably as a result of the research of French physicist Augustin Fresnel.<sup>47</sup> How could a tested theory of impeccable scientific credentials, its luster burnished by the name of Newton, be so thoroughly routed – not merely generalized or simplified? Was scientific progress so inexorable, so durable after all?

The response of scientists was to retreat to the level of the description of facts, in order to salvage a stable core of knowledge from the ebb and flow of theories. As Ernst Mach put it in 1872, history of science taught the Heraclitean lesson of *panta rhei*, for revolutions in science had become perpetual: “The attempts to hold fast to the beautiful moment through textbooks have always been futile. One gradually accustoms oneself [to the fact] that science is in-

complete, mutable.”<sup>48</sup> Mach held up Joseph Fourier’s heat theory as a “model theory” in science because it wasn’t really a theory at all, being founded only on “observable fact.”<sup>49</sup> The expectations for scientific progress voiced by Kant and others had not been disappointed; rather, they had been fulfilled with a vengeance. Never before had science bustled and flourished as it did in the latter half of the nineteenth century. Scientists multiplied in number, and with them, new theories, observations, and experiments. With these efforts, however, science not only grew; it also changed, and changed at a rate that could be measured in months rather than generations. No theory was safe from this breakneck progress, not even Newtonian celestial mechanics. By the 1890s Henri Poincaré was calling for ever more precise techniques of approximation in order to test whether Newton’s law alone could explain all observed astronomical phenomena.<sup>50</sup>

Within this maelstrom of change, only facts seemed to hold out the hope of definitive achievement in science. Like diamonds, scientific facts not only hardened but grew more precious to scientists in the nineteenth century – hence the fervor of proponents of mechanical objectivity in fending off all possible adulterations and distortions of facts by judgment or, especially, imagination. Eighteenth-century savants had revered facts but had believed them to be the alpha, not the omega, of scientific achieve-

46 Rachel Laudan, “Histories of Sciences and Their Uses: A Review to 1913,” *History of Science* 31 (1993): 5–12.

47 For a detailed account of this episode, see Jed Z. Buchwald, *The Rise of the Wave Theory of Light. Optical Theory and Experiment in the Early Nineteenth Century* (Chicago: University of Chicago Press, 1989).

48 Ernst Mach, *Die Geschichte und die Wurzel des Satzes von der Erhaltung der Arbeit* [1872], 2nd ed. (Leipzig: Johann Ambrosius Barth, 1879), 1.

49 Ernst Mach, *Die Principien der Wärmelehre* (Leipzig: Johann Ambrosius Barth, 1896), 115.

50 Henri Poincaré, *Les Méthodes nouvelles de la mécanique céleste*, vol. 1 (Paris: Gauthier-Villars, 1892–1899), 3–4.

ment. (It should be noted that in eighteenth-century classifications of knowledge the custodians of fact were not natural scientists, per se, but rather civil and natural historians.) Moreover, they were confident that facts mangled by the *esprit de système* or an errant imagination would ultimately be corrected by theory. Their nineteenth-century successors, caught up in the gallop of progress, had lost this innocent trust in the corrective power of theories that came and went like mayflies. Pure facts, severed from theory and sheltered from the imagination, were the last, best hope for permanence in scientific achievement. As anthropologists teach us, loathing stems from some breach of purity, some sacred boundary transgressed. The wild imagination potentially contaminated the purity of facts, and this is why it came not only to be feared but also loathed.

There is a rusting irony in the reversed fortunes of art and science, already visible in the mid-nineteenth-century writings of scientists. Alexander von Humboldt sadly reflected in 1844 on the contrast between ephemeral science and enduring literature, saying, "It has often been a discouraging consideration, that while purely literary products of the mind are rooted in the depth of feelings and creative imagination, all that is connected with empiricism and with fathoming of phenomena and physical law takes on a new aspect in a few decades, ... so that, as one commonly says, outdated scientific writings fall into oblivion as [no longer] readable."<sup>51</sup> By 1917 Max Weber could regard the opposition of transitory science to stable art to be a platitude, one that made it difficult to understand what sense it made to pursue science as a career. Near the end of World War I, addressing an audience of

Munich students who desperately wanted him to explain how science illuminated the meaning of life, Weber flatly asserted that science provided no such answers; science could hardly answer the question of what the meaning of a scientific career was. Why should one devote a lifetime of labor to producing a result that "in 10, 20, 50 years is outdated"? Subjective art endured, but objective science evaporated. Weber's own answer crowned this irony with yet one more. The spiritual motivation and reward for a lifetime devoted to science was exactly the same as for a lifetime devoted to art: science for science's sake, art for art's sake, the immolation of the personality in the service of "the pure object alone."<sup>52</sup> Having disavowed the artistic imagination and having lost the permanence of artistic achievement, science nonetheless aspired to the ascetic single-mindedness of art.

<sup>51</sup> Humboldt, *Cosmos*, vol. 1, xxiv.

<sup>52</sup> Max Weber, "Wissenschaft als Beruf" [1917], in *Max Weber Gesamtausgabe*, vol. 17, ed. Wolfgang J. Mommsen and Wolfgang Schluchter with Birgitt Morgenbrod (Tübingen: J. C. B. Mohr, 1992), 84–87.

# Steven Weinberg

## *Physics & history*

I am one of the few contributors to this issue of *Dædalus* who is not in any sense a historian. I work and live in the country of physics, but history is the place that I love to visit as a tourist. Here I wish to consider, from the perspective of a physicist, the uses that history has for physics, and the dangers both pose to each other.

I should begin by observing that one of the best uses of the history of physics is to help us teach physics to nonphysicists. Although many of them are very nice people, nonphysicists are rather odd. Physicists get tremendous pleasure out of being able to calculate all sorts of things, everything from the shape of a cable in a suspension bridge to the flight of a projectile or the energy of the hydrogen atom. Nonphysicists, for some rea-

son, do not appear to experience a comparable thrill in considering such matters. This is sad but true. It poses a problem, because if one intends to teach nonphysicists the machinery by which these calculations are done, one is simply not going to get a very receptive class. History offers a way around this pedagogical problem: everyone loves a story. For example, a professor can tell the story (as I did in a book and in courses at Harvard and Texas) of the discovery of the subatomic particles – the electron, the proton, and all the others.<sup>1</sup> In the course of learning this history, the student – in order to understand what was going on in the laboratories of J. J. Thomson, Ernest Rutherford, and our other heroes – has to learn something about how particles move under the influence of various forces, about energy and momentum, and about electric and magnetic fields. Thus, in order to understand the stories, they need to learn some of the physics we think they should know. It was Gerald Holton's 1952 book *Introduction to Concepts and Theories of Physical Science* that first utilized precisely this method of teaching physics; Holton told the story of the development of modern phys-

---

*Steven Weinberg, a Fellow of the American Academy since 1968 and winner of the Nobel Prize in Physics in 1979, wrote this essay for the Winter 1998 issue of "Dædalus." Since 1982, Weinberg has been Josey Regental Professor of Science at the University of Texas at Austin. He currently also holds the Jack S. Josey-Welch Foundation Chair in Science and directs the Theory Research Group at the University of Texas at Austin.*

---

© 2005 by the American Academy of Arts & Sciences

<sup>1</sup> Steven Weinberg, *The Discovery of Subatomic Particles* (San Francisco: Scientific American/Freeman, 1982).

ics, all the while using it as a vehicle for teaching physics. Unfortunately, despite his efforts and those of many who came after him, the problem of teaching physics to nonphysicists remains unsolved. It is still one of the great problems facing education – how to communicate “hard sciences” to an unwilling public. In many colleges throughout the country the effort has been given up completely. Visiting small liberal-arts colleges, one often finds that the only course offered in physics at all is the usual course for premedical students. Many undergraduates will thus never get the chance to encounter a book like Holton’s.

History plays a special role for elementary particle physicists like myself. In a sense, our perception of history resembles that of Western religions, Christianity and Judaism, as compared to the historical view of other branches of science, which are more like that of Eastern religious traditions. Christianity and Judaism teach that history is moving toward a climax – the day of judgment; similarly, many elementary particle physicists think that our work in finding deeper explanations of the nature of the universe will come to an end in a final theory toward which we are working. An opposing perception of history is held by those faiths that believe that history will go on forever, that we are bound to the wheel of endless reincarnation; likewise, particle physicists’ vision of history is quite different from that of most of the sciences. Other scientists look forward to an endless future of finding interesting problems – understanding consciousness, or turbulence, or high temperature superconductivity – that will go on forever. In elementary particle physics our aim is to put ourselves out of business. This gives a historical dimension to our choice of the sort of work on which to concentrate. We tend

to seek out problems that will further this historic goal – not just work that is interesting, useful, or that influences other fields, but work that is historically progressive, that moves us toward the goal of a final theory.

In this quest for a final theory, problems get bypassed. Things that once were at the frontier, as nuclear physics was in the 1930s, no longer are. This has happened recently to the theory of strong interactions. We now understand the strong forces that hold the quarks together inside the nuclear particles in terms of a quantum field theory called quantum chromodynamics. When I say that we understand these forces, I do not mean that we can do every calculation we might wish to do; we are still unable to solve some of the classic problems of strong interaction physics, such as calculating the mass of the proton (the nucleus of the hydrogen atom). A silly letter in *Physics Today* recently asked why we bother to talk about speculative fundamental theories like string theory when the longstanding problem of calculating the mass of the proton remains to be solved. Such criticism misses the point of research focused on a historical goal. We have solved enough problems using quantum chromodynamics to know that the theory is right; it is not necessary to mop up all the islands of unsolved problems in order to make progress toward a final theory. Our situation is a little like that of the U.S. Navy in World War II: bypassing Japanese strong points like Truk or Rabaul, the Navy instead moved on to take Saipan, which was closer to its goal of the Japanese home islands. We too must learn that we can bypass some problems. This is not to say that these problems are not worth working on; in fact, some of my own recent work has been in the application of quantum chromodynamics to nuclear physics.

Nuclear forces present a classic problem – one on which I was eager to work. But I am not under the illusion that this work is part of the historical progress toward a final theory. Nuclear forces present a problem that remains interesting, but not as part of the historical mission of fundamental physics.

If history has its value, it has its dangers as well. The danger in history is that in contemplating the great work of the past, the great heroic ideas – relativity, quantum mechanics, and so on – we develop such reverence for them that we become unable to reassess their place in what we envision as a final physical theory. An example of this is general relativity. As developed by Einstein in 1915, general relativity appears almost logically inevitable. There was a fundamental principle, Einstein's principle of the equivalence of gravitation and inertia, which says that there is no difference between gravity and the effects of inertia (like centrifugal force). The principle of equivalence can be reformulated as the principle that gravity is just an effect of the curvature of space and time – a beautiful principle from which Einstein's theory of gravitation follows almost uniquely. But there is an "almost" here. To arrive at the equations of general relativity, Einstein in 1915 had to make an additional assumption; he assumed that the equations of general relativity would be of a particular type, known as second-order partial differential equations. This is not the place to explain precisely what a second-order partial differential equation is – roughly speaking, it is an equation in which appear not only things like gravitational fields, and the rates at which these things change with time and position, but also second-order rates, the rates at which the rates change. It does not include higher order rates, for instance, third-order rates – the

rates at which the rates that are changing are changing. This may seem like a technicality, and it is certainly not a grand principle like the principle of equivalence. It is just a limit on the sorts of equations that will be allowed in the theory. So why did Einstein make this assumption – this very technical assumption, with no philosophical underpinnings? For one thing, people were used to such equations at the time: the equations of Maxwell that govern electromagnetic fields and the wave equations that govern the propagation of sound are all second-order differential equations. For a physicist in 1915, therefore, it was a natural assumption. If a theorist does not know what else to do, it is a good tactic to assume the simplest possibility; this is more likely to produce a theory that one could actually solve, providing at least the chance to decide whether or not it agrees with experiment. In Einstein's case, the tactic worked.

But this kind of pragmatic success does not in itself provide a rationale that would satisfy, of all people, Einstein. Einstein's goal was never simply to find theories that fit the data. Remember, it was Einstein who said that the purpose of the kind of physics he did was "not only to know how nature is and how her transactions are carried through, but also to reach as far as possible the utopian and seemingly arrogant aim of knowing why nature is thus and not otherwise . . . ." He certainly was not achieving that goal when he arbitrarily assumed that the equations for general relativity were second-order differential equations. He could have made them fourth-order differential equations, but he did not.

Our perspective on this today, which has been developing gradually over the last fifteen or twenty years, is different



from that of Einstein. Many of us now regard general relativity as nothing but an effective field theory – that is to say, a field theory that provides an approximation to a more fundamental theory, an approximation valid in the limit of large distances, probably including any distances that are larger than the scale of an atomic nucleus. Indeed, if one supposes that there really are terms in the Einstein equations that involve rates of fourth or higher order, such terms would still play no significant role at sufficiently large distances. This is why Einstein's tactic worked. There is a rational reason for assuming the equations are second-order differential equations, which is that any terms in the equations involving higher order rates would not make much of a difference in any astronomical observations. As far as I know, however, this was not Einstein's rationale.

This may seem rather a minor point to raise here, but in fact the most interesting work today in the study of gravitation is precisely in contexts in which the presence of higher-order rates in the field equations would make a big difference. The most important problem in the quantum theory of gravity arises from the fact that when one does various calculations – as, for instance, in attempting to calculate the probability that a gravitational wave will be scattered by another gravitational wave – one gets answers that turn out to be infinite. Another problem in the classical theory of gravitation arises from the presence of singularities: matter can apparently collapse to a point in space with infinite energy density and infinite space-time curvature. These absurdities, which have been exercising the attention of physicists for many decades, are precisely problems that involve gravity at very short distances – not the large distances of astronomy, but distances

much smaller than the size of an atomic nucleus.

From the point of view of modern effective field theory, there are no infinities in the quantum theory of gravity. The infinities are cancelled in exactly the same way that they are in all our other theories, by just being absorbed into a redefinition of parameters in the field equations; but this works only if we include terms involving rates of fourth order and all higher orders. (John Donaghue of the University of Massachusetts at Amherst has done more than anyone in showing how this works.) The old problems of infinities and singularities in the theory of gravitation cannot be dealt with by taking Einstein's theory seriously as a fundamental theory. From the modern point of view – if you like, from my point of view – Einstein's theory is nothing but an approximation valid at long distances, which cannot be expected to deal successfully with infinities and singularities. Yet some professional quantum gravitationalists (if that is the word) spend their whole careers studying the applications of the original Einstein theory, the one that only involves second-order differential equations, to problems involving infinities and singularities. Elaborate formalisms have been developed that aim to look at Einstein's theory in a more sophisticated way, in the hope that doing so will somehow or other make the infinities or singularities go away. This ill-placed loyalty to general relativity in its original form persists because of the enormous prestige the theory earned from its historic successes.

But it is precisely in this way that the great heroic ideas of the past can weigh upon us, preventing us from seeing things in a fresh light. Said another way, it is those ideas that were most successful of which we should be most wary. Otherwise we become like the French

army, which in 1914 tried to imitate the successes of Napoleon and almost lost the war – and then in 1940 tried to imitate the 1916 success of Marshall Petain in defending Verdun, only to suffer decisive defeat. Such examples exist in the history of physics as well. For instance, there is an approach to quantum field theory called second quantization, which fortunately no longer plays a significant role in research but continues to play a role in the way that textbooks are written. Second quantization goes back to a paper written in 1927 by Jordan and Klein that put forth the idea that after one has introduced a wave function in quantizing a theory of particles, you should then quantize the wave function. Surprisingly, many people think that this is the way to look at quantum field theory; it is not.

We have to expect the same fate for our present theories. The standard model of weak, electromagnetic, and strong forces, used to describe nature under conditions that can be explored in today's accelerators, may itself neither disappear nor be proved wrong but instead be looked at in quite a different way. Most particle physicists now think of the standard model as only an effective field theory that provides a low-energy approximation of a more fundamental theory.

Enough about the danger of history to science; let us now take up the danger of scientific knowledge to history. This arises from a tendency to imagine that discoveries are made according to our present understandings. Gerald Holton has done as much as anyone in trying to point out these dangers and puncture these misapprehensions. In his papers about Einstein he shows, for example, that the natural deduction of the special theory of relativity from the experiment of Michelson and Morley, which de-

monstrated that there is no motion through the ether, is not at all the way Einstein actually came to special relativity. Holton has also written about Kepler. At one point in my life I was one of those people who thought that Kepler deduced his famous three laws of planetary motion by studying the data of Tycho Brahe. But Holton points out how much else besides data, how much of the spirit of the Middle Ages and of the Greek world, went into Kepler's thinking – how many things that we now no longer associate with planetary motion were on Kepler's mind. By assuming that scientists of the past thought about things the way we do, we make mistakes; what is worse, we lose appreciation for the difficulties, for the intellectual challenges, that they faced.

Once, at the Tate Gallery in London, I heard a lecturer talking to a tour group about the Turner paintings. Turner was very important, said the guide, because he foreshadowed the Impressionists of the later nineteenth century. I had thought Turner was an important painter because he painted beautiful pictures; Turner did not know that he was foreshadowing anything. One has to look at things as they really were in their own time. This also applies, of course, to political history. Consider the term "Whig interpretation of history," which was invented by Herbert Butterfield in a lecture in 1931. As Butterfield explained it, "The Whig historian seems to believe that there is an unfolding logic in history." He went on to attack the person he regarded as the archetypal Whig historian, Lord Acton, who wished to use history as a way to pass moral judgments on the past. Acton wanted history to serve as the "arbiter of controversy, the upholder of that moral standard which the powers of earth and religion itself tend constantly to depress.... It is the office

of historical science to maintain morality as the sole impartial criterion of men and things.” Butterfield went on to say:

If history can do anything it is to remind of us of those complications that undermine our certainties, and to show us that all our judgments are merely relative to time and circumstance . . . . We can never assert that history has proved any man right in the long run. We can never say that the ultimate issue, the succeeding course of events, or the lapse of time have proved that Luther was right against the pope or that Pitt was wrong against Charles James Fox.<sup>2</sup>

This is the point at which the historian of science and the historian of politics should part company. The passage of time has shown that, for example, Darwin was right against Lamarck, the atomists were right against Ernst Mach, and Einstein was right against the experimentalist Walter Kaufman, who presented data contradicting special relativity. To put it another way, Butterfield was correct; there is no sense in which Whig morality (much less the Whig Party) existed at the time of Luther. But nevertheless it is true that natural selection was working during the time of Lamarck, and the atom did exist in the days of Mach, and fast electrons behaved according to the laws of relativity even before Einstein. Present scientific knowledge has the potentiality of being relevant in the history of science in the way that the present moral and political judgments may not be relevant in political or social history.

Many historians, sociologists, and philosophers of science have taken the desire for historicism, the worry about falling into a Whig interpretation of history, to extremes. To quote Holton,

2 Herbert Butterfield, *The Whig Interpretation of History* (New York: Scribners, 1951), 75.

“Much of the recent philosophical literature claims that science merely staggers from one fashion, conversion, revolution, or incommensurable exemplar to the next in a kind of perpetual, senseless Brownian motion, without discernible direction or goal.”<sup>3</sup> I made a similar observation in an address to the American Academy of Arts and Sciences about a year and a half ago, noting in passing that there are people who see scientific theories as nothing but social constructions. The talk was circulated by the Academy, as is their practice, and a copy of it fell into the hands of someone who over twenty years ago had been closely associated with a development known as the Sociology of Scientific Knowledge (SSK). He wrote me a long and unhappy letter; among other things, he complained about my remark that the Strong Program initiated at the University of Edinburgh embodied a radical social-constructivist view, in which scientific theories are nothing but social constructions. He sent me a weighty pile of essays, saying that they demonstrated that he and his colleagues do recognize that reality plays a role in our world. I took this criticism to heart and decided that I would read the essays. I also looked back over some old correspondence that I had had with Harry Collins, who for many years led the well-known Sociology of Scientific Knowledge group at the University of Bath. My purpose in all of this was to look at these materials from as sympathetic a point of view as I could, try to understand what they were saying, and assume that they must be saying something that is not absurd.

I did find described (though not espoused) in an article by David Bloor, who is one of the Edinburgh group, and

3 Gerald Holton, *Einstein, History, and Other Passions* (Reading, Mass.: Addison-Wesley, 1996), 22.

also in my correspondence with Harry Collins, a point of view that on the face of it is not absurd. As I understand it, there is a position called “methodological idealism” or “methodological antirealism,” which holds that historians or sociologists should take no position on what is ultimately true or real. Instead of using today’s scientific knowledge as a guiding principle for their work, the argument goes, they should try to look at nature as it must have been viewed by the scientists under study at the time that those scientists were working. In itself, this is not an absurd position. In particular, it can help to guard us against the kind of silliness that (for instance) I was guilty of when I interpreted Kepler’s work in terms of what we now know about planetary motion.

Even so, the attitude of methodological antirealism bothered me, though for a while I could not point to what I found wrong with it. In preparing this essay I have tried to think this through, and I have come to the conclusion that there are a number of minor things wrong with methodological antirealism: it can cripple historical research, it is often boring, and it is basically impossible. More significantly, however, it has a major drawback – in an almost literal sense, it misses the point of the history of science.

Let us first address the minor points. If it were really possible to reconstruct everything that happened during some past scientific discovery, then it might be helpful to forget everything that has happened since; but in fact much of what occurred will always be unknown to us. Consider just one example. J. J. Thomson, in the experiments that made him known as the discoverer of the electron, was measuring a certain crucial quantity, the ratio of the electron’s mass to its charge. As always happens,

he found a range of values. Although he quoted various values in his published work, the values he would always refer to as his favorite results were those at the high end of the range. Why did Thomson quote the high values as his favorite values? It is possible that Thomson knew that on the days those results had been obtained he had been more careful; perhaps those were the days he had not bumped into the laboratory table, or before which he’d had a good night’s sleep. But the possibility also exists that perhaps his first values had been at the high end of the range, and he was determined to show that he had been right at the beginning. Which explanation is correct? There is simply no way of reconstructing the past. Not his notebooks, not his biography – nothing will allow us now to reconstitute those days in the Cavendish Laboratory and find out on which days Thomson was more clumsy or felt more sleepy than usual. There is one thing that we do know, however: the actual value of the ratio of the electron’s mass to its charge, which was the same in Thomson’s time as in our own. We know, in fact, that the actual value is not at the high end but, rather, at the low end of the range of Thomson’s experimental values, which strongly suggests that when Thomson’s measurements gave high values they were not actually more careful – and that therefore it is more likely that Thomson quoted these values because he was just trying to justify his first measurements.

This is a trivial example of the use of present scientific knowledge in the history of science, because here we are just talking about a number, not a natural law or an ontological principle. I chose this example simply because it shows so clearly that to decide to ignore present scientific knowledge is often to throw away a valuable historical tool.

A second minor drawback of methodological antirealism is that a reader who does not know anything about our present understanding of nature is likely to find the history of science terribly boring. For instance, a historian might describe how in 1911 the Dutch physicist Kamerlingh Onnes was measuring the electrical resistance of a sample of cold mercury and thought that he had found a short circuit. The historian could go on for pages and pages describing how Onnes searched for the short circuit, and how he took apart the wiring and put it back together again without any success in finding the source of the short circuit. Could anything be more boring than to read this description if one did not know in advance that there *was* no short circuit – that what Onnes was observing was in fact the vanishing of the resistance of mercury when cooled to a certain temperature, and that this was nothing less than the discovery of superconductivity? Of course, it is impossible today for a physicist or a historian of physics not to know about superconductivity. Indeed, we are quite incapable while reading about the experiments of Kamerlingh Onnes of imagining that his problem was nothing but a short circuit. Even if one had never heard of superconductivity, the reader would know that there was something going on besides a short circuit; why else would the historian bother with these experiments? Plenty of experimental physicists have found short circuits, and no one studies them.

But these are minor issues. The main drawback of methodological antirealism is that it misses the point about the history of science that makes it different from other kinds of history: Even though a scientific theory is in a sense a social consensus, it is unlike any other sort of consensus in that it is culture-free and permanent.

This is just what many sociologists of science deny. David Bloor stated in a talk at Berkeley a year ago that “the important thing is that reality underdetermines the scientists’ understanding.” I gather he means that although he recognizes that reality has some effect on what scientists do – so that scientific theories are not “nothing but” social constructions – scientific theories are also not what they are simply because that is the way nature is. In a similar spirit, Stanley Fish, in a recent article in the *New York Times*, argued that the laws of physics are like the rules of baseball. Both are certainly conditioned by external reality – after all, if baseballs moved differently under the influence of Earth’s gravity, the rules would call for the bases to be closer together or further apart – but the rules of baseball also reflect the way that the game developed historically and the preferences of players and fans.<sup>4</sup>

Now, what Bloor and Fish say about the laws of nature does apply while these laws are being discovered. Holton’s work on Einstein, Kepler, and superconductivity has shown that many cultural and psychological influences enter into scientific work. But the laws of nature are not like the rules of baseball. They are culture-free and they are permanent – not as they are being developed, not as they were in the mind of the scientist who first discovers them, not in the course of what Latour and Woolgar call “negotiations” over what theory is going to be accepted, but in their final form, in which cultural influences are refined away. I will even use the dangerous words “nothing but”: aside from inessentials like the mathematical notation we use, the laws of physics as we understand them now are nothing but a description of reality.

4 Stanley Fish, “Professor Sokal’s Bad Joke,” *New York Times*, May 21, 1996, op-ed section.



I cannot prove that the laws of physics in their mature form are culture-free. Physicists live embedded in the Western culture of the late twentieth century, and it is natural to be skeptical if we say that our understanding of Maxwell's equations, quantum mechanics, relativity, or the standard model of elementary particles is culture-free. I am convinced of this because the purely scientific arguments for these theories seem to me overwhelmingly convincing. I can add that as the typical background of physicists has changed, in particular, as the number of women and Asians in physics has increased, the nature of our understanding of physics has not changed. These laws in their mature form have a toughness that resists cultural influence.

The history of science is further distinguished from political or artistic history (in such a way as to reinforce my remarks about the influence of culture) in that the achievements of science become permanent. This assertion may seem to contradict a statement at the beginning of this essay – that we now look at general relativity in a different way than Einstein did, and that even now we are beginning to look at the standard model differently than we did when it was first being developed. But what changes is our understanding of both why the theories are true and their scope of validity. For instance, at one time we thought there was an exact symmetry in nature between left and right, but then it was discovered that this is only true in certain contexts and to a certain degree of approximation. But the symmetry between right and left was not a simple mistake, nor has it been abandoned; we simply understand it better. Within its scope of validity, this symmetry has become a permanent part of science, and I cannot see that this will ever change.

In holding that the social constructivists missed the point, I have in mind a phenomenon known in mathematical physics as the approach to a fixed point. Various problems in physics deal with motion in some sort of space. Such problems are governed by equations dictating that wherever one starts in the space, one always winds up at the same point, known as a fixed point. Ancient geographers had something similar in mind when they said that all roads led to Rome. Physical theories are like fixed points, toward which we are attracted. Starting points may be culturally determined, paths may be affected by personal philosophies, but the fixed point is there nonetheless. It is something toward which any physical theory moves; when we get there we know it, and then we stop.

The kind of physics I have done for most of my life, working in the theory of fields and elementary particles, is moving toward a fixed point. But this fixed point is unlike any other in science. That final theory toward which we are moving will be a theory of unrestricted validity, a theory applicable to all phenomena throughout the universe – a theory that, when finally reached, will be a permanent part of our knowledge of the world. Then our work as elementary particle physicists will be done, and will become nothing but history.

# Robert N. Bellah

## *Civil religion in America*

While some have argued that Christianity is the national faith, and others that church and synagogue celebrate only the generalized religion of “the American Way of Life,” few have realized that there actually exists alongside of and rather clearly differentiated from the churches an elaborate and well-institutionalized civil religion in America. This article argues not only that there is such a thing, but also that this religion – or perhaps better, this religious dimension – has its own seriousness and integrity and requires the same care in understanding that any other religion does.<sup>1</sup>

Kennedy’s inaugural address of January 20, 1961 serves as an example and a clue with which to introduce this complex subject. That address began:

We observe today not a victory of party  
but a celebration of freedom – symboliz-

---

*Robert N. Bellah, Elliott Professor of Sociology Emeritus at the University of California, Berkeley, has been a Fellow of the American Academy since 1967. This essay appeared in the Winter 1967 issue of “Daedalus.” At the time of its publication, Bellah was professor of sociology at Harvard University.*

---

© 2005 by the American Academy of Arts & Sciences

ing an end as well as a beginning – signifying renewal as well as change. For I have sworn before you and Almighty God the same solemn oath our forebears prescribed nearly a century and three quarters ago.

1 Why something so obvious should have escaped serious analytical attention is in itself an interesting problem. Part of the reason is probably the controversial nature of the subject. From the earliest years of the nineteenth century, conservative religious and political groups have argued that Christianity is, in fact, the national religion. Some of them have from time to time and as recently as the 1950s proposed constitutional amendments that would explicitly recognize the sovereignty of Christ. In defending the doctrine of separation of church and state, opponents of such groups have denied that the national polity has, intrinsically, anything to do with religion at all. The moderates on this issue have insisted that the American state has taken a permissive and indeed supportive attitude toward religious groups (tax exemption, etc.), thus favoring religion but still missing the positive institutionalization with which I am concerned. But part of the reason this issue has been left in obscurity is certainly due to the peculiarly Western concept of religion as denoting a single type of collectivity of which an individual can be a member of one and only one at a time. The Durkheimian notion that every group has a religious dimension, which would be seen as obvious in southern or eastern Asia, is foreign to us. This obscures the recognition of such dimensions in our society.

The world is very different now. For man holds in his mortal hands the power to abolish all forms of human poverty and to abolish all forms of human life. And yet the same revolutionary beliefs for which our forebears fought are still at issue around the globe – the belief that the rights of man come not from the generosity of the state but from the hand of God.

And it concluded:

Finally, whether you are citizens of America or of the world, ask of us the same high standards of strength and sacrifice that we shall ask of you. With a good conscience our only sure reward, with history the final judge of our deeds, let us go forth to lead the land we love, asking His blessing and His help, but knowing that here on earth God's work must truly be our own.

These are the three places in this brief address in which Kennedy mentioned the name of God. If we could understand why he mentioned God, the way in which he did it, and what he meant to say in those three references, we would understand much about American civil religion. But this is not a simple or obvious task, and American students of religion would probably differ widely in their interpretation of these passages.

Let us consider first the placing of the three references. They occur in the two opening paragraphs and in the closing paragraph, thus providing a sort of frame for the more concrete remarks that form the middle part of the speech. Looking beyond this particular speech, we would find that similar references to God are almost invariably to be found in the pronouncements of American presidents on solemn occasions, though usually not in the working messages that the president sends to Congress on various concrete issues. How, then, are we to interpret this placing of references to God?

It might be argued that the passages quoted reveal the essentially irrelevant role of religion in the very secular society that is America. The placing of the references in this speech as well as in public life generally indicates that religion has “only a ceremonial significance”; it gets only a sentimental nod which serves largely to placate the more unenlightened members of the community, before a discussion of the really serious business with which religion has nothing whatever to do. A cynical observer might even say that an American president has to mention God or risk losing votes. A semblance of piety is merely one of the unwritten qualifications for the office, a bit more traditional than but not essentially different from the present-day requirement of a pleasing television personality.

But we know enough about the function of ceremony and ritual in various societies to make us suspicious of dismissing something as unimportant because it is “only a ritual.” What people say on solemn occasions need not be taken at face value, but it is often indicative of deep-seated values and commitments that are not made explicit in the course of everyday life. Following this line of argument, it is worth considering whether the very special placing of the references to God in Kennedy's address may not reveal something rather important and serious about religion in American life.

It might be countered that the very way in which Kennedy made his references reveals the essentially vestigial place of religion today. He did not refer to any religion in particular. He did not refer to Jesus Christ, or to Moses, or to the Christian church; certainly he did not refer to the Catholic church. In fact, his only reference was to the concept of God, a word which almost all Americans

can accept but which means so many different things to so many different people that it is almost an empty sign. Is this not just another indication that in America religion is considered vaguely to be a good thing, but that people care so little about it that it has lost any content whatever? Isn't Eisenhower reported to have said, "Our government makes no sense unless it is founded in a deeply felt religious faith – and I don't care what it is,"<sup>2</sup> and isn't that a complete negation of any real religion?

These questions are worth pursuing because they raise the issue of how civil religion relates to the political society, on the one hand, and to private religious organization, on the other. President Kennedy was a Christian, more specifically a Catholic Christian. Thus, his general references to God do not mean that he lacked a specific religious commitment. But why, then, did he not include some remark to the effect that Christ is the Lord of the world or some indication of respect for the Catholic church? He did not because these are matters of his own private religious belief and of his relation to his own particular church; they are not matters relevant in any direct way to the conduct of his public office. Others with different religious views and commitments to different churches or denominations are equally qualified participants in the political process. The principle of separation of church and state guarantees the freedom of religious belief and association but at the same time clearly segregates the religious sphere, which is considered to be essentially private, from the political one.

Considering the separation of church and state, how is a president justified in

using the word *God* at all? The answer is that the separation of church and state has not denied the political realm a religious dimension. Although matters of personal religious belief, worship, and association are considered to be strictly private affairs, there are, at the same time, certain common elements of religious orientation that the great majority of Americans share. These have played a crucial role in the development of American institutions and still provide a religious dimension for the whole fabric of American life, including the political sphere. This public religious dimension is expressed in a set of beliefs, symbols, and rituals that I am calling the American civil religion. The inauguration of a president is an important ceremonial event in this religion. It reaffirms, among other things, the religious legitimation of the lightest political authority.

Let us look more closely at what Kennedy actually said. First he said, "I have sworn before you and Almighty God the same solemn oath our forebears prescribed nearly a century and three quarters ago." The oath is the oath of office, including the acceptance of the obligation to uphold the Constitution. He swears it before the people (you) and God. Beyond the Constitution, then, the president's obligation extends not only to the people but to God. In American political theory, sovereignty rests, of course, with the people, but implicitly, and often explicitly, the ultimate sovereignty has been attributed to God. This is the meaning of the motto "In God we trust," as well as the inclusion of the phrase "under God" in the pledge to the flag. What difference does it make that sovereignty belongs to God? Though the will of the people as expressed in majority vote is carefully institutionalized as the operative source of political authori-

<sup>2</sup> Quoted in Will Herberg, *Protestant-Catholic-Jew* (Garden City, N.Y.: Doubleday, 1955), 97.

ty, it is deprived of an ultimate significance. The will of the people is not itself the criterion of right and wrong. There is a higher criterion in terms of which this will can be judged; it is possible that the people may be wrong. The president's obligation extends to the higher criterion.

When Kennedy says that "the rights of man come not from the generosity of the state but from the hand of God," he is stressing this point again. It does not matter whether the state is the expression of the will of an autocratic monarch or of the "people"; the rights of man are more basic than any political structure and provide a point of revolutionary leverage from which any state structure may be radically altered. That is the basis for his reassertion of the revolutionary significance of America.

But the religious dimension in political life as recognized by Kennedy not only provides a grounding for the rights of man which makes any form of political absolutism illegitimate; it also provides a transcendent goal for the political process. This is implied in his final words that "here on earth God's work must truly be our own." What he means here is, I think, more clearly spelled out in a previous paragraph, the wording of which, incidentally, has a distinctly biblical ring:

Now the trumpet summons us again – not as a call to bear arms, though arms we need – not as a call to battle, though embattled we are – but a call to bear the burden of a long twilight struggle, year in and year out, "rejoicing in hope, patient in tribulation" – a struggle against the common enemies of man: tyranny, poverty, disease and war itself.

The whole address can be understood as only the most recent statement of a theme that lies very deep in the Ameri-

can tradition, namely the obligation, both collective and individual, to carry out God's will on earth. This was the motivating spirit of those who founded America, and it has been present in every generation since. Just below the surface throughout Kennedy's inaugural address, it becomes explicit in the closing statement that God's work must be our own. That this very activist and non-contemplative conception of the fundamental religious obligation, which has been historically associated with the Protestant position, should be enunciated so clearly in the first major statement of the first Catholic president seems to underline how deeply established it is in the American outlook. Let us now consider the form and history of the civil religious tradition in which Kennedy was speaking.

The phrase *civil religion* is, of course, Rousseau's. In Chapter 8, Book 4, of *The Social Contract*, he outlines the simple dogmas of the civil religion: the existence of God, the life to come, the reward of virtue and the punishment of vice, and the exclusion of religious intolerance. All other religious opinions are outside the cognizance of the state and may be freely held by citizens. While the phrase *civil religion* was not used, to the best of my knowledge, by the founding fathers, and I am certainly not arguing for the particular influence of Rousseau, it is clear that similar ideas, as part of the cultural climate of the late eighteenth century, were to be found among the Americans. For example, Franklin writes in his autobiography:

I never was without some religious principles. I never doubted, for instance, the existence of the Deity; that he made the world and govern'd it by his Providence; that the most acceptable service of God



was the doing of good to men; that our souls are immortal; and that all crime will be punished, and virtue rewarded either here or hereafter. These I esteemed the essentials of every religion; and, being to be found in all the religions we had in our country, I respected them all, tho' with different degrees of respect, as I found them more or less mix'd with other articles, which, without any tendency to inspire, promote or confirm morality, serv'd principally to divide us, and make us unfriendly to one another.

It is easy to dispose of this sort of position as essentially utilitarian in relation to religion. In Washington's Farewell Address (though the words may be Hamilton's) the utilitarian aspect is quite explicit:

Of all the dispositions and habits which lead to political prosperity, Religion and Morality are indispensable supports. In vain would that man claim the tribute of Patriotism, who should labour to subvert these great Pillars of human happiness, these firmest props of the duties of men and citizens. The mere politician, equally with the pious man ought to respect and cherish them. A volume could not trace all their connections with private and public felicity. Let it simply be asked where is the security for property, for reputation, for life, if the sense of religious obligation *desert* the oaths, which are the instruments of investigation in the Courts of justice? And let us with caution indulge the supposition, that morality can be maintained without religion. Whatever may be conceded to the influence of refined education on minds of peculiar structure, reason and experience both forbid us to expect that National morality can prevail in exclusion of religious principle.

But there is every reason to believe that religion, particularly the idea of God,

played a constitutive role in the thought of the early American statesmen.

Kennedy's inaugural pointed to the religious aspect of the Declaration of Independence, and it might be well to look at that document a bit more closely. There are four references to God. The first speaks of the "Laws of Nature and of Nature's God" which entitle any people to be independent. The second is the famous statement that all men "are endowed by their Creator with certain inalienable Rights." Here Jefferson is locating the fundamental legitimacy of the new nation in a conception of "higher law" that is itself based on both classical natural law and biblical religion. The third is an appeal to "the Supreme Judge of the world for the rectitude of our intentions," and the last indicates "a firm reliance on the protection of divine Providence." In these last two references, a biblical God of history who stands in judgment over the world is indicated.

The intimate relation of these religious notions with the self-conception of the new republic is indicated by the frequency of their appearance in early official documents. For example, we find in Washington's first inaugural address of April 30, 1789:

It would be peculiarly improper to omit in this first official act my fervent supplications to that Almighty Being who rules over the universe, who presides in the councils of nations, and whose providential aids can supply every defect, that His benediction may consecrate to the liberties and happiness of the people of the United States a Government instituted by themselves for these essential purposes, and may enable every instrument employed in its administration to execute with success the functions allotted to his charge.

No people can be bound to acknowledge and adore the Invisible Hand which conducts the affairs of man more than those of the United States. Every step by which we have advanced to the character of an independent nation seems to have been distinguished by some token of providential agency . . . .

The propitious smiles of Heaven can never be expected on a nation that disregards the eternal rules of order and right which Heaven itself has ordained . . . . The preservation of the sacred fire of liberty and the destiny of the republican model of government are justly considered, perhaps, as *deeply*, as *finally*, staked on the experiment entrusted to the hands of the American people.

Nor did these religious sentiments remain merely the personal expression of the president. At the request of both houses of Congress, Washington proclaimed on October 3 of that same first year as president that November 26 should be “a day of public thanksgiving and prayer,” the first Thanksgiving Day under the Constitution.

The words and acts of the founding fathers, especially the first few presidents, shaped the form and tone of the civil religion as it has been maintained ever since. Though much is selectively derived from Christianity, this religion is clearly not itself Christianity. For one thing, neither Washington nor Adams nor Jefferson mentions Christ in his inaugural address; nor do any of the subsequent presidents, although not one of them fails to mention God.<sup>3</sup> The God of

3 God is mentioned or referred to in all inaugural addresses but Washington’s second, which is a very brief (two paragraphs) and perfunctory acknowledgment. It is not without interest that the actual word *God* does not appear until Monroe’s second inaugural, March 5, 1821. In his first inaugural, Washington refers to God as

the civil religion is not only rather “unitarian”; he is also on the austere side, much more related to order, law, and right than to salvation and love. Even though he is somewhat deist in cast, he is by no means simply a watchmaker God. He is actively interested and involved in history, with a special concern for America. Here the analogy has much less to do with natural law than with ancient Israel; the equation of America with Israel in the idea of the “American Israel” is not infrequent.<sup>4</sup> What was implicit in the words of Washington already quoted becomes explicit in Jefferson’s second inaugural when he said, “I shall need, too, the favor of that Being

---

“that Almighty Being who rules the universe,” “Great Author of every public and private good,” “Invisible Hand,” and “benign Parent of the Human Race.” John Adams refers to God as “Providence,” “Being who is supreme over all,” “Patron of Order,” “Fountain of Justice,” and “Protector in all ages of the world of virtuous liberty.” Jefferson speaks of “that Infinite Power which rules the destinies of the universe,” and “that Being in whose hands we are.” Madison speaks of “that Almighty Being whose power regulates the destiny of nations,” and “Heaven.” Monroe uses “Providence” and “the Almighty” in his first inaugural and finally “Almighty God” in his second. See *Inaugural Addresses of the Presidents of the United States from George Washington 1789 to Harry S. Truman 1949*, 82d Congress, 2d Session, House Document No. 540, 1952.

4 For example, Abiel Abbot, pastor of the First Church in Haverhill, Massachusetts, delivered a Thanksgiving sermon in 1799, *Traits of Resemblance in the People of the United States of America to Ancient Israel*, in which he said, “It has been often remarked that the people of the United States come nearer to a parallel with Ancient Israel, than any other nation upon the globe. Hence OUR AMERICAN ISRAEL is a term frequently used; and common consent allows it apt and proper.” Cited in Hans Kohn, *The Idea of Nationalism* (New York: Macmillan Publishing Co., 1961), 665.

in whose hands we are, who led our fathers, as Israel of old, from their native land and planted them in a country flowing with all the necessities and comforts of life." Europe is Egypt; America, the promised land. God has led his people to establish a new sort of social order that shall be a light unto all the nations.<sup>5</sup>

This theme, too, has been a continuous one in the civil religion. We have already alluded to it in the case of the Kennedy inaugural. We find it again in President Johnson's inaugural address:

They came here – the exile and the stranger, brave but frightened – to find a place where a man could be his own man. They made a covenant with this land. Conceived in justice, written in liberty, bound in union, it was meant one day to inspire the hopes of all mankind; and it binds us still. If we keep its terms, we shall flourish.

What we have, then, from the earliest years of the republic is a collection of beliefs, symbols, and rituals with respect to sacred things and institutionalized in a collectivity. This religion – there seems no other word for it – while not antithetical to and indeed sharing much in common with Christianity, was neither sec-

tarian nor in any specific sense Christian. At a time when the society was overwhelmingly Christian, it seems unlikely that this lack of Christian reference was meant to spare the feelings of the tiny non-Christian minority. Rather, the civil religion expressed what those who set the precedents felt was appropriate under the circumstances. It reflected their private as well as public views. Nor was the civil religion simply "religion in general." While generality was undoubtedly seen as a virtue by some, as in the quotation from Franklin above, the civil religion was specific enough when it came to the topic of America. Precisely because of this specificity, the civil religion was saved from empty formalism and served as a genuine vehicle of national religious self-understanding.

But the civil religion was not, in the minds of Franklin, Washington, Jefferson, or other leaders, with the exception of a few radicals like Tom Paine, ever felt to be a substitute for Christianity. There was an implicit but quite clear division of function between the civil religion and Christianity. Under the doctrine of religious liberty, an exceptionally wide sphere of personal piety and voluntary social action was left to the churches.

But the churches were neither to control the state nor to be controlled by it. The national magistrate, whatever his private religious views, operates under the rubrics of the civil religion as long as he is in his official capacity, as we have already seen in the case of Kennedy. This accommodation was undoubtedly the product of a particular historical moment and of a cultural background dominated by Protestantism of several varieties and by the Enlightenment, but it has survived despite subsequent changes in the cultural and religious climate.

5 That the Mosaic analogy was present in the minds of leaders at the very moment of the birth of the republic is indicated in the designs proposed by Franklin and Jefferson for a seal of the United States of America. Together with Adams, they formed a committee of three delegated by the Continental Congress on July 4, 1776, to draw up the new device. "Franklin proposed as the device Moses lifting up his wand and dividing the Red Sea while Pharaoh was overwhelmed by its waters, with the motto 'Rebellion to tyrants is obedience to God.' Jefferson proposed the children of Israel in the wilderness 'led by a cloud by day and a pillar of fire at night.'" Anson Phelps Stokes, *Church and State in the United States*, vol. 1 (New York: Harper, 1950), 467 – 468.

Until the Civil War, the American civil religion focused above all on the event of the Revolution, which was seen as the final act of the Exodus from the old lands across the waters. The Declaration of Independence and the Constitution were the sacred scriptures and Washington the divinely appointed Moses who led his people out of the hands of tyranny. The Civil War, which Sidney Mead calls “the center of American history,”<sup>6</sup> was the second great event that involved the national self-understanding so deeply as to require expression in the civil religion. In 1835, Tocqueville wrote that the American republic had never really been tried, that victory in the Revolutionary War was more the result of British preoccupation elsewhere and the presence of a powerful ally than of any great military success of the Americans. But in 1861 the time of testing had indeed come. Not only did the Civil War have the tragic intensity of fratricidal strife, but it was one of the bloodiest wars of the nineteenth century; the loss of life was far greater than any previously suffered by Americans.

The Civil War raised the deepest questions of national meaning. The man who not only formulated but in his own person embodied its meaning for Americans was Abraham Lincoln. For him the issue was not in the first instance slavery but “whether that nation, or any nation so conceived, and so dedicated, can long endure.” He had said in Independence Hall in Philadelphia on February 22, 1861:

All the political sentiments I entertain have been drawn, so far as I have been able to draw them, from the sentiments which originated in and were given to the world from this Hall. I have never had a feeling,

6 Sidney Mead, *The Lively Experiment* (New York: Harper & Row, 1963), 12.

politically, that did not spring from the sentiments embodied in the Declaration of Independence.<sup>7</sup>

The phrases of Jefferson constantly echo in Lincoln’s speeches. His task was, first of all, to save the Union – not for America alone but for the meaning of America to the whole world so unforgettably etched in the last phrase of the Gettysburg Address.

But inevitably the issue of slavery as the deeper cause of the conflict had to be faced. In the second inaugural, Lincoln related slavery and the war in an ultimate perspective:

If we shall suppose that American slavery is one of those offenses which, in the providence of God, must needs come, but which, having continued through His appointed time, He now wills to remove, and that He gives both to the North and South this terrible war as the woe due to those by whom the offense came, shall we discern therein any departure from those divine attributes which the believers in a living God always ascribe to Him? Fondly do we hope, fervently do we pray, that this mighty scourge of war may speedily pass away. Yet, if God wills that it continue until all the wealth piled by the bondsman’s two hundred and fifty years of unrequited toil shall be sunk, and until every drop of blood drawn with the lash shall be paid by another drawn with the sword, as was said three thousand years ago, so still it must be said “the judgements of the Lord are true and righteous altogether.”

But he closes on a note if not of redemption then of reconciliation – “With malice toward none, with charity for all.”

With the Civil War, a new theme of death, sacrifice, and rebirth enters the

7 Quoted by Arthur Lehman Goodhart in Allan Nevins, ed., *Lincoln and the Gettysburg Address* (Urbana: University of Illinois Press, 1964), 39.



civil religion. It is symbolized in the life and death of Lincoln. Nowhere is it stated more vividly than in the Gettysburg Address, itself part of the Lincolnian “New Testament” among the civil scriptures. Robert Lowell has recently pointed out the “insistent use of birth images” in this speech explicitly devoted to “these honored dead”: “brought forth,” “conceived,” “created,” “a new birth of freedom.” He goes on to say:

The Gettysburg Address is a symbolic and sacramental act. Its verbal quality is resonance combined with a logical, matter of fact, prosaic brevity. . . . In his words, Lincoln symbolically died, just as the Union soldiers really died – and as he himself was soon ready to die. By his words, he gave the field of battle a symbolic significance that it had lacked. For us and our country, he left Jefferson’s ideals of freedom and equality joined to the Christian sacrificial act of death and rebirth. I believe this is a meaning that goes beyond sect or religion and beyond peace and war, and is now part of our lives as a challenge, obstacle and hope.<sup>8</sup>

Lowell is certainly right in pointing out the Christian quality of the symbolism here, but he is also right in quickly disavowing any sectarian implication. The earlier symbolism of the civil religion had been Hebraic without being in any specific sense Jewish. The Gettysburg symbolism (“. . . those who here gave their lives, that that nation might live”) is Christian without having anything to do with the Christian church.

The symbolic equation of Lincoln with Jesus was made relatively early. Herndon, who had been Lincoln’s law partner, wrote:

For fifty years God rolled Abraham Lincoln through his fiery furnace. He did it

8 Ibid., 88 – 89.

to try Abraham and to purify him for his purposes. This made Mr. Lincoln humble, tender, forbearing, sympathetic to suffering, kind, sensitive, tolerant; broadening, deepening and widening his whole nature; making him the noblest and loveliest character since Jesus Christ. . . . I believe that Lincoln was God’s chosen one.<sup>9</sup>

With the Christian archetype in the background, Lincoln, “our martyred president,” was linked to the war dead, those who “gave the last full measure of devotion.” The theme of sacrifice was indelibly written into the civil religion.

The new symbolism soon found both physical and ritualistic expression. The great number of the war dead required the establishment of a number of national cemeteries. Of these, the Gettysburg National Cemetery, which Lincoln’s famous address served to dedicate, has been overshadowed only by the Arlington National Cemetery. Begun somewhat vindictively on the Lee estate across the river from Washington, partly with the end that the Lee family could never reclaim it,<sup>10</sup> it has subsequently become the most hallowed monument of the civil religion. Not only was a section set aside for the Confederate dead, but it has received the dead of each succeeding American war. It is the site of the one important new symbol to come out of World War I, the Tomb of the Unknown Soldier; more recently it has become the site of the tomb of another martyred president and its symbolic eternal flame.

Memorial Day, which grew out of the Civil War, gave ritual expression to the

9 Quoted in Sherwood Eddy, *The Kingdom of God and the American Dream* (New York: Harper & Brothers, 1941), 162.

10 Karl Decker and Angus McSween, *Historic Arlington* (Washington, D.C.: Decker and McSween Publishing Co., 1892), 60 – 67.



themes we have been discussing. As Lloyd Warner has so brilliantly analyzed it, the Memorial Day observance, especially in the towns and smaller cities of America, is a major event for the whole community involving a rededication to the martyred dead, to the spirit of sacrifice, and to the American vision.<sup>11</sup> Just as Thanksgiving Day, which incidentally was securely institutionalized as an annual national holiday only under the presidency of Lincoln, serves to integrate the family into the civil religion, so Memorial Day has acted to integrate the local community into the national cult. Together with the less overtly religious Fourth of July and the more minor celebrations of Veterans Day and the birthdays of Washington and Lincoln, these two holidays provide an annual ritual

11 How extensive the activity associated with Memorial Day can be is indicated by Warner: "The sacred symbolic behavior of Memorial Day, in which scores of the town's organizations are involved, is ordinarily divided into four periods. During the year separate rituals are held by many of the associations for their dead, and many of these activities are connected with later Memorial Day events. In the second phase, preparations are made during the last three or four weeks for the ceremony itself, and some of the associations perform public rituals. The third phase consists of scores of rituals held in all the cemeteries, churches, and halls of the associations. These rituals consist of speeches and highly ritualized behavior. They last for two days and are climaxed by the fourth and last phase, in which all the separate celebrants gather in the center of the business district on the afternoon of Memorial Day. The separate organizations, with their members in uniform or with fitting insignia, march through the town, visit the shrines and monuments of the hero dead, and, finally, enter the cemetery. Here dozens of ceremonies are held, most of them highly symbolic and formalized." During these various ceremonies Lincoln is continually referred to and the Gettysburg Address recited many times. W. Lloyd Warner, *American Life* (Chicago: University of Chicago Press, 1962), 8-9.

calendar for the civil religion. The public school system serves as a particularly important context for the cultic celebration of the civil rituals.

In reifying and giving a name to something that, though pervasive enough when you look at it, has gone on only semiconsciously, there is risk of severely distorting the data. But the reification and the naming have already begun. The religious critics of "religion in general," or of the "religion of the 'American Way of Life,'" or of "American Shinto" have really been talking about the civil religion. As usual in religious polemic, they take as criteria the best in their own religious tradition and as typical the worst in the tradition of the civil religion. Against these critics, I would argue that the civil religion at its best is a genuine apprehension of universal and transcendent religious reality as seen in or, one could almost say, as revealed through the experience of the American people. Like all religions, it has suffered various deformations and demonic distortions. At its best, it has neither been so general that it has lacked incisive relevance to the American scene nor so particular that it has placed American society above universal human values. I am not at all convinced that the leaders of the churches have consistently represented a higher level of religious insight than the spokesmen of the civil religion. Reinhold Niebuhr has this to say of Lincoln, who never joined a church and who certainly represents civil religion at its best:

An analysis of the religion of Abraham Lincoln in the context of the traditional religion of his time and place and of its polemical use on the slavery issue, which corrupted religious life in the days before and during the Civil War, must lead to the conclusion that Lincoln's religious conviction

tions were superior in depth and purity to those, not only of the political leaders of his day, but of the religious leaders of the era.<sup>12</sup>

Perhaps the real animus of the religious critics has been not so much against the civil religion in itself but against its pervasive and dominating influence within the sphere of church religion. As S. M. Lipset has recently shown, American religion at least since the early nineteenth century has been predominantly activist, moralistic, and social rather than contemplative, theological, or innerly spiritual.<sup>13</sup> Tocqueville spoke of American church religion as “a political institution which powerfully contributes to the maintenance of a democratic republic among the Americans”<sup>14</sup> by supplying a strong moral consensus amidst continuous political change. Henry Bary in 1902 spoke of American church religion as “*la poésie du civisme*.”<sup>15</sup>

12 Reinhold Niebuhr, “The Religion of Abraham Lincoln,” in Nevins, *Lincoln and the Gettysburg Address*, 72. William J. Wolfe of the Episcopal Theological School in Cambridge, Massachusetts, has written: “Lincoln is one of the greatest theologians of America – not in the technical meaning of producing a system of doctrine, certainly not as the defender of some one denomination, but in the sense of seeing the hand of God intimately in the affairs of nations. Just so the prophets of Israel criticized the events of their day from the perspective of the God who is concerned for history and who reveals His will within it. Lincoln now stands among God’s latter-day prophets.” *The Religion of Abraham Lincoln* (New York: n.p., 1963), 24.

13 Seymour Martin Lipset, “Religion and American Values,” in *The First New Nation* (New York: Basic Books, 1963).

14 Alexis de Tocqueville, *Democracy in America*, vol. I (New York: Vintage Books, 1954), 310.

15 Henry Bary, *La Religion dans la société aux Etats-Unis* (Paris: A. Colin, 1902), 31.

It is certainly true that the relation between religion and politics in America has been singularly smooth. This is in large part due to the dominant tradition. As Tocqueville wrote:

The greatest part of British America was peopled by men who, after having shaken off the authority of the Pope, acknowledged no other religious supremacy: they brought with them into the New World a form of Christianity which I cannot better describe than by styling it a democratic and republican religion.<sup>16</sup>

The churches opposed neither the Revolution nor the establishment of democratic institutions. Even when some of them opposed the full institutionalization of religious liberty, they accepted the final outcome with good grace and without nostalgia for an *ancien régime*. The American civil religion was never anticlerical or militantly secular. On the contrary, it borrowed selectively from the religious tradition in such a way that the average American saw no conflict between the two. In this way, the civil religion was able to build up without any bitter struggle with the church powerful symbols of national solidarity and to mobilize deep levels of personal motivation for the attainment of national goals.

Such an achievement is by no means to be taken for granted. It would seem that the problem of a civil religion is quite general in modern societies and

16 Tocqueville, *Democracy in America*, 311. Later he says, “In the United States even the religion of most of the citizens is republican, since it submits the truths of the other world to private judgment, as in politics the care of their temporal interests is abandoned to the good sense of the people. Thus every man is allowed freely to take that road which he thinks will lead him to heaven, just as the law permits every citizen to have the right of choosing his own government” (436).

that the way it is solved or not solved will have repercussions in many spheres. One needs only to think of France to see how differently things can go. The French Revolution was anticlerical to the core and attempted to set up an anti-Christian civil religion. Throughout modern French history, the chasm between traditional Catholic symbols and the symbolism of 1789 has been immense.

American civil religion is still very much alive. Just three years ago we participated in a vivid reenactment of the sacrifice theme in connection with the funeral of our assassinated president. The American Israel theme is clearly behind both Kennedy's New Frontier and Johnson's Great Society. Let me give just one recent illustration of how the civil religion serves to mobilize support for the attainment of national goals. On March 15, 1965 President Johnson went before Congress to ask for a strong voting-rights bill. Early in the speech he said:

Rarely are we met with the challenge, not to our growth or abundance, or our welfare or our security – but rather to the values and the purposes and the meaning of our beloved nation.

The issue of equal rights for American Negroes is such an issue. And should we defeat every enemy, and should we double our wealth and conquer the stars and still be unequal to this issue, then we will have failed as a people and as a nation.

For with a country as with a person, "What is a man profited, if he shall gain the whole world, and lose his own soul?"

And in conclusion he said:

Above the pyramid on the great seal of the United States it says in Latin, "God has favored our undertaking."

God will not favor everything that we do. It is rather our duty to divine his will. I cannot help but believe that He truly understands and that He really favors the undertaking that we begin here tonight.<sup>17</sup>

The civil religion has not always been invoked in favor of worthy causes. On the domestic scene, an American Legion type of ideology that fuses God, country, and flag has been used to attack nonconformist and liberal ideas and groups of all kinds. Still, it has been difficult to use the words of Jefferson and Lincoln to support special interests and undermine personal freedom. The defenders of slavery before the Civil War came to reject the thinking of the Declaration of Independence. Some of the most consistent of them turned against not only Jeffersonian democracy but Reformation religion; they dreamed of a South dominated by medieval chivalry and divine-right monarchy.<sup>18</sup> For all the overt religiosity of the radical right today, their relation to the civil religious consensus is tenuous, as when the John Birch Society attacks the central American symbol of democracy itself.

With respect to America's role in the world, the dangers of distortion are greater and the built-in safeguards of the tradition weaker. The theme of the American Israel was used, almost from the beginning, as a justification for the shameful treatment of the Indians so characteristic of our history. It can be overtly or implicitly linked to the idea of manifest destiny which has been used to legitimate several adventures in imperialism since the early nineteenth century.

<sup>17</sup> House, *U.S. Congressional Record*, March 15, 1965, 4924, 4926.

<sup>18</sup> See Louis Hartz, "The Feudal Dream of the South," in *The Liberal Tradition in America* (New York: Harcourt, Brace, 1955).

Never has the danger been greater than today. The issue is not so much one of imperial expansion, of which we are accused, as of the tendency to assimilate all governments or parties in the world which support our immediate policies or call upon our help by invoking the notion of free institutions and democratic values. Those nations that are for the moment "on our side" become "the free world." A repressive and unstable military dictatorship in South Vietnam becomes "the free people of South Vietnam and their government." It is then part of the role of America as the New Jerusalem and "the last hope of earth" to defend such governments with treasure and eventually with blood. When our soldiers are actually dying, it becomes possible to consecrate the struggle further by invoking the great theme of sacrifice. For the majority of the American people who are unable to judge whether the people in South Vietnam (or wherever) are "free like us," such arguments are convincing. Fortunately, President Johnson has been less ready to assert that "God has favored our undertakings" in the case of Vietnam than with respect to civil rights. But others are not so hesitant. The civil religion has exercised long-term pressure for the humane solution of our greatest domestic problem, the treatment of the Negro American. It remains to be seen how relevant it can become for our role in the world at large, and whether we can effectually stand for "the revolutionary beliefs for which our forebears fought," in John F. Kennedy's words.

The civil religion is obviously involved in the most pressing moral and political issues of the day. But it is also caught in another kind of crisis, theoretical and theological, of which it is at the moment largely unaware. "God" has clearly been

a central symbol in the civil religion from the beginning and remains so today. This symbol is just as central to the civil religion as it is to Judaism or Christianity. In the late eighteenth century this posed no problem; even Tom Paine, contrary to his detractors, was not an atheist. From left to right and regardless of church or sect, all could accept the idea of God. But today, as even *Time* has recognized, the meaning of the word *God* is by no means so clear or so obvious. There is no formal creed in the civil religion. We have had a Catholic president; it is conceivable that we could have a Jewish one. But could we have an agnostic president? Could a man with conscientious scruples about using the word *God* the way Kennedy and Johnson have used it be elected chief magistrate of our country? If the whole God symbolism requires reformulation, there will be obvious consequences for the civil religion, consequences perhaps of liberal alienation and of fundamentalist ossification that have not so far been prominent in this realm. The civil religion has been a point of articulation between the profoundest commitments of the Western religious and philosophical tradition and the common beliefs of ordinary Americans. It is not too soon to consider how the deepening theological crisis may affect the future of this articulation.

In conclusion it may be worthwhile to relate the civil religion to the most serious situation that we as Americans now face, what I call the third time of trial. The first time of trial had to do with the question of independence, whether we should or could run our own affairs in our own way. The second time of trial was over the issue of slavery, which in turn was only the most salient aspect of the more general problem of the full in-



stitutionalization of democracy within our country. This second problem we are still far from solving though we have some notable successes to our credit. But we have been overtaken by a third great problem which has led to a third great crisis, in the midst of which we stand. This is the problem of responsible action in a revolutionary world, a world seeking to attain many of the things, material and spiritual, that we have already attained. Americans have, from the beginning, been aware of the responsibility and the significance our republican experiment has for the whole world. The first internal political polarization in the new nation had to do with our attitude toward the French Revolution. But we were small and weak then, and "foreign entanglements" seemed to threaten our very survival. During the last century, our relevance for the world was not forgotten, but our role was seen as purely exemplary. Our democratic republic rebuked tyranny by merely existing. Just after World War I we were on the brink of taking a different role in the world, but once again we turned our back.

Since World War II the old pattern has become impossible. Every president since Roosevelt has been groping toward a new pattern of action in the world, one that would be consonant with our power and our responsibilities. For Truman and for the period dominated by John Foster Dulles that pattern was seen to be the great Manichaeian confrontation of East and West, the confrontation of democracy and "the false philosophy of communism" that provided the structure of Truman's inaugural address. But with the last years of Eisenhower and with the successive two presidents, the pattern began to shift. The great problems came to be seen as caused not solely by the evil intent of any one group of men, but as stemming from much more

complex and multiple sources. For Kennedy, it was not so much a struggle against particular men as against "the common enemies of man: tyranny, poverty, disease and war itself."

But in the midst of this trend toward a less primitive conception of ourselves and our world, we have somehow, without anyone really intending it, stumbled into a military confrontation where we have come to feel that our honor is at stake. We have in a moment of uncertainty been tempted to rely on our overwhelming physical power rather than on our intelligence, and we have, in part, succumbed to this temptation. Bewildered and unnerved when our terrible power fails to bring immediate success, we are at the edge of a chasm the depth of which no man knows.

I cannot help but think of Robinson Jeffers, whose poetry seems more apt now than when it was written, when he said:

Unhappy country, what wings you  
have! . . .  
Weep (it is frequent in human affairs),  
weep for the terrible magnificence of  
the means,  
The ridiculous incompetence of the  
reasons, the bloody and shabby  
Pathos of the result.

But as so often before in similar times, we have a man of prophetic stature, without the bitterness or misanthropy of Jeffers, who, as Lincoln before him, calls this nation to its judgment:

When a nation is very powerful but lacking in self-confidence, it is likely to behave in a manner that is dangerous both to itself and to others.

Gradually but unmistakably, America is succumbing to that arrogance of power which has afflicted weakened and in some cases destroyed great nations in the past.



If the war goes on and expands, if that fatal process continues to accelerate until America becomes what it is not now and never has been, a seeker after unlimited power and empire, then Vietnam will have had a mighty and tragic fallout indeed.

I do not believe that will happen. I am very apprehensive but I still remain hopeful, and even confident, that America, with its humane and democratic traditions, will find the wisdom to match its power.<sup>19</sup>

Without an awareness that our nation stands under higher judgment, the tradition of the civil religion would be dangerous indeed. Fortunately, the prophetic voices have never been lacking. Our present situation brings to mind the Mexican-American war that Lincoln, among so many others, opposed. The spirit of civil disobedience that is alive today in the civil rights movement and the opposition to the Vietnam war was already clearly outlined by Henry David Thoreau when he wrote, "If the law is of such a nature that it requires you to be an agent of injustice to another, then I say, break the law." Thoreau's words "I would remind my countrymen that they are men first, and Americans at a late and convenient hour"<sup>20</sup> provide an essential standard for any adequate thought and action in our third time of trial. As Americans, we have been well favored in the world, but it is as men that we will be judged.

Out of the first and second times of trial have come, as we have seen, the ma-

lor symbols of the American civil religion. There seems little doubt that a successful negotiation of this third time of trial – the attainment of some kind of viable and coherent world order – would precipitate a major new set of symbolic forms. So far the flickering flame of the United Nations burns too low to be the focus of a cult, but the emergence of a genuine transnational sovereignty would certainly change this. It would necessitate the incorporation of vital international symbolism into our civil religion, or, perhaps a better way of putting it, it would result in American civil religion becoming simply one part of a new civil religion of the world. It is useless to speculate on the form such a civil religion might take, though it obviously would draw on religious traditions beyond the sphere of biblical religion alone. Fortunately, since the American civil religion is not the worship of the American nation but an understanding of the American experience in the light of ultimate and universal reality, the reorganization entailed by such a new situation need not disrupt the American civil religion's continuity. A world civil religion could be accepted as a fulfillment and not a denial of American civil religion. Indeed, such an outcome has been the eschatological hope of American civil religion from the beginning. To deny such an outcome would be to deny the meaning of America itself.

Behind the civil religion at every point lie biblical archetypes: Exodus, Chosen People, Promised Land, New Jerusalem, Sacrificial Death and Rebirth. But it is also genuinely American and genuinely new. It has its own prophets and its own martyrs, its own sacred events and sacred places, its own solemn rituals and symbols. It is concerned that America be a society as perfectly in accord with

19 Speech of Senator J. William Fulbright of April 28, 1966, as reported in *The New York Times*, April 29, 1966.

20 Quoted in Yehoshua Arieli, *Individualism and Nationalism in American Ideology* (Cambridge, Mass.: Harvard University Press, 1964), 274.

the will of God as men can make it, and a light to all the nations.

It has often been used and is being used today as a cloak for petty interests and ugly passions. It is in need – as is any living faith – of continual reformation, of being measured by universal standards. But it is not evident that it is incapable of growth and new insight.

It does not make any decision for us. It does not remove us from moral ambiguity, from being, in Lincoln's fine phrase, an "almost chosen people." But it is a heritage of moral and religious experience from which we still have much to learn as we formulate the decisions that lie ahead.

# Clifford Geertz

## *Deep play: notes on the Balinese cockfight*

Early in April of 1958, my wife and I arrived, malarial and diffident, in a Balinese village we intended, as anthropologists, to study. A small place, about five hundred people, and relatively remote, it was its own world. We were intruders, professional ones, and the villagers dealt with us as Balinese seem always to deal with people not part of their life who yet press themselves upon them: as though we were not there. For them, and to a degree for ourselves, we were nonpersons, specters, invisible men.

We moved into an extended family compound (that had been arranged before through the provincial government) belonging to one of the four major factions in village life. But except for our landlord and the village chief, whose cousin and brother-in-law he was, everyone ignored us in a way only a Balinese can do. As we wandered around, un-

certain, wistful, eager to please, people seemed to look right through us with a gaze focused several yards behind us on some more actual stone or tree. Almost nobody greeted us; but nobody scowled or said anything unpleasant to us either, which would have been almost as satisfactory. If we ventured to approach someone (something one is powerfully inhibited from doing in such an atmosphere), he moved, negligently but definitively, away. If, seated or leaning against a wall, we had him trapped, he said nothing at all, or mumbled what for the Balinese is the ultimate nonword – “yes.” The indifference, of course, was studied; the villagers were watching every move we made and they had an enormous amount of quite accurate information about who we were and what we were going to be doing. But they acted as if we simply did not exist, which, in fact, as this behavior was designed to inform us, we did not, or anyway not yet.

This is, as I say, general in Bali. Everywhere else I have been in Indonesia, and more latterly in Morocco, when I have gone into a new village people have poured out from all sides to take a very close look at me, and, often, an all-too-probing feel as well. In Balinese villages, at least those away from the tourist circuit, nothing happens at all. People go

---

*Clifford Geertz, a Fellow of the American Academy since 1966, is professor emeritus at the Institute for Advanced Study in Princeton, New Jersey. At the time of this essay's publication in the Winter 1972 issue of "Daedalus," Geertz was professor of social sciences at the Institute for Advanced Study.*

---

© 2005 by the American Academy of Arts & Sciences

on pounding, chatting, making offerings, staring into space, carrying baskets about while one drifts around feeling vaguely disembodied. And the same thing is true on the individual level. When you first meet a Balinese, he seems virtually not to relate to you at all: he is, in the term Gregory Bateson and Margaret Mead made famous, “away.”<sup>1</sup> Then – in a day, a week, a month (with some people the magic moment never comes) – he decides, for reasons I have never been quite able to fathom, that you *are* real, and then he becomes a warm, gay, sensitive, sympathetic, though, being Balinese, always precisely controlled person. You have crossed, somehow, some moral or metaphysical shadow line. Though you are not exactly taken as a Balinese (one has to be born to that), you are at least regarded as a human being rather than a cloud or a gust of wind. The whole complexion of your relationship dramatically changes to, in the majority of cases, a gentle, almost affectionate one – a low-keyed, rather playful, rather mannered, rather bemused geniality.

My wife and I were still very much in the gust of wind stage, a most frustrating, and even, as you soon begin to doubt whether you are really real after all, unnerving one, when, ten days or so after our arrival, a large cockfight was held in the public square to raise money for a new school.

Now, a few special occasions aside, cockfights are illegal in Bali under the Republic (as, for not altogether unrelated reasons, they were under the Dutch), largely as a result of the pretensions to puritanism radical nationalism tends to bring with it. The elite, which is not

itself so very puritan, worries about the poor, ignorant peasant gambling all his money away, about what foreigners will think, about the waste of time better devoted to building up the country. It sees cockfighting as “primitive,” “backward,” “unprogressive,” and generally unbecoming an ambitious nation. And, as with those other embarrassments – opium smoking, begging, or uncovered breasts – it seeks, rather unsystematically, to put a stop to it.

Of course, like drinking during Prohibition or, today, smoking marihuana, cockfights, being a part of “The Balinese Way of Life,” nonetheless go on happening, and with extraordinary frequency. And, like Prohibition or marihuana, from time to time the police (who, in 1958 at least, were almost all not Balinese but Javanese) feel called upon to make a raid, confiscate the cocks and spurs, fine a few people, and even now and then expose some of them in the tropical sun for a day as object lessons which never, somehow, get learned, even though occasionally, quite occasionally, the object dies.

As a result, the fights are usually held in a secluded corner of a village in semi-secrecy, a fact which tends to slow the action a little – not very much, but the Balinese do not care to have it slowed at all. In this case, however, perhaps because they were raising money for a school that the government was unable to give them, perhaps because raids had been few recently, perhaps, as I gathered from subsequent discussion, there was a notion that the necessary bribes had been paid, they thought they could take a chance on the central square and draw a larger and more enthusiastic crowd without attracting the attention of the law.

They were wrong. In the midst of the third match, with hundreds of people,

1 Gregory Bateson and Margaret Mead, *Balinese Character: A Photographic Analysis* (New York: New York Academy of Sciences, 1942), 68.

including, still transparent, myself and my wife, fused into a single body around the ring, a superorganism in the literal sense, a truck full of policemen armed with machine guns roared up. Amid great screeching cries of “pulisi! pulisi!” from the crowd, the policemen jumped out, and, springing into the center of the ring, began to swing their guns around like gangsters in a motion picture, though not going so far as actually to fire them. The superorganism came instantly apart as its components scattered in all directions. People raced down the road, disappeared head-first over walls, scrambled under platforms, folded themselves behind wicker screens, scuttled up coconut trees. Cocks armed with steel spurs sharp enough to cut off a finger or run a hole through a foot were running wildly around. Everything was dust and panic.

On the established anthropological principle, When in Rome, my wife and I decided, only slightly less instantaneously than everyone else, that the thing to do was run too. We ran down the main village street, northward, away from where we were living, for we were on that side of the ring. About halfway down another fugitive ducked suddenly into a compound – his own, it turned out – and we, seeing nothing ahead of us but rice fields, open country, and a very high volcano, followed him. As the three of us came tumbling into the courtyard, his wife, who had apparently been through this sort of thing before, whipped out a table, a tablecloth, three chairs, and three cups of tea, and we all, without any explicit communication whatsoever, sat down, commenced to sip tea, and sought to compose ourselves.

A few moments later, one of the policemen marched importantly into the yard, looking for the village chief. (The chief had not only been at the fight, he

had arranged it. When the truck drove up he ran to the river, stripped off his sarong, and plunged in so he could say, when at length they found him sitting there pouring water over his head, that he had been away bathing when the whole affair had occurred and was ignorant of it. They did not believe him and fined him three hundred rupiah, which the village raised collectively.) Seeing my wife and I, “White Men,” there in the yard, the policeman performed a classic double take. When he found his voice again he asked, approximately, what in the devil did we think we were doing there. Our host of five minutes leaped instantly to our defense, producing an impassioned description of who and what we were, so detailed and so accurate that it was my turn, having barely communicated with a living human being save my landlord and the village chief for more than a week, to be astonished. We had a perfect right to be there, he said, looking the Javanese upstart in the eye. We were American professors; the government had cleared us; we were there to study culture; we were going to write a book to tell Americans about Bali. And we had all been there drinking tea and talking about cultural matters all afternoon and did not know anything about any cockfight. Moreover, we had not seen the village chief all day, he must have gone to town. The policeman retreated in rather total disarray. And, after a decent interval, bewildered but relieved to have survived and stayed out of jail, so did we.

The next morning the village was a completely different world for us. Not only were we no longer invisible, we were suddenly the center of all attention, the object of a great outpouring of warmth, interest, and, most especially, amusement. Everyone in the village knew we had fled like everyone else.



They asked us about it again and again (I must have told the story, small detail by small detail, fifty times by the end of the day), gently, affectionately, but quite insistently teasing us: “Why didn’t you just stand there and tell the police who you were?” “Why didn’t you just say you were only watching and not betting?” “Were you really afraid of those little guns?” As always, kinesthetically minded and, even when fleeing for their lives (or, as happened eight years later, surrendering them), the world’s most poised people, they gleefully mimicked, also over and over again, our graceless style of running and what they claimed were our panic-stricken facial expressions. But above all, everyone was extremely pleased and even more surprised that we had not simply “pulled out our papers” (they knew about those too) and asserted our Distinguished Visitor status, but had instead demonstrated our solidarity with what were now our covillagers. (What we had actually demonstrated was our cowardice, but there is fellowship in that too.) Even the Brahmana priest, an old, grave, halfway-to-Heaven type who because of its associations with the underworld would never be involved, even distantly, in a cockfight, and was difficult to approach even to other Balinese, had us called into his courtyard to ask us about what had happened, chuckling happily at the sheer extraordinariness of it all.

In Bali, to be teased is to be accepted. It was the turning point so far as our relationship to the community was concerned, and we were quite literally “in.” The whole village opened up to us, probably more than it ever would have otherwise (I might actually never have gotten to that priest, and our accidental host became one of my best informants), and certainly very much faster. Getting caught, or almost caught, in a vice raid

is perhaps not a very generalizable recipe for achieving that mysterious necessity of anthropological field work, rapport, but for me it worked very well. It led to a sudden and unusually complete acceptance into a society extremely difficult for outsiders to penetrate. It gave me the kind of immediate, inside-view grasp of an aspect of “peasant mentality” that anthropologists not fortunate enough to flee headlong with their subjects from armed authorities normally do not get. And, perhaps most important of all, for the other things might have come in other ways, it put me very quickly onto a combination emotional explosion, status war, and philosophical drama of central significance to the society whose inner nature I desired to understand. By the time I left I had spent about as much time looking into cockfights as into witchcraft, irrigation, caste, or marriage.

Bali, mainly because it is Bali, is a well-studied place. Its mythology, art, ritual, social organization, patterns of child rearing, forms of law, even styles of trance, have all been microscopically examined for traces of that elusive substance Jane Belo called “The Balinese Temper.”<sup>2</sup> But, aside from a few passing remarks, the cockfight has barely been noticed, although as a popular obsession of consuming power it is at least as important a revelation of what being a Balinese “is really like” as these more celebrated phenomena.<sup>3</sup> As much of America surfaces in a ballpark, on a

2 Jane Belo, “The Balinese Temper,” in Jane Belo, ed., *Traditional Balinese Culture* (New York: Columbia University Press, 1970; originally published in 1935), 85–110.

3 The best discussion of cockfighting is again Bateson and Mead’s (*Balinese Character*, 24–25, 140), but it, too, is general and abbreviated.

golf links, at a race track, or around a poker table, much of Bali surfaces in a cock ring. For it is only apparently cocks that are fighting there. Actually, it is men.

To anyone who has been in Bali any length of time, the deep psychological identification of Balinese men with their cocks is unmistakable. The double entendre here is deliberate. It works in exactly the same way in Balinese as it does in English, even to producing the same tired jokes, strained puns, and uninventive obscenities. Bateson and Mead have even suggested that, in line with the Balinese conception of the body as a set of separately animated parts, cocks are viewed as detachable, self-operating penises, ambulant genitals with a life of their own.<sup>4</sup> And while I do not have the kind of unconscious material either to confirm or disconfirm this intriguing notion, the fact that they are masculine symbols *par excellence* is about as indubitable, and to the Balinese about as evident, as the fact that water runs downhill.

The language of everyday moralism is shot through, on the male side of it, with roosterish imagery. *Sabung*, the word for

cock (and one which appears in inscriptions as early as A.D. 922), is used metaphorically to mean “hero,” “warrior,” “champion,” “man of parts,” “political candidate,” “bachelor,” “dandy,” “lady-killer,” or “tough guy.” A pompous man whose behavior presumes above his station is compared to a tailless cock who struts about as though he had a large, spectacular one. A desperate man who makes a last, irrational effort to extricate himself from an impossible situation is likened to a dying cock who makes one final lunge at his tormentor to drag him along to a common destruction. A stingy man, who promises much, gives little, and begrudges that is compared to a cock which, held by the tail, leaps at another without in fact engaging him. A marriageable young man still shy with the opposite sex or someone in a new job anxious to make a good impression is called “a fighting cock caged for the first time.”<sup>5</sup> Court trials, wars, political contests, inheritance disputes, and street arguments are all compared to cock-fights.<sup>6</sup> Even the very island itself is perceived from its shape as a small, proud cock, poised, neck extended, back taut, tail raised, in eternal challenge to large, feckless, shapeless Java.<sup>7</sup>

4 Ibid., 25 – 26. The cockfight is unusual within Balinese culture in being a single sex public activity from which the other sex is totally and expressly excluded. Sexual differentiation is culturally extremely played down in Bali and most activities, formal and informal, involve the participation of men and women on equal ground, commonly as linked couples. From religion, to politics, to economics, to kinship, to dress, Bali is a rather “unisexual” society, a fact both its customs and its symbolism clearly express. Even in contexts where women do not in fact play much of a role – music, painting, certain agricultural activities – their absence, which is only relative in any case, is more a mere matter of fact than socially enforced. To this general pattern, the cockfight, entirely of, by, and for men (women – at least Balinese women – do not even watch), is the most striking exception.

5 Christiaan Hooykaas, *The Lay of the Jaya Prana* (London: Luzac, 1958), 39. The lay has a stanza (no. 17) with the reluctant bridegroom use. Jaya Prana, the subject of a Balinese Uriah myth, responds to the lord who has offered him the loveliest of six hundred servant girls: “Godly King, my Lord and Master / I beg you, give me leave to go / such things are not yet in my mind; / like a fighting cock encaged / indeed I am on my mettle / I am alone / as yet the flame has not been fanned.”

6 For these, see V. E. Korn, *Het Adatrecht van Bali*, 2d ed. (s’Gravenhage: G. Naeff, 1932), index under *toh*.

7 There is indeed a legend to the effect that the separation of Java and Bali is due to the action

But the intimacy of men with their cocks is more than metaphorical. Balinese men, or anyway a large majority of Balinese men, spend an enormous amount of time with their favorites, grooming them, feeding them, discussing them, trying them out against one another, or just gazing at them with a mixture of rapt admiration and dreamy self-absorption. Whenever you see a group of Balinese men squatting idly in the council shed or along the road in their hips down, shoulders forward, knees up fashion, half or more of them will have a rooster in his hands, holding it between his thighs, bouncing it gently up and down to strengthen its legs, ruffling its feathers with abstract sensuality, pushing it out against a neighbor's rooster to rouse its spirit, withdrawing it toward his loins to calm it again. Now and then, to get a feel for another bird, a man will fiddle this way with someone else's cock for a while, but usually by moving around to squat in place behind it, rather than just having it passed across to him as though it were merely an animal.

In the houseyard, the high-walled enclosures where the people live, fighting cocks are kept in wicker cages, moved frequently about so as to maintain the optimum balance of sun and shade. They are fed a special diet, which varies somewhat according to individual theories but which is mostly maize, sifted for impurities with far more care than it is when mere humans are going to eat it and offered to the animal kernel by kernel. Red pepper is stuffed down their beaks and up their anuses to give them

spirit. They are bathed in the same ceremonial preparation of tepid water, medicinal herbs, flowers, and onions in which infants are bathed, and for a prize cock just about as often. Their combs are cropped, their plumage dressed, their spurs trimmed, their legs massaged, and they are inspected for flaws with the squinted concentration of a diamond merchant. A man who has a passion for cocks, an enthusiast in the literal sense of the term, can spend most of his life with them, and even those, the overwhelming majority, whose passion though intense has not entirely run away with them, can and do spend what seems not only to an outsider, but also to themselves, an inordinate amount of time with them. "I am cock crazy," my landlord, a quite ordinary *afficianado* by Balinese standards, used to moan as he went to move another cage, give another bath, or conduct another feeding. "We're all cock crazy."

The madness has some less visible dimensions, however, because although it is true that cocks are symbolic expressions or magnifications of their owner's self, the narcissistic male ego writ out in Aesopian terms, they are also expressions – and rather more immediate ones – of what the Balinese regard as the direct inversion, aesthetically, morally, and metaphysically, of human status: animality.

The Balinese revulsion against any behavior regarded as animal-like can hardly be overstressed. Babies are not allowed to crawl for that reason. Incest, though hardly approved, is a much less horrifying crime than bestiality. (The appropriate punishment for the second is death by drowning, for the first being forced to live like an animal.<sup>8</sup>) Most de-

---

of a powerful Javanese religious figure who wished to protect himself against a Balinese culture hero (the ancestor of two Ksatria castes) who was a passionate cockfighting gambler. See Christiaan Hooykaas, *Agama Tirtha* (Amsterdam: Noord-Hollandsche, 1964), 184.

8 An incestuous couple is forced to wear pig yokes over their necks and crawl to a pig trough and eat with their mouths there. On this, see

mons are represented – in sculpture, dance, ritual, myth – in some real or fantastic animal form. The main puberty rite consists in filing the child's teeth so they will not look like animal fangs. Not only defecation but eating is regarded as a disgusting, almost obscene activity, to be conducted hurriedly and privately, because of its association with animality. Even falling down or any form of clumsiness is considered to be bad for these reasons. Aside from cocks and a few domestic animals – oxen, ducks – of no emotional significance, the Balinese are averse to animals and treat their large number of dogs not merely callously but with a phobic cruelty. In identifying with his cock, the Balinese man is identifying not just with his ideal self, or even his penis, but also, and at the same time, with what he most fears, hates, and ambivalence being what it is, is fascinated by – The Powers of Darkness.

The connection of cocks and cockfighting with such Powers, with the animalistic demons that threaten constantly to invade the small, cleared-off space in which the Balinese have so carefully built their lives and devour its inhabitants is quite explicit. A cockfight, any cockfight, is in the first instance a blood sacrifice offered, with the appropriate chants and oblations, to the demons in order to pacify their ravenous, cannibal hunger. No temple festival should be conducted until one is made. (If it is omitted someone will inevitably fall into a trance and command with the voice of an angered spirit that the oversight be immediately corrected.) Collective responses to natural evils – illness, crop

failure, volcanic eruptions – almost always involve them. And that famous holiday in Bali, The Day of Silence (*Njepi*), when everyone sits silent and immobile all day long in order to avoid contact with a sudden influx of demons chased momentarily out of hell, is preceded the previous day by large-scale cockfights (in this case legal) in almost every village on the island.

In the cockfight, man and beast, good and evil, ego and id, the creative power of aroused masculinity and the destructive power of loosened animality fuse in a bloody drama of hatred, cruelty, violence, and death. It is little wonder that when, as is the invariable rule, the owner of the winning cock takes the carcass of the loser – often torn limb from limb by its enraged owner – home to eat, he does so with a mixture of social embarrassment, moral satisfaction, aesthetic disgust, and cannibal joy. Or that a man who has lost an important fight is sometimes driven to wreck his family shrines and curse the gods, an act of metaphysical (and social) suicide. Or that in seeking earthly analogues for heaven and hell the Balinese compare the former to the mood of a man whose cock has just won, the latter to that of a man whose cock has just lost.

Cockfights (*tetadjen*; *sabungan*) are held in a ring about fifty feet square. Usually they begin toward late afternoon and run three or four hours until sunset. About nine or ten separate matches (*sehet*) comprise a program. Each match is precisely like the others in general pattern: there is no main match, no connection between individual matches, no variation in their format, and each is arranged on a completely ad hoc basis. After a fight has ended and the emotional debris is cleaned away – the bets paid, the curses cursed, the carcasses pos-

---

Jane Belo, "Customs Pertaining to Twins in Bali," in Belo, ed., *Traditional Balinese Culture*, 49; on the abhorrence of animality generally, Bateson and Mead, *Balinese Character*, 22.

sessed – seven, eight, perhaps even a dozen men slip negligently into the ring with a cock and seek to find there a logical opponent for it. This process, which rarely takes less than ten minutes, and often a good deal longer, is conducted in a very subdued, oblique, even dissembling manner. Those not immediately involved give it at best but disguised, sidelong attention; those who, embarrassedly, are, attempt to pretend somehow that the whole thing is not really happening.

A match made, the other hopefuls retire with the same deliberate indifference, and the selected cocks have their spurs (*tadji*) affixed – razor sharp, pointed steel swords, four or five inches long. This is a delicate job which only a small proportion of men, a half-dozen or so in most villages, know how to do properly. The man who attaches the spurs also provides them, and if the rooster he assists wins its owner awards him the spur-leg of the victim. The spurs are affixed by winding a long length of string around the foot of the spur and the leg of the cock. For reasons I shall come to presently, it is done somewhat differently from case to case, and is an obsessively deliberate affair. The lore about spurs is extensive – they are sharpened only at eclipses and the dark of the moon, should be kept out of the sight of women, and so forth. And they are handled, both in use and out, with the same curious combination of fussiness and sensuality the Balinese direct toward ritual objects generally.

The spurs affixed, the two cocks are placed by their handlers (who may or may not be their owners) facing one another in the center of the ring.<sup>9</sup> A

<sup>9</sup> Except for unimportant, small-bet fights (on the question of fight “importance,” see below) spur affixing is usually done by someone other than the owner. Whether the owner handles his

coconut pierced with a small hole is placed in a pail of water, in which it takes about twenty-one seconds to sink, a period known as a *tjeng* and marked at beginning and end by the beating of a slit gong. During these twenty-one seconds the handlers (*pengangkeb*) are not permitted to touch their roosters. If, as sometimes happens, the animals have not fought during this time, they are picked up, fluffed, pulled, prodded, and otherwise insulted, and put back in the center of the ring and the process begins again. Sometimes they refuse to fight at all, or one keeps running away, in which case they are imprisoned together under a wicker cage, which usually gets them engaged.

Most of the time, in any case, the cocks fly almost immediately at one another in a wing-beating, head-thrusting, leg-kicking explosion of animal fury so pure, so absolute, and in its own way so beautiful, as to be almost abstract, a Platonic concept of hate. Within moments one or the other drives home a solid blow with his spur. The handler whose cock has delivered the blow immediately picks it up so that it will not get a return blow, for if he does not the match is likely to end in a mutually mortal tie as the two birds wildly hack each other to pieces. This is particularly true if, as often happens, the spur

---

own cock or not more or less depends on how skilled he is at it, a consideration whose importance is again relative to the importance of the fight. When spur affixers and cock handlers are someone other than the owner, they are almost always a quite close relative – a brother or cousin – or a very intimate friend of his. They are thus almost extensions of his personality, as the fact that all three will refer to the cock as “mine,” say “I” fought So-and-So, and so on, demonstrates. Also, owner-handler-affixer triads tend to be fairly fixed, though individuals may participate in several and often exchange roles within a given one.



sticks in its victim's body, for then the aggressor is at the mercy of his wounded foe.

With the birds again in the hands of their handlers, the coconut is now sunk three times after which the cock which has landed the blow must be set down to show that he is firm, a fact he demonstrates by wandering idly around the rink for a coconut sink. The coconut is then sunk twice more and the fight must recommence.

During this interval, slightly over two minutes, the handler of the wounded cock has been working frantically over it, like a trainer patching a mauled boxer between rounds, to get it in shape for a last, desperate try for victory. He blows in its mouth, putting the whole chicken head in his own mouth and sucking and blowing, fluffs it, stuffs its wounds with various sorts of medicines, and generally tries anything he can think of to arouse the last ounce of spirit which may be hidden somewhere within it. By the time he is forced to put it back down he is usually drenched in chicken blood, but, as in prizefighting, a good handler is worth his weight in gold. Some of them can virtually make the dead walk, at least long enough for the second and final round.

In the climactic battle (if there is one; sometimes the wounded cock simply expires in the handler's hands or immediately as it is placed down again), the cock who landed the first blow usually proceeds to finish off his weakened opponent. But this is far from an inevitable outcome, for if a cock can walk he can fight, and if he can fight, he can kill, and what counts is which cock expires first. If the wounded one can get a stab in and stagger on until the other drops, he is the official winner, even if he himself topples over an instant later.

Surrounding all this melodrama – which the crowd packed tight around the ring follows in near silence, moving their bodies in kinesthetic sympathy with the movement of the animals, cheering their champions on with wordless hand motions, shiftings of the shoulders, turnings of the head, falling back *en masse* as the cock with the murderous spurs careens toward one side of the ring (it is said that spectators sometimes lose eyes and fingers from being too attentive), surging forward again as they glance off toward another – is a vast body of extraordinarily elaborate and precisely detailed rules.

These rules, together with the developed lore of cocks and cockfighting which accompanies them, are written down in palm-leaf manuscripts (*lontar*; *rontal*) passed on from generation to generation as part of the general legal and cultural tradition of the villages. At a fight, the umpire (*saja komong*; *djuru kembar*) – the man who manages the coconut – is in charge of their application and his authority is absolute. I have never seen an umpire's judgment questioned on any subject, even by the more despondent losers, nor have I ever heard, even in private, a charge of unfairness directed against one, or, for that matter, complaints about umpires in general. Only exceptionally well-trusted, solid, and, given the complexity of the code, knowledgeable citizens perform this job, and in fact men will bring their cocks only to fights presided over by such men. It is also the umpire to whom accusations of cheating, which, though rare in the extreme, occasionally arise, are referred; and it is he who in the not infrequent cases where the cocks expire virtually together decides which (if either, for, though the Balinese do not care for such an outcome, there can be ties) went

first. Likened to a judge, a king, a priest, and a policeman, he is all of these, and under his assured direction the animal passion of the fight proceeds within the civic certainty of the law. In the dozens of cockfights I saw in Bali, I never once saw an altercation about rules. Indeed, I never saw an open altercation, other than those between cocks, at all.

This crosswise doubleness of an event which, taken as a fact of nature, is rage untrammelled and, taken as a fact of culture, is form perfected, defines the cockfight as a sociological entity. A cockfight is what, searching for a name for something not vertebrate enough to be called a group and not structureless enough to be called a crowd, Erving Goffman has called a “focused gathering” – a set of persons engrossed in a common flow of activity and relating to one another in terms of that flow.<sup>10</sup> Such gatherings meet and disperse; the participants in them fluctuate; the activity that focuses them is discreet – a particulate process that reoccurs rather than a continuous one that endures. They take their form from the situation that evokes them, the floor on which they are placed, as Goffman puts it; but it is a form, and an articulate one, nonetheless. For the situation, the floor is itself created, in jury deliberations, surgical operations, block meetings, sit-ins, cockfights, by the cultural preoccupations – here, as we shall see, the celebration of status rivalry – which not only specify the focus but, assembling actors and arranging scenery, bring it actually into being.

In classical times (that is to say, prior to the Dutch invasion of 1908), when there were no bureaucrats around to improve popular morality, the staging of a cockfight was an explicitly societal mat-

ter. Bringing a cock to an important fight was, for an adult male, a compulsory duty of citizenship; taxation of fights, which were usually held on market day, was a major source of public revenue; patronage of the art was a stated responsibility of princes; and the cock ring, or *wantilan*, stood in the center of the village near those other monuments of Balinese civility – the council house, the origin temple, the marketplace, the signal tower, and the banyan tree. Today, a few special occasions aside, the newer rectitude makes so open a statement of the connection between the excitements of collective life and those of blood sport impossible, but, less directly expressed, the connection itself remains intimate and intact. To expose it, however, it is necessary to turn to the aspect of cockfighting around which all the others pivot, and through which they exercise their force, an aspect I have thus far studiously ignored. I mean, of course, the gambling.

The Balinese never do anything in a simple way that they can contrive to do in a complicated one, and to this generalization cockfight wagering is no exception.

In the first place, there are two sorts of bets, or *toh*.<sup>11</sup> There is the single axial bet

11 This word, which literally means an indelible stain or mark, as in a birthmark or a vein in a stone, is used as well for a deposit in a court case, for a pawn, for security offered in a loan, for a stand-in for someone else in a legal or ceremonial context, for an earnest advanced in a business deal, for a sign placed in a field to indicate its ownership is in dispute, and for the status of an unfaithful wife from whose lover her husband must gain satisfaction or surrender her to him. See Korn, *Het Adatrecht van Bali*; Theodoor Pigeaud, *Javaans-Nederlands Handwoordenboek* (Groningen: Wolters, 1938); H. H. Juynboll, *Oudjavaansche-Nederlandsche Woordenlijst* (Leiden: Brill, 1923).

10 Erving Goffman, *Encounters: Two Studies in the Sociology of Interaction* (Indianapolis: Bobbs-Merrill, 1961), 9 – 10.

in the center between the principals (*toh ketengah*), and there is the cloud of peripheral ones around the ring between members of the audience (*toh kesasi*). The first is typically large; the second typically small. The first is collective, involving coalitions of bettors clustering around the owner; the second is individual, man to man. The first is a matter of deliberate, very quiet, almost furtive arrangement by the coalition members and the umpire huddled like conspirators in the center of the ring; the second is a matter of impulsive shouting, public offers, and public acceptances by the excited throng around its edges. And most curiously, and as we shall see most revealingly, *where the first is always, without exception, even money, the second, equally without exception, is never such*. What is a fair coin in the center is a biased one on the side.

The center bet is the official one, hedged in again with a webwork of rules, and is made between the two cock owners, with the umpire as overseer and public witness.<sup>12</sup> This bet, which, as I say, is always relatively and sometimes very large, is never raised simply by the owner in whose name it is made, but by him together with four or five, sometimes seven or eight, allies – kin, village mates, neighbors, close friends. He may, if he is not especially well-to-do, not even be the major contributor, though, if only to show that he is not involved in any chicanery, he must be a significant one.

12 The center bet must be advanced in cash by both parties prior to the actual fight. The umpire holds the stakes until the decision is rendered and then awards them to the winner, avoiding, among other things, the intense embarrassment both winner and loser would feel if the latter had to pay off personally following his defeat. About 10 percent of the winner's receipts are subtracted for the umpire's share and that of the fight sponsors.

Of the 57 matches for which I have exact and reliable data on the center bet, the range is from 15 ringgits to 500, with a mean at 85 and with the distribution being rather noticeably trimodal: small fights (15 ringgits either side of 35) accounting for about 45 percent of the total number; medium ones (20 ringgits either side of 70) for about 25 percent; and large (75 ringgits either side of 175) for about 20 percent, with a few very small and very large ones out at the extremes. In a society where the normal daily wage of a manual laborer – a brickmaker, an ordinary farmworker, a market porter – was about 3 ringgits a day, and considering the fact that fights were held on the average about every 2.5 days in the immediate area I studied, this is clearly serious gambling, even if the bets are pooled rather than individual efforts.

The side bets are, however, something else altogether. Rather than the solemn, legalistic pactmaking of the center, wagering takes place rather in the fashion in which the stock exchange used to work when it was out on the curb. There is a fixed and known odds paradigm which runs in a continuous series from 10 – 9 at the short end to 2 – 1 at the long: 10 – 9, 9 – 8, 8 – 7, 7 – 6, 6 – 5, 5 – 4, 4 – 3, 3 – 2, 2 – 1. The man who wishes to back the *underdog cock* (leaving aside how favorites, *kebut*, and underdogs, *ngai*, are established for the moment) shouts the short-side number indicating the odds he wants *to be given*. That is, if he shouts *gasal*, “five,” he wants the underdog at 5 – 4 (or, for him, 4 – 5); if he shouts “four,” he wants it at 4 – 3 (again, he putting up the “three”), if “nine,” at 9 – 8, and so on. A man backing the favorite, and thus considering giving odds if he can get them short enough, indicates the fact by crying out the color type of that

cock – “brown,” “speckled,” or whatever.<sup>13</sup>

As odds-takers (backers of the underdog) and odds-givers (backers of the favorite) sweep the crowd with their

13 Actually, the typing of cocks, which is extremely elaborate (I have collected more than twenty classes, certainly not a complete list), is not based on color alone, but on a series of independent, interacting, dimensions, which include, beside color, size, bone thickness, plumage, and temperament. (But *not* pedigree. The Balinese do not breed cocks to any significant extent, nor, so far as I have been able to discover, have they ever done so. The *asil*, or jungle cock, which is the basic fighting strain everywhere the sport is found, is native to southern Asia, and one can buy a good example in the chicken section of almost any Balinese market for anywhere from 4 or 5 ringgits up to 50 or more.) The color element is merely the one normally used as the type name, except when the two cocks of different types – as on principle they must be – have the same color, in which case a secondary indication from one of the other dimensions (“large speckled” v. “small speckled,” etc.) is added. The types are coordinated with various cosmological ideas which help shape the making of matches, so that, for example, you fight a small, headstrong, speckled brown-on-white cock with flat-lying feathers and thin legs from the east side of the ring on a certain day of the complex Balinese calendar, and a large, cautious, all-black cock with tufted feathers and stubby legs from the north side on another day, and so on. All this is again recorded in palm-leaf manuscripts and endlessly discussed by the Balinese (who do not all have identical systems), and full-scale componential-cum-symbolic analysis of cock classifications would be extremely valuable both as an adjunct to the description of the cockfight and in itself. But my data on the subject, though extensive and varied, do not seem to be complete and systematic enough to attempt such an analysis here. For Balinese cosmological ideas more generally see Belo, ed., *Traditional Balinese Culture*, and J. L. Swellengrebel, ed., *Bali: Studies in Life, Thought, and Ritual* (The Hague: W. van Hoeve, 1960); for calendrical ones, Clifford Geertz, *Person, Time, and Conduct in Bali: An Essay in Cultural Analysis* (New Haven, Conn.: Southeast Asia Studies, Yale University, 1966), 45 – 53.

shouts, they begin to focus in on one another as potential betting pairs, often from far across the ring. The taker tries to shout the giver into longer odds, the giver to shout the taker into shorter ones.<sup>14</sup> The taker, who is the wooer in this situation, will signal how large a bet he wishes to make at the odds he is shouting by holding a number of fingers up in front of his face and vigorously waving them. If the giver, the wooed, replies in kind, the bet is made; if he does not, they unlock gazes and the search goes on.

The side betting, which takes place after the center bet has been made and its size announced, consists then in a rising crescendo of shouts as backers of the underdog offer their propositions to anyone who will accept them, while those who are backing the favorite but do not like the price being offered, shout equally frenetically the color of the cock to show they too are desperate to bet but want shorter odds.

Almost always odds-calling, which tends to be very consensual in that at any one time almost all callers are calling the same thing, starts off toward the long end of the range (5 – 4 or 4 – 3) and then moves, also consensually, toward the short end with greater or lesser speed and to a greater or lesser degree. Men crying “five” and finding themselves answered only with cries of “brown” start crying “six,” either drawing the other callers fairly quickly with them or retiring from the scene as their

14 For purposes of ethnographic completeness, it should be noted that it is possible for the man backing the favorite – the odds-giver – to make a bet in which he wins if his cock wins or there is a tie, a slight shortening of the odds (I do not have enough cases to be exact, but ties seem to occur about once every fifteen or twenty matches). He indicates his wish to do this by shouting *sapih* (“tie”) rather than the cock-type, but such bets are in fact infrequent.



too-generous offers are snapped up. If the change is made and partners are still scarce, the procedure is repeated in a move to “seven,” and so on, only rarely, and in the very largest fights, reaching the ultimate “nine” or “ten” levels. Occasionally, if the cocks are clearly mismatched, there may be no upward movement at all, or even a movement down the scale to 4 – 3, 3 – 2, very, very rarely 2 – 1, a shift which is accompanied by a declining number of bets as a shift upward is accompanied by an increasing number. But the general pattern is for the betting to move a shorter or longer distance up the scale toward the, for side bets, nonexistent pole of even money, with the overwhelming majority of bets falling in the 4 – 3 to 8 – 7 range.<sup>15</sup>

As the moment for the release of the cocks by the handlers approaches, the screaming, at least in a match where the center bet is large, reaches almost frenzied proportions as the remaining un-

15 The precise dynamics of the movement of the betting is one of the most intriguing, most complicated, and, given the hectic conditions under which it occurs, most difficult to study, aspects of the fight. Motion picture recording plus multiple observers would probably be necessary to deal with it effectively. Even impressionistically – the only approach open to a lone ethnographer caught in the middle of all this – it is clear that certain men lead both in determining the favorite (that is, making the opening cock-type calls which always initiate the process) and in directing the movement of the odds, these “opinion leaders” being the more accomplished cockfighters-cum-solid-citizens to be discussed below. If these men begin to change their calls, others follow; if they begin to make bets, so do others and – though there is always a large number of frustrated bettors crying for shorter or longer odds to the end – the movement more or less ceases. But a detailed understanding of the whole process awaits what, alas, it is not very likely ever to get: a decision theorist armed with precise observations of individual behavior.

fulfilled bettors try desperately to find a last-minute partner at a price they can live with. (Where the center bet is small, the opposite tends to occur: betting dies off, trailing into silence, as odds lengthen and people lose interest.) In a large-bet, well-made match – the kind of match the Balinese regard as “real cockfighting” – the mob scene quality, the sense that sheer chaos is about to break loose, with all those waving, shouting, pushing, clambering men is quite strong, an effect which is only heightened by the intense stillness that falls with instant suddenness, rather as if someone had turned off the current, when the slit gong sounds, the cocks are put down, and the battle begins.

When it ends, anywhere from fifteen seconds to five minutes later, *all bets are immediately paid*. There are absolutely no IOUs, at least to a betting opponent. One may, of course, borrow from a friend before offering or accepting a wager, but to offer or accept it you must have the money already in hand and, if you lose, you must pay it on the spot, before the next match begins. This is an iron rule, and as I have never heard of a disputed umpire’s decision (though doubtless there must sometimes be some), I have also never heard of a welshed bet, perhaps because in a worked-up cockfight crowd the consequences might be, as they are reported to be sometimes for cheaters, drastic and immediate.

It is, in any case, this formal asymmetry between balanced center bets and unbalanced side ones that poses the critical analytical problem for a theory which sees cockfight wagering as the link connecting the fight to the wider world of Balinese culture. It also suggests the way to go about solving it and demonstrating the link.

The first point that needs to be made in this connection is that the higher the



center bet, the more likely the match will in actual fact be an even one. Simple considerations of rationality suggest that. If you are betting 15 ringgits on a cock, you might be willing to go along with even money even if you feel your animal somewhat the less promising. But if you are betting 500 you are very, very likely to be loathe to do so. Thus, in large-bet fights, which of course involve the better animals, tremendous care is taken to see that the cocks are about as evenly matched as to size, general condition, pugnacity, and so on as is humanly possible. The different ways of adjusting the spurs of the animals are often employed to secure this. If one cock seems stronger, an agreement will be made to position his spur at a slightly less advantageous angle – a kind of handicapping, at which spur affixers are, so it is said, extremely skilled. More care will be taken, too, to employ skillful handlers and to match them exactly as to abilities.

In short, in a large-bet fight the pressure to make the match a genuinely fifty-fifty proposition is enormous, and is consciously felt as such. For medium fights the pressure is somewhat less, and for small ones less yet, though there is always an effort to make things at least approximately equal, for even at 15 ringgits (5 days' work) no one wants to make an even-money bet in a clearly unfavorable situation. And, again, what statistics I have tend to bear this out. In my 57 matches, the favorite won 33 times overall, the underdog 24, a 1.4 to 1 ratio. But if one splits the figures at 60 ringgits center bets, the ratios turn out to be 1.1 to 1 (12 favorites, 11 underdogs) for those above this line, and 1.6 to 1 (21 and 13) for those below it. Or, if you take the extremes, for very large fights, those with center bets over 100 ringgits the ratio is 1 to 1 (7 and 7); for very small

fights, those under 40 ringgits, it is 1.9 to 1 (19 and 10).<sup>16</sup>

Now, from this proposition – that the higher the center bet the more exactly a fifty-fifty proposition the cockfight is – two things more or less immediately follow: (1) the higher the center bet, the greater is the pull on the side betting toward the short-odds end of the wagering spectrum and vice versa; (2) the higher the center bet, the greater the volume of side betting and vice versa.

The logic is similar in both cases. The closer the fight is in fact to even money, the less attractive the long end of the odds will appear and, therefore, the shorter it must be if there are to be takers. That this is the case is apparent from mere inspection, from the Balinese's own analysis of the matter, and from what more systematic observations I was able to collect. Given the difficulty of making precise and complete recordings of side betting, this argument is hard to cast in numerical form, but in all my cases the odds-giver, odds-taker consensual point, a quite pronounced minimax saddle where the bulk (at a guess, two-thirds to three-quarters in most cases) of the bets are actually made, was three or four points further along the scale toward the shorter end

16 Assuming only binomial variability, the departure from a fifty-fifty expectation in the 60 ringgits and below case is 1.38 standard deviations, or (in a one-direction test) an 8 in 100 possibility by chance alone; for the below 40 ringgits case it is 1.65 standard deviations, or about 5 in 100. The fact that these departures though real are not extreme merely indicates, again, that even in the smaller fights the tendency to match cocks at least reasonably evenly persists. It is a matter of relative relaxation of the pressures toward equalization, not their elimination. The tendency for high-bet contests to be coin-flip propositions is, of course, even more striking, and suggests the Balinese know quite well what they are about.

for the large-center-bet fights than for the small ones, with medium ones generally in between. In detail, the fit is not, of course, exact, but the general pattern is quite consistent: the power of the center bet to pull the side bets toward its own even-money pattern is directly proportional to its size, because its size is directly proportional to the degree to which the cocks are in fact evenly matched. As for the volume question, total wagering is greater in large-center-bet fights because such fights are considered more “interesting,” not only in the sense that they are less predictable, but, more crucially, that more is at stake in them – in terms of money, in terms of the quality of the cocks, and consequently, as we shall see, in terms of social prestige.<sup>17</sup>

The paradox of fair coin in the middle, biased coin on the outside is thus a merely apparent one. The two betting systems, though formally incongruent,

17 The reduction in wagering in smaller fights (which, of course, feeds on itself; one of the reasons people find small fights uninteresting is that there is less wagering in them, and contrariwise for large ones) takes place in three mutually reinforcing ways. First, there is a simple withdrawal of interest as people wander off to have a cup of coffee or chat with a friend. Second, the Balinese do not mathematically reduce odds, but bet directly in terms of stated odds as such. Thus, for a 9 – 8 bet, one man wagers 9 ringgits, the other 8; for 5 – 4, one wagers 5, the other 4. For any given currency unit, like the ringgit, therefore, 6.3 times as much money is involved in a 10 – 9 bet as in a 2 – 1 bet, for example, and, as noted, in small fights betting settles toward the longer end. Finally, the bets which are made tend to be one- rather than two-, three-, or in some of the very largest fights, four- or five-finger ones. (The fingers indicate the *multiples* of the stated bet odds at issue, not absolute figures. Two fingers in a 6 – 5 situation means a man wants to wager 10 ringgits on the underdog against 12, three in an 8 – 7 situation, 21 against 24, and so on.)

are not really contradictory to one another, but part of a single larger system in which the center bet is, so to speak, the “center of gravity,” drawing, the larger it is the more so, the outside bets toward the short-odds end of the scale. The center bet thus “makes the game,” or perhaps better, defines it, signals what, following a notion of Jeremy Bentham’s, I am going to call its “depth.”

The Balinese attempt to create an interesting, if you will, “deep,” match by making the center bet as large as possible so that the cocks matched will be as equal and as fine as possible, and the outcome, thus, as unpredictable as possible. They do not always succeed. Nearly half the matches are relatively trivial, relatively uninteresting – in my borrowed terminology, “shallow” – affairs. But that fact no more argues against my interpretation than the fact that most painters, poets, and playwrights are mediocre argues against the view that artistic effort is directed toward profundity and, with a certain frequency, approximates it. The image of artistic technique is indeed exact: the center bet is a means, a device, for creating “interesting,” “deep” matches, *not* the reason, or at least not the main reason, *why* they are interesting, the source of their fascination, the substance of their depth. The question why such matches are interesting – indeed, for the Balinese, exquisitely absorbing – takes us out of the realm of formal concerns into more broadly sociological and social-psychological ones, and to a less purely economic idea of what “depth” in gaming amounts to.<sup>18</sup>

18 Besides wagering there are other economic aspects of the cockfight, especially its very close connection with the local market system which, though secondary both to its motivation and to its function, are not without importance. Cockfights are open events to which anyone who

Bentham's concept of "deep play" is found in his *The Theory of Legislation*.<sup>19</sup> By it he means play in which the stakes are so high that it is, from his utilitarian standpoint, irrational for men to engage in it at all. If a man whose fortune is a thousand pounds (or ringgits) wagers five hundred of it on an even bet, the marginal utility of the pound he stands to win is clearly less than the marginal disutility of the one he stands to lose.

---

wishes may come, sometimes from quite distant areas, but well over 90 percent, probably over 95, are very local affairs, and the locality concerned is defined not by the village, nor even by the administrative district, but by the rural market system. Bali has a three-day market week with the familiar "solar-system" type rotation. Though the markets themselves have never been very highly developed, small morning affairs in a village square, it is the microregion such rotation rather generally marks out – ten or twenty square miles, seven or eight neighboring villages (which in contemporary Bali is usually going to mean anywhere from five to ten or eleven thousand people) from which the core of any cockfight audience, indeed virtually all of it, will come. Most of the fights are in fact organized and sponsored by small combines of petty rural merchants under the general premise, very strongly held by them and indeed by all Balinese, that cockfights are good for trade because "they get money out of the house, they make it circulate." Stalls selling various sorts of things as well as assorted sheer-chance gambling games (see below) are set up around the edge of the area so that this even takes on the quality of a small fair. This connection of cockfighting with markets and market sellers is very old, as, among other things, their conjunction in inscriptions (Roelof Goris, *Prasasti Bali*, 2 vols. [Bandung: N. V. Masa Baru, 1954]) indicates. Trade has followed the cock for centuries in rural Bali and the sport has been one of the main agencies of the island's monetization.

19 The phrase is found in the Hildreth translation, International Library of Psychology, 1931, note to 106; see L. L. Fuller, *The Morality of Law* (New Haven, Conn.: Yale University Press, 1964), 6ff.

In genuine deep play, this is the case for both parties. They are both in over their heads. Having come together in search of pleasure they have entered into a relationship which will bring the participants, considered collectively, net pain rather than net pleasure. Bentham's conclusion was, therefore, that deep play was immoral from first principles and, a typical step for him, should be prevented legally.

But more interesting than the ethical problem, at least for our concerns here, is that despite the logical force of Bentham's analysis men do engage in such play, both passionately and often, and even in the face of law's revenge. For Bentham and those who think as he does (nowadays mainly lawyers, economists, and a few psychiatrists), the explanation is, as I have said, that such men are irrational – addicts, fetishists, children, fools, savages, who need only to be protected against themselves. But for the Balinese, though naturally they do not formulate it in so many words, the explanation lies in the fact that in such play money is less a measure of utility, had or expected, than it is a symbol of moral import, perceived or imposed.

It is, in fact, in shallow games, ones in which smaller amounts of money are involved, that increments and decrements of cash are more nearly synonyms for utility and disutility, in the ordinary, unexpanded sense – for pleasure and pain, happiness and unhappiness. In deep ones, where the amounts of money are great, much more is at stake than material gain: namely, esteem, honor, dignity, respect – in a word, though in Bali a profoundly freighted word, status.<sup>20</sup> It is at stake symbolical-

20 Of course, even in Bentham, utility is not normally confined as a concept to monetary losses and gains, and my argument here might

ly, for (a few cases of ruined addict gamblers aside) no one's status is actually altered by the outcome of a cockfight; it is only, and that momentarily, affirmed or insulted. But for the Balinese, for whom nothing is more pleasurable than an affront obliquely delivered or more painful than one obliquely received – particularly when mutual acquaintances, undeceived by surfaces, are watching – such appraisive drama is deep indeed.

This, I must stress immediately, is *not* to say that the money does not matter, or that the Balinese is no more concerned about losing 500 ringgits than 15. Such a conclusion would be absurd. It is because money *does*, in this hardly unmaterialistic society, matter and matter very much that the more of it one risks the more of a lot of other things, such as one's pride, one's poise, one's dispassion, one's masculinity, one also risks, again only momentarily but again very publicly as well. In deep cockfights an owner and his collaborators, and, as we shall see, to a lesser but still quite real extent also their backers on the outside, put their money where their status is.

It is in large part *because* the marginal disutility of loss is so great at the higher levels of betting that to engage in such betting is to lay one's public self, allusively and metaphorically, through the medium of one's cock, on the line. And though to a Benthamite this might seem merely to increase the irrationality of the enterprise that much further, to the Balinese what it mainly increases is the meaningfulness of it all. And as (to follow Weber rather than Bentham) the

---

be more carefully put in terms of a denial that for the Balinese, as for any people, utility (pleasure, happiness . . . ) is merely identifiable with wealth. But such terminological problems are in any case secondary to the essential point: the cockfight is not roulette.

imposition of meaning on life is the major end and primary condition of human existence, that access of significance more than compensates for the economic costs involved.<sup>21</sup> Actually, given the even-money quality of the larger matches, important changes in material fortune among those who regularly participate in them seem virtually nonexistent, because matters more or less even out over the long run. It is, actually, in the smaller, shallow fights, where one finds the handful of more pure, addict-type gamblers involved – those who *are* in it mainly for the money – that “real” changes in social position, largely downward, are affected. Men of this sort, plungers, are highly dispraised by “true cockfighters” as fools who do not understand what the sport is all about, vulgarians who simply miss the point of it all. They are, these addicts, regarded as fair game for the genuine enthusiasts, those who do understand, to take a little money away from, something that is easy enough to do by luring them, through the force of their greed, into irrational bets on mismatched cocks. Most of them do indeed manage to ruin

21 Max Weber, *The Sociology of Religion* (Boston: Beacon Press, 1963). There is nothing specifically Balinese, of course, about deepening significance with money, as Whyte's description of corner boys in a working-class district of Boston demonstrates: “Gambling plays an important role in the lives of Cornerville people. Whatever game the corner boys play, they nearly always bet on the outcome. When there is nothing at stake, the game is not considered a real contest. This does not mean that the financial element is all-important. I have frequently heard men say that the honor of winning was much more important than the money at stake. The corner boys consider playing for money the real test of skill and, unless a man performs well when money is at stake, he is not considered a good competitor.” W. F. Whyte, *Street Corner Society*, 2d ed. (Chicago: University of Chicago Press, 1955), 140.



themselves in a remarkably short time, but there always seems to be one or two of them around, pawning their land and selling their clothes in order to bet, at any particular time.<sup>22</sup>

This graduated correlation of “status gambling” with deeper fights and, inversely, “money gambling” with shallower ones is in fact quite general. Bettors themselves form a sociomoral hierarchy in these terms. As noted earlier, at most cockfights there are, around the very edges of the cockfight area, a large number of mindless, sheer-chance type gambling games (roulette, dice throw, coin-spin, pea-under-the-shell) operated by concessionaires. Only women, children, adolescents, and various other sorts of people who do not (or not yet) fight cocks – the extremely poor, the socially despised, the personally idiosyncratic – play at these games, at, of course, penny ante levels. Cockfighting men would be ashamed to go anywhere near them. Slightly above these people in standing are those who, though they do not themselves fight cocks, bet on the

smaller matches around the edges. Next, there are those who fight cocks in small, or occasionally medium matches, but have not the status to join in the large ones, though they may bet from time to time on the side in those. And finally, there are those, the really substantial members of the community, the solid citizenry around whom local life revolves, who fight in the larger fights and bet on them around the side. The focusing element in these focused gatherings, these men generally dominate and define the sport as they dominate and define the society. When a Balinese male talks, in that almost venerative way, about “the true cockfighter,” the *bebatoh* (“bettor”) or *djuru kurung* (“cage keeper”), it is this sort of person, not those who bring the mentality of the pea-and-shell game into the quite different, inappropriate context of the cockfight, the driven gambler (*potét*, a word which has the secondary meaning of thief or reprobate), and the wistful hanger-on, that they mean. For such a man, what is really going on in a match is something rather closer to an *affaire d’honneur* (though, with the Balinese talent for practical fantasy, the blood that is spilled is only figuratively human) than to the stupid, mechanical crank of a slot machine.

What makes Balinese cockfighting deep is thus not money in itself, but what, the more of it that is involved the more so, money causes to happen: the migration of the Balinese status hierarchy into the body of the cockfight. Psychologically an Aesopian representation of the ideal/demonic, rather narcissistic, male self, sociologically it is an equally Aesopian representation of the complex fields of tension set up by the controlled, muted, ceremonial, but for all that deeply felt, interaction of those selves in the context of everyday life. The cocks may

22 The extremes to which this madness is conceived on occasion to go – and the fact that it is considered madness – is demonstrated by the Balinese folktale *I Tuhung Kuning*. A gambler becomes so deranged by his passion that, leaving on a trip, he orders his pregnant wife to take care of the prospective newborn if it is a boy but to feed it as meat to his fighting cocks if it is a girl. The mother gives birth to a girl, but rather than giving the child to the cocks she gives them a large rat and conceals the girl with her own mother. When the husband returns the cocks, crowing a jingle, inform him of the deception and, furious, he sets out to kill the child. A goddess descends from heaven and takes the girl up to the skies with her. The cocks die from the food given them, the owner’s sanity is restored, the goddess brings the girl back to the father who reunites him with his wife. The story is given as “Geel Komkommertje” in Jacoba Hooykaas-van Leeuwen Boomkamp, *Sprookjes en Verhalen van Bali* (s’Gravenhage: Van Hoeve, 1956), 19 – 25.



be surrogates for their owners' personalities, animal mirrors of psychic form, but the cockfight is – or more exactly, deliberately is made to be – a simulation of the social matrix, the involved system of crosscutting, overlapping, highly corporate groups – villages, kingroups, irrigation societies, temple congregations, “castes” – in which its devotees live.<sup>23</sup> And as prestige, the necessity to affirm it, defend it, celebrate it, justify it, and just plain bask in it (but not, given the strongly ascriptive character of Balinese stratification, to seek it), is perhaps the central driving force in the society, so also – ambulant penises, blood sacrifices, and monetary exchanges aside – is it of the cockfight. This apparent amusement and seeming sport is, to take another phrase from Erving Goffman, “a status bloodbath.”<sup>24</sup>

The easiest way to make this clear, and at least to some degree to demonstrate it, is to invoke the village whose cockfighting activities I observed the closest – the one in which the raid occurred and from which my statistical data are taken.

As all Balinese villages, this one – Tihingan, in the Klungkung region of southeast Bali – is intricately organized, a labyrinth of alliances and oppositions. But, unlike many, two sorts of corporate groups, which are also status groups, particularly stand out, and we may con-

23 For a fuller description of Balinese rural social structure, see Clifford Geertz, “Form and Variation in Balinese Village Structure,” *American Anthropologist* 61 (1959): 94–108; “Tihingan, A Balinese Village,” in R. M. Koentjaraningrat, *Villages in Indonesia* (Ithaca, N.Y.: Cornell University Press, 1967), 210–243; and, though it is a bit off the norm as Balinese villages go, V. E. Korn, *De Dorpsrepubliek tnganan Pagringsingan* (Santpoort [Netherlands]: C. A. Mees, 1933).

24 Goffman, *Encounters*, 78.

centrate on them, in a part-for-whole way, without undue distortion.

First, the village is dominated by four large, patrilineal, partly endogamous descent groups which are constantly vying with one another and form the major factions in the village. Sometimes they group two and two, or rather the two larger ones versus the two smaller ones plus all the unaffiliated people; sometimes they operate independently. There are also subfactions within them, subfactions within the subfactions, and so on to rather fine levels of distinction. And second, there is the village itself, almost entirely endogamous, which is opposed to all the other villages round about in its cockfight circuit (which, as explained, is the market region), but which also forms alliances with certain of these neighbors against certain others in various supravillage political and social contexts. The exact situation is thus, as everywhere in Bali, quite distinctive; but the general pattern of a tiered hierarchy of status rivalries between highly corporate but various based groupings (and, thus, between the members of them) is entirely general.

Consider, then, as support of the general thesis that the cockfight, and especially the deep cockfight, is fundamentally a dramatization of status concerns, the following facts, which to avoid extended ethnographic description I will simply pronounce to be facts – though the concrete evidence-examples, statements, and numbers that could be brought to bear in support of them is both extensive and unmistakable:

1. A man virtually never bets against a cock owned by a member of his own kin-group. Usually he will feel obliged to bet for it, the more so the closer the kin tie and the deeper the fight. If he is certain

in his mind that it will not win, he may just not bet at all, particularly if it is only a second cousin's bird or if the fight is a shallow one. But as a rule he will feel he must support it and, in deep games, nearly always does. Thus the great majority of the people calling "five" or "speckled" so demonstratively are expressing their allegiance to their kinsman, not their evaluation of his bird, their understanding of probability theory, or even their hopes of unearned income.

2. This principle is extended logically. If your kingroup is not involved you will support an allied kingroup against an unallied one in the same way, and so on through the very involved networks of alliances which, as I say, make up this, as any other, Balinese village.

3. So, too, for the village as a whole. If an outsider cock is fighting any cock from your village you will tend to support the local one. If, what is a rarer circumstance but occurs every now and then, a cock from outside your cockfight circuit is fighting one inside it you will also tend to support the "home bird."

4. Cocks which come from any distance are almost always favorites, for the theory is the man would not have dared to bring it if it was not a good cock, the more so the further he has come. His followers are, of course, obliged to support him, and when the more grand-scale legal cockfights are held (on holidays, and so on) the people of the village take what they regard to be the best cocks in the village, regardless of ownership, and go off to support them, although they will almost certainly have to give odds on them and to make large bets to show that they are not a cheapskate village. Actually, such "away games," though infrequent, tend to mend the ruptures between village members that the constantly occurring "home games," where

village factions are opposed rather than united, exacerbate.

5. Almost all matches are sociologically relevant. You seldom get two outsider cocks fighting, or two cocks with no particular group backing, or with group backing which is mutually unrelated in any clear way. When you do get them, the game is very shallow, betting very slow, and the whole thing very dull, with no one save the immediate principals and an addict gambler or two at all interested.

6. By the same token, you rarely get two cocks from the same group, even more rarely from the same subfaction, and virtually never from the same sub-subfaction (which would be in most cases one extended family) fighting. Similarly, in outside village fights two members of the village will rarely fight against one another, even though, as bitter rivals, they would do so with enthusiasm on their home grounds.

7. On the individual level, people involved in an institutionalized hostility relationship, called *puik*, in which they do not speak or otherwise have anything to do with each other (the causes of this formal breaking of relations are many: wife-capture, inheritance arguments, political differences) will bet very heavily, sometimes almost maniacally, against one another in what is a frank and direct attack on the very masculinity, the ultimate ground of his status, of the opponent.

8. The center bet coalition is, in all but the shallowest games, *always* made up by structural allies – no "outside money" is involved. What is "outside" depends upon the context, of course, but given it, no outside money is mixed in with the main bet; if the principals cannot raise it, it is not made. The center bet, again especially in deeper games, is thus the most direct and open expression of social opposition,

which is one of the reasons why both it and match making are surrounded by such an air of unease, furtiveness, embarrassment, and so on.

9. The rule about borrowing money – that you may borrow *for* a bet but not *in* one – stems (and the Balinese are quite conscious of this) from similar considerations: you are never at the *economic* mercy of your enemy that way. Gambling debts, which can get quite large on a rather short-term basis, are always to friends, never to enemies, structurally speaking.

10. When two cocks are structurally irrelevant or neutral so far as *you* are concerned (though, as mentioned, they almost never are to each other) you do not even ask a relative or a friend whom he is betting on, because if you know how he is betting and he knows you know, and you go the other way, it will lead to strain. This rule is explicit and rigid; fairly elaborate, even rather artificial precautions are taken to avoid breaking it. At the very least you must pretend not to notice what he is doing, and he what you are doing.

11. There is a special word for betting against the grain, which is also the word for “pardon me” (*mpura*). It is considered a bad thing to do, though if the center bet is small it is sometimes all right as long as you do not do it too often. But the larger the bet and the more frequently you do it, the more the “pardon me” tack will lead to social disruption.

12. In fact, the institutionalized hostility relation, *puik*, is often formally initiated (though its causes always lie elsewhere) by such a “pardon me” bet in a deep fight, putting the symbolic fat in the fire. Similarly, the end of such a relationship and resumption of normal social intercourse is often signaled (but, again, not actually brought about) by one or the other of the enemies supporting the other’s bird.

13. In sticky, cross-loyalty situations, of which in this extraordinarily complex social system there are of course many, where a man is caught between two more or less equally balanced loyalties, he tends to wander off for a cup of coffee or something to avoid having to bet, a form of behavior reminiscent of that of American voters in similar situations.<sup>25</sup>

14. The people involved in the center bet are, especially in deep fights, virtually always leading members of their group – kinship, village, or whatever. Further, those who bet on the side (including these people) are, as I have already remarked, the more established members of the village – the solid citizens. Cockfighting is for those who are involved in the everyday politics of prestige as well, not for youth, women, subordinates, and so forth.

15. So far as money is concerned, the explicitly expressed attitude toward it is that it is a secondary matter. It is not, as I have said, of no importance; Balinese are no happier to lose several weeks’ income than anyone else. But they mainly look on the monetary aspects of the cockfight as self-balancing, a matter of just moving money around, circulating it among a fairly well-defined group of serious cockfighters. The really important wins and losses are seen mostly in other terms, and the general attitude toward wagering is not any hope of cleaning up, of making a killing (addict gamblers again excepted), but that of the horseplayer’s prayer: “Oh, God, please let me break even.” In prestige terms, however, you do not want to break even, but, in a momentary, punctuate sort of way, win utterly. The talk (which goes on all the time) is about fights against such-and-such a cock of So-and-So which

25 B. R. Berelson, P. F. Lazarsfeld, and W. N. McPhee, *Voting: A Study of Opinion Formation in a Presidential Campaign* (Chicago: University of Chicago Press, 1954).

your cock demolished, not on how much you won, a fact people, even for large bets, rarely remember for any length of time, though they will remember the day they did in Pan Loh's finest cock for years.

16. You must bet on cocks of your own group aside from mere loyalty considerations, for if you do not people generally will say, "What! Is he too proud for the likes of us? Does he have to go to Java or Den Pasar [the capital town] to bet, he is such an important man?" Thus there is a general pressure to bet not only to show that you are important locally, but that you are not so important that you look down on everyone else as unfit even to be rivals. Similarly, home team people must bet against outside cocks or the outsiders will accuse it – a serious charge – of just collecting entry fees and not really being interested in cockfighting, as well as again being arrogant and insulting.

17. Finally, the Balinese peasants themselves are quite aware of all this and can and, at least to an ethnographer, do state most of it in approximately the same terms as I have. Fighting cocks, almost every Balinese I have ever discussed the subject with has said, is like playing with fire only not getting burned. You activate village and kingroup rivalries and hostilities, but in "play" form, coming dangerously and entrancingly close to the expression of open and direct interpersonal and intergroup aggression (something which, again, almost never happens in the normal course of ordinary life), but not quite, because, after all, it is "only a cockfight."

More observations of this sort could be advanced, but perhaps the general point is, if not made, at least well-delineated, and the whole argument thus far can be usefully summarized in a formal paradigm:

#### THE MORE A MATCH IS . . .

1. Between near status equals (and/or personal enemies)
2. Between high status individuals

#### THE DEEPER THE MATCH.

#### THE DEEPER THE MATCH . . .

1. The closer the identification of cock and man (or: more properly, the deeper the match the more the man will advance his best, most closely-identified-with cock).
2. The finer the cocks involved and the more exactly they will be matched.
3. The greater the emotion that will be involved and the more the general absorption in the match.
4. The higher the individual bets center and outside, the shorter the outside bet odds will tend to be, and the more betting there will be overall.
5. The less an "economic" and the more a "status" view of gaming will be involved, and the "solider" the citizens who will be gaming.<sup>26</sup>

Inverse arguments hold for the shallower the fight, culminating, in a reversed-signs sense, in the coin-spinning and dice-throwing amusements. For deep fights there are no absolute upper limits, though there are of course practical ones, and there are a great many legend-like tales of great Duel-in-the-Sun combats between lords and princes in classical times (for cockfighting has always been as much an elite concern as a popular one), far deeper than anything

26 As this is a formal paradigm, it is intended to display the logical, not the causal, structure of cockfighting. Just which of these considerations leads to which, in what order, and by what mechanisms, is another matter – one I have attempted to shed some light on in the general discussion.

anyone, even aristocrats, could produce today anywhere in Bali.

Indeed, one of the great culture heroes of Bali is a prince, called after his passion for the sport, "The Cockfighter," who happened to be away at a very deep cockfight with a neighboring prince when the whole of his family – father, brothers, wives, sisters – were assassinated by commoner usurpers. Thus spared, he returned to dispatch the upstarts, regain the throne, reconstitute the Balinese high tradition, and build its most powerful, glorious, and prosperous state. Along with everything else that the Balinese see in fighting cocks – themselves, their social order, abstract hatred, masculinity, demonic power – they also see the archetype of status virtue, the arrogant, resolute, honor-mad player with real fire, the ksatria prince.<sup>27</sup>

27 In another of Hooykaas-van Leeuwen Boomkamp's folk tales ("De Gast," *Sprookjes en Verhalen van Bali*, 172–180), a low caste *Sudra*, a generous, pious, and carefree man who is also an accomplished cockfighter, loses, despite his accomplishment, fight after fight until he is not only out of money but down to his last cock. He does not despair, however – "I bet," he says, "upon the Unseen World."

His wife, a good and hardworking woman, knowing how much he enjoys cockfighting, gives him her last "rainy day" money to go and bet. But, filled with misgivings due to his run of ill luck, he leaves his own cock at home and bets merely on the side. He soon loses all but a coin or two and repairs to a food stand for a snack, where he meets a decrepit, odorous, and generally unappetizing old beggar leaning on a staff. The old man asks for food, and the hero spends his last coins to buy him some. The old man then asks to pass the night with the hero, which the hero gladly invites him to do. As there is no food in the house, however, the hero tells his wife to kill the last cock for dinner. When the old man discovers this fact, he tells the hero he has three cocks in his own mountain hut and says the hero may have one of them for fighting. He also asks for the hero's son to accompany him as a servant, and, after the son agrees, this is done.

"Poetry makes nothing happen," Auden says in his elegy of Yeats, "it survives in the valley of its saying . . . a way of happening, a mouth." The cockfight too, in this colloquial sense, makes nothing happen. Men go on allegorically humiliating one another and being allegorically humiliated by one another, day after day, glorying quietly in the experience if they have triumphed, crushed only slightly more openly by it if they have not. *But no one's status really changes*. You cannot ascend the status ladder by winning cockfights; you cannot, as an individual, really ascend it at all. Nor can you descend it

---

The old man turns out to be Siva and, thus, to live in a great palace in the sky, though the hero does not know this. In time, the hero decides to visit his son and collect the promised cock. Lifted up into Siva's presence, he is given the choice of three cocks. The first crows: "I have beaten fifteen opponents." The second crows, "I have beaten twenty-five opponents." The third crows, "I have beaten the King." "That one, the third, is my choice," says the hero, and returns with it to earth.

When he arrives at the cockfight, he is asked for an entry fee and replies, "I have no money; I will pay after my cock has won." As he is known never to win, he is let in because the king, who is there fighting, dislikes him and hopes to enslave him when he loses and cannot pay off. In order to ensure that this happens, the king matches his finest cock against the hero's. When the cocks are placed down, the hero's flees, and the crowd, led by the arrogant king, hoots in laughter. The hero's cock then flies at the king himself, killing him with a spur stab in the throat. The hero flees. His house is encircled by the king's men. The cock changes into a Garuda, the great mythic bird of Indic legend, and carries the hero and his wife to safety in the heavens.

When the people see this, they make the hero king and his wife queen and they return as such to earth. Later their son, released by Siva, also returns and the hero-king announces his intention to enter a hermitage. ("I will fight no more cockfights. I have bet on the Unseen and won.") He enters the hermitage and his son becomes king.



that way.<sup>28</sup> All you can do is enjoy and savor, or suffer and withstand, the concocted sensation of drastic and momentary movement along an aesthetic semblance of that ladder, a kind of behind-the-mirror status jump which has the look of mobility without its actuality.

As any art form – for that, finally, is what we are dealing with – the cockfight renders ordinary, everyday experience comprehensible by presenting it in terms of acts and objects which have had their practical consequences removed and been reduced (or, if you prefer, raised) to the level of sheer appearances, where their meaning can be more powerfully articulated and more exactly perceived. The cockfight is “really real” only to the cocks – it does not kill anyone, castrate anyone, reduce anyone to animal status, alter the hierarchical relations among people, nor refashion the hierarchy; it does not even redistribute income in any significant way. What it does is what, for other peoples with other temperaments and other conventions, *Lear* and *Crime and Punishment* do; it catches up these themes – death, masculinity, rage, pride, loss, beneficence, chance – and, ordering them into an encompassing structure, presents them in such a way as to throw into relief a particular view of their essential nature. It puts a construction on them, makes

them, to those historically positioned to appreciate the construction, meaningful – visible, tangible, graspable – “real,” in an ideational sense. An image, fiction, a model, a metaphor, the cockfight is a means of expression; its function is neither to assuage social passions nor to heighten them (though, in its play-with-fire way, it does a bit of both), but, in a medium of feathers, blood, crowds, and money, to display them.

The question of how it is that we perceive qualities in things – paintings, books, melodies, plays – that we do not feel we can assert literally to be there has come, in recent years, into the very center of aesthetic theory.<sup>29</sup> Neither the sentiments of the artist, which remain his, nor those of the audience, which remain theirs, can account for the agitation of one painting or the serenity of another. We attribute grandeur, wit, despair, exuberance to strings of sounds; lightness, energy, violence, fluidity to blocks of stone. Novels are said to have strength, buildings eloquence, plays momentum, ballets repose. In this realm of eccentric predicates, to say that the cockfight, in its perfected cases at least, is “disquietful” does not seem at all unnatural, merely, as I have just denied it practical consequence, somewhat puzzling.

The disquietfulness arises, “somehow,” out of a conjunction of three attributes of the fight: its immediate dramatic shape; its metaphoric content; and its social context. A cultural

28 Addict gamblers are really less declassed (for their status is, as everyone else's, inherited) than merely impoverished and personally disgraced. The most prominent addict gambler in my cockfight circuit was actually a very high caste *satria* who sold off most of his considerable lands to support his habit. Though everyone privately regarded him as a fool and worse (some, more charitable, regarded him as sick), he was publicly treated with the elaborate deference and politeness due his rank. On the independence of personal reputation and public status in Bali, see Geertz, *Person, Time, and Conduct*, 28 – 35.

29 For four, somewhat variant, treatments, see Susanne Langer, *Feeling and Form* (New York: Scribners, 1953); Richard Wollheim, *Art and Its Objects* (New York: Harper and Row, 1968); Nelson Goodman, *Languages of Art* (Indianapolis: Bobbs-Merrill, 1968); Maurice Merleau-Ponty, “The Eye and the Mind,” in his *The Primacy of Perception* (Evanston, Ill.: Northwestern University Press, 1964), 159 – 190.

figure against a social ground, the fight is at once a convulsive surge of animal hatred, a mock war of symbolical selves, and a formal simulation of status tensions, and its aesthetic power derives from its capacity to force together these diverse realities. The reason it is disquietful is not that it has material effects (it has some, but they are minor); the reason that it is disquietful is that, joining pride to selfhood, selfhood to cocks, and cocks to destruction, it brings to imaginative realization a dimension of Balinese experience normally well-observed from view. The transfer of a sense of gravity into what is in itself a rather blank and unvarious spectacle, a commotion of beating wings and throbbing legs, is effected by interpreting it as expressive of something unsettling in the way its authors and audience live, or, even more ominously, what they are.

As a dramatic shape, the fight displays a characteristic that does not seem so remarkable until one realizes that it does not have to be there: a radically atomistic structure.<sup>30</sup> Each match is a world

30 British cockfights (the sport was banned there in 1840) indeed seem to have lacked it, and to have generated, therefore, a quite different family of shapes. Most British fights were "mains," in which a preagreed number of cocks were aligned into two teams and fought serially. Score was kept and wagering took place both on the individual matches and on the main as a whole. There were also "battle Royales," both in England and on the Continent, in which a large number of cocks were let loose at once with the one left standing at the end the victor. And in Wales, the so-called "Welsh main" followed an elimination pattern, along the lines of a present-day tennis tournament, winners proceeding to the next round. As a genre, the cockfight has perhaps less compositional flexibility than, say, Latin comedy, but it is not entirely without any. On cockfighting more generally, see Arch Rupert, *The Art of Cockfighting* (New York: Devin-Adair, 1949); G. R. Scott, *History of Cockfighting* (London: Charles Skil-

unto itself, a particulate burst of form. There is the match making, there is the betting, there is the fight, there is the result – utter triumph and utter defeat – and there is the hurried, embarrassed passing of money. The loser is not consoled. People drift away from him, look through him, leave him to assimilate his momentary descent into nonbeing, reset his face, and return, scarless and intact, to the fray. Nor are winners congratulated, or events rehashed; once a match is ended the crowd's attention turns totally to the next, with no looking back. A shadow of the experience no doubt remains with the principals, perhaps even with some of the witnesses, of a deep fight, as it remains with us when we leave the theater after seeing a powerful play well-performed; but it quite soon fades to become at most a schematic memory – a diffuse glow or an abstract shudder – and usually not even that. Any expressive form lives only in its own present – the one it itself creates. But, here, that present is severed into a string of flashes, some more bright than others, but all of them disconnected, aesthetic quanta. Whatever the cockfight says, it says in spurts.

But, as I have argued lengthily elsewhere, the Balinese live in spurts.<sup>31</sup> Their life, as they arrange it and perceive it, is less a flow, a directional movement out of the past, through the present, toward the future than an on-off pulsation of meaning and vacuity, an ar-

---

ton, 1957); and Lawrence Fitz-Barnard, *Fighting Sports* (London: Odhams Press, 1921).

31 *Person, Time, and Conduct*, esp. 42ff. I am, however, not the first person to have argued it: see G. Bateson, "Bali, the Value System of a Steady State," and "An Old Temple and a New Myth," in Belo, ed., *Traditional Balinese Culture*, 384 – 402, 111 – 136.

rhythmic alternation of short periods when “something” (that is, something significant) is happening and equally short ones where “nothing” (that is, nothing much) is – between what they themselves call “full” and “empty” times, or, in another idiom, “junctures” and “holes.” In focusing activity down to a burning-glass dot, the cockfight is merely being Balinese in the same way in which everything from the monadic encounters of everyday life, through the clanging pointillism of *gamelan* music, to the visiting-day-of-the-gods temple celebrations are. It is not an imitation of the punctuateness of Balinese social life, nor a depiction of it, nor even an expression of it; it is an example of it, carefully prepared.<sup>32</sup>

If one dimension of the cockfight’s structure, its lack of temporal directionality, makes it seem a typical segment of the general social life, however, the other, its flat-out, head-to-head (or spur-to-spur) aggressiveness, makes it seem a contradiction, a reversal, even a subversion of it. In the normal course of things, the Balinese are shy to the point of obsessiveness of open conflict. Oblique, cautious, subdued, controlled, masters of indirection and dissimulation – what they call *alus*, “polished,” “smooth” – they rarely face what they can turn away from, rarely resist what they can evade. But here they portray themselves as wild and murderous, manic explosions of instinctual cruelty. A powerful rendering of life as the Balinese most deeply do not want it (to adapt a phrase Frye has used of Gloucester’s blinding) is set in the context of a sample of it as they do

in fact have it.<sup>33</sup> And, because the context suggests that the rendering, if less than a straightforward description is nonetheless more than an idle fancy, it is here that the disquietfulness – the disquietfulness of the *fight*, not (or, anyway, not necessarily) its patrons, who seem in fact rather thoroughly to enjoy it – emerges. The slaughter in the cock ring is not a depiction of how things literally are among men, but, what is almost worse, of how, from a particular angle, they imaginatively are.<sup>34</sup>

The angle, of course, is stratificatory. What, as we have already seen, the cockfight talks most forcibly about is status relationships, and what it says about them is that they are matters of life and death. That prestige is a pro-

33 Northrop Frye, *The Educated Imagination* (Bloomington: University of Indiana Press, 1964), 99.

34 There are two other Balinese values and dis-values which, connected with punctuate temporality on the one hand and unbridled aggressiveness on the other, reinforce the sense that the cockfight is at once continuous with ordinary social life and a direct negation of it: what the Balinese call *ramé*, and what they call *paling*. *Ramé* means crowded, noisy, and active, and is a highly sought-after social state: crowded markets, mass festivals, busy streets are all *ramé* as, of course, is, in the extreme, a cockfight. *Ramé* is what happens in the “full” times (its opposite, *sepi*, “quiet,” is what happens in the “empty” ones). *Paling* is social vertigo, the dizzy, disoriented, lost, turned-around feeling one gets when one’s place in the coordinates of social space is not clear, and it is a tremendously disfavored, immensely anxiety-producing state. Balinese regard the exact maintenance of spatial orientation (“not to know where north is” is to be crazy), balance, decorum, status relationships, and so forth, as fundamental to ordered life (*krama*) and *paling*, the sort of whirling confusion of position the scrambling cocks exemplify as its profoundest enemy and contradiction. On *ramé* see Bateson and Mead, *Balinese Character*, 3, 64; on *paling*, *ibid.*, 11, and Belo, ed., *Traditional Balinese Culture*, 90ff.

32 For the necessity of distinguishing among “description,” “representation,” “exemplification,” and “expression” (and the irrelevance of “imitation” to all of them) as modes of symbolic reference, see Goodman, *Languages of Art*, 6–10, 45–91, 225–241.

foundly serious business is apparent everywhere one looks in Bali – in the village, the family, the economy, the state. A peculiar fusion of Polynesian title ranks and Hindu castes, the hierarchy of pride is the moral backbone of the society. But only in the cockfight are the sentiments upon which that hierarchy rests revealed in their natural colors. Enveloped elsewhere in a haze of etiquette, a thick cloud of euphemism and ceremony, gesture and allusion, they are here expressed in only the thinnest disguise of an animal mask, a mask which in fact demonstrates them far more effectively than it conceals them. Jealousy is as much a part of Bali as poise, envy as grace, brutality as charm; but without the cockfight the Balinese would have a much less certain understanding of them, which is, presumably, why they value it so highly.

Any expressive form works (when it works) by disarranging semantic contexts in such a way that properties conventionally ascribed to certain things are unconventionally ascribed to others, which are then seen actually to possess them. To call the wind a cripple, as Stevens does, to fix tone and manipulate timbre, as Schoenberg does, or, closer to our case, to picture an art critic as a dissolute bear, as Hogarth does, is to cross conceptual wires; the established conjunctions between objects and their qualities are altered and phenomena – fall weather, melodic shape, or cultural journalism – are clothed in signifiers which normally point to other referents.<sup>35</sup> Similarly, to connect – and con-

35 The Stevens reference is to his “The Motive for Metaphor,” (“You like it under the trees in autumn, / Because everything is half dead. / The wind moves like a cripple among the leaves / And repeats words without meaning”); the Schoenberg reference is to the third of his *Five Orchestral Pieces* (Opus 16), and is borrowed

nect, and connect – the collision of roosters with the divisiveness of status is to invite a transfer of perceptions from the former to the latter, a transfer which is at once a description and a judgment. (Logically, the transfer could, of course, as well go the other way; but, like most of the rest of us, the Balinese are a great deal more interested in understanding men than they are in understanding cocks.)

What sets the cockfight apart from the ordinary course of life, lifts it from the realm of everyday practical affairs, and surrounds it with an aura of enlarged importance is not, as functionalist sociology would have it, that it reinforces status discriminations (such reinforcement is hardly necessary in a society where every act proclaims them), but that it provides a metasocial commentary upon the whole matter of assorting human beings into fixed hierarchical ranks and then organizing the major part of collective existence around that assortment. Its function, if you want to call it that, is interpretive: it is a Balinese reading of Balinese experience; a story they tell themselves about themselves.

To put the matter this way is to engage in a bit of metaphorical refocusing of

---

from H. H. Drager, “The Concept of ‘Tonal Body,’” in Susanne Langer, ed., *Reflections on Art* (New York: Oxford University Press, 1961), 174. On Hogarth, and on this whole problem – there called “multiple matrix matching” – see E. H. Gombrich, “The Use of Art for the Study of Symbols,” in James Hogg, ed., *Psychology and the Visual Arts* (Baltimore: Penguin Books, 1969), 149 – 170. The more usual term for this sort of semantic alchemy is “metaphorical transfer,” and good technical discussions of it can be found in M. Black, *Models and Metaphors* (Ithaca, N.Y.: Cornell University Press, 1962), 25ff; Goodman, *Language as Art*, 44ff; and W. Percy, “Metaphor as Mistake,” *Sewanee Review* 66 (1958): 78 – 99.

one's own, for it shifts the analysis of cultural forms from an endeavor in general parallel to dissecting an organism, diagnosing a symptom, deciphering a code, or ordering a system – the dominant analogies in contemporary anthropology – to one in general parallel with penetrating a literary text. If one takes the cockfight, or any other collectively sustained symbolic structure, as a means of “saying something of something” (to invoke a famous Aristotelian tag), then one is faced with a problem not in social mechanics but social semantics.<sup>36</sup> For the anthropologist, whose concern is with formulating sociological principles, not with promoting or appreciating cockfights, the question is, what does one learn about such principles from examining culture as an assemblage of texts?

Such an extension of the notion of a text beyond written material, and even beyond verbal, is, though metaphorical, not, of course, all that novel. The *interpretatio naturae* tradition of the Middle Ages, which, culminating in Spinoza, attempted to read nature as Scripture, the Nietzschean effort to treat value systems as glosses on the will to power (or the Marxian one to treat them as glosses on property relations), and the Freudian replacement of the enigmatic text of the manifest dream with the plain one of the latent, all offer precedents, if not equally recommendable ones.<sup>37</sup> But the idea remains theoretically undeveloped; and the more profound

36 The tag is from the second book of the *Organon*, *On Interpretation*. For a discussion of it, and for the whole argument for freeing “the notion of text . . . from the notion of scripture or writing,” and constructing, thus, a general hermeneutics, see Paul Ricoeur, *Freud and Philosophy* (New Haven, Conn.: Yale University Press, 1970), 20ff.

37 Ibid.

corollary, so far as anthropology is concerned, that cultural forms can be treated as texts, as imaginative works built out of social materials, has yet to be systematically exploited.<sup>38</sup>

In the case at hand, to treat the cockfight as a text is to bring out a feature of it (in my opinion, the central feature of it) that treating it as a rite or a pastime, the two most obvious alternatives, would tend to obscure: its use of emotion for cognitive ends. What the cockfight says it says in a vocabulary of sentiment – the thrill of risk, the despair of loss, the pleasure of triumph. Yet what it says is not merely that risk is exciting, loss depressing, or triumph gratifying, banal tautologies of affect, but that it is of these emotions, thus exemplified, that society is built and individuals put together. Attending cockfights and participating in them is, for the Balinese, a kind of sentimental education. What he learns there is what his culture's ethos and his private sensibility (or, anyway, certain aspects of them) look like when spelled out externally in a collective text; that the two are near enough alike to be articulated in the symbolics of a single such text; and – the disquieting part – that the text in which this revelation is accomplished consists of a

38 Lévi-Strauss's “structuralism” might seem an exception. But it is only an apparent one, for, rather than taking myths, totem rites, marriage rules, or whatever as texts to interpret, Lévi-Strauss takes them as ciphers to solve, which is very much not the same thing. He does not seek to understand symbolic forms in terms of how they function in concrete situations to organize perceptions (meanings, emotions, concepts, attitudes); he seeks to understand them entirely in terms of their internal structure, *indépendent de tout sujet, de tout objet, et de toute contexte*. For my own view of this approach – that is suggestive and indefensible – see Clifford Geertz, “The Cerebral Savage: On the Work of Lévi-Strauss,” *Encounter* 48 (1967): 25 – 32.



chicken hacking another mindlessly to bits.

Every people, the proverb has it, loves its own form of violence. The cockfight is the Balinese reflection on theirs: on its look, its uses, its force, its fascination. Drawing on almost every level of Balinese experience, it brings together themes – animal savagery, male narcissism, opponent gambling, status rivalry, mass excitement, blood sacrifice – whose main connection is their involvement with rage and the fear of rage, and, binding them into a set of rules which at once contains them and allows them play, builds a symbolic structure in which, over and over again, the reality of their inner affiliation can be intelligibly felt. If, to quote Northrop Frye again, we go to see *Macbeth* to learn what a man feels like after he has gained a kingdom and lost his soul, Balinese go to cockfights to find out what a man, usually composed, aloof, almost obsessively self-absorbed, a kind of moral autocosm, feels like when, attacked, tormented, challenged, insulted, and driven in result to the extremes of fury, he has totally triumphed or been brought totally low. The whole passage, as it takes us back to Aristotle (though to the *Poetics* rather than the *Hermeneutics*), is worth quotation:

But the poet [as opposed to the historian], Aristotle says, never makes any real statements at all, certainly no particular or specific ones. The poet's job is not to tell you what happened, but what happens: not what did take place, but the kind of thing that always does take place. He gives you the typical, recurring, or what Aristotle calls universal event. You wouldn't go to *Macbeth* to learn about the history of Scotland – you go to it to learn what a man feels like after he's gained a kingdom and lost his soul. When you meet

such a character as Micawber in Dickens, you don't feel that there must have been a man Dickens knew who was exactly like this: you feel that there's a bit of Micawber in almost everybody you know, including yourself. Our impressions of human life are picked up one by one, and remain for most of us loose and disorganized. But we constantly find things in literature that suddenly co-ordinate and bring into focus a great many such impressions, and this is part of what Aristotle means by the typical or universal human event.<sup>39</sup>

It is this kind of bringing of assorted experiences of everyday life to focus that the cockfight, set aside from that life as “only a game” and reconnected to it as “more than a game,” accomplishes, and so creates what, better than typical or universal, could be called a paradigmatic human event – that is, one that tells us less what happens than the kind of thing that would happen if, as is not the case, life were art and could be as freely shaped by styles of feeling as *Macbeth* and *David Copperfield* are.

Enacted and reenacted, so far without end, the cockfight enables the Balinese, as, read and reread, *Macbeth* enables us, to see a dimension of his own subjectivity. As he watches fight after fight with the active watching of an owner and a bettor (for cockfighting has no more interest as a pure spectator sport than croquet or dog racing do), he grows familiar with it and what it has to say to him, much as the attentive listener to string quartets or the absorbed viewer of still lifes grows slowly more familiar with them in a way which opens his subjectivity to himself.<sup>40</sup>

39 Frye, *The Educated Imagination*, 63–64.

40 The use of the, to Europeans, “natural” visual idiom for perception – “see,” “watches,”

Yet, because – in another of those paradoxes, along with painted feelings and unsequenced acts, which haunt aesthetics – that subjectivity does not properly exist until it is thus organized, art forms generate and regenerate the very subjectivity they pretend only to display. Quartets, still lifes, and cockfights are not merely reflections of a preexisting sensibility analogically represented; they are positive agents in the creation and maintenance of such a sensibility. If we see ourselves as a pack of Micawbers it is from reading too much Dickens (if we see ourselves as unillusioned realists, it is from reading too little); and similarly for Balinese, cocks, and cockfights. It is in such a way, coloring experience with the light they cast it in, rather than through whatever material effects they may have, that the arts play their role, as arts, in social life.<sup>41</sup>

---

and so forth – is more than usually misleading here, for the fact that, as mentioned earlier, Balinese follow the progress of the fight as much (perhaps, as fighting cocks are actually rather hard to see except as blurs of motion, more) with their bodies as with their eyes, moving their limbs, heads, and trunks in gestural mimicry of the cocks' maneuvers, means that much of the individual's experience of the fight is kinesthetic rather than visual. If ever there was an example of Kenneth Burke's definition of a symbolic act as "the dancing of an attitude" (*The Philosophy of Literary Form*, rev. ed. [New York: Vintage Books, 1957], 9) the cockfight is it. On the enormous role of kinesthetic perception in Balinese life, Bateson and Mead, *Balinese Character*, 84–88; on the active nature of aesthetic perception in general, Goodman, *Language of Art*, 241–244.

41 All this coupling of the occidental great with the oriental lowly will doubtless disturb certain sorts of aestheticians as the earlier efforts of anthropologists to speak of Christianity and totemism in the same breath disturbed certain sorts of theologians. But as ontological questions are (or should be) bracketed in the sociology of religion, judgmental ones are (or should

In the cockfight, then, the Balinese forms and discovers his temperament and his society's temper at the same time. Or, more exactly, he forms and discovers a particular face of them. Not only are there a great many other cultural texts providing commentaries on status hierarchy and self-regard in Bali, but there are a great many other critical sectors of Balinese life besides the stratificatory and the agonistic that receive such commentary. The ceremony consecrating a Brahmana priest, a matter of breath control, postural immobility, and vacant concentration upon the depths of being, displays a radically different, but to the Balinese equally real, property of social hierarchy – its reach toward the numinous transcendent. Set not in the matrix of the kinetic emotionality of animals, but in that of the static passionlessness of divine mentality, it expresses tranquility not disquiet. The mass festivals at the village temples, which mobilize the whole local population in elaborate hostings of visiting gods – songs, dances, compliments, gifts – assert the spiritual unity of village mates against their status inequality and project a mood of amity and trust.<sup>42</sup> The cock-

---

be) bracketed in the sociology of art. In any case, the attempt to deprovincialize the concept of art is but part of the general anthropological conspiracy to deprovincialize all important social concepts – marriage, religion, law, rationality – and though this is a threat to aesthetic theories which regard certain works of art as beyond the reach of sociological analysis, it is no threat to the conviction, for which Robert Graves claims to have been reprimanded at his Cambridge trips, that some poems are better than others.

42 For the consecration ceremony, see V. E. Korn, "The Consecration of the Priest," in Swellengrebel, ed., *Bali*, 131–154; for (somewhat exaggerated) village communion, Roelof Goris, "The Religious Character of the Balinese Village," *ibid.*, 79–100.

fight is not the master key to Balinese life, any more than bullfighting is to Spanish. What it says about that life is not unqualified nor even unchallenged by what other equally eloquent cultural statements say about it. But there is nothing more surprising in this than in the fact that Racine and Molière were contemporaries, or that the same people who arrange chrysanthemums cast swords.<sup>43</sup>

The culture of a people is an ensemble of texts, themselves ensembles, which the anthropologist strains to read over the shoulders of those to whom they properly belong. There are enormous difficulties in such an enterprise, methodological pitfalls to make a Freudian quake, and some moral perplexities as well. Nor is it the only way that symbolic forms can be sociologically handled.

43 That what the cockfight has to say about Bali is not altogether without perception and the disquiet it expresses about the general pattern of Balinese life is not wholly without reason is attested by the fact that in two weeks of December 1965, during the upheavals following the unsuccessful coup in Djakarta, between forty and eighty thousand Balinese (in a population of about two million) were killed, largely by one another – the worst outburst in the country. (John Hughes, *Indonesian Upheaval* [New York: McKay, 1967], 173–183. Hughes's figures are, of course, rather casual estimates, but they are not the most extreme.) This is not to say, of course, that the killings were caused by the cockfight, could have been predicted on the basis of it, or were some sort of enlarged version of it with real people in the place of the cocks – all of which is nonsense. It is merely to say that if one looks at Bali not just through the medium of its dances, its shadowplays, its sculpture, and its girls, but – as the Balinese themselves do – also through the medium of its cockfight, the fact that the massacre occurred seems, if no less appalling, less like a contradiction to the laws of nature. As more than one real Gloucester has discovered, sometimes people actually get life precisely as they most deeply do not want it.

Functionalism lives, and so does psychologism. But to regard such forms as “saying something of something,” and saying it to somebody, is at least to open up the possibility of an analysis which attends to their substance rather than to reductive formulas professing to account for them.

As in more familiar exercises in close reading, one can start anywhere in a culture's repertoire of forms and end up anywhere else. One can stay, as I have here, within a single, more or less bounded form and circle steadily within it. One can move between forms in search of broader unities or informing contrasts. One can even compare forms from different cultures to define their character in reciprocal relief. But whatever the level at which one operates, and however intricately, the guiding principle is the same: societies, like lives, contain their own interpretations. One has only to learn how to gain access to them.

Robert M. Solow

*How did economics get that way  
& what way did it get?*

My exposure to economics as a discipline began in September 1940 when I enrolled as a freshman in the elementary economics course at Harvard College. I will try in this essay to make sense of the evolution of economics over a span of more than fifty years.

An analogy that comes to mind is from *The Boston Globe*. The Sunday edition occasionally publishes pairs of photographs of urban landscapes. They are taken from the same spot, looking in the same direction, but are at least thirty, forty, or fifty years apart. One shows a corner of the city as it looked then and the other as it looks now. Some buildings have disappeared, some new ones have been built, and some of the old ones are still there but with altered facades. This description is also true of the landscape and structure of econom-

ics, and I would like to provide a few then-and-now snapshots. The difference, however, is that with economics something more is called for; the pictures have to be connected. I would like to tell a story about how and why the architecture of economics changed. It will be a sort of Whig history but without the smugness.

There were three textbooks that were used in the 1940 economics course at Harvard. One was a standard principles text by Frederic Garver and Alvin Hansen. Hansen had been at Minnesota with Garver but by 1940 was a professor at Harvard and – although we freshmen had no inkling – the leading figure in bringing the ideas of John Maynard Keynes’s *General Theory of Employment, Interest and Money*<sup>1</sup> into American economics. The second text was a large introductory book called *Modern Economic Society* by Sumner Slichter,<sup>2</sup> also a member of the Harvard faculty and usually referred to as the dean of American labor economists. The book was more about

---

Robert M. Solow, Institute Professor Emeritus and Professor of Economics Emeritus at the Massachusetts Institute of Technology, has been a Fellow of the American Academy since 1956. In 1987 he received the Nobel Prize in Economics. At the time of this essay’s publication in the Winter 1997 issue of “*Dædalus*,” Solow was Institute Professor Emeritus at the Massachusetts Institute of Technology.

---

© 2005 by the American Academy of Arts & Sciences

1 John Maynard Keynes, *The General Theory of Employment, Interest and Money* (London: Macmillan, 1936).

2 Sumner H. Slichter, *Modern Economic Society* (New York: H. Holt and Company, 1935).

economic institutions and their functioning than about theory. The third text was a little green volume by Luthringer, Chandler, and Cline about money and banking, one of a series of little green books. (Lester Chandler of Princeton was the only one of the authors whose name we ever heard again.) It was a pretty boring text, as I remember, but fortunately we only had to read bits of it. This is actually an important point, and I will come back to it later.

Even a quick physical comparison of a good contemporary elementary text with Garver and Hansen and Slichter tells us something. Leaving aside the typographical changes – color, wider margins, larger type – the modern text is sprinkled with diagrams, tables, even simple equations, whereas the older ones present page after page of unbroken prose. In some seven hundred pages, Garver and Hansen have fewer than forty tables or figures. Some of them represent the working-out of numerical examples of simple propositions, and the rest, maybe half, contain data about the U.S. economy. Similarly, there are fifty-five graphs, again divided between a small number of analytical diagrams and a larger number of graphical presentations of actual data. Slichter is not radically different in his nine hundred pages.

The modern counterpart, while no more intellectually demanding for the student (perhaps even less so), is full of diagrams, tables, and equations. The use of analytical diagrams is probably ten times as intense, and the volume of real-world data presented is correspondingly greater. Propositions are often stated in the form of equations, but these are almost always simple statements (i.e., two intuitively understandable quantities must be equal); there is not a lot of heavy mathematics in these texts. (The

older books mention one equation, the Quantity Equation.) The numerical example, hallowed in economics since the days of David Ricardo, is still in use, but it is no longer the analytical workhorse.

The older books are long on classifications – kinds of goods, kinds of industries, kinds of labor – and on descriptions of public and private institutions. The first 260 pages in Slichter's text are exclusively descriptive of the U.S. economy as it then was. I would guess that fewer than one hundred of the next six hundred pages are devoted to the development of analysis or to the application of analysis. Most provide more institutional descriptions, very sensible discussions of economic policy, and serious looks at recent history as it would be seen by an economist. No one should underestimate the value of these historical reflections. They are, in a way, the application of analytical ideas. But there is a not-so-subtle difference. The modern textbook presents and uses economic analysis as a tool to be directly applied to contemporary or historical situations. The student is shown how to map real events into the categories that appear on the axes of the diagrams or the terms in the equations. The older texts are simply more discursive. The underlying ideas are treated more like categories that resonate to this or that bit of history or policy; the authors ruminate more than they analyze.

One sees this clearly in the way these two books present the idea of supply and demand. This is the one piece of analysis that gets careful treatment. Characteristically, however, Garver and Hansen are very good on how one should think about different kinds of commodities – perishable or not, bought frequently or seldom, standardized or not – but the student is not encouraged to make literal use of the apparatus of supply and de-



mand curves. Both books spend time discussing monopolistic elements in real-world markets, but most of the discussion is institutional. There is, of course, no serious treatment of monopoly price because there was very little known at the time.

I do not want to be misunderstood. Garver and Hansen and Slichter were serious people. Their reflections on the workings of the economy are worth reading. They inspire bursts of nostalgia; words like “civilized” came to mind. The point is that the modern text takes a different approach. Of course it explains more; the intervening sixty years of economic research have not been wasted. But it is the tone that I want to emphasize. The modern text treats economics as a collection of analytical tools to be applied quite directly to observable situations.

It is plain from this comparison that there was a significant change between 1940 and 1990 in economics as a discipline and also in the way it sees itself. Perhaps this sea change deserves to be called a transformation. One way to describe it is to say that economics became a self-consciously technical subject, no longer a fit occupation for the gentleman-scholar. And I mean that literally: nowadays economists arrive at their conclusions by using an evolving collection of analytical techniques, most of them nonintuitive, the sort that have to be learned laboriously. The shift of the center of gravity from Great Britain to the United States (and to the G.I. Bill veterans at that) may have helped the process along. Judicious discussion is no longer the way serious economics is carried out. Of course, that is not all that happened in fifty years. A lot of new knowledge was acquired, most of it by virtue of those analytical techniques. New branches of economics appeared,

some of them because new facts and institutions emerged, some of them for internal intellectual reasons. Not many subfields seem to have disappeared, though there was some rearrangement as a more unified macroeconomics absorbed segments like “business cycles.” At the most general level, however, the change in tone was as I have described it.

Many outside observers and some critics from within the profession have interpreted this development as a sweeping victory for “formalism” in economics. The intended implication is that economics has lost touch with everyday life, that it has become more self-involved and less relevant to social concerns as it became more formal (and more mathematical). I think that this view of the discipline rests on a misconception about the change in the way mainstream economists go about their work. Barking may well be justified, but not up the wrong tree.

If “formalist economics” means anything, it must mean economic theory constructed more or less after the model of Euclid’s geometry. One starts with a few axioms, as close to “self-evident” as they can be – although this is harder to do when the subject matter is more complicated than points and lines in a plane – and then tries to work out all the logical implications of those axioms. Formalist economics starts with a small number of assumptions about the behavior of individual economic agents, and a few more about their interactions with each other, and goes on to study what can then be said about the resulting economic system.

The past fifty years have indeed seen formalist economics grow and prosper. But it has not grown very much. Only a small minority within the profession practices economic theory in this style. To tell the truth, not many more pay any

attention at all to formalist theory. Generally speaking, formalists write for one another. The formalist school contains some extraordinarily able people, and of course it attracts economists who not only are talented at mathematics of a certain kind but enjoy it. It is not surprising, therefore, that outsiders think that there is a lot of formalism in economics, just as half a cup of blood spread around a bathroom can make it look like a scene from *Psycho*. Nevertheless, it is an illusion. Modern mainstream economics is not all that formal.

What the outsider really sees is *model-building*, which is an altogether different sort of activity. In college classrooms in the 1940s, whole semesters could go by without anyone talking about building or testing a model. Today, if you ask a mainstream economist a question about almost any aspect of economic life, the response will be: suppose we model that situation and see what happens. It is important, then, to understand what a model is and what it is not.

A model is a deliberately simplified representation of a much more complicated situation. (I have no reference for this, but I think I remember that the philosopher J. L. Austin wrote somewhere that “one would be tempted to describe oversimplification as the occupational disease of philosophers if it were not their occupation.” Exactly.)

The idea is to focus on one or two causal or conditioning factors, exclude everything else, and hope to understand how just these aspects of reality work and interact. There are thousands of examples; the point is that modern mainstream economics consists of little else but examples of this process.

What follows are three of them, described in the sketchiest terms. Suppose we are interested in the effects of taxa-

tion on the willingness to work. (God knows that is a reasonable thing to be interested in.) The usual approach goes something like this: Imagine a typical person of working age who enjoys both consumer goods and leisure, and whose tastes for them can be described in a simple and well-behaved way. This person has a certain amount of nonwage income, from property or from transfer payments of various kinds. He has the option of working any number of hours at a wage rate determined by the market. Part of his income is taxed away according to some known schedule. We have to assume that this person does the best he can to satisfy his tastes for leisure and for the goods that his after-tax income can buy.

We now ask the question that led to this model in the first place. How will he respond to higher tax rates – by working more or fewer hours? If he makes no adjustment, he will have the same amount of leisure time but have fewer goods. That may suggest that he work longer hours, giving up some leisure time for more goods. With the higher tax rates, however, each hour worked brings less in the way of goods, suggesting that work has become less attractive. He may choose, therefore, to work fewer hours. It may make a difference whether the tax system imposes different rates on wage and nonwage income. Perhaps it depends on the details of his preferences; not every person need react in the same way. This model asks for some deeper analysis, which it gets.

Notice all the casual oversimplifications. Not everyone can choose how many hours to work. People do not buy “consumer goods” in general; they buy hundreds of different things, some of which go particularly well with leisure. Some people, but not others, have some control over the intensity with which

they work. There are customs and norms that affect the behavior of different groups. All of this sort of talk is cheap. The point of the exercise is to simplify and see where it leads. Alternative simplifications are possible, and making those choices is the art of the model-builder. How do we judge success? It is a good question, and I will return to it soon.

Here is a different type of example. Anyone who has looked at the history of business cycles knows that net investment in inventories by businesses is highly volatile and can easily account for most of the top-to-bottom change in production during a recession. It is therefore a matter of some importance that we understand the nature of inventory fluctuations. There are plenty of reasons for firms to hold inventories and to change the amount of inventories they hold. Production schedules are efficient when they are smooth, but sales can fluctuate unpredictably (or predictably, as from season to season). Inventories of finished goods provide a buffer, enabling firms to meet a fluctuating demand with smooth production. Inventories of goods-in-process and, to a lesser extent, raw materials and components may be tied fairly closely to current production. Some firms build up inventories in anticipation of future sales, or they may try to run their inventories down if they expect sales to be slack. Inventories of raw materials may provide a way to speculate on the prices of raw materials, buying more than needed when the price is low and using up the surplus when the current price is high. Finally, firms may find themselves with inventories that are lower or higher than they actually want: higher because sales have been disappointing, lower if sales have been unexpectedly strong. Even this list is not a complete inventory of

reasons for holding and changing inventories. And there are potentially important conditioning factors that have been completely left out: relations with suppliers and customers and financial constraints, for instance.

Modeling inventory fluctuations is a matter of finding a way to represent some or all of these motives so that they can be weighed against one another in much the same way that a profit-seeking firm will have to weigh them as it decides what to do. Notice that last month's unintended inventory fluctuations will have an effect on this month's plans, so that the behavior to be described has a dynamics of its own. How do we judge success? Good question, and I will come to it soon.

Lastly, I give yet a third example because it illustrates a quite different point. Ten years ago, Elhanan Helpman modeled a group of countries trading with one another under very special circumstances. Each country specialized completely in producing a single variety of good. In the eyes of consumers, each country's "own" variety served as a symmetrically imperfect substitute for each other country's variety. Consumers, however, all had the same set of tastes, no matter where they lived. Under these restrictive assumptions and a few others, he showed that there would be a simple formula relating the volume of a country's trade to its size. In reality, countries do not specialize in producing a single good, and consumers do not have the same tastes wherever they are. Nevertheless, Helpman's formula seemed to work quite well for a group of OECD (i.e., advanced) countries. The moral might be that, in reality, production patterns are a lot more specialized than tastes.

Recently, however, other economists tried out the Helpman formula on a

*How did economics get that way & what way did it get?*

group of non-OECD countries, including some in Latin America and Africa. It seemed to work pretty well for them too. Paradoxically, perhaps that success casts some doubt on the Helpman model: one would not expect the less advanced countries to exhibit the same specialization in production and commonality of tastes that is plausible for OECD countries. After all, there may be quite different models that imply a similar relation between the size of a country and the volume of its trade. It appears that measuring success may not be a simple notion.

A good model makes the right strategic simplifications. In fact, a really good model is one that generates a lot of understanding from focusing on a very small number of causal arrows. Model-building is not a mechanical process. Some people are better at this sort of thing than others. Economic models are usually stated mathematically, but they do not have to be. They can be described in words, as I have been doing, or in diagrammatic form, or in computer flow charts for that matter. But mathematics turns out to be a very efficient way to express the structure of a simplified model and it is, of course, a marvelous tool for discovering the implications of a particular model. That is probably why outsiders tend to think of model-building as just more formalism. That is a mistake. The mere use of mathematics does not constitute formalism. Maybe the sharpest way to make this point is to say that the mathematics in these models is almost never deep. There are exceptions, of course. Nevertheless I venture the estimate (safe because it is unverifiable) that there is little or no correlation in fact between the difficulty or mathematical depth of an economic model and its value as science. God is in the details, or perhaps in the absence of

details. There is something to be said for both.

The interesting question is *why* economics stopped being clubbable and became technical sometime in the 1940s and 1950s, and *why* model-building took over as the standard intellectual exercise. I think one has to allow for the possibility that it is, after all, the best way to do economics, and we are just seeing the survival of the fittest in another context. That would be Whig history with a vengeance. I confess to some sympathy with that view, but only within limits. I would add that the model-building approach is peculiarly vulnerable to unproductive controversy of a particular kind. I will discuss this later when I talk about measuring success (one model at a time).

I have a different hypothesis to suggest – that technique and model-building came along with the expanding availability of data, and each reinforces the other. Each new piece of information about the economy, especially if it is quantitative information, practically sits there and begs for explanation. Someone will eventually be clever enough to see that it is now feasible to construct a model. Reciprocally, alternative models have to compete on some basis. They are not usually fancy enough to compete on the basis of elegance or depth or the intellectual equivalent of pectoral development. They compete on the basis of their ability to give a satisfying account of some facts. Facts ask for explanations, and explanations ask for new facts.

There is another partner in this evolutionary spiral: the development of new methods of data analysis and statistical inference. The other highly visible change in the style of academic economics since 1940 has been the explosion of econometrics from an esoteric minority taste to an essential part of a Ph.D. edu-

cation – at least one chapter in most of the dissertations produced in a major department. I will say a little about this vertex of the triangle later.

The spread of model-building coincided in time with the development and diffusion of Keynesian economics. This was an accident, but an accident with consequences: the heyday of Keynesian economics provides a wonderful example of the interplay among theory, the availability of data, and the econometric method. *The General Theory* dates from 1936; Simon Kuznets's book on national income accounting appeared in 1938.<sup>3</sup> Both were no doubt related to the depression of the 1930s, but that is just history. The point is that Keynesian theory needed the national income and product accounts to make contact with reality, and the availability of national income and product accounts made Keynesian macroeconomics fruitful (and helped to shape it).

When I mentioned at the very beginning that my freshman textbook of money and banking was a bore, this is what I had in mind. It was the only vestige of macroeconomics that we were taught – although the unemployment rate in 1940 was still about 14 percent! – and it consisted of a few details about the fractional-reserve banking system and the way it provides credit and generates cash. There was such a thing as business-cycle theory, but it was regarded as a sort of special topic. The textbook writers before 1940 had neither the theory nor the data required to give a coherent account of macroeconomics as part of the core of the subject.

Keynes more or less invented macroeconomics. He was not much of a mod-

el-builder himself, but he opened up a gold mine for those who came after. Suddenly there were models of aggregate consumption and aggregate investment, small but complete models of aggregate output and employment, and data against which they could be tested and perhaps improved. Econometricians had new problems of statistical inference to solve. And it all seemed so *important*.

*The General Theory* was and is a very difficult book to read. It contains several distinct lines of thought that are never quite made mutually consistent. It was an extraordinarily influential book for my generation of students (along with John Hicks's *Value and Capital*<sup>4</sup> and Paul Samuelson's *Foundations of Economic Analysis*<sup>5</sup>), but we learned not as much from it – it was, as I said, almost unreadable – as from a number of explanatory articles that appeared on all our graduate-school reading lists. These articles reduced one or two of those trains of thought to an intelligible *model*, which for us became “Keynesian economics.” The most important of those articles were by John Hicks and Oskar Lange, but there was a whole series of them, by Brian Reddaway, David Champernowne, and others. This story provides a different sort of illustration of the clarifying power of the model-building method.

It is very likely that the war, as much as the depression, worked in the same direction. The panoply of wartime policy – Treasury finance, price and production controls, logistics of various kinds – involved economists in social engineer-

*How did economics get that way & what way did it get?*

<sup>3</sup> Simon Smith Kuznets, *National Income and Capital Formation, 1919 – 1935* (New York: National Bureau of Economic Research, 1938).

<sup>4</sup> Sir John Richard Hicks, *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory* (Oxford: Clarendon Press, 1939).

<sup>5</sup> Paul Anthony Samuelson, *Foundations of Economic Analysis* (Cambridge, Mass.: Harvard University Press, 1947).



ing. Any routinization of policy, even nonintrusive policy, leads inevitably to technical questions. What will be the consequences if we do A? Vague generalities will not do for an answer; demands for quantification are just around the corner. Policy A can usually be undertaken with more or less intensity. The powers that be will not only want to quantify the consequences but they will have the power to measure where no one had measured before. Models happen.

If they happen in connection with the availability of data, as I have suggested, then success will be measured by the ability to “explain” the data. Fitness is goodness of fit. (I put “explain” in quotes to emphasize that there need be no claim to fundamental explanation. A model of inventory accumulation will likely eventuate in an equation that relates inventory spending to a small number of observable variables. If that equation actually holds to a fair degree of approximation, the model explains the data.)

There is, however, a twist, and I think it is important. If the logic of model-building, in economics anyway, is a drastic simplification, then one cannot expect any model to fit the facts in every detail. There are examples of models – and not only in economics – that have been judged to be very successful because they manage to account for fairly gross, large-scale patterns that are actually observed and measured. In practice, two consequences seem to follow.

The first is a persistent temptation to add explanatory variables in order to improve the fit. The variables do not follow from the model, or else they would already be there. But it is usually easy to think of reasonable auxiliaries, things that “should” plausibly affect inventory investment even if they were not includ-

ed in the original, narrowly focused model. Trouble arises because data are scarce in economics; more cannot be generated by experimentation. So there is a danger of “overfitting”: adding variables that work in the data at hand but will turn out to be irrelevant in the next batch, making them therefore deeply irrelevant.

This plays into the second and more fundamental problem. In the nature of the case it will often happen that two quite different models can fit the facts just about equally as well. No doubt the right way to proceed is to think of circumstances in which the two models give widely different predictions and to look around for real-life situations that offer the opportunity to discriminate between them. But that may not be possible. (Chemists can do experiments. There is a movement that does experimental economics, but I cannot guess how far it can go.) So naturally the temptation becomes irresistible to compete by adding variables, making slight changes in formulation, looking around for especially favorable data, and otherwise using the tricks of the trade. It can become very difficult ever to displace an entrenched model by a better one. Clever and motivated – including ideologically motivated – people can fight a rearguard battle that would make Robert E. Lee look like an amateur. (And, of course, they may turn out to be right.) Old models never die; they just fade away. So the model-building approach to economics has its problems. But it is what we have: not formalism, and not the more discursive approach that began to break up in the 1940s and is now long gone.

As this description suggests, model-building economists tend to be natural-born, loose-fitting positivists. Progress will come from weeding out empirically

unsuccessful models and improving and extending those that survive empirical tests. This is not to say that mainstream economists think explicitly about method. Philosophical tendencies may come and go. They are attended to only by a tiny fringe of economists who care about formal methodology. Their arguments make no dent in the mainstream, which goes on making and testing models.

It may be useful if I tuck in here a brief commentary on recent and current controversy within academic macroeconomics, as seen from the point of view advanced in this essay. (David Kreps's discussion is similar, but not identical.) This was not part of my original plan, but the controversy is highly visible. The conference held at the Huntington Library in March of 1995 seemed to have a lively interest in the details. Most intriguingly, however, the controversy is often presented as a dispute between formalists and informalists. It would not damage the argument I have been making if there were an element of truth in that characterization. Maybe there is, a little. But in fact I think the story is ultimately a strong confirmation of the thesis of this essay.

In the mid-1970s, the standard textbook treatment of macroeconomics was recognizably "Keynesian" or "American Keynesian." It aimed specifically to provide an aggregative model of the whole economy that could give some sort of analytical account of unemployment, excess capacity, and recession (and their opposites) as pathologies of the market economy. Opposition comes from a school of thought that did in fact invoke the equally standard formal theory of a capitalist economy (the theory of general competitive equilibrium). This school pointed out that this (microeconomic) model had no room for unemployment, excess capacity, and recession,

and made the (somewhat) formalist appeal that mainstream economists were in the position of teaching on Tuesday and Thursday a macroeconomics that was fundamentally incompatible with the microeconomics taught on Monday and Wednesday. I do not think that the appeal to "microfoundations" amounted to much. Macroeconomic hypotheses had always been justified by some sort of appeal to microeconomic reasoning.

Nevertheless it was in part an appeal to formal criteria, and there was considerable force to this logic. But the "New Classical Macroeconomics" then faced the problem of explaining, or explaining away, the fluctuations in aggregate income and employment that constitute the everyday history of prosperity and recession. And this task had to be performed within the framework of a formal theory that seemed to exclude even the possibility of the events to be explained. I will not recount the ingenious proposals that were invented to perform this feat because they were ultimately felt to be implausible. No empirical successes were forthcoming. This approach languished.

The original New Classical Macroeconomics evolved into, or was superseded by, a related style of modeling called "Real Business Cycle Theory" ("real" means "nonmonetary"). And now we cut to what I take to be the chase in this narrative. The goal of Real Business Cycle Theory was the same: to show that the everyday experience of economic fluctuations could indeed be accounted for within the framework of formal general equilibrium theory, without the "impure," "ad hoc," "Keynesian" violations of standard principles. In doing so it proceeded to abandon formalism in all but name by canonizing one very simple, very special, and very maneuverable

version of competitive general equilibrium – in fact, by adopting a highly specific model. It is a model of an economy populated by a single immortal family with perfect foresight. The industrial and market structure of the economy is such that it carries out, step by step, the infinite-horizon optimal plan of the single “representative consumer.” The generality that is the hallmark of formalism is gone.

The gimmick is that the economy is disturbed by irregular, unforeseeable changes in the preferences of the representative consumer and/or in the technology available to the industrial sector. Economic fluctuations are thus not pathological at all; they are the best that can be done by way of adapting to these pleasant and unpleasant surprises. The model itself was already there to be used. The bulk of the intellectual effort goes into the ways of showing that the data of observed fluctuations are compatible with the demands of the model. This is not easy because the key driving forces – irregular changes in tastes and technology – are not directly observable.

So this is formalism, in more or less a window-dressing sense. In practice it is a little bit of model-building and a lot of fairly sophisticated data analysis. It is not a revolution or transformation in the way macroeconomics is done.

To be sure, there has been a dramatic change in doctrine. One genealogy of models has replaced another. One set of implications has replaced another. This was a genuine shift of ideas, perhaps related to events of the 1970s that were, at least temporarily, hard to explain with older models, perhaps related to the general mood of conservatism and suspicion of government action that affected economists as well as others. Such shifts occur from time to time, in macroeconomics and elsewhere. This one did not

amount to a significant move toward formalism. The new doctrine does try to appropriate an air of “rigor” – a standard ploy – but this is mostly advertising.

My reading of the current state of affairs is much like Kreps’s. In the course of massaging the model to make it conform to the facts, the more adventurous advocates of Real Business Cycle Theory have found it necessary to modify many of the clean but extreme assumptions that give formal general equilibrium theory its artificial vanilla flavor. As the representative-consumer-with-perfect-foresight model has been extended to allow for elements of imperfect information, imperfect competition, imperfect flexibility of prices, incomplete markets, and a teeny bit of heterogeneity among the inhabitants, it has come closer and closer to the more or less “Keynesian” model it was supposed to discredit. It is possible – though surely not inevitable – that in another decade all that will be left are some purely technical differences in modeling strategies plus an underlying difference in spirit, with one side regarding all those imperfections as (removable?) flaws in the system and the other side regarding them as the essence of the system itself.

Seen this way, the macroeconomic controversy (made more intense by ideological freight) only makes the model-building tradition seem pretty irrevocable. But then what about a historical approach to economics? Is such a thing viable? The issue is worth discussing, and it will shed some further light on the main argument. In one sense, economics *is* history. I have been insisting that the modern approach to economics is mostly about accounting for data. It is hard to imagine where else data can come from but the past. So economics is about accounting for the past. Most of the time

it is the recent past, but there is no reason why the more distant past cannot be treated in the same way, if only the relevant data are available or can be reconstructed.

When I studied economic history as a graduate student at the end of the 1940s, my teacher was A. P. Usher. I read his *History of Mechanical Inventions*,<sup>6</sup> Clapham on British economic history, and sections of various works on monetary history. They were long on narrative and short on analysis, a lot like the elementary textbooks of a decade earlier. It did not occur to me then, as it has since, that the more distant past provides something potentially valuable to the model-building economist. A good model embodies *accurately* a representation of the institutions, norms, and attitudes that govern economic behavior in a particular time and place. There is no reason to presuppose that a successful model of the supply of labor in the second half of the twentieth century will apply unchanged to the nineteenth century when institutions, norms, and attitudes were different. Long runs of history offer the economist or historian or economic historian the chance to figure out how changes in the “noneconomic” background factors have an influence on behavior in the narrowly economic realm. It is a little like being able to extend the range of temperatures or pressures available in a laboratory.

I am not sure that is how it has worked in practice. One thing is certain: the same progression from discursiveness to model-building has happened in economic history as in economics. Economic historians have become a lot more like economists than economists have become like economic historians. Today’s

economic historians are very likely to be model-builders. I have the disappointing impression that they are far too willing to accept the models devised by late twentieth-century economists and apply them uncritically to the data of other times and places. There are certainly some sterling exceptions who do properly exploit the advantages offered by exotic data; my insider informant says they are few and far between. There are a few who use the study of the evolution of institutions as a laboratory for economic analysis under unconventional assumptions. There are even some who continue to do narrative economic history. On the whole one has to report that the historical approach to economics has lost most of its distinctiveness and is losing the rest. This seems to be a case of not being able to lick ’em, or not wanting to.

The case is much the same with respect to the other social sciences. I am tempted to guess that economics has drawn further away from the other social sciences in the past half-century. But the truth is that there was little or no interchange even in 1940 or 1950. Despite the existence of the occasional sport like Richard Thaler or George Akerlof, who learned from the other social sciences, most of the flow of ideas is in the other direction. There are subcultures in political science and sociology that seem to want to adopt the methods, the terminology, and sometimes the assumptions of economics. Those *Wahlverwandtschaften* are best discussed from within the other disciplines. Richard Swedberg’s *Economics and Sociology* is full of interesting material,<sup>7</sup> but I am less optimistic than he is about any systematic development.

*How did economics get that way & what way did it get?*

6 Abbott Payson Usher, *A History of Mechanical Inventions* (New York: McGraw-Hill, Inc., 1929).

7 Richard Swedberg, *Economics and Sociology: Redefining Their Boundaries* (Princeton, N.J.: Princeton University Press, 1990).

Some sociologists and political scientists are drawn to the way economics uses rationality – in effect, constrained maximization – as an organizing principle and as a source of ideas for model-building. You could do a lot worse. But there is an irony tucked away in that remark. Some economists, though not many, would like to look to sociology as a way of escaping from the narrow idea of rationality. Actually, that way of putting it is not quite right, so I shall try again. The program of constrained maximization has to rest on a careful statement of what is being maximized and what the constraints are. Mainstream economics takes a narrow view of both; some hardy souls would like to try out a wider range of assumptions. They look to sociology and social psychology as a source of alternative ideas. On the whole they do not find what they are looking for, though again there are notable exceptions.

This is not anybody's fault. The writings of people like Jon Elster, Mark Granovetter, Arthur Stinchcombe, and Aage Sørensen – just to take those who are closest to the economists' wavelength – are full of interest to those all too few economists who read them. But they do not provide the usable raw material (or intermediate product) that is being sought. Even a book like Elster's *The Cement of Society*,<sup>8</sup> intelligent as it is and on exactly the right subject, does not send an economist racing to the drawing board. I suppose, though without much confidence, that this failure to connect may arise because the other social sciences have not adopted the model-building philosophy that motivates and guides economists. Experience has

taught me that I should say explicitly that I have no neocolonialist designs: sociology may be right to stay away from model-building as a mode of thought. Adjacent territories may adopt different track gauges for good and sufficient reasons, but their railroads will have problems at the border crossings.

It might be useful for me to say some fairly informal things about the analogies between economics and the natural sciences. It is an uncomfortable task. I have read the usual quota of layman books and, after forty-seven years on the faculty at MIT, I have a lot of friends who are physicists, chemists, and biologists, not to mention engineers. But it is perfectly clear to me that I have no real sense of what goes on in a physicist's or biologist's mind. Still, it is a topic that often comes up in cross-disciplinary discussion.

There is no doubt that economists are attracted to the style of explanation they see (or think they see) in physics. This is at least clear in the externals. Economists feel at home with equilibrium conditions deduced from first principles or from reliable empirical statements. Similarly, they are used to deducing dynamics from local assumptions or generalizations; economics is full of differential or finite-difference equations. All this seems fairly harmless, as long as it works. It will occasionally turn out that some piece of economics is mathematically identical to some piece of utterly unrelated physics. (This has actually happened to me, although I know absolutely nothing about physics.) I think this has no methodological significance but arises merely because everyone playing this sort of game tends to follow the line of least mathematical resistance. I know that Philip Mirowski believes that deeper aspects of mainstream economic theory are the product

8 Jon Elster, *The Cement of Society: A Study of Social Order* (Cambridge: Cambridge University Press, 1989).



of a profound imitation of nineteenth-century physical theory. That thesis strikes me as false, but I would not claim expert knowledge.

To the extent that economists have the ambition to behave like physicists, they face two dangerous pitfalls. The first is the temptation to believe that the laws of economics are like the laws of physics: exactly the same everywhere on earth and at every moment since Hector was a pup. That is certainly true about the behavior of heat and light. But the part of economics that is independent of history and social context is not only small but dull.

I want to suggest that a second pitfall comes with the imitation of theoretical physics: there is a tendency to undervalue keen observation and shrewd generalization, virtues that I think are more usually practiced by biologists. There has long been a tendency in economics to promote biology as an analogy. Even a genuinely great man like Marshall took this line. Most of what is said on this subject is a hopelessly vague use of unexamined analogy, uninformed by biological theory. I am making a much weaker point, that there is a lot to be said in favor of staring at the piece of reality you are studying and asking, just what is going on here? Economists who are enamored of the physics style seem to bypass that stage, to their disadvantage.

There is another respect in which a broader biological analogy might be relevant. Many economists have noted that the evolutionary paradigm ought to be a useful way of doing business. In isolated instances it has already been valuable in economics, but perhaps a little less so than might have been expected. I attribute this to the absence of any close parallel to the quantitatively analyzable transmission mechanism provided in biology by population genetics. Now, with the

rapid development of evolutionary game theory, there may be an opening for real progress. The loop closes, because what is now needed is a body of data to be exploited by the evolutionary game-modeler.

Nothing that has been said in this essay is complicated enough to require summary. Since part of my aim has been to dispel a misperception, I should conclude by making another pass at explaining how the misperception has come to be so widely accepted. Many observers in the other social sciences and in the wide, wide world perceive that economics has become formalistic, abstract, negligent of the real world. The truth is, I think, that economics has become technical, which is quite different. (Nobody regards computer-aided tomography as formalistic.) Far from being unworldly, modern model-builders are obsessed with data.

How could this confusion arise? I have already suggested that it may be the trappings of mathematical model-building that gives the wrong impression to outsiders. Now I will try out another thought. There is a tendency for theory to outrun data. (This includes statistical theory as well as economic theory.) Theory is cheap, and data are expensive. Much the same thing seems to happen in high-energy physics. I am told that the very latest ideas in particle theory could not come close to being tested with any current accelerator or even with the superconducting supercollider if it were to be built. No one even knows how enough energy could be mobilized to do the experiments that might confirm today's most advanced speculations.

In economics, model-builders' busywork is to refine their ideas to ask questions to which the available data cannot

*How did economics get that way & what way did it get?*

*Robert M.  
Solow*

give the answer. Econometric theorists invent methods to estimate parameters about which the data have no information. And, of course, people are recruited whose talent is for just these activities, whose interest is more in method than in substance. As the models become more refined, the signal-to-noise ratio in the data becomes very attenuated. Since no empirical verdict is forthcoming, the student goes back to the drawing board – and refines the idea even more. Oscar Wilde described a fox hunt as the unspeakable in pursuit of the inedible. Perhaps here we have the overeducated in pursuit of the unknowable. But it sure beats the alternatives.

# Thomas C. Schelling

## *Reciprocal measures for arms stabilization*

There has been a widespread change in the thinking on arms control in the last year or so. Much of it is due to the focus of attention on “measures to safeguard against surprise attack” (to use the official terminology). Although this subject is still listed anachronistically under “disarmament,” it is differently oriented. It assumes deterrence as the keystone of our security policy, and tries to improve it. It accepts a retaliatory capability as something to be enhanced, not degraded – something to be made more secure, less accident-prone, less in need of striking quickly to avoid its own destruction, less capable of gaining advantage from a sudden attack of its own. An anomaly of this approach to arms control is that it does not necessarily involve “disarmament” in the literal sense.

---

*Thomas C. Schelling, a Fellow of the American Academy since 1967, is Distinguished University Professor at the School of Public Policy at the University of Maryland. At the time of this essay's publication in the Fall 1960 issue of "Dædalus," Schelling was professor of economics and associate of the Center for International Affairs at Harvard University.*

---

© 2005 by the American Academy of Arts & Sciences

Another anomaly, which rather shakes the disarmament tradition, is that weapons may be more stabilizing and less aggressive if they are capable of civilian reprisal rather than of military engagement. A standoff between two retaliatory forces is in some ways equivalent to an exchange of hostages; and “inhuman” weapons, capable of inflicting damage but not able to go after the enemy’s strategic forces, acquire virtue because of their clearly deterrent function and the lack of temptation they give either side to strike first.

More important, though, is the fact that schemes to avert surprise attack are manifestly compatible with a national military policy, not a renunciation of it. They emphasize the possibility that one can simultaneously think seriously and sympathetically about our military posture and about collaborating with our enemies to improve it. To propose, as does the notion of “measures to safeguard against surprise attack,” that military cooperation with potential enemies may offer opportunities to improve our military posture, opens a new field for imaginative scientific and military thinking, and may eventually enlist the support of the military services themselves.

Most of this progress is still ahead of us; the revolution in thinking about arms control is barely started. Officially we have taken only the most hesitant steps in defining arms control in a way that does not contradict our national security policies. We still talk officially as though “disarmament” can only save money, without noticing that under the new philosophy it could cost more. We still work officially with an image of disarmament that makes it solely a peacetime (cold-wartime) process of negotiating explicit detailed agreements in a multinational context for the reduction or elimination of weapons, without adequately recognizing that, as in limiting war, limiting the arms race can be a more tacit and less formal process than the “treaty” idea implies. More important, the prevalent image of disarmament is still one that gives the process a uniquely defined end point – the point of no arms at all, or virtually none except in the hands of some international authority or synthetic state that would have the power to police the world against international violence but against nothing else.

The cautious and the skeptical, the pessimists and the realists, have doubts about how rapidly that end point can be approached, whether it will be approached at all, and whether the process once started may not be reversed. But the ultimate goal is rarely challenged except by those who have no interest in arms control. And by far the most frequent argument raised in favor of particular limited measures of arms control, perhaps the most widely persuasive, is that these limited measures are at least “steps toward” the goal of ultimate disarmament. We have not faced up to the implications of the anomaly that “measures to safeguard against surprise attack” are designed to preserve a

nuclear striking power, and are not easily construed as just another “step toward” ultimate disarmament.<sup>1</sup>

We still talk about “levels” of armament or disarmament, as though there were only two directions in which to go, up and down, the arms race going in one direction and arms control in the other. We have not yet admitted that, even in the framework of arms control, it could be an open question whether we ought to be negotiating with our enemies for more arms, less arms, different kinds of arms, or arrangements superimposed on existing armaments. We have given little thought even to the weapon system that would be required by that ultimate international authority that might police the world against armed violence, and to whether it, too, would be embarrassed by a “massive retaliation” doctrine that would lack credibility; whether it, too, might be subject to surprise attack; whether it, too, would lack resolution (as some think NATO might lack resolution) to reach an awful collective decision in response to nibbling aggression or bland violation.

The point of this paper is that there is a vast new area to be explored once we break out of the traditional confinement of “disarmament” – the entire area of military collaboration with potential enemies to reduce the likelihood of war or to reduce its scope and violence. It is an area worth exploring because our present military policies and prospects, however we feel about the adequacy of current programs, cannot promise security from a major thermonuclear

1 See T. C. Schelling, “Surprise Attack and Disarmament,” in Klaus Knorr, ed., *NATO and American Security* (Princeton, N.J.: Princeton University Press, 1959), or the shorter version in T. C. Schelling, *The Strategy of Conflict* (Cambridge, Mass.: Harvard University Press, 1960), chap. 10.

war; and even modest improvements achieved through cooperation with the Soviets should be welcome.

It is not true that in the modern world a gain for the Russians is necessarily a loss for us, and vice versa. We can both suffer losses, and this fact provides scope for cooperation. We both have – unless the Russians have already determined to launch an attack and are preparing for it – a common interest in reducing the advantage of striking first, simply because that very advantage, even if common to both sides, increases the likelihood of war. If at the expense of some capability for launching surprise attack one can deny that capability to the other, it may be a good bargain. We both have a common interest in avoiding the kind of false alarm, panic, misunderstanding, or loss of control, that may lead to an unpremeditated war, in a situation aggravated by the recognition on both sides that it is better to go first than to go second. We have a common interest in not getting drawn or provoked or panicked into war by the actions of a third party (whether that party intends the result or not). And we may have an interest in saving some money by not doing on both sides the things that, if we both do them, tend to cancel out.

This common interest does not depend on trust and good faith. In fact it seems likely that unless thoroughgoing distrust can be acknowledged on both sides, it may be hard to reach any real understanding on the subject. The intellectual clarity required to recognize the nature of the common interest may be incompatible with the pretense that we trust each other, or that there is any sequence of activities in the short run by which either side could demonstrate its good faith to the other.

Ancient despotisms may have understood better than we do how to tranquil-

ize relations between them while hating and distrusting. They exchanged hostages, drank wine from the same glass, met in public to inhibit the massacre of one by the other, and even deliberately exchanged spies to facilitate transmittal of authentic information. And perhaps, having exchanged a son for a daughter in the cold-blooded interest of contract enforcement, they may have reduced tension sufficiently to permit a little affection to grow up in later generations.

The premise underlying my point of view is that a main determinant of the likelihood of war is the nature of present military technology. We and the Russians are trapped by our military technology. Weapon developments of the last fifteen years, especially of the last seven or eight, have themselves been responsible for the most alarming aspects of the present strategic situation. They have enhanced the advantage, in the event war should come, of being the one to start it. They have inhumanly compressed the time available to make the most terrible decisions. They have almost eliminated any belief that a really big war either could be or should be limited in scope or brought to a close by any process other than the sheer exhaustion of weapons. They have greatly reduced the confidence of either side that it can predict the weapons its enemy has or will have in the future. In these and other ways the evolution of military technology has exacerbated whatever propensities toward war are inherent in the political conflict between us and our enemies. It might be naïve to say that this is an unmixed evil for both us and the Soviets, since it powerfully affects the bilateral contest between us; nevertheless, it is hard to escape the judgment that nature might have been kinder in the way she let our military technology



unfold itself over the last decade and a half.

It is interesting – more than that, it is useful – to ask what technological achievements (available both to us and to our enemies) we wish had never occurred, and what technological failures we wish had turned out otherwise. Do we wish the hydrogen bomb had never come along to make intercontinental missiles economical? Do we wish that nuclear-powered aircraft had made airborne alert so cheap that retaliatory aircraft could stay aloft rather than be vulnerable on the ground to a missile attack? Do we hope that no one ever discovers an economical means of nullifying ballistic-missile submarines, so that neither side can hope to preclude retaliation by sudden attack? Do we wish that warning systems were so nearly perfect that “false alarm” were virtually impossible, or so poor that we could never be tempted to rely on them? Do we wish that missiles had never become so accurate that they could be used to destroy an enemy’s missiles in an effort to negate an enemy’s retaliatory threat? Do we wish that radioactive fallout could not occur, or do we welcome it as a peculiarly retaliatory (and hence deterrent) weapon effect that is of little use in a preemptive attack? Do we wish that secrecy about weapons and weapon production were much more difficult to maintain than it is, or welcome certain kinds of secrecy as a form of mutually appreciated security against surprise attack?

The reason why it is productive to speculate on these questions, rather than merely fanciful, is that arms control can usefully be thought of as a way of changing some of the answers. In addition to what we can do unilaterally to improve our warning, to maintain close control over our forces, to make

our forces more secure against attack, to avoid the need for precipitate decisions, and to avoid accidents or the mistaken decisions that they might cause, there may be opportunities to exchange facilities or understandings with our enemies, or to design and deploy our forces differently by agreement with our enemies who do likewise, in a way that enhances those aspects of technology we like and that helps to nullify those that we do not.

If we wish that radar were better and cheaper and less limited by the Earth’s curvature, we might make it so by exchanging real estate with the Russians for the construction by each of us of observation posts on each other’s soil. If we hope that no one can ever predict with confidence how his own missiles would do, in a surprise attack, against the hardened missile sites of his opponent, we might deny each other the necessary knowledge by banning tests of large weapons in the era in which anyone actually has a missile in a hard underground site that he could use in a weapon-effects test. If instead we wish that each side might preserve the privacy of its railroad lines for mobile missiles, we might jointly eschew certain surveillance techniques; and if we thought that antimissile defenses of missile sites might be more feasible, and retaliatory forces correspondingly less vulnerable, with the further testing of nuclear weapons and their effects, we might look with more favor on continued weapon testing. These considerations are by no means the whole story in arms control, but they do remind us that we and our enemies can both jointly welcome, or jointly deplore, certain technological developments (like the improved accuracy of long-range missiles) and may possibly find ways, jointly, to enhance them or to offset them,

over and above the things that we can do unilaterally.

These examples suggest some of the criteria that can be applied to limited arms-control schemes, and some of the difficulties in implementing them. As to criteria, the first thing to emphasize is that it takes a good deal of strategic analysis to decide whether a particular limitation or augmentation of weapons or facilities is a good one or a bad one. Viewing limited measures on their individual merits, and not as steps in a comprehensive program that can be justified only by a long sequence of steps to follow, one has to ask whether the technological and economic consequences of a particular scheme are or are not conducive to military stability; and the answer is very unlikely to be closely correlated with whether more weapons or fewer weapons are involved, bigger weapons or smaller ones, or even whether notions of “more” and “less,” “bigger” and “smaller,” can be applied. Whether we would like to see reconnaissance satellites banned or encouraged may depend, for example, on whether we think they will mainly provide targeting information to the initiator of war or mainly provide warning to a potential defender so that a potential attacker is the more deterred. Whether we like big missiles or not may depend on whether we believe, as so many believed a few years ago, that missiles would be simple and sturdy and hard to destroy in their underground sites or believe as so many fear now that increased accuracies and yields make the present generation of missiles better for a first strike than for a second strike. Whether we wish missile technology to be advanced or retarded may depend on whether or not we believe, as many do, that the next generation of missiles will be easier to protect,

easier to hide, or easier to keep moving, and therefore less insecure. Whether one welcomes nuclear-powered ballistic-missile submarines on both sides or deplores them depends on whether they seem to be peculiarly good at surviving and retaliating, and hence “deterrent,” or peculiarly good at getting up close for a no-warning strike on an enemy’s retaliatory power. And if it were somehow possible to enforce a ban on “dirty” bombs, there would still be a genuine strategic question of whether or not we wish deterrent capabilities to be enhanced by the greater punitive power of dirty bombs, recognizing that comparatively slow-acting fallout may be of much less utility to a potential attacker, whose main interest is to minimize retaliation on himself.

The fact that developments such as these require strategic analysis before it can be decided whether they are good or bad is, aside from being true, discouraging. It means that even among the experts there will be disagreement about the consequences of any particular prohibition or exchange of military facilities; it may be next to impossible to get widespread understanding of the relevant arguments, even within governments. And if fairly detailed analysis is required, and careful distinctions have to be made, prohibitions might have to be specified in equally careful detail and with equally fine distinctions. This is certainly an obstacle to negotiation. Furthermore, any analysis – and any prohibition or agreement or exchange of facilities that is justified on the basis of such analysis – is subject to rapid obsolescence. The friendly warning satellite appears, a year later, as a vicious targeting aid to the surprise attacker; the network of warning systems originally designed for mutual reassurance proves in

operation to have too high a false-alarm rate; the missile-guidance systems that we deplored because of their extreme accuracy and the advantage they would give the attacker may prove, after we outlaw them, to have been the main hope for mobile missile systems desired for their invulnerability and hence for their stability. By the time we reach agreement on precisely what to allow in our satellites, where to place our radar, or what missiles to ban, new evidence or new analysis comes along to suggest that the justification of the particular scheme we are about to subscribe to is all wrong.

Finally, by the time we look at individual schemes in sufficient detail to judge whether their strategic implications are “good” for both us and our enemies, we may have narrowed them down to the point where they are intolerably biased. It is probably a mathematically sound principle that the more measures we put in a package, the more their bilateral biases will cancel out, and hence the greater will be the joint gain relative to the competitive advantage. This may mean that once a potential arms-control system is dissected into sufficiently small pieces to apply the right kind of analysis, we shall have more individual bargaining counters too small and too biased for the negotiating process.

The recent negotiations on weapon tests may prove to be typical. First, there has been almost no public discussion of whether the further testing of weapons and weapon effects would really be conducive to the development of greater bilateral military stability or instability over the coming years.<sup>2</sup> Even if the pub-

2 That is, whether further testing would mainly facilitate the development of more secure retaliatory weapon systems with better communication and control, less subject to accident and false alarm, or instead would mainly en-

lic could be got interested in this crucial question, it would be unlikely to have the information it would need to judge the answer. (There has been a good deal of public discussion of the merits and possible demerits of preventing the further spread of nuclear weapons to small countries, but remarkably little discussion of just how a test ban would obstruct the spread.) Second, while it may seem a mischievous stroke of fortune that somebody discovered, between the two conferences, facts or ideas that made the policing of a test ban appear more difficult than it had appeared the year before, this may be exactly what we have to expect in every case. If today we had “completely solved” the new technical problems introduced by the “decoupling” technique, we should still have to be prepared for somebody’s discovering next year a new possibility that had been overlooked, one that contemporary detection technology could not yet cope with.

The test-ban discussions also illustrate that, when an issue has been narrowed down, the bias in the advantages may seem to outweigh the joint advantages. There is more controversy, and understandably so, over whether a prohibition on small-weapon tests is in the American interest, than on whether a prohibition covering the whole spectrum is.

But of all the characteristics of the present test-ban negotiations, the most

---

hance the potency of weapons for preemptive attack and aggravate the urge, when in doubt, to strike quickly and without restraint. The answer is by no means obvious for the period immediately ahead. It should be noted that tests involve not only new-weapon performance but weapon *effects* on previously untested targets, and the latter may be especially relevant to such things as anti-ICBM defense, civil defense, and the vulnerability of fixed or mobile weapons, warning systems, and communication and control systems.

significant may be that we have had a moratorium for some time without a formal agreement. (We do not, of course, have rights of inspection; so we cannot be sure that the moratorium has been kept; but it likely has been, except possibly for the most easily disguised tests.) And this moratorium resulted from no detailed negotiations, no careful specifications, and no written documents to be initialed and ratified. I do not think this result can be wholly explained by the pressure of public opinion. Part of the motivation must be that, whatever one side is sacrificing in improved technology, the other side is also foregoing tests, and each would probably resume them if the other did. Thus the main sanction of an arms-control agreement – the expectation that each will abstain only if the other does – is probably present in this case. It is therefore a genuine instance of “arms control.” If it suffers from being tentative, temporary, qualified, and conditional, so might any arms-control agreement, even if duly negotiated and signed; furthermore, who can say yet that the present “agreement,” if such we may call it, will not be of some duration?

Here, I think, we have an important clue to a process by which arms control may be reached, and the kinds of arms control that can be reached by that process. Maybe arms control is destined to be something more informal than is suggested by the great diplomatic deployments in Geneva. Maybe limited measures of arms control can be arrived at by quite indirect and incomplete communication; maybe they will take the form of a proposal embodied in unilateral action (or abstention from action), which continues if matched by corresponding action on the other side and only for so long as it is. Maybe instead of *arguing*

about what we should do, we will simply do it and dare the other side to do likewise, or do it and quietly suggest that we would like to keep it up, but only if they find it in their interest to do something comparable.

But if arms control is to be arrived at by a more tacit and informal process, and if we are going to call “arms control” any of the military things that we and the Russians abstain from because of an awareness that as long as each abstains the other probably will too, we should look around and see whether we do not already have a good deal of arms control. If we have, we should look at it closely to see what lessons we can draw.

Offhand, it appears (but a more imaginative examination might prove otherwise) that the tacit understandings we have with the Russians concern what we do with our weapons more than what we possess.<sup>3</sup> We seem to have some understandings about traffic rules for patrolling bombers; there are apparently certain lines we stay on this side of, lines the Russians presumably can recognize, the crossing of which they can probably monitor to some extent. This is certainly a restraint that we unilaterally observe in the interest of reducing misunderstandings and alarms. As far as I know, the traffic rules are communicated, not explicitly, but simply by behav-

<sup>3</sup> A possible exception is civil defense. The extraordinary aversion to civil defense in the U.S. government must be complex in its explanation; but an element is very likely a belief that a genuine civil defense program might open up a new dimension of the arms race, leading either to a “civil-defense race” with the USSR or just to an aggravation of the arms competition. The same may be true in the USSR. An interesting question is how much “clandestine” civil defense the Russians are undertaking, and their reasons for keeping it private. (In pointing this out, the author is not trying to justify the aversion to civil defense.)

ing in accordance with them (perhaps *conspicuously* in accordance with them) and possibly by having chosen the dividing lines in such a way that their significance is recognizable. We both abstain from harassing actions on each other's strategic forces; we do not jam each other's military communications, scare each other with fallout from weapons tests, or wage surreptitious peacetime undersea wars of attrition.<sup>4</sup> We may yet develop tacit understandings about zones and traffic rules for submarines, and may (or may not) develop a tradition for leaving each other's reconnaissance satellites alone. We both very obviously abstain from assassination. The Russians recently "negotiated" (by a process of nudging) a sharper understanding about sharing the Pacific for target practice. It remains to be seen whether the U-2 incident causes certain tacit or latent understandings to come unstuck.<sup>5</sup>

In all likelihood we may abstain from the use of nuclear weapons in some limited war, though both sides often seem to denounce officially the notion that a serious limited war should be, or could be, fought without nuclear weapons. Here is an interesting case of an arms limitation that may be tacitly recognized by both sides, and recognized only because each thinks the other may observe

4 Not yet, that is, or not very much. Preserving some of the mutual restraints we now enjoy may be as important an "arms-control" objective as creating more.

5 It seems a correct interpretation that there is still some element of implicit understanding about not transferring nuclear weapons to other countries. Its status is presently a great deal more ambiguous than the author expected a couple of years ago; nevertheless there must be a general awareness on both sides that the restraint of either will be weakened or dissolved by promiscuousness on the other's part.

it too, yet one that is not only not formally agreed on but even denounced and denied by both sides. It seems doubtful whether this tacit understanding could be made much stronger by a written document.<sup>6</sup> A restraint on the use of nuclear weapons may be more persuasive if it seems to rest on the enemy's own self-interest – on his understanding that if he abstains we may too, but only if he does – than if it pretends to rest on the power of a written agreement or on a fiction of "good faith."

In fact, all of the tacitly agreed limits that do apply, or may apply, in limited war can be construed as a kind of informal arms control tacitly arrived at. My impression is that we and the Russians will go to some length to avoid having American and Russian troops directly engage each other in a limited war, simply because such an engagement might create extremely unstable expectations about whether the war could remain limited. We and the Russians both recognize many legalistic limitations in war, such as the distinction between North Koreans and Chinese, between volunteers and regulars, between the provision of materials to an ally and the provision of manpower, between doing an ally's reconnaissance for him and doing his bombing, perhaps even the distinction between local airfields that are fair game because they are on the ground within a disputed country and the decks of carriers offshore that might for some reason be construed as "sanctuary."

6 It could be made much stronger by various unilateral actions. One would be to increase our *capability* to get along without nuclears in limited war. Another would be to add symbolic support to the understanding; the test-ban negotiations – especially if a formal agreement is reached – almost certainly do this, whether they are intended to or not.



Most of these limits are arbitrary, conventional, and casuistic – purely matters of tradition and precedent. For that reason they are uncertain and insecure; nobody is even nominally committed to honor them. But they demonstrate that it is possible for potential enemies to arrive tacitly, or by indirect communication, at a meeting of minds about some rules, and about how to interpret intentions through the way one operates and deploys his resources. Most important, the limits that can be observed in limited war are a powerful demonstration that sheer self-interest – the recognition of a need to collaborate with an enemy in wartime, to reach understandings that transcend the formalities of explicit communication; the recognition of a mutual interest in avoiding accidents, incidents, misunderstandings and unnecessary alarms, and in holding to any constraints that can be found – can provide potent sanctions that need not rest on explicit negotiation and formal agreements.

We may, then, increase our understanding of the nature of arms control, what it rests on and how it may come about, by recognizing limited war as a kind of arms control in itself. And perhaps it differs from peacetime (i.e., cold-war) arms control less than we customarily think. Perhaps the psychology and the sanctions and the mode of communication, the kinds of reasoning involved, the lack of formal agreement or even acknowledgment, that typify limited war, represent a more central and typical process of international negotiation than we usually give it credit for.

There is another aspect of limited war that deserves emphasis in this connection. The limits in limited war are arrived at not by verbal bargaining, but by maneuver, by actions, and by statements and declarations that are not direct com-

munication to the enemy. Each side tends to act in some kind of recognizable pattern, so that any limits that it is actually observing can be appreciated by the enemy; and each tries to perceive what restraints the other is observing. For that reason the limits themselves must be clear-cut, must be of an “obvious” character, must be based on qualitative distinctions rather than matters of degree. They must not be too selective, too gerrymandered in discriminating between what is inside and what is outside the limit. They must attach themselves to benchmarks, demarcation lines, and distinctions that come naturally. They must have simplicity. They must take advantage of conventions and traditions and precedents that exist, even if the precedents and traditions are biased between the two sides or a nuisance to both sides. Often they must involve all-or-none distinctions, or across-the-board distinctions like that between land and water, between material and manpower, between two sides of a border, or even some arbitrary but potent and highly suggestive feature like a parallel of latitude.<sup>7</sup>

This is certainly true in the case of the use of nuclear weapons in limited war. It is enormously more likely that a limit against any use of nuclear weapons could be recognized, sensed, and adhered to by both sides on condition that each other observe it, than that any particular quantitative limitation, target limitation, fission vs. fusion limitation, or limitation based on who is the “aggressor,” could be jointly and tacitly converged on by the participants.

<sup>7</sup> For an extensive analysis of tacit bargaining, with special reference to limited war, see chapters 3 and 4 and Appendix A of T. C. Schelling, *The Strategy of Conflict* (Cambridge, Mass.: Harvard University Press, 1960).

But the same is certainly true of a test suspension. A tacitly reached moratorium on testing nuclear weapons – mutual and reciprocal but essentially unilateral on both sides – is much more likely to be stable and durable, much less likely to be eroded by ambiguous behavior, than a selective moratorium. If we and the Russians are very selective in our unilateral restraints, each choosing the particular yields, altitudes, fission-fusion combinations, and localities for tests, it seems unlikely either that both sides will hit on the same limitations and maintain them with confidence, or that both will hit on “equivalent” though different restraints.

To some extent, then, the gains and losses of a particular agreement, i.e., the way any particular understanding that is reached may discriminate between the two parties (or among more than two parties), are likely to be dictated somewhat by the elements of the problem, and not altogether by the detailed preferences of the parties to the understanding or their bargaining skill. An absolute ban on weapon tests, for example, or any other across-the-board prohibition, is somewhat arbitrary in the way it distributes the advantages; but perhaps some of its appeal is precisely in the fact that it is somewhat arbitrary, somewhat determined by chance or by the very structure of the problem, dictated by circumstances rather than by either side to the other.

If an important part of our arms control – or let us call it “mutual arms accommodation” – with our enemies is going to be tacit and informal, a matter of reciprocated unilateral actions and abstentions, we need to take seriously the problem of communicating with our enemies about what we are doing, and of reaching understandings with them. In some respects informal communication

is easier, in some ways harder; the process is different from that of formal, explicit, detailed negotiation, and imposes different requirements. Informal communication is usually ambiguous; a government speaks by hint as well as by overt statement and proposal, it speaks indirectly through the medium of press conferences, leaks of information, and remarks to third parties. It speaks with many voices, in the executive branch, in the congress, and even in private articles and news stories that are “inspired” or are inferred to be so. And it speaks through the actions it takes.<sup>8</sup>

The differences should not be exaggerated; even when large teams of professional diplomats and technical experts are assembled in Geneva, much of the communication takes these other forms. Nevertheless, the strategy of communication is different, particularly because of the greater need in informal negotiations to reach a real understanding. In formal and explicit negotiation, what eventually matters is to a large extent what gets written down and agreed to; even if there was not a meeting of minds, there may have been a meeting of words that provides a record of the expectations of both sides and the obligations perceived. In informal negotiation the ultimate sanction depends less on a piece of paper than on the clarity of the understanding reached. If one behaves in a particular way, in anticipation of the other’s reciprocation, there is a need to make clear precisely how one is behaving, with what mutual purpose in mind, so that the other can read the proposal in it, infer what would constitute reciprocation, and design its own behavior accordingly.

8 In a sense, the abortive summit conference of May 1960 did not involve less “negotiation” just because the meeting never took place.

There is furthermore a greater need to be persuasive. In explicit negotiation, it may be possible to reach an agreement whose terms are reasonably well understood without agreement on principles or any reciprocal understanding of each other's motives. If the letter of the agreement is clear, the spirit can remain somewhat in doubt. In informal negotiation, the spirit bears most of the burden; and if the *idea* behind what we think we are doing is not perceived by our partner (enemy), what we expect of him – or what we may reasonably be expected to expect of him – may be too dimly perceived to be the basis for genuine reciprocation.

Suppose we decide to put more emphasis on ballistic-missile submarines, for example, in the belief that they are peculiarly “stable” weapons because of their lesser susceptibility to destruction in case of a surprise attack and because they are not so much under obligation to strike quickly in the event of an ambiguous warning (or war itself), or else because their smaller warheads, with possibly a lesser degree of accuracy as compared with ground-based missiles, makes them less of a threat to the enemy's retaliatory forces and more of a genuine deterrent. Suppose we decide that we could afford to do this only if the enemy himself oriented his own strategic program toward similarly “stable” weapon systems. It might not be at all clear to the Russians what our motives are, or what the conditions were for our going through with the program. Or suppose we have a crash program for the development of a more secure ground-based missile force, this program to be financed by a sharp increase in the defense budget, with a good deal of expenditure on command, control, and communication arrangements so as to reduce both the vulnerability of our weap-

ons and their sensitivity to accident or false alarm. In particular, suppose that our budget rises because of increased outlays associated with our desire for a *slow* reacting force, rather than one that must react rapidly. In such circumstances, our actions may be stabilizing or destabilizing, depending on whether the enemy can perceive that we are making the world safer for him rather than increasing his need (and ours) to jump the gun in a crisis. If we institute an airborne alert, it may be important to do so in a way that enhances the apparent as well as the real security and stability of our retaliatory weapon systems. This might mean that we would have to choose deliberately, say, flight patterns that manifestly enhance the security of our forces rather than the speed with which they could initiate a surprise attack of their own.

By far the most important prerequisite is that we understand our own motives well enough to take actions that are consistent with a deterrent philosophy, and well enough so that we can articulate it to ourselves. If we have such a philosophy, and if our actions are consistent with it, and if for our own purposes we articulate that philosophy in explaining our budget decisions here at home, we are probably well on the way to conveying that philosophy persuasively to our enemy, if he is at all receptive. A special problem here is that our overt position on disarmament must not be too inconsistent with the philosophy that we are trying to display and get across to our enemy. If, for example, we really believed in a policy of collaborating with the Russians to develop a stable situation of mutual deterrence, and if we determined to make important changes, to this end, in the configuration of our weapons but these changes were not in the direction of general disarmament,

we would put a double burden on our communication if the front we presented on arms-control questions bore no relation to that philosophy. This does not necessarily mean that we have to speak in our formal disarmament diplomacy in a manner that is sincere and consistent with what we are fundamentally trying to get across to the Russians. It may just mean that our insincerity should be as manifest as the inconsistency, so that when we do contradict ourselves the Russians know that this is for show and that they should look for the real message elsewhere. Still, it would help if we could find the diplomatic courage to shift even the formal discussions of arms control more into accord with our basic military policy, at the same time as we try to adapt that military policy in directions that the Russians can appreciate and reciprocate, so that disarmament negotiations can help a little, or at least hinder as little as possible, the development of a genuine understanding.

Even so, it is still an unanswered question whether the Russians are at all disposed to participate in any “mutual arms accommodation” with us, beyond what we already do in a tacit way. And it is a difficult technical question whether, even if they are disposed to cooperate with us and appreciate the principle of stable retaliatory systems with minimum proclivity toward false alarm and minimum temptation toward surprise attack, there are any promising actions to be undertaken. Weapon systems can rarely be classified indisputably as first-strike or second-strike weapons, as “accident-prone” or “accident-proof”; a good deal of technical analysis has to lie behind a judgment, many of the technical judgments may not be made equally by us and our enemies, the judgment has to be made in the context of an evolving

weapon system for which facts are really only forecasts, and what is known today may no longer be true tomorrow. It is, furthermore, too much to expect the massive bureaucracy of our defense establishment and our foreign service, and the partisan conflicts in Congress, to produce and maintain a coherent philosophy and transmit it with high fidelity to a suspicious enemy whose receptivity and reasoning processes we can only poorly evaluate. But it is worth trying.

One possibility, already adverted to, is to design our military forces conspicuously and deliberately in the direction of deterrence, stability, and slow reaction. That is, to articulate as a policy the design of a strategic force that is peculiarly good at waiting out crises, at surviving a surprise attack, and at punishing an attacker *ex post facto*, and not particularly good at initiating a preventive attack, not in need of responding rapidly to warning.

This may not be a bad policy to follow unilaterally; but the advantage of pursuing it is greater if the enemy pursues it too. The more each side perceives the other as designing his force for a sudden preemptive attack in a crisis, or for a premeditated surprise attack, the more one is tempted himself to develop a quick-reacting system, one that is peculiarly suited to catching the enemy's military forces before they have left the ground. Thus to some extent such a policy is a conditional policy; the motive is greater if the principle is reciprocated by the enemy.

It would be extraordinarily difficult, perhaps impossible, to negotiate a detailed understanding of precisely what kinds of weapons in what configurations, and how deployed, would meet the “stability” criterion. For that reason the idea may not be one that lends

itself to explicit detailed negotiated agreements. But that does not rule out the possibility that both sides may perceive value in pursuing such policies in a general way, and may recognize that their own behavior not only helps the other side pursue a similar policy but helps to induce it by the tacit promise of reciprocation. As mentioned above, we already do this in such matters as the traffic rules we both unilaterally observe and reciprocate; there may be a good deal of room for gradually extending this kind of reciprocal unilateral action, even though the subject may never appear on the agenda of a diplomatic negotiation.

Compared with a *peaceful* world disarmed, schemes to stabilize mutual deterrence are a poor second best; judged against the prospect of war, measures to make it less likely may be attractive. This point of view will not appeal to any who believe that war results from the sheer existence of arms and the temptation to use them, or from the influence of militarists in modern society whose prestige increases in proportion to the arms budget, and who believe that distrust is only aggravated by people's acting as though distrust exists. History shows, it is said, that man cannot live in a world with arms without using them. History rarely shows anything quite that universal; but even granting it, the question is not whether it is asking much of man to learn to live in a world with arms and not to use them excessively. The question is whether it takes more skill and wisdom for man to learn to live in a world with arms and not to use them than it does for man to disarm himself so totally that he can't have war even if he wants it (or can't want it any longer). If modern social institutions are capable of achieving disarmament in the first place, and of avoiding arms races in per-

petuity thereafter, perhaps they are capable of supporting a world with arms without war. Those who argue that peace with arms is impossible but act as though peace and disarmament are not, may be using a double standard.

And it must be remembered that total disarmament, even if achieved, does not by itself preclude subsequent arms races; nor does a good start toward total disarmament preclude a violent reversal. To the extent that an arms advantage is more easily obtained when the level of armaments on both sides is low – to the extent that the consequences of cheating are greater in a world with few arms – arms races might become more violent, the lower the level of armament from which they start. Particularly in a world in which the pace of scientific progress is rapid but jerky, uneven as between countries, and full of opportunities and uncertainties for weapons development, it is not at all clear that the world would be less uneasy about arms advantages if each side continually thought of itself as nearly naked. What can explain the complacency of the American response to the first Soviet sputnik except a feeling (superbly rationalized) that the existing level of arms provided so much security that no single new achievement, or even a revision of the comparative time schedules by a year or two, could quite upset the balance.

Another area of possible cooperation is in damping the arms race through the exchange of information. I am not much impressed with the budgetary fury of our participation in the arms race, but it is not hard to imagine that the budgetary arms race might get into much higher gear. If it does, part of the motivation (at least in this country) may be due to uncertainty about the level of armament on the other side. The “missile gap” that



one estimates, or feels obliged to assume to exist in the absence of information, may exceed the actual missile gap, causing a more frantic increase in armaments than would be undertaken with better information. And it may induce reciprocal action on the other side, which also wishes to avoid an intolerably unfavorable imbalance.

To illustrate: suppose that either side felt reasonably secure against sudden attack as long as its enemy's numerical superiority in missiles never reached, say, 2 to 1. In this case, just knowing what each other possesses and is producing could make possible a stable equilibrium at a modest level of strategic armaments, while ignorance of the enemy's strength might seem to require an unlimited effort to avoid falling too far behind. With actual weapons such simple calculations are of course impossible; but the principle is valid.

An important difficulty of applying it, though, is that the ways by which one can get authentic information about the other's present and projected strength may provide more strategic information than the other side can tolerate.<sup>9</sup> A special difficulty is that the Soviets may already know most of what they need to know for this purpose; it is mainly we who do not.

But it is interesting that they might possibly prefer that we know the truth. If in fact we are on the verge of a crash program based on an exaggerated estimate of what they have already done, it could cost them money (and perhaps an increase in the risk of war) to keep

<sup>9</sup> Also, one side yields a bluffing or bargaining advantage if it reveals that its weaponry is less impressive than may have been thought. It loses, too, the possibility of surreptitiously achieving a dominant superiority. But losses of this kind are the price of arms control in the first place.

up with us. It is also interesting that the truth is probably not something that they could readily reveal on their own. They have to find some way of giving us evidence for believing the truth (or a less exaggerated estimate of the truth) and give it in a way that does not yield targeting and other information that they would find intolerable. The fact that this intelligence gap is mainly on our side does not preclude Soviet interest in some means of conveying the information to us, and it does not obviate the need for cooperative techniques for receiving it.

Measures to prevent "accidental war," war by misunderstanding, war by false alarm, are another possibility. One aspect of this has been mentioned: the reciprocal development of the kinds of forces and modes of behavior that minimize accidents or their consequences, minimize alarms and misunderstandings, minimize the need to react quickly in the face of ambiguous evidence. But there is another type of joint or reciprocal activity that could help. It would be to arrange in advance, even if crudely and informally, communication procedures, exchange of information, and inspection facilities, for use in the event of an accident, alarm, or misunderstanding that created a crisis. Part of this is just procedural – making sure that we and the Russians have the same idea about who gets in touch with whom when communication or bargaining is suddenly required. Part of it is intellectual – thinking ahead of time about how one would go about reassuring the Russians in the event they had a false alarm, and what we could demand of them for our own reassurance if we ever got ambiguous evidence. Part of it is physical – making sure that, if we should need inspectors on a particular scene within a

few hours to verify that something was an accident, or to verify that the Russians were calm, or to verify that the Russians were not taking actions we thought they were taking, the necessary inspectors and equipment would be available within a few hours' travel time from where we would need them. Just having some Russians available at strategic points around the United States, able to see things with their own eyes if we suddenly wanted them to and able to report home instantly through authentic channels, might be useful someday. And if we ever want them, we may want them in a hurry; there may not be time to identify them, brief them, ship them over here, and train them for their job, once the accident occurs or the crisis is on or the misinformation filters through the Russian warning system.<sup>10</sup>

There is a more ambitious possibility. Neither we nor the Russians at the present time take arms control terribly seriously; we do not view it as an alternative to a war that is imminent. But it is not impossible to imagine crises in which the likelihood of immediate war would become a grave preoccupation. Once the threat of imminent war rises above some threshold, the mere consciousness that each side is preoccupied with it – and with the importance of being the one to start it, if it should come – will aggravate the propensities that already exist. It is perfectly conceivable that in a real crisis there would be a sudden and drastic change in the attitudes of both sides toward arms control. “Preventive arms control” might begin to look like a risky but attractive alternative to a possibly

inevitable preemptive war. Sudden and drastic “measures to safeguard against surprise attack” might have to be negotiated on an acutely demanding time schedule.

If so, success may depend on whether one or both sides is intellectually prepared for the contingency, whether some understandings have been reached in advance, and whether certain facilities can be improvised to monitor whatever arrangements might be forthcoming. One of the important “limited” arms-control measures that we might take in advance of such a crisis, either by ourselves or with our enemies, either informally or explicitly, is a development of understandings, procedures, personnel, and equipment, of an imaginative and adaptable sort, capable of going into action at such time as we and the Russians both decide that now is the time for arms control and we can't wait.

A final possibility, a pessimistic but a serious one and one suggested by the analogy between arms control and limited war, is the role of arms control in general war if general war occurs. We usually think of arms control or deterrence as having failed if war breaks out; and so it has, but it can fail worse if we give up at that point. It is not entirely clear that a general war – a war between the USA and the USSR, involving their strategic forces on a large scale – would necessarily be unlimited either in the way it would be fought or in the way it would be concluded. Particularly as we come to think about an inadvertent war – one that results by some kind of accident or misunderstanding, or one that is reluctantly initiated by the Russians or by us in the belief that it is urgent to preempt at once – it is worthwhile to consider whether fury is the only guide we need in conducting the war, and whether the

<sup>10</sup> A more extensive discussion of this point will appear in T. C. Schelling, “Arms Control: Proposal for a Special Surveillance Force,” *World Politics* (October 1960).

exhaustion of weapons on both sides is the only condition for terminating it.

It is commonly taken for granted that if the Russians initiate a general war it would be in a vicious effort to exterminate us both as a nation and as a people, and that they would be so impatient to do this as to spend valuable weapons to create civil damage at the outset. But it is not obvious that a coldly calculating enemy would afford himself the luxury of going after cities and people when there are more urgent targets that he has to destroy in order to reduce the scale of our retaliation. Nor is it obvious that an impetuous attacker, one whose motivation is partly the fear that if he does not strike first he will be second, would be immune to the thought that he might want to surrender if the thing went badly, to accept our surrender if it went well, or to negotiate a truce between those extremes. If there is no immediate strategic need to kill our people, it may occur to him that they are worth more alive than dead; the threat of killing them gives him something to bargain with in the course of the war or at its termination. Similarly for us: if the war was a mistake we might be more interested in minimizing the consequences of the error, whosever error it was, and in maintaining the possibility of a negotiated outcome that limited damage on both sides. For this bargaining purpose, live Russians and our unspent weapons are assets, and about the only ones we'd have.

The subject is a complicated one and cannot be decided here. It has to be acknowledged that there are dangers in suggesting to the Russians that we are even aware of the possibility that an attack on us might not be cataclysmic for us both. But the possibility is so universally unmentioned and so terribly important that it deserves to be brought

into the open for study. Its relation to arms control is that the mere possibility of limiting a general war between us and our principal enemy may depend on some understanding, tacit and informal as it may be, that we share ahead of time. There may be little national advantage in abstaining from certain targets in the event of war, or in attempting to communicate, unless the enemy can be alert to what is going on.

**T**erminating a war through anything other than the sheer exhaustion of weapons on both sides would require some form of arms control. It is a noteworthy characteristic of a possible World War III that even unconditional surrender may be physically impossible. How do the Russians persuade us that they have destroyed (or are prepared to destroy or deliver us) some or all of their significant weapons and are prepared to submit to our political demands? We cannot even trust them not to test weapons under a test-suspension agreement; in circumstances infinitely more desperate, when a one-hour pause in the war may be of strategic benefit to somebody, if they send us an urgent message acknowledging their guilt in the war and proposing that we preserve our world by letting them surrender to us, are we likely to be able to do anything? If they are fooling, and if we are fooled, the cost will be tremendous; if they are not fooling and we choose to ignore them, the cost will be tremendous. Can we think of what they might do to prove that they mean it? Have we got the facilities to monitor them and to police them? Have we incorporated in our strategic forces, and in the operating doctrine of those forces, recognition of their potential role in policing the disarmament by which the war might be brought to a close?

Actually “surrender” is a poor word here. Anywhere between the two extremes of unconditional surrender by one side or the other, the truce or understanding or scheme for bringing the war to a close might better be described as “disarmament” or “arms control.” Historically one might have allowed an enemy, when he “conditionally” surrendered, to keep some purely defensive weapons as a hedge against the victor’s violating his promise. This is a kind of asymmetrical disarmament scheme. In the future, at the close of a general war, one might have to allow the conditionally surrendering enemy to retain some retaliatory weapons, these being the only kind that two major powers can use to enforce promises from each other. In effect, “measures to safeguard against surprise attack,” possibly one-sided, possibly bilateral, and certainly more drastic than any that have yet been considered, might be the minimum requirement of a conditionally surrendering enemy.

Thus anywhere between the two extremes of total surrender, the outcome should be viewed as a disarmament process, with the asymmetry presumably reflecting the degree of victory or defeat. But as remarked above, even the extremes of unconditional surrender require much the same kind of procedure for mutual relaxation, cessation of hostilities, inspection, enforcement, and so forth. Any general war that is terminated by a bilateral understanding, by anything other than the independent exhaustion of weapons on both sides, requires something in the nature of an enormous, complex and dynamic scheme for arms control.

If this possibility is to be left open, we need to anticipate it in the design of our strategic forces and in our plans for their use. It may require special facilities and

equipment to bring a war to a close, of a kind not necessarily provided for in a plan that considers only the contingency of an all-out war to the finish. But it also requires some mutual awareness ahead of time, on the part of both our enemy and ourselves, and perhaps some crude and tacit, if not careful and explicit, understanding about the modes and techniques of negotiation in the event of war.

# John Hope Franklin

## *The two worlds of race: a historical view*

Measured by universal standards the history of the United States is indeed brief. But during the brief span of three and one-half centuries of colonial and national history Americans developed traditions and prejudices which created the two worlds of race in modern America. From the time that Africans were brought as indentured servants to the mainland of English America in 1619, the enormous task of rationalizing and justifying the forced labor of peoples on the basis of racial differences was begun; and even after legal slavery was ended, the notion of racial differences persisted as a basis for maintaining segregation and discrimination. At the same time, the effort to establish a more healthy basis for the new world social order was begun, thus launching the continuing battle between the two worlds of

race, on the one hand, and the world of equality and complete human fellowship, on the other.

For a century before the American Revolution the status of Negroes in the English colonies had become fixed at a low point that distinguished them from all other persons who had been held in temporary bondage. By the middle of the eighteenth century, laws governing Negroes denied to them certain basic rights that were conceded to others. They were permitted no independence of thought, no opportunity to improve their minds or their talents or to worship freely, no right to marry and enjoy the conventional family relationships, no right to own or dispose of property, and no protection against miscarriages of justice or cruel and unreasonable punishments. They were outside the pale of the laws that protected ordinary humans. In most places they were to be governed, as the South Carolina code of 1712 expressed it, by special laws “as may restrain the disorders, rapines, and inhumanity to which they are naturally prone and inclined . . . .” A separate world for them had been established by law and custom. Its dimensions and the conduct of its inhabitants were determined by those living in a quite different world.

---

*John Hope Franklin, James B. Duke Professor Emeritus of History at Duke University, has been a Fellow of the American Academy since 1964. This essay appeared in the Fall 1965 issue of “Dædalus.” At the time of its publication, Franklin was professor of American history at the University of Chicago.*

---

© 2005 by the American Academy of Arts & Sciences



By the time that the colonists took up arms against their mother country in order to secure their independence, the world of Negro slavery had become deeply entrenched and the idea of Negro inferiority well established. But the dilemmas inherent in such a situation were a source of constant embarrassment. "It always appeared a most iniquitous scheme to me," Mrs. John Adams wrote her husband in 1774, "to fight ourselves for what we are daily robbing and plundering from those who have as good a right to freedom as we have." There were others who shared her views, but they were unable to wield much influence. When the fighting began General George Washington issued an order to recruiting officers that they were not to enlist "any deserter from the ministerial army, nor any stroller, negro, or vagabond, or person suspected of being an enemy to the liberty of America nor any under eighteen years of age." In classifying Negroes with the dregs of society, traitors, and children, Washington made it clear that Negroes, slave or free, were not to enjoy the high privilege of fighting for political independence. He would change that order later, but only after it became clear that Negroes were enlisting with the "ministerial army" in droves in order to secure their own freedom. In changing his policy if not his views, Washington availed himself of the services of more than 5,000 Negroes who took up arms against England.<sup>1</sup>

Many Americans besides Mrs. Adams were struck by the inconsistency of their stand during the War for Independence, and they were not averse to making

moves to emancipate the slaves. Quakers and other religious groups organized antislavery societies, while numerous individuals manumitted their slaves. In the years following the close of the war most of the states of the East made provisions for the gradual emancipation of slaves. In the South, meanwhile, the antislavery societies were unable to effect programs of statewide emancipation. When the Southerners came to the Constitutional Convention in 1787 they succeeded in winning some representation on the basis of slavery, in securing federal support of the capture and rendition of fugitive slaves, and in preventing the closing of the slave trade before 1808.

Even where the sentiment favoring emancipation was pronounced, it was seldom accompanied by a view that Negroes were the equals of whites and should become a part of one family of Americans. Jefferson, for example, was opposed to slavery; and if he could have had his way, he would have condemned it in the Declaration of Independence. It did not follow, however, that he believed Negroes to be the equals of whites. He did not want to "degrade a whole race of men from the work in the scale of beings which their Creator may *perhaps* have given them . . . I advance it therefore, as a suspicion only, that the blacks, whether originally a distinct race, or made distinct by time and circumstance, are inferior to the whites in the endowment both of body and mind." It is entirely possible that Jefferson's later association with the extraordinarily able Negro astronomer and mathematician, Benjamin Banneker, resulted in some modification of his views. After reading a copy of Banneker's almanac, Jefferson told him that it was "a document to which your whole race had a right for its justifications against the

<sup>1</sup> Benjamin Quarles, *The Negro in the American Revolution* (Chapel Hill: Published for the Institute of Early American History and Culture, Williamsburg, Va., by University of North Carolina Press, 1961), 15 - 18.

doubts which have been entertained of them.”<sup>2</sup>

In communities such as Philadelphia and New York, where the climate was more favorably disposed to the idea of Negro equality than in Jefferson’s Virginia, few concessions were made, except by a limited number of Quakers and their associates. Indeed, the white citizens in the City of Brotherly Love contributed substantially to the perpetuation of two distinct worlds of race. In the 1780s, the white Methodists permitted Negroes to worship with them, provided the Negroes sat in a designated place in the balcony. On one occasion, when the Negro worshippers occupied the front rows of the balcony, from which they had been excluded, the officials pulled them from their knees during prayer and evicted them from the church. Thus, in the early days of the Republic and in the place where the Republic was founded, Negroes had a definite “place” in which they were expected at all times to remain. The white Methodists of New York had much the same attitude toward their Negro fellows. Soon, there were separate Negro churches in these and other communities. Baptists were very much the same. In 1809 thirteen Negro members of a white Baptist church in Philadelphia were dismissed, and they formed a church of their own. Thus, the earliest Negro religious institutions emerged as the result of the rejection by white communicants of their darker fellow worshippers. Soon there would be other institutions – schools, newspapers, benevolent societies – to serve those who lived in a world apart.

Those Americans who conceded the importance of education for Negroes

2 John Hope Franklin, *From Slavery to Freedom: A History of American Negroes* (New York: Knopf, 1956), 156 – 157.

tended to favor some particular type of education that would be in keeping with their lowly station in life. In 1794, for example, the American Convention of Abolition Societies recommended that Negroes be instructed in “those mechanic arts which will keep them most constantly employed and, of course, which will less subject them to idleness and debauchery, and thus prepare them for becoming good citizens of the United States.” When Anthony Benezet, a dedicated Pennsylvania abolitionist, died in 1784 his will provided that on the death of his wife the proceeds of his estate should be used to assist in the establishment of a school for Negroes. In 1787 the school of which Benezet had dreamed was opened in Philadelphia, where the pupils studied reading, writing, arithmetic, plain accounts, and sewing.

Americans who were at all interested in the education of Negroes regarded it as both natural and normal that Negroes should receive their training in separate schools. As early as 1773 Newport, Rhode Island, had a colored school, maintained by a society of benevolent clergymen of the Anglican Church. In 1798 a separate private school for Negro children was established in Boston; and two decades later the city opened its first public primary school for the education of Negro children. Meanwhile, New York had established separate schools, the first one opening its doors in 1790. By 1814 there were several such institutions that were generally designated as the New York African Free Schools.<sup>3</sup>

Thus, in the most liberal section of the country, the general view was that Negroes should be kept out of the mainstream of American life. They were

3 Carter G. Woodson, *The Education of the Negro Prior to 1861* (Washington, D.C.: Associated Publishers, Inc., 1919), 93 – 97.

forced to establish and maintain their own religious institutions, which were frequently followed by the establishment of separate benevolent societies. Likewise, if Negroes were to receive any education, it should be special education provided in separate educational institutions. This principle prevailed in most places in the North throughout the period before the Civil War. In some Massachusetts towns, however, Negroes gained admission to schools that had been maintained for whites. But the School Committee of Boston refused to admit Negroes, arguing that the natural distinction of the races, which “no legislature, no social customs, can efface renders a promiscuous intermingling in the public schools disadvantageous both to them and to the whites.” Separate schools remained in Boston until the Massachusetts legislature in 1855 enacted a law providing that in determining the qualifications of students to be admitted to any public school no distinction should be made on account of the race, color, or religious opinion of the applicant.

Meanwhile, in the Southern states, where the vast majority of the Negroes lived, there were no concessions suggesting equal treatment, even among the most liberal elements. One group that would doubtless have regarded itself as liberal on the race question advocated the deportation of Negroes to Africa, especially those who had become free. Since free Negroes “neither enjoyed the immunities of freemen, nor were they subject to the incapacities of slaves,” their condition and “unconquerable prejudices” prevented amalgamation with whites, one colonization leader argued. There was, therefore, a “peculiar moral fitness” in restoring them to “the land of their fathers.” Men like Henry Clay, Judge Bushrod Washington,

and President James Monroe thought that separation – expatriation – was the best thing for Negroes who were or who would become free.<sup>4</sup>

While the colonization scheme was primarily for Negroes who were already free, it won, for a time, a considerable number of sincere enemies of slavery. From the beginning Negroes were bitterly opposed to it, and only infrequently did certain Negro leaders, such as Dr. Martin Delany and the Reverend Henry M. Turner, support the idea. Colonization, however, retained considerable support in the most responsible quarters. As late as the Civil War, President Lincoln urged Congress to adopt a plan to colonize Negroes, as the only workable solution to the race problem in the United States. Whether the advocates of colonization wanted merely to prevent the contamination of slavery by free Negroes or whether they actually regarded it as the just and honorable thing to do, they represented an important element in the population that rejected the idea of the Negro’s assimilation into the mainstream of American life.

Thus, within fifty years after the Declaration of Independence was written, the institution of slavery, which received only a temporary reversal during the Revolutionary era, contributed greatly to the emergence of the two worlds of race in the United States. The natural rights philosophy appeared to have little effect on those who became committed, more and more, to seeking a rationalization for slavery. The search was apparently so successful that even in areas where slavery was declining, the support for maintaining two worlds of race was strong. Since the Negro church and

<sup>4</sup> P. J. Staudenraus, *The African Colonization Movement, 1816 – 1865* (New York: Columbia University Press, 1961), 22 – 32.

school emerged in Northern communities where slavery was dying, it may be said that the free society believed almost as strongly in racial separation as it did in racial freedom.

The generation preceding the outbreak of the Civil War witnessed the development of a set of defenses of slavery that became the basis for much of the racist doctrine to which some Americans have subscribed from then to the present time. The idea of the inferiority of the Negro enjoyed wide acceptance among Southerners of all classes and among many Northerners. It was an important ingredient in the theory of society promulgated by Southern thinkers and leaders. It was organized into a body of systematic thought by the scientists and social scientists of the South, out of which emerged a doctrine of racial superiority that justified any kind of control over the slave. In 1826 Dr. Thomas Cooper said that he had not the slightest doubt that Negroes were an "inferior variety of the human species; and not capable of the same improvement as the whites." Dr. S. C. Cartwright of the University of Louisiana insisted that the capacities of the Negro adult for learning were equal to those of a white infant; and the Negro could properly perform certain physiological functions only when under the control of white men. Because of the Negro's inferiority, liberty and republican institutions were not only unsuited to his temperament, but actually inimical to his well-being and happiness.

Like racists in other parts of the world, Southerners sought support for their ideology by developing a common bond with the less privileged. The obvious basis was race; and outside the white race there was to be found no favor from God, no honor or respect from man. By

the time that Europeans were reading Gobineau's *Inequality of Races*, Southerners were reading Cartwright's *Slavery in the Light of Ethnology*. In admitting all whites into the pseudo-nobility of race, Cartwright won their enthusiastic support in the struggle to preserve the integrity and honor of the race. Professor Thomas R. Dew of the College of William and Mary comforted the lower-class whites by indicating that they could identify with the most privileged and affluent of the community. In the South, he said, "no white man feels such inferiority of rank as to be unworthy of association with those around him. Color alone is here the badge of distinction, the true mark of aristocracy, and all who are white are equal in spite of the variety of occupation."<sup>5</sup>

Many Northerners were not without their own racist views and policies in the turbulent decades before the Civil War. Some, as Professor Louis Filler has observed, displayed a hatred of Negroes that gave them a sense of superiority and an outlet for their frustrations. Others cared nothing one way or the other about Negroes and demanded only that they be kept separate.<sup>6</sup> Even some of the abolitionists themselves were ambivalent on the question of Negro equality. More than one antislavery society was agitated by the suggestion that Negroes be invited to join. Some members thought it reasonable for them to attend, but not to be put on an "equality with ourselves." The New York abolitionist, Lewis Tappan, admitted "that when the subject of acting out our profound prin-

5 John Hope Franklin, *The Militant South, 1800 - 1861* (Cambridge, Mass.: Belknap Press of Harvard University Press, 1956), 83 - 86.

6 Louis Filler, *The Crusade Against Slavery, 1830 - 1860* (New York: Harper, 1960), 142 - 145.

ciples in treating men irrespective of color is discussed heat is always produced.”<sup>7</sup>

In the final years before the beginning of the Civil War, the view that the Negro was different, even inferior, was widely held in the United States. Leaders in both major parties subscribed to the view, while the more extreme racists deplored any suggestion that the Negro could ever prosper as a free man. At Peoria, Illinois, in October 1854, Abraham Lincoln asked what stand the opponents of slavery should take regarding Negroes. “Free them, and make them politically and socially, our equals? My own feelings not admit of this; and if mine would, we well know that those of the great mass of white people will not. Whether this feeling accords with justice and sound judgment, is not the sole question, if indeed, it is any part of it. A universal feeling, whether well or ill founded, cannot be safely disregarded. We cannot, then, make them equals.”

The Lincoln statement was forthright, and it doubtless represented the views of most Americans in the 1850s. Most of those who heard him or read his speech were of the same opinion as he. In later years, the Peoria pronouncement would be used by those who sought to detract from Lincoln’s reputation as a champion of the rights of the Negro. In 1964, the White Citizens’ Councils reprinted portions of the speech in large advertisements in the daily press and insisted that Lincoln shared their views on the desirability of maintaining two distinct worlds of race.

Lincoln could not have overcome the nation’s strong predisposition toward racial separation if he had tried. And he did not try very hard. When he called

for the enlistment of Negro troops, after issuing the Emancipation Proclamation, he was content not only to set Negroes apart in a unit called “U.S. Colored Troops,” but also to have Negro privates receive \$10 per month including clothing, while whites of the same rank received \$13 per month plus clothing. Only the stubborn refusal of many Negro troops to accept discriminatory pay finally forced Congress to equalize compensation for white and Negro soldiers.<sup>8</sup> The fight for union that became also a fight for freedom never became a fight for equality or for the creation of one racial world.

The Lincoln and Johnson plans for settling the problems of peace and freedom never seriously touched on the concomitant problem of equality. To be sure, in 1864 President Lincoln privately raised with the governor of Louisiana the question of the franchise for a limited number of Negroes, but when the governor ignored the question the President let the matter drop. Johnson raised a similar question in 1866, but he admitted that it was merely to frustrate the design of radical reformers who sought a wider franchise for Negroes. During the two years following Appomattox Southern leaders gave not the slightest consideration to permitting any Negroes, regardless of their service to the Union or their education or their property, to share in the political life of their communities. Not only did every Southern state refuse to permit Negroes to vote, but they also refused to provide Negroes with any of the educational opportunities that they were providing for the whites.

The early practice of political disfranchisement and of exclusion from public educational facilities helped to deter-

7 Leon F. Litwack, *North of Slavery; The Negro in the Free States, 1790 – 1860* (Chicago: University of Chicago Press, 1961), 216 – 217.

8 Benjamin Quarles, *The Negro in the Civil War* (Boston: Little, Brown, 1953), 200.



mine subsequent policies that the South adopted regarding Negroes. While a few leaders raised their voices against these policies and practices, it was Negroes themselves who made the most eloquent attacks on such discriminations. As early as May 1865, a group of North Carolina Negroes told President Johnson that some of them had been soldiers and were doing everything possible to learn how to discharge the higher duties of citizenship. "It seems to us that men who are willing on the field of battle to carry the muskets of the Republic, in the days of peace ought to be permitted to carry the ballots; and certainly we cannot understand the justice of denying the elective franchise to men who have been fighting *for* the country, while it is freely given to men who have just returned from *four years* fighting against it." Such pleas fell on deaf ears, however; and it was not until 1867, when Congress was sufficiently outraged by the inhuman black codes, widespread discriminations in the South, and unspeakable forms of violence against Negroes, that new federal legislation sought to correct the evils of the first period of Reconstruction.

The period that we know as Radical Reconstruction had no significant or permanent effect on the status of the Negro in American life. For a period of time, varying from one year to fifteen or twenty years, some Negroes enjoyed the privileges of voting. They gained political ascendancy in a very few communities only temporarily, and they never even began to achieve the status of a ruling class. They made no meaningful steps toward economic independence or even stability; and in no time at all, because of the pressures of the local community and the neglect of the federal government, they were brought under the complete economic subservience of

the old ruling class. Organizations such as the Ku Klux Klan were committed to violent action to keep Negroes "in their place" and, having gained respectability through sponsorship by Confederate generals and the like, they proceeded to wreak havoc in the name of white supremacy and protection of white womanhood.<sup>9</sup>

Meanwhile, various forms of segregation and discrimination, developed in the years before the Civil War in order to degrade the half million free Negroes in the United States, were now applied to the four million Negroes who had become free in 1865. Already the churches and the military were completely segregated. For the most part the schools, even in the North, were separate. In the South segregated schools persisted, even in the places where the radicals made a halfhearted attempt to desegregate them. In 1875 Congress enacted a Civil Rights Act to guarantee the enjoyment of equal rights in carriers and all places of public accommodation and amusement. Even before it became law Northern philanthropists succeeded in forcing the deletion of the provision calling for desegregated schools. Soon, because of the massive resistance in the North as well as in the South and the indifferent manner in which the federal government enforced the law, it soon became a dead letter everywhere. When it was declared unconstitutional by the Supreme Court in 1883, there was universal rejoicing, except among the Negroes, one of whom declared that they had been "baptized in ice water."

Neither the Civil War nor the era of Reconstruction made any significant step toward the permanent elimination of racial barriers. The radicals of the

9 John Hope Franklin, *Reconstruction: After the Civil War* (Chicago: University of Chicago Press, 1961), 154–158.

post-Civil War years came no closer to the creation of one racial world than the patriots of the Revolutionary years. When Negroes were, for the first time, enrolled in the standing army of the United States, they were placed in separate Negro units. Most of the liberals of the Reconstruction era called for and worked for separate schools for Negroes. Nowhere was there any extensive effort to involve Negroes in the churches and other social institutions of the dominant group. Whatever remained of the old abolitionist fervor, which can hardly be described as unequivocal on the question of true racial equality, was rapidly disappearing. In its place were the sentiments of the businessmen who wanted peace at any price. Those having common railroad interests or cropmarketing interests or investment interests could and did extend their hands across sectional lines and joined in the task of working together for the common good. In such an atmosphere the practice was to accept the realities of two separate worlds of race. Some even subscribed to the view that there were significant economic advantages in maintaining the two worlds of race.

The post-Reconstruction years witnessed a steady deterioration in the status of Negro Americans. These were the years that Professor Rayford Logan has called the “nadir” of the Negro in American life and thought. They were the years when Americans, weary of the crusade that had, for the most part, ended with the outbreak of the Civil War, displayed almost no interest in helping the Negro to achieve equality. The social Darwinists decried the very notion of equality for Negroes, arguing that the lowly place they occupied was natural and normal. The leading literary journals vied with each other in describ-

ing Negroes as lazy, idle, improvident, immoral, and criminal.<sup>10</sup> Thomas Dixon’s novels, *The Klansman* and *The Leopard’s Spots*, and D. W. Griffith’s motion picture, “The Birth of A Nation,” helped to give Americans a view of the Negro’s role in American history that “proved” that he was unfit for citizenship, to say nothing of equality. The dictum of William Graham Sumner and his followers that “stateways cannot change folkways” convinced many Americans that legislating equality and creating one great society where race was irrelevant was out of the question.

But many Americans believed that they *could* legislate inequality; and they proceeded to do precisely that. Beginning in 1890, one Southern state after another revised the suffrage provisions of its constitution in a manner that made it virtually impossible for Negroes to qualify to vote. The new literacy and “understanding” provisions permitted local registrars to disqualify Negroes while permitting white citizens to qualify. Several states, including Louisiana, North Carolina, and Oklahoma, inserted “grandfather clauses” in their constitutions in order to permit persons, who could not otherwise qualify, to vote if their fathers or grandfathers could vote in 1866. (This was such a flagrant discrimination against Negroes, whose ancestors could not vote in 1866, that the United States Supreme Court in 1915 declared the “grandfather clause” unconstitutional.) Then came the Democratic white primary in 1900 that made it impossible for Negroes to participate in local elections in the South, where, by this time, only the Democratic party had any appreciable strength. (After more than a generation of assaults on it,

10 Rayford W. Logan, *The Negro in American Life and Thought: The Nadir, 1877–1901* (New York: Dial Press, 1954), 239–274.

the white primary was finally declared unconstitutional in 1944.)

Inequality was legislated in still another way. Beginning in the 1880s, many states, especially but not exclusively in the South, enacted statutes designed to separate the races. After the Civil Rights Act was declared unconstitutional in 1883 state legislatures were emboldened to enact numerous segregation statutes. When the United States Supreme Court, in the case of *Plessy v. Ferguson*, set forth the “separate but equal” doctrine in 1896, the decision provided a new stimulus for laws to separate the races and, of course, to discriminate against Negroes. In time, Negroes and whites were separated in the use of schools, churches, cemeteries, drinking fountains, restaurants, and all places of public accommodation and amusement. One state enacted a law providing for the separate warehousing of books used by white and Negro children. Another required the telephone company to provide separate telephone booths for white and Negro customers. In most communities housing was racially separated by law or practice.<sup>11</sup>

Where there was no legislation requiring segregation, local practices filled the void. Contradictions and inconsistencies seemed not to disturb those who sought to maintain racial distinctions at all costs. It mattered not that one drive-in snack bar served Negroes only on the inside, while its competitor across the street served Negroes only on the outside. Both were committed to making racial distinctions; and in communities where practices and mores had the force of law, the distinction was everything. Such practices were greatly strength-

11 John Hope Franklin, “History of Racial Segregation in the United States,” *Annals of the Academy of Political and Social Science* 304 (March 1956): 1–9.

ened when, in 1913, the federal government adopted policies that segregated the races in its offices as well as in its eating and restroom facilities.

By the time of World War I, Negroes and whites in the South and in parts of the North lived in separate worlds, and the apparatus for keeping the worlds separate was elaborate and complex. Negroes were segregated by law in the public schools of the Southern states, while those in the Northern ghettos were sent to predominantly Negro schools, except where their numbers were insufficient. Scores of Negro newspapers sprang up to provide news of Negroes that the white press consistently ignored. Negroes were as unwanted in the white churches as they had been in the late eighteenth century; and Negro churches of virtually every denomination were the answer for a people who had accepted the white man’s religion even as the white man rejected his religious fellowship.

Taking note of the fact that they had been omitted from any serious consideration by the white historians, Negroes began in earnest to write the history of their own experiences as Americans. There had been Negro historians before the Civil War, but none of them had challenged the white historians’ efforts to relegate Negroes to a separate, degraded world. In 1882, however, George Washington Williams published his *History of the Negro Race in America* in order to “give the world more correct ideas about the colored people.” He wrote, he said, not “as a partisan apologist, but from a love for the truth of history.”<sup>12</sup> Soon there were other historical works by Negroes describing their progress and their contributions and arguing that

12 George W. Williams, *History of the Negro Race in America from 1619 to 1880* (New York: G. P. Putnam’s Sons, 1882), x.

they deserved to be received into the full fellowship of American citizens.

It was in these post-Reconstruction years that some of the most vigorous efforts were made to destroy the two worlds of race. The desperate pleas of Negro historians were merely the more articulate attempts of Negroes to gain complete acceptance in American life. Scores of Negro organizations joined in the struggle to gain protection and recognition of their rights and to eliminate the more sordid practices that characterized the treatment of the Negro world by the white world. Unhappily, the small number of whites who were committed to racial equality dwindled in the post-Reconstruction years, while government at every level showed no interest in eliminating racial separatism. It seemed that Negro voices were indeed crying in the wilderness, but they carried on their attempts to be heard. In 1890 Negroes from twenty-one states and the District of Columbia met in Chicago and organized the Afro-American League of the United States. They called for more equitable distribution of school funds, fair and impartial trial for accused Negroes, resistance "by all legal and reasonable means" to mob and lynch law, and enjoyment of the franchise by all qualified voters. When a group of young Negro intellectuals, led by W. E. B. Du Bois, met at Niagara Falls, Ontario, in 1905, they made a similar call as they launched their Niagara Movement.

However eloquent their pleas, Negroes alone could make no successful assault on the two worlds of race. They needed help – a great deal of help. It was the bloody race riots in the early years of the twentieth century that shocked civic-minded and socially conscious whites into answering the Negro's pleas for support. Some whites began to take the view that the existence of two societies whose

distinction was based solely on race was inimical to the best interests of the entire nation. Soon, they were taking the initiative and in 1909 organized the National Association for the Advancement of Colored People. They assisted the following year in establishing the National Urban League. White attorneys began to stand with Negroes before the United States Supreme Court to challenge the "grandfather clause," local segregation ordinances, and flagrant miscarriages of justice in which Negroes were the victims. The patterns of attack developed during these years were to become invaluable later. Legal action was soon supplemented by picketing, demonstrating, and boycotting, with telling effect particularly in selected Northern communities.<sup>13</sup>

The two world wars had a profound effect on the status of Negroes in the United States and did much to mount the attack on the two worlds of race. The decade of World War I witnessed a very significant migration of Negroes. They went in large numbers – perhaps a half million – from the rural areas of the South to the towns and cities of the South and North. They were especially attracted to the industrial centers of the North. By the thousands they poured into Pittsburgh, Cleveland, and Chicago. Although many were unable to secure employment, others were successful and achieved a standard of living they could not have imagined only a few years earlier. Northern communities were not altogether friendly and hospitable to the newcomers, but the opportunities for education and the enjoyment of political self-respect were the greatest they had ever seen. Many of them felt that they were entirely justified in their re-

13 Franklin, *From Slavery to Freedom*, 437 – 443.

newed hope that the war would bring about a complete merger of the two worlds of race.

Those who held such high hopes, however, were naive in the extreme. Already the Ku Klux Klan was being revived – this time in the North as well as in the South. Its leaders were determined to develop a broad program to unite “native-born white Christians for concerted action in the preservation of American institutions and the supremacy of the white race.” By the time that the war was over, the Klan was in a position to make capital of the racial animosities that had developed during the conflict itself. Racial conflicts had broken out in many places during the war; and before the conference at Versailles was over race riots in the United States had brought about what can accurately be described as the “long, hot summer” of 1919.

If anything, the military operations which aimed to save the world for democracy merely fixed more permanently the racial separation in the United States. Negro soldiers not only constituted entirely separate fighting units in the United States Army, but, once overseas, were assigned to fighting units with the French Army. Negroes who sought service with the United States Marines or the Air Force were rejected, while the Navy relegated them to menial duties. The reaction of many Negroes was bitter, but most of the leaders, including Du Bois, counseled patience and loyalty. They continued to hope that their show of patriotism would win for them a secure place of acceptance as Americans.

Few Negro Americans could have anticipated the wholesale rejection they experienced at the conclusion of World War I. Returning Negro soldiers were lynched by hanging and burning, even while still in their military uniforms.

The Klan warned Negroes that they must respect the rights of the white race “in whose country they are permitted to reside.” Racial conflicts swept the country, and neither federal nor state governments seemed interested in effective intervention. The worlds of race were growing further apart in the postwar decade. Nothing indicated this more clearly than the growth of the Universal Negro Improvement Association, led by Marcus Garvey. From a mere handful of members at the end of the war, the Garvey movement rapidly became the largest secular Negro group ever organized in the United States. Although few Negroes were interested in settling in Africa – the expressed aim of Garvey – they joined the movement by the hundreds of thousands to indicate their resentment of the racial duality that seemed to them to be the central feature of the American social order.<sup>14</sup>

More realistic and hardheaded were the Negroes who were more determined than ever to engage in the most desperate fight of their lives to destroy racism in the United States. As the editor of the *Crisis* said in 1919, “We return from fighting. We return fighting. Make way for Democracy! We saved it in France, and by the Great Jehovah, we will save it in the U.S.A., or know the reason why.” This was the spirit of what Alain Locke called “The New Negro.” He fought the Democratic white primary, made war on the whites who consigned him to the ghetto, attacked racial discrimination in employment, and pressed for legislation to protect his rights. If he was seldom successful during the postwar decade and the depression, he made it quite clear that he was unalterably op-

<sup>14</sup> Edmund David Cronon, *Black Moses, The Story of Marcus Garvey and the Universal Negro Improvement Association* (Madison: University of Wisconsin Press, 1955), 202 – 206.



posed to the un-American character of the two worlds of race.

Hope for a new assault on racism was kindled by some of the New Deal policies of Franklin D. Roosevelt. As members of the economically disadvantaged group, Negroes benefited from relief and recovery legislation. Most of it, however, recognized the existence of the two worlds of race and accommodated itself to it. Frequently bread lines and soup kitchens were separated on the basis of race. There was segregation in the employment services, while many new agencies recognized and bowed to Jim Crow. Whenever agencies, such as the Farm Security Administration, fought segregation and sought to deal with people on the basis of their needs rather than race they came under the withering fire of the racist critics and seldom escaped alive. Winds of change, however slight, were discernible, and nowhere was this in greater evidence than in the new labor unions. Groups like the Congress of Industrial Organizations, encouraged by the support of the Wagner Labor Relations Act, began to look at manpower resources as a whole and to attack the old racial policies that viewed labor in terms of race.

As World War II approached, Negroes schooled in the experiences of the 1920s and 1930s were unwilling to see the fight against Nazism carried on in the context of an American racist ideology. Some white Americans were likewise uncomfortable in the role of freeing Europe of a racism which still permeated the United States; but it was the Negroes who dramatized American inconsistency by demanding an end to discrimination in employment in defense industries. By threatening to march on Washington in 1941 they forced the President to issue an order forbidding such discrimination. The opposition was loud and

strong. Some state governors denounced the order, and some manufacturers skillfully evaded it. But it was a significant step toward the elimination of the two worlds.

During World War II the assault on racism continued. Negroes, more than a million of whom were enlisted in the armed services, bitterly fought discrimination and segregation. The armed services were, for the most part, two quite distinct racial worlds. Some Negro units had white officers, and much of the officer training was desegregated. But it was not until the final months of the war that a deliberate experiment was undertaken to involve Negro and white enlisted men in the same fighting unit. With the success of the experiment and with the warm glow of victory over Nazism as a backdrop, there was greater inclination to recognize the absurdity of maintaining a racially separate military force to protect the freedoms of the country.<sup>15</sup>

During the war there began the greatest migration in the history of Negro Americans. Hundreds of thousands left the South for the industrial centers of the North and West. In those places they met hostility, but they also secured employment in aviation plants, automobile factories, steel mills, and numerous other industries. Their difficulties persisted as they faced problems of housing and adjustment. But they continued to move out of the South in such large numbers that by 1965 one-third of the twenty million Negroes in the United States lived in twelve metropolitan centers of the North and West. The ramifications of such large-scale migration were numerous. The concentration of Negroes in communities where they suffered no

<sup>15</sup> Lee Nichols, *Breakthrough on the Color Front* (New York: Random House, 1954), 221 – 226.

political disabilities placed in their hands an enormous amount of political power. Consequently, some of them went to the legislatures, to Congress, and to positions on the judiciary. In turn, this won for them political respect as well as legislation that greatly strengthened their position as citizens.

Following World War II there was a marked acceleration in the war against the two worlds of race in the United States. In 1944 the Supreme Court ruled against segregation in interstate transportation, and three years later it wrote the final chapter in the war against the Democratic white primary. In 1947 the President's Committee on Civil Rights called for the "elimination of segregation, based on race, color, creed, or national origin, from American life."<sup>16</sup> In the following year President Truman asked Congress to establish a permanent Fair Employment Practices Commission. At the same time he took steps to eliminate segregation in the armed services. These moves on the part of the judicial and executive branches of the federal government by no means destroyed the two worlds of race, but they created a more healthy climate in which the government and others could launch an attack on racial separatism.

The attack was greatly strengthened by the new position of world leadership that the United States assumed at the close of the war. Critics of the United States were quick to point to the inconsistencies of an American position that spoke against racism abroad and countenanced it at home. New nations, brown and black, seemed reluctant to follow

<sup>16</sup> *To Secure These Rights, The Report of the President's Committee on Civil Rights* (New York: Simon and Schuster, 1947), 166.

the lead of a country that adhered to its policy of maintaining two worlds of race – the one identified with the old colonial ruling powers and the other with the colonies now emerging as independent nations. Responsible leaders in the United States saw the weakness of their position, and some of them made new moves to repair it.

Civic and religious groups, some labor organizations, and many individuals from the white community began to join in the effort to destroy segregation and discrimination in American life. There was no danger, after World War II, that Negroes would ever again stand alone in their fight. The older interracial organizations continued, but they were joined by new ones. In addition to the numerous groups that included racial equality in their overall programs, there were others that made the creation of one racial world their principal objective. Among them were the Congress of Racial Equality, the Southern Christian Leadership Conference, and the Student Non-Violent Coordinating Committee. Those in existence in the 1950s supported the court action that brought about the decision against segregated schools. The more recent ones have taken the lead in pressing for new legislation and in developing new techniques to be used in the war on segregation.

The most powerful direct force in the maintenance of the two worlds of race has been the state and its political subdivisions. In states and communities where racial separation and discrimination are basic to the way of life, the elected officials invariably pledge themselves to the perpetuation of the duality. Indeed, candidates frequently vie with one another in their effort to occupy the most extreme segregationist position possible on the race question. Ap-

pointed officials, including the constabulary and, not infrequently, the teachers and school administrators, become auxiliary guardians of the system of racial separation. In such communities Negroes occupy no policy-making positions, exercise no influence over the determination of policy, and are seldom even on the police force. State and local resources, including tax funds, are at the disposal of those who guard the system of segregation and discrimination; and such funds are used to enforce customs as well as laws and to disseminate information in support of the system.

The white community itself acts as a guardian of the segregated system. Schooled in the specious arguments that assert the supremacy of the white race and fearful that a destruction of the system would be harmful to their own position, they not only "go along" with it but, in many cases, enthusiastically support it. Community sanctions are so powerful, moreover, that the independent citizen who would defy the established order would find himself not only ostracized but, worse, the target of economic and political reprisals.

Within the community many self-appointed guardians of white supremacy have emerged at various times. After the Civil War and after World War I it was the Ku Klux Klan, which has shown surprising strength in recent years. After the desegregation decision of the Supreme Court in 1954 it was the White Citizens' Council, which one Southern editor has called the "uptown Klu Klux Klan." From time to time since 1865, it has been the political demagogue, who has not only made capital by urging his election as a sure way to maintain the system but has also encouraged the less responsible elements of the community to take the law into their own hands.

Violence, so much a part of American history and particularly of Southern history, has been an important factor in maintaining the two worlds of race. Intimidation, terror, lynchings, and riots have, in succession, been the handmaidens of political entities whose officials have been unwilling or unable to put an end to it. Violence drove Negroes from the polls in the 1870s and has kept them away in droves since that time. Lynchings, the spectacular rope and faggot kind or the quiet kind of merely "doing away" with some insubordinate Negro, have served their special purpose in terrorizing whole communities of Negroes. Riots, confined to no section of the country, have demonstrated how explosive the racial situation can be in urban communities burdened with the strain of racial strife.

The heavy hand of history has been a powerful force in the maintenance of a segregated society and, conversely, in the resistance to change. Americans, especially Southerners whose devotion to the past is unmatched by that of any others, have summoned history to support their arguments that age-old practices and institutions cannot be changed overnight, that social practices cannot be changed by legislation. Southerners have argued that desegregation would break down long-established customs and bring instability to a social order that, if left alone, would have no serious racial or social disorders. After all, Southern whites "know" Negroes; and their knowledge has come from many generations of intimate association and observation, they insist.

White Southerners have also summoned history to support them in their resistance to federal legislation designed to secure the civil rights of Negroes. At every level – in local groups, state governments, and in Congress – white

Southerners have asserted that federal civil rights legislation is an attempt to turn back the clock to the Reconstruction era, when federal intervention, they claim, imposed a harsh and unjust peace.<sup>17</sup> To make effective their argument, they use such emotion-laden phrases as “military occupation,” “Negro rule,” and “blackout of honest government.” Americans other than Southerners have been frightened by the Southerners’ claim that civil rights for Negroes would cause a return to the “evils” of Reconstruction. Insecure in their own knowledge of history, they have accepted the erroneous assertions about the “disaster” of radical rule after the Civil War and the vengeful punishment meted out to the South by the Negro and his white allies. Regardless of the merits of these arguments that seem specious on the face of them – to say nothing of their historical inaccuracy – they have served as effective brakes on the drive to destroy the two worlds of race.

One suspects, however, that racial bigotry has become more expensive in recent years. It is not so easy now as it once was to make political capital out of the race problem, even in the deep South. Local citizens – farmers, laborers, manufacturers – have become a bit weary of the promises of the demagogue that he will preserve the integrity of the races if he is, at the same time, unable to persuade investors to build factories and bring capital to their communities. Some Southerners, dependent on tourists, are not certain that their vaunted racial pride is so dear, if it keeps visitors away and brings depression to their economy. The cities that see them-

17 John Hope Franklin, “As For Our History,” in Charles G. Sellers, ed., *The Southerner as American* (Chapel Hill: University of North Carolina Press, 1960), 1–18.

selves bypassed by a prospective manufacturer because of their reputation in the field of race relations might have some sober second thoughts about the importance of maintaining their two worlds. In a word, the economics of segregation and discrimination is forcing, in some quarters, a reconsideration of the problem.

It must be added that the existence of the two worlds of race has created forces that cause some Negroes to seek its perpetuation. Some Negro institutions, the product of a dual society, have vested interests in the perpetuation of that society. And Negroes who fear the destruction of their own institutions by desegregation are encouraged by white racists to fight for their maintenance. Even where Negroes have a desire to maintain their institutions because of their honest commitment to the merits of cultural pluralism, the desire becomes a strident struggle for survival in the context of racist forces that seek with a vengeance to destroy such institutions. The firing of a few hundred Negro schoolteachers by a zealous, racially oriented school board forces some second thoughts on the part of the Negroes regarding the merits of desegregation.

The drive to destroy the two worlds of race has reached a new, dramatic, and somewhat explosive stage in recent years. The forces arrayed in behalf of maintaining these two worlds have been subjected to ceaseless and powerful attacks by the increasing numbers committed to the elimination of racism in American life. Through techniques of demonstrating, picketing, sitting-in, and boycotting they have not only harassed their foes but marshaled their forces. Realizing that another ingredient was needed, they have pressed for new and better laws and the active sup-

port of government. At the local and state levels they began to secure legislation in the 1940s to guarantee the civil rights of all, eliminate discrimination in employment, and achieve decent public and private housing for all.

While it is not possible to measure the influence of public opinion in the drive for equality, it can hardly be denied that over the past five or six years public opinion has shown a marked shift toward vigorous support of the civil rights movement. This can be seen in the manner in which the mass-circulation magazines as well as influential newspapers, even in the South, have stepped up their support of specific measures that have as their objective the elimination of at least the worst features of racism. The discussion of the problem of race over radio and television and the use of these media in reporting newsworthy and dramatic events in the world of race undoubtedly have had some impact. If such activities have not brought about the enactment of civil rights legislation, they have doubtless stimulated the public discussion that culminated in such legislation.

The models of city ordinances and state laws and the increased political influence of civil rights advocates stimulated new action on the federal level. Civil rights acts were passed in 1957, 1960, and 1964 – after almost complete federal inactivity in this sphere for more than three-quarters of a century. Strong leadership on the part of the executive and favorable judicial interpretations of old as well as new laws have made it clear that the war against the two worlds of race now enjoys the sanction of the law and its interpreters. In many respects this constitutes the most significant development in the struggle against racism in the present century.

The reading of American history over the past two centuries impresses one

with the fact that ambivalence on the crucial question of equality has persisted almost from the beginning. If the term “equal rights for all” has not always meant what it appeared to mean, the inconsistencies and the paradoxes have become increasingly apparent. This is not to say that the view that “equal rights for some” has disappeared or has even ceased to be a threat to the concept of real equality. It is to say, however, that the voices supporting inequality, while no less strident, have been significantly weakened by the very force of the numbers and elements now seeking to eliminate the two worlds of race.



Jill K. Conway

*Politics, pedagogy & gender*

In the mid-nineteenth century the public education system of the United States drew its corps of teachers from the nation's population of young women. In contrast, European public education remained a male-dominated enterprise until well into the twentieth century. Traditionally, the United States' early and extensive recruitment of female teachers has been interpreted as a sign of enlightened attitudes about women and their place in society. Horace Mann's innovative Massachusetts normal schools, which trained young women to be teachers, are customarily cited as examples of feminism in action. So, until recently, was the career of Catherine Beecher, the archetypal proselytizer for the female teaching profession. The development of a public elementary school system before the Civil War and

the extension of that system through the establishment of secondary schools in the last quarter of the nineteenth century provide a happy ending to the traditional story of the establishment of the first "women's" profession.<sup>1</sup>

Underlying this popular history of women in teaching is the assumption that access to new work opportunities has the same meaning for everyone. If we stop to ask what gender meant for the nineteenth-century founders of American public education, however, the story takes on new levels of meaning. Some of its themes speak directly to our educational dilemmas today. Its interest lies not in the sex of the teachers who staffed America's one-room schools but in the political and psychological images that men and women held regarding the gender of those teachers. The story of women's opportunities to enter teaching as a respectable occupation for single women outside the home is a case study in the meaning of access. Examination of the case of women teachers' recruitment in the mid-nineteenth century should make us rethink the incremental model of change that is presumed to characterize the liberal state.

---

*Jill K. Conway, chairperson of Lend Lease Corporation, has been a Fellow of the American Academy since 1991. When this essay appeared in the Fall 1987 issue of "Daedalus," she was a visiting scholar in the Program in Science, Technology and Society at the Massachusetts Institute of Technology. Conway was president of Smith College from 1975 to 1985.*

---

© 2005 by the American Academy of Arts & Sciences

1 Horace Mann, *Eleventh Annual Report* (Massachusetts Normal Schools, 1845), 24.

The number of women involved in this recruitment is certainly striking. By 1848 women greatly outnumbered men as annual entrants to the teaching profession; in absolute numbers their predominance was established. In that year 2,424 men taught in the public (or common) schools of America beside 5,510 women.<sup>2</sup> During the 1850s the same pattern was replicated in the Midwest. After 1864 one of the impositions of the victorious North on the southern states during Reconstruction was the establishment of a predominantly female cadre of elementary school teachers. In the last three decades of the nineteenth century the same pattern emerged in the public high schools. By 1890, 65 percent of all teachers in the United States were women. Members of the new female profession were remarkably youthful, averaging from twenty-one to twenty-five years of age in different regions of the country.

Popular attitudes encouraged single women to become teachers but discouraged their presence in the schools once they married. The country's teachers were predominantly daughters of the native-born, from rural families. In comparison with European teachers, American teachers were not well educated. As late as the 1930s only 12 percent of elementary teachers in the United States had earned bachelor's degrees.<sup>3</sup> In the nineteenth century many entrants to the profession had not even completed high school. Because so many teachers were drawn from rural farm families, most had not traveled more than 100 miles from their place of birth. Their experi-

ence of high culture was minimal. Surveys carried out at the turn of the century recorded that most teachers had never seen reproductions of works of art during their own schooling. As adults their only reading was an occasional novel and the standard popular magazines of the day. To compensate for these deficiencies, the normal schools offered teaching programs that were largely remedial.<sup>4</sup>

The woman teacher, whether rural or urban, earned about 60 percent of the salary paid to men in the same school system. Around 1900 the average woman teacher's salary was \$350 per year. Higher earnings were available to women in the textile industry and in most other industrial settings. In some states mechanics and clerks earned twice the annual wages of male teachers, whose earnings were more than a third higher than those of their female counterparts. The universal custom of "boarding out" was a major factor in depressing the level of teachers' earnings: nineteenth-century school districts held down the cost of elementary schools by housing teachers in rotation with families whose children were currently school pupils. This dubious hospitality was motivated partly by economic considerations and partly by the prevailing sentiment that young single women should not be allowed to live outside a family setting. The school district's room and board carried with it a censorious social control that young single women could resist only at their peril. In short, the young teacher's social status was marginal.<sup>5</sup>

2 Redding R. Sugg, Jr., *Motherteacher: The Feminization of American Education* (Charlottesville: University Press of Virginia, 1978), 37.

3 Lindley J. Stiles, ed., *The Teacher's Role in American Society* (New York: Harper & Row, 1957), 279.

4 Lotus D. Coffman, *The Social Composition of the Teaching Population* (New York: Bureau of Publications, Teacher's College, Columbia University, 1911), 550.

5 *Ibid.*, 550. See also Myra H. Strober and Audri Gordon Lanford, "The Feminization of Public

This marginality was not borne for long; rates of turnover were very high. Most women elementary teachers taught for only three or four years. Although 90 percent of the elementary instructors by the 1920s were women, their rapid turnover meant that they did not develop as school leaders or as curriculum planners.<sup>6</sup> Men did not remain teachers for long either; they did not form strong bonds to the occupation of teacher as they did to the professions of medicine and engineering. Yet male teachers were seven times more likely to become school administrators than their female colleagues. Despite the social changes that have raised women's work aspirations in recent decades, these early trends have continued unaltered. Today men hold 99.4 percent of all school superintendencies. The only area of school administration in which women predominate is librarianship. Clearly gender shapes one's status within the teaching profession, even though teaching has traditionally been singled out for its supposed hospitality to women. What, then, are we to make of women's early access to teaching in the United States? What values shaped the establishment of the common schools in America, and what was the operative significance of ideas about gender in that process? To paraphrase William James, what was the meaning of the ideas being translated into action when people like Horace Mann began to recruit women for teacher training?

If we look at the political debates that preceded the establishment of the public

---

School Teaching: A Cross-Sectional Analysis, 1850–1880," *Signs* 11 (2) (1986): 212–235, and Willard S. Ellsbree, *The American Teacher: Evolution of a Profession in a Democracy* (New York: American Book Company, 1933), 281.

6 Ellsbree, *The American Teacher*, 206.

education system in the 1830s and 1840s, we see that political forces divided over the level of intellectual aspiration desired as an outcome of state-supported education and over the place of elites of education and talent within the young republic. One thing that united Jefferson and his Federalist opponents was the value they saw in an educated elite drawn from the best talent of their new society. Jefferson wanted his elite to be democratically recruited, its education publicly supported; he expected the result to be the highest intellectual achievement.

One of the major shifts of value in the Jacksonian era was the rejection of the idea of a socially valuable elite formed by education and high culture. Instead, Americans of that era favored a popular education that was broadly accessible and limited in its intellectual goals. As Michael Katz has shown in his study of the development of public education in Massachusetts, some of the old Federalist elites found popular education attractive not so much as a means of training the mind but as a way of providing instruction in behavior.<sup>7</sup> Many New England moralists who sought to control the excesses of frontier behavior thought that this goal might be achieved through the common schools. Their intellectual aspirations for the students who were expected to attend these schools were minimal.

We know from recent studies of the legislative decisions approving the establishment of the common schools that Federalists and Jacksonians alike sought to develop public education as inexpensively as possible. The compromise that led to agreement on tax-supported public education combined the older Jeffer-

7 Michael Katz, *The Irony of Early School Reform* (Cambridge, Mass.: Harvard University Press, 1968).

sonian ideal of wide access to public education with Federalist and Jacksonian concerns for limited education at minimal cost to the taxpayer. The goal of cost containment made the recruitment of women completely logical because all parties to the educational debate agreed that women lacked acquisitive drives and would serve at subsistence salaries. The potentially explosive conflict over the intellectual goals of public education could also be avoided by choosing women as teachers. Their access to education was slight, so that male control over the normal schools that trained teachers insured control over the content of the curriculum. Furthermore, beliefs about the female temperament promised that the pedagogical style of women teachers would be emotional and value-oriented rather than rational and critical. Thus neither Jacksonians nor Federalists needed to make resolution of their conflicts over the goals of education an explicit part of their political agenda.<sup>8</sup> The resolution of fundamental contradictions about a strategic institution for the evolving society could safely be postponed as long as women teachers presented no threat to the objectives of low cost and strictly utilitarian public education.

The following three quotations demonstrate gender stereotyping at work in the public-education policy discussions of late nineteenth-century legislators and public officials. Each of the speakers favored the recruitment of women teachers. These passages illustrate the important components of the gender ideology accepted by all parties to the dispute over the goals of education.

[Women] manifest a livelier interest, more contentment in the work, have al-

together superior success in managing and instructing young children, and I know of instances, where by the silken cord of affection, have led many a stubborn will, and wild ungoverned impulse, into habits of obedience and study even in the large winter schools (Henry Barnard, *Second Annual Report* [Connecticut School, 1840], 27 – 28).

[Women] are endowed by nature with stronger parental impulses, and this makes the society of children delightful, and turns duty into a pleasure. Their minds are less withdrawn from their employment, by the active scenes of life; and they are less intent and scheming for future honors and emoluments. As a class, they never look forward, as young men almost invariably do, to a period of legal emancipation from parental control . . . . They are also of purer morals (*Fourth Annual Report* [Boston Board of Education, 1841], 45 – 46).

In childhood the intellectual faculties are but partially developed – the affections much more fully. At that early age the affections are the key of the whole being. The female teacher readily possesses herself of that key, and thus having access to the heart, the mind is soon reached and operated upon (Assemblyman Hurlburd, *New York State Education Exhibit* [World's Columbian Exposition, Chicago, 1893], 45 – 46).

At the center of the cluster of ideas that made up each writer's picture of women we see a belief in women's capacity to influence children's behavior through the emotions. Barnard's "silken cord of affection" and Hurlburd's "access to the heart" were characteristic themes in discourse about women as teachers. The writer of the Boston Board of Education's annual report associated

8 Sugg, *Motherteacher*, 4 – 25.

women's ability to establish emotional links with children with women's lack of acquisitiveness and acceptance of dependence. These presumed qualities made women ideal candidates to teach in elementary schools, the purpose of which was to instill principles of behavior and convey basic literacy at a minimum cost to the public purse. Women were favored and actively recruited as elementary teachers because their presence in the schools satisfied a larger political agenda. Their perceived gender characteristics and their lack of academic preparation were positive advantages in the eyes of early public education officials; with a corps of women teachers there was no danger that investment in public education might foster the creation of new elites.

What, then, were the consequences of this congruence of ideology and economic concerns that served to give women preferred access to the teaching profession in the United States? The first consequence, extensively commented on by foreign visitors, was that discipline in American schools was very different from any known in European classrooms. As women were not thought suited to administering corporal punishment, the rod was virtually absent from America's schools. Maintaining discipline and conveying knowledge became more a matter of persuasion than an exercise of power based on authority. One learned because one liked the teacher, not out of respect for the learning that the teacher represented, as was the case in the French lycée or the German gymnasium. The climate in the American schoolroom was wholly different; the classroom was considered an extension of the home.

This should not be taken to mean that the stereotype of the steely-eyed New England schoolmarm was incorrect;

there were many such outstanding women. What it did mean, however, was that maleness involved rebellion against the values for which the schoolmarm stood. Many celebrations of maleness in American culture have retained overtones of adolescent rebellion against a female cultural presence that ostensibly cannot be easily incorporated into a strong adult male identity.

We may speculate about the consequences of subsuming school and home within a maternal, domestic culture rather than having the school serve as an impersonal agent of cultural authority, much like the church or the army. How would *Huckleberry Finn* read if the journey on the raft were an escape from male institutions? *Huckleberry Finn*'s journey raises many profound questions about American culture. One critical question is whether the overrepresentation of one gender in the early stages of schooling permits either boys or girls to develop the balanced identities we associate with creativity. For the purpose of understanding American educational institutions, another question that requires answering is this: If the school exists in opposition to male values and frontier life, how are we to understand higher education? In what ways is there a cultural imperative to redress the balance between maternal and masculine values at different levels of the system? What has that cultural requirement meant for American intellectual life?

Teaching through love made the school a setting in which many ideas about child development were played out; it was never an agency for strenuous effort to discipline and develop young intellectual talents. Thus, the traditional twelve years of schooling did not bring the young American student to the levels of learning aimed at



by the lycée or the gymnasium. Instead, and increasingly, American education came to require a further four years of intellectual exploration at the college level before the young person was considered to be in a position to make adult career commitments. Moreover, because of American public schools' identification with maternal functions, colleges and universities have distanced themselves from schools and stressed the "masculine" tough-mindedness of American scholarship. This difference remains an enduring puzzle to Europeans, who see both schools and universities in a continuum of intellectual endeavor, and who value intellectual playfulness.

We may interpret this impulse to distance higher learning from schools as a natural response to some of the major nineteenth-century curricular debates. Because the schools operated as agents of maternal values, school curricula were organized along the lines of accepted models of child development. G. Stanley Hall's celebrated theories of child development, which held that the child recapitulated the various stages of human evolutionary development, required that the teacher act as a helpful director as the pupil traversed these stages. It is unlikely that Hall would have designed so unintellectual a teaching role had he assumed that most elementary school teachers would be men. His ideas about child development were revolutionary in their largely successful redefinition of childhood as a series of developmental stages rather than as a time when the "imp of Satan" had to be disciplined; however, his view of the teacher was based on earlier nineteenth-century assumptions about the female temperament.

John Dewey's Progressive schools discarded the notion of a fixed body of in-

tellectual skills to be acquired entirely in school. Progressive pedagogy asked that the teacher help the young to discover the world through their innate intelligence. It took individuals with an almost superhuman capacity for nurturing to manage this kind of schoolroom. Few teachers could completely repress the desire to instruct, as Dewey's theories required. Many rueful survivors of Progressive schools testified to the demoralizing nature of such self-abnegation. It is reasonable to ask whether educational theorists would have designed teaching roles of such preternatural maternal patience had they expected their male colleagues to take principal responsibility for such instruction. Had the standard levels of education required for elementary school teachers been higher, educational reformers of the Progressive variety might have found earlier curricular ideals less easy to disregard. It was because the minds of young teachers were seen as *tabulae rasae* that older notions of learning could be easily ignored. Certainly if one assesses Dewey's pedagogy from the standpoint of the gender stereotypes enshrined within it, its conservatism is striking. Dewey advanced a new theory of learning and stated new political goals for American schools, but his assumptions about the temperamental and intellectual characteristics of teachers differed little from the assumptions made by Henry Barnard and his colleagues in the 1840s.<sup>9</sup>

While many of the goals of Progressive education were admirable, the fact that the overwhelming majority of teachers in the American elementary school sys-

9 On G. Stanley Hall's educational theories, see Dorothy Ross, *G. Stanley Hall, The Psychologist as Prophet* (Chicago: University of Chicago Press, 1972). On Dewey and Progressive education, see John Dewey, *On Education* (New York: The Modern Library, 1964).

tem were young women was a substantial influence on the way reformers thought about the role of teacher. Because of the persistence of the idea that women related to children primarily through the emotions, reformers prescribed intellectually demeaning roles for teachers – roles that often ignored the teacher's intellectual capacity in relation to the child's.

Similarly, the fact that most teachers did not have the right to vote affected the dynamics of the political relationship between the common schools and the larger society. From its inception the public education system operated at the center of a vortex of political forces, many of which were intrinsically unrelated to pedagogical issues. The schools were affected by political battles over such issues as patronage rights, appointments to teaching staffs and desirable jobs on maintenance staffs, which districts would be granted the economic benefits of building contracts, and which merchants should benefit from the purchasing power of students and their families. Moreover, it was taken for granted that parents, who had an abiding interest in the curriculum and its relationship to employment opportunities, and whose taxes paid teachers' salaries, had a democratic right to influence what was and was not taught to their children. These interests found expression in city and state politics, but women teachers were disfranchised until 1919 and consequently were unable to directly participate in the political process that shaped and established priorities for public education. Fathers and men teachers could mobilize voter support for school policies through their lodges or friendly societies, or later through Rotary, Kiwanis, or Lions Clubs; women could not. This situation affected women's status as teach-

ers and indirectly affected the political importance of schools: an important component of professionalization in all modern societies is the degree to which would-be professionals are able to persuade economically or politically powerful elites that their services are important enough to command special rewards. Women teachers, unable to undertake this effort effectively, found their logical political allies in the ranks of organized labor.

The history that produced this logic is vividly illustrated in the disputes affecting the Chicago school system in the 1880s and 1890s. The city's total population was 500,000 and there were 59,000 pupils in the public schools, which expended a budget of over \$1 million. The school system was the biggest employer in the city. The school board was appointed by the mayor, and it controlled or influenced three sets of resources critical to Chicago's economic future: land voted to support the public schools, contracts for school buildings, and tax abatements for corporations occupying land within the city. The major issues of concern to teachers were security of tenure, pension rights, and professional evaluation for promotion.<sup>10</sup> Women teachers felt considerable social distance from the exclusively male school superintendents in the city, who were themselves political appointees. In the campaign to secure teachers' pension rights, the female-led Chicago Federation of Teachers found that it carried no weight with the municipal government, so it waged battle in the courts. In her autobiography, Margaret A. Haley, the founder

10 Robert J. Braun, *Teachers and Power: The Story of the American Federation of Teachers* (New York: Simon and Schuster, 1972), 21–27. See also Robert I. Reid, ed., *Battleground: The Autobiography of Margaret A. Haley* (Urbana: University of Illinois Press, 1982).

of the federation, records the process by which she came to conclude that, because of women's limited voting rights, her union's predominantly female membership would gain political leverage by affiliating with a strong political organization – the Chicago Federation of Labor. She recognized that laws were only enacted in response to the political pressure of voters. "Except in a few western states," she wrote, "the women of the nation had practically no voting power."<sup>11</sup>

The early choice of unionization was a natural one for nonvoting workers; its consequences were profound. As early as the Chicago Federation of Teachers' 1902 decision to affiliate with the Chicago Federation of Labor, the city's elementary teachers were in a confrontational relationship with political and social elites. The male school principals and superintendents, who identified with management in the labor-versus-management model of the school and the teacher's role within it, were even more distanced from teachers. The working peers of the school administrators were the political actors who had selected and appointed them. The place of the school in political priorities reflected the fact that most of its constituency could not vote and that its spokesmen were distant from the classroom. Decisions about educational policy were usually based entirely on the budgetary priorities of individual districts and regions. Economic considerations favored the selection of women teachers and, by the late nineteenth century, women principals; women's salaries in such positions did not reflect high esteem for their professional achievements. Jessie May Short, an assistant professor of mathematics at Reed Col-

lege in Portland, Oregon, described her experience in an Oregon high school in the 1920s.

A personal experience will illustrate the discriminations that are considered normal in the smaller schools. . . . For five years I was principal of a high school in a delightful county-seat town. During the five years the high school enrollment doubled, a new building was erected, I had salary increases each year. I resigned for graduate study although I was offered a small salary increase if I would remain. The man who took my place was freely given a salary fifty percent higher than I had received. Before his first year had closed he was literally taken from the school and thrown into a snow bank. The school board asked me to return and made me what they considered a generous offer, a ten percent increase over my former salary. I suggested that I might consider the appointment at the fifty percent increase the board had willingly given the man who could not handle the situation. The idea of compensating the service without regard to the sex of the one rendering the service was, as I had anticipated, beyond their comprehension.<sup>12</sup>

Short's experience strikingly illustrates that the public's view of the worth of the predominantly female teaching profession and of the predominantly male management of the public schools was fundamentally shaped by the gender of those who served in the system. Because there was little popular respect for the function of the teacher, most important professional prerogatives were gained only after protracted battle. The early decades of unionizing and struggling against low social esteem focused

<sup>12</sup> Jessie May Short, *Women in the Teaching Profession: Or Running as Fast as You Can to Stay in the Same Place* (Portland, Oreg.: Reed College, June 1939), 10.

<sup>11</sup> Reid, *Battleground*, 90.

teachers' concerns on job security to the neglect of curricular issues. The cherished right of tenure, sought since the 1880s, was not achieved until the 1950s, when the postwar baby boom and the cold-war mentality of the Sputnik era gave schools and teachers national importance.

The public's low esteem of the profession was also related to the youthfulness of women teachers. As most of them remained teachers for no more than three or four years, it was easy for local school boards to disregard their opinions. The assumption that young women need protection gave school boards and committees ample justification to scrutinize teachers' conduct and to represent such activity to be in the teachers' best interest. The small minority of men teachers acquired the status of their women colleagues by association. Because society accorded such scant respect to the role of teacher, it was considered perfectly appropriate to pay teachers wages equivalent to those of unskilled labor. By 1900 teacher turnover was as high as 10 percent a year; every year 40,000 new recruits had to be brought into the common school system.<sup>13</sup> The high annual rates of change in teaching personnel throughout the first century of the profession made teachers seem much more like transient workers than career professionals (teaching was not accepted as a lifetime career for women until the Second World War). School reformers even today struggle with the consequences of Margaret Haley's accurate perception that to bargain successfully, women teachers had to unionize like industrial laborers.

13 B. A. Hinsdale, "The Training of Teachers," in *Education in the United States: A Series of Monographs Prepared for the United States Exhibit at the Paris Exposition, 1900*, ed. Nicholas Murray Butler (Albany, N.Y.: J. Lyon, 1900), 16.

If we compare the public esteem accorded to teaching in the late nineteenth century with that held for other emerging professions, we begin to see that the difference lies in the fact that most of the people recruited into public education were women. Consider, for instance, attitudes toward the engineer – the male professional who emerged to meet national needs in transportation, communication, and industrial technology over the same one hundred years that saw the establishment of public education. In the United States the social origins of engineers were almost identical to those of teachers. Engineers too came from rural and blue-collar families. Initially, their training was not highly theoretical and their tasks were strictly utilitarian. Yet engineers were held in high public esteem.

Clearly, gender categories and cultural values had a tremendous influence on the process of professionalization. We have only to read Henry Adams's assessment of the new technology in his commentaries on *The Virgin and the Dynamo*, or Thorstein Veblen's description of the engineer in *The Engineers and the Price System* (1919), to see what a difference gender made. "These technological specialists," Veblen wrote, "whose constant supervision is indispensable to the due working of the industrial system, constitute the general staff of industry, whose work is to control the strategy of production at large and keep an oversight of the tactics of production in detail."<sup>14</sup> During the Depression, when married women teachers were dismissed by school systems to create openings for unemployed men, Lewis Mumford wrote, "The establishment of the class of engineers in its proper characteristics

14 Thorstein Veblen, *The Engineers and the Price System* (New York: Heubsch, 1921), 52 – 53.

is the more important because this class will, without doubt, constitute the direct and necessary instrument of coalition between men of science and industrialists, by which alone the new social order can commence.”<sup>15</sup> No one thought to exclaim on how much the new social order might depend on the labors of “the class of teachers.” Engineers, of course, pursued their training at the college level and developed a professional culture of aggressive masculinity. Their skills were of critical and immediate importance to the business elites of American society – but then so were the skills of teachers, although no one recognized their value.

Gender stereotypes helped to account for the differences in social mobility experienced by women and men drawn from the same social background. If we look at the gender composition of the teaching profession cross-culturally, we see that the American pattern established at the time of the creation of the public school system was unique. In 1930 – 1931, a national survey of American teachers showed that women outnumbered men by 19 to 1 in elementary education and by 3 to 1 in secondary education. In contrast, men held 65 percent of the elementary teaching posts in Norway and 69 percent of the secondary teaching positions there. In Germany 75 percent of the primary school teachers and 71 percent of the secondary school teachers were men; the ratios for France were similar.<sup>16</sup> These figures reflect the conditions that existed in societies that

had had relatively stable populations when the public system of elementary and secondary education was being established, and that made strongly centralized educational planning a high national priority.

In these European countries, lifetime careers of steady progression through the different levels of the public school system were established; entry-level positions based on long and strict academic preparation were accepted as the norm. In France, for instance, completion of the baccalaureate was required to become a lycée teacher; further progress in the system required an advanced degree. Besides contributing substantially to the intellectual level of the schools, this pattern of recruitment defined the teacher as an agent of the nation’s culture, not simply a representative of its maternal values.

When the possibility of recruiting more men to the profession or requiring teachers to undergo more rigorous academic preparation was broached in the United States, it was generally discarded as prohibitively costly. In 1906 – 1907, for instance, the New York City school superintendent acknowledged the desirability of having a cadre of teachers more balanced in gender composition. In a report, he commented that the achievement of this goal would require equalizing the pay scales of the gender groups and raising all salary levels. This, he calculated, was politically impossible. It would add between \$8 million and \$11 million to the annual school system budget. To propose such a budget increase in the absence of popular demand would be political suicide, and there was not the slightest popular sentiment for such action.<sup>17</sup>

15 Lewis Mumford, *Technics and Civilization* (New York: Harcourt Brace, 1934), 219 – 220.

16 Edward S. Evenden, Guy C. Gamble, and Harold G. Blue, “Teacher Personnel in the United States,” in *National Survey of the Education of Teachers*, vol II (Washington, D.C.: U.S. Government Printing Office, 1933), 20.

17 Sugg, *Motherteacher*, 122.



Gender was a highly significant factor in the way American society mobilized its resources to develop its public education system. Assumptions about female temperament and motivation dovetailed with the often contradictory ideals and values of the public school system's creators. Stereotypes about women coincided neatly with the economic priorities that dictated how much money was appropriated for public education, and reinforced popular preferences regarding the purpose of public schooling. Assumptions about the gender and intellectual level of the typical teacher influenced successive waves of curricular reform. Culturally, these gender stereotypes had a tremendous impact on everyone involved in the schools – teachers, pupils, principals, superintendents, school board members. These assumptions played a part in what it meant to grow up male or female in America. Their enduring power explains the continued inability of our affluent society to muster either the will or the resources to create and maintain schools that are intellectually demanding and that accord the profession of teaching sufficient dignity to engender high teacher morale.

Much has been made of the degree to which teaching offered American women the opportunity to move out of family subordination and into an independent existence. The memoirs of some of America's greatest women reformers tell us that this new life outside the family was a heady experience. Frances Willard, for example, wrote of learning to live without reliance on her parents as a very young teacher. Through her struggles with unruly children in rural one-room schools, she came to see herself as an agent for improving society. Dozens of other young women documented similar experiences. Service as teachers inspired many young women to seek other

active careers. Both as individuals and as a group, women proved themselves capable of creating and sustaining demanding intellectual tasks when they were given adequate preparation and appropriate remuneration. It was not the sex of women teachers that created problems in the school system and made the status of teachers so lowly; it was the gender identity that women carried into the schools with them. It is the terms on which women enter occupations that govern their opportunities. The mere fact of entry does not create opportunities. Horace Mann and Henry Barnard, two of America's greatest educational reformers, actively admired women and thought that by employing them as teachers they could secure both a better society and important advantages for women. They bore women no ill will whatsoever. Their assumptions about women, however, established the terms on which women entered the teaching profession, and those terms were far more consequential than the great numbers of women who were invited to teach in the public schools. Those terms still matter today. So too does our ambivalence about the goals of public education. This piece of unfinished business from the politics of the Jacksonian era matters as much today as it did in the 1840s. We cannot conclude it satisfactorily without taking into account the unintended consequences of our assumptions about the gender of teachers. They matter not only to women but to our whole society.

# Tu Wei-ming

## *Cultural China: the periphery as the center*

The inscription of the Tang's basin reads, "If one day you truly renew yourself, day after day you will renew yourself; indeed, renew yourself every day." In the "Announcement to the Prince of Kang" it is said, "You shall give rise to a renewed people." In the *Book of Poetry* it is said, "Though Zhou is an old state, the Mandate it holds is new." For this purpose, the profound person exerts himself to the utmost in everything.

– The Great Learning<sup>1</sup>

China, one of the longest continuous civilizations in human history, "may be visualized as a majestic flowing stream."<sup>2</sup> Chinese culture, the generic term symbolizing the vicissitudes

of the material and spiritual accomplishments of the Chinese people, has undergone major interpretive phases in recent decades and is now entering a new era of critical self-reflection. The meaning of being Chinese is intertwined with China as a geopolitical concept and Chinese culture as a living reality.

For China, Chinese people, and Chinese culture, the image of the twentieth century as an atrocious collective experience of destructiveness and violence emerges with fulgent salience as we approach the fin de siècle rumination. Stability has often meant a delicate balance for a few years; even a decade of peaceful coexistence evokes memories of permanence. The fluctuating Chinese political landscape, precipitated by external events unprecedented in Chinese history since the mid-nineteenth century, has become so restless in the last decades that not only the players but the rules of the game have constantly changed. For instance, in the eight decades since the end of the Qing dynasty in 1911, a suc-

---

*Tu Wei-ming, director of the Harvard-Yenching Institute and Harvard-Yenching Professor of Chinese History and Philosophy and of Confucian Studies, wrote this essay for the Spring 1991 issue of "Dædalus." At the time of its publication, he was director of the Institute of Culture and Communication at the East-West Center in Honolulu and professor of Chinese history and philosophy at Harvard University. He has been a Fellow of the American Academy since 1988.*

---

© 2005 by the American Academy of Arts & Sciences

1 Daniel K. Gardner, *Chu Hsi and the Ta-hsueh: Neo-Confucian Reflection on the Confucian Canon* (Cambridge, Mass.: Harvard University Press, 1987), 96–97 with minor modifications.

2 Herbert Passin, "The Occupation – Some Reflections," *Dædalus* 119 (3) (Summer 1990): 125.

cession of different versions of the state constitution were drafted and promulgated in both the Republic of China and the People's Republic of China. Not revisions or amendments, each new version superseded the previous one.<sup>3</sup> Indeed, virtually no institution of significance (university, church, press, professional society, or civic organization) has lasted for more than a generation. The two major parties (the Nationalist and the Communist) seem to have endured in form, but they both have been so substantially and radically restructured that a sense of cynicism and uncertainty prevails among their members. The most devastating rupture, however, occurred within the intellectual community.

Although China has never been subjected to the kind of comprehensive colonial rule experienced by India, China's semicolonial status severely damaged her spiritual life and her ability to tap indigenous symbolic resources. Chinese intellectuals have been much more deprived than their Indian counterparts ever were. While Indian intellectuals have continued to draw from the wellsprings of their spiritual lives, despite two centuries of British colonialism, the Western impact fundamentally dislodged the Chinese intellectuals from their Confucian haven. Having loosened their moorings in a society which had provided a secure and respected anchorage for their predecessors for more than two millennia, they desperately tried to find a niche in a cruel new

3 In the history of the People's Republic of China, four radically different constitutions were implemented between 1954 and 1982. For an indictment against the Chinese Communist Party's abuse of constitutional authority, see Yan Jiaqi, "China Is Not Actually a Republic" and "May 17th Manifesto," in Yan Jiaqi, *Toward a Democratic China: My Intellectual Autobiography* (Honolulu: University of Hawaii Press, forthcoming).

world defined in terms of power with or without wealth. Their sense of impotence, frustration, and humiliation, prompted by a curious mixture of political nationalism and cultural iconoclasm, framed the context for their quest for identity not only as Chinese but as thinking and reflective Chinese in an increasingly alienating and dehumanizing world.

The question of Chineseness, as it first emerged in the "axial age" half a millennium prior to the birth of Confucius in 551 B.C., entails both geopolitical and cultural dimensions. While the place of China has substantially expanded over time, the idea of a cultural core area first located in the Wei River Valley, a tributary of the Yellow River, and later encompassing parts of the Yangtze River has remained potent and continuous in the Chinese consciousness. Educated Chinese know reflexively what China proper refers to; they may not be clear about the periphery but they know for sure that the center of China, whether Xi'an or Beijing, is in the north near the Yellow River. The archaeological finds in recent decades have significantly challenged the thesis that China grew from the Wei Valley like a light source radiating from the center. Even in neolithic periods, there were several centers spreading across present-day China. The Middle Kingdom came into being as a confederation of several equally developed cultural areas rather than growing out of an ever-expanding core.<sup>4</sup> Yet, regardless of this scholarly persuasive explanation of the origins of Chinese civilization, the impression that geopolitical China evolved through a long pro-

4 Kwang-chih Chang, *The Archaeology of Ancient China*, 4th ed., rev. and enl. (New Haven, Conn.: Yale University Press, 1986).

cess centering around a definable core remains deeply rooted.

If the presumed core area was instrumental in forming a distinctive Chinese identity, Chinese culture symbolizing a living historical presence made the sense of being Chinese even more pronounced; it signified a unique form of life profoundly different from other styles of living often condemned as barbarian. The expression *hua* or *huaxia*, meaning Chinese, connotes culture and civilization. Those who lived in China proper were, inter alia, cultured and civilized, clearly differentiable from those barbarians in the periphery who had yet to learn the proper ways of dressing, eating, dwelling, and traveling. On the surface, the classical distinction between Chinese and barbarians was predicated on the divergence of two drastically different modes of life: the agrarian community of the central plain and the nomadic tribes of the steppes.<sup>5</sup> But the rise of Chinese cultural consciousness was occasioned by primordial ties defined in ethnic, territorial, linguistic, and ethical-religious terms. Although it is often noted that culture, rather than ethnicity, features prominently in defining Chineseness, the cultured and civilized Chinese, as the myth goes, claim a common ancestry. Indeed, the symbol of the “children of the Yellow Emperor”<sup>6</sup> is constantly reenacted in Chinese literature and evokes feelings of ethnic pride.

This idea of being Chinese, geopolitically and culturally defined, is further reinforced by a powerful historical consciousness informed by one of the most

voluminous veritable documents in human history. Indeed, the chronological annals have flowed uninterrupted since 841 B.C. This cumulative tradition is preserved in Chinese characters, a script separable from and thus unaffected by phonological transmutations. Whether or not it is simply a false sense of continuity, the Chinese refer to the Han (206 B.C. – A.D. 220) and Tang (618 – 907) dynasties as if their greatness still provides practicable standards for contemporary Chinese culture and politics.

The Middle Kingdom syndrome, or the Middle Kingdom complex,<sup>7</sup> may have made it psychologically difficult for the Chinese leadership to abandon its sense of superiority as the center, but we must also remember that China had never been thoroughly challenged by an alien equal – if not superior – civilization until the penetration of the West in the mid-nineteenth century. The “Buddhist Conquest of China”<sup>8</sup> entailed the introduction, domestication, maturation, and development of Indian spirituality in China for more than six hundred years, culminating in the intense Sinicization of Buddhist teachings in distinctively Chinese schools of Tiantai, Huayan, and Chan.<sup>9</sup> The military and political domination of the Middle Kingdom by the Jurchens, the Khitans, the Mongols, and the Manchus in the last millennium was compensated, in cultural terms, by the Sinicization of Jin, Liao,

5 Owen Lattimore, *Inner Asian Frontiers of China* (London: Oxford University Press, 1940).

6 Lynn Pan, *Sons of the Yellow Emperor: A History of the Chinese Diaspora* (Boston: Little, Brown, 1990).

7 For a recent reference to the “Middle Kingdom Complex,” see Lucian Pye, “China: Erratic State, Frustrated Society,” *Foreign Affairs* 69 (4) (Fall 1990): 62.

8 Erik Zürcher, *The Buddhist Conquest of China* (Leiden: Brill, 1959).

9 Arthur Wright, *Buddhism in Chinese History* (Stanford, Calif.: Stanford University Press, 1953).

Yuan, and Qing into legitimate Chinese dynasties. China survived these “conquests” as a geopolitical entity and Chinese culture flourished. Nevertheless, if we take seriously the image of “a majestic flowing stream,” we must acknowledge that these great outside influences altered this stream at various points. In accordance with this, China, or Chinese culture, has never been a static structure but a dynamic, constantly changing landscape.

In the Chinese historical imagination, the coming of the West, however, could be seen as more “decentering”; it was as if the Buddhist conquest and the Mongol invasion had been combined and compressed into one generation. It is understandable, therefore, that it has thoroughly destroyed the “pattern of the Chinese past”<sup>10</sup> and fundamentally redefined the *Problematik* for the Chinese intellectual. The convulsive disturbances that geopolitical China has suffered since the Opium War (1839 – 1942) are well documented,<sup>11</sup> but the effervescences in Chinese culture which eventually brought about the intriguing paradox of iconoclasm and nationalism of the May Fourth (1919) generation (as well as subsequent generations) are so elusive that scholars of modern thought are still groping for a proper explanatory model to probe them.<sup>12</sup>

10 Mark Elvin, *The Pattern of the Chinese Past* (Stanford, Calif.: Stanford University Press, 1973).

11 Peter Ward Fay, *The Opium War, 1840 – 1842* (Chapel Hill: University of North Carolina Press, 1975).

12 Chow Tse-tsung, *The May Fourth Movement: Intellectual Revolution in Modern China* (Cambridge, Mass.: Harvard University Press, 1960). For a thought-provoking attempt to come to terms with the May Fourth intellectual ethos, see Vera Schwarcz, *The Chinese Enlightenment: Intellectuals and the Legacy of the May Fourth*

A radical manifestation of this ambivalent May Fourth legacy is the recent advocacy of comprehensive modernization qua Westernization in the People’s Republic of China after the official closure of the devastating Cultural Revolution decade (1966 – 1976). This new rhetoric is deceptively simple: since China’s backwardness, fully acknowledged by the Chinese intelligentsia as occasioned by the open-door policy of the reform, had deep roots in the Chinese polity, society, and culture, a total transformation of Chineseness is a precondition for China’s modernization. Strategically, the most painful and yet effective method of this total transformation is to invite the modern West with all of its fruitful ambiguities to “decenter” the Chinese mentality. This wishful thinking – liberation through a willing and willful confrontation with radical otherness – has become a powerful countercultural thrust against both ossified Marxism-Leninism and the still vibrant “habits of the heart”<sup>13</sup> molded by the Confucian tradition.

The “River Elegy,” a controversial, interpretive, six-part television series on Chinese cultural roots and ethos, straightforwardly advocated the necessity of embracing the blue ocean as the only way to save the “Yellow Earth.”<sup>14</sup>

---

*Movement of 1919* (Berkeley: University of California Press, 1986).

13 A seminal term in Alexis de Tocqueville’s *Democracy in America* is used by Robert Bellah and his coauthors to characterize the individual ethos of contemporary American culture. See Robert Bellah et al., *The Habits of the Heart* (Berkeley: University of California Press, 1985).

14 The contrast of the blue ocean and the “Yellow Earth” is used in the “River Elegy” to show the open, dynamic, and exploratory spirit of the West and the closed, stagnant, and insular mentality of China.



Aired twice in 1988, the “River Elegy” provoked a heated nationwide debate on tradition, modernity, change, China, and the West.<sup>15</sup> From top Party leaders and intellectuals to workers, soldiers, and farmers, from the metropolitan areas of Beijing, Shanghai, and Wuhan to the sparsely populated Great Northwest, several hundred million citizens were affected by the central message: China, behind even western Africa in per capita income, would soon be disfranchised as a player in the international game.<sup>16</sup> The intellectual community was stunned by the poignancy of the question, Whither China? Overwhelmingly siding with the radical Westernizers, they have accepted that reform requires the courage to restructure China fundamentally by importing proven models of success. The sacred symbols of the ancestral land stand condemned. The dragon, the symbol of Chinese ancestry, is condemned as outmoded imperial authoritarianism;<sup>17</sup> the Great Wall, the symbol of historical continuity, is condemned as a manifestation of closed-minded conservatism; and the Yellow River, long regarded as the cradle of Chinese civilization, is condemned as unmitigated violence against innocent people. The

15 Frederic Wakeman, Jr., “All the Rage in China,” *The New York Review of Books*, March 2, 1989, 19 – 21.

16 In 1988 mainland China’s per capita income was US\$330, while the average for sixteen countries in western Africa was \$340. “1990 World Population Data Sheet,” Population Reference Bureau, Inc., Washington, D.C.

17 For example, Yan Jiaqi, a prominent leader of the Alliance for a Democratic China, is an articulate critic of the “dragon mentality.” In his *World Economic Herald* (*Shijie jingji daobao*, Shanghai) interview with Dai Qing, “China Is No Longer a Dragon” (March 21, 1988), he specifically berates the symbolism of the dragon.

unstated message, obvious to most, gives a warning, and indeed an outright challenge to the power holders of the Party: speed up the reform or else! Chineseness, under scathing assault, is ironically made to stand for the modus operandi of an authoritarian, conservative, and brutal ruling minority.

The paradox embedded in the message of the “River Elegy” evokes memories of the May Fourth intellectual dilemma: the intertwining of nationalism (patriotism) and iconoclasm (antitraditionalism).<sup>18</sup> This leads inevitably to a whole set of thought-provoking questions. If Chinese intellectuals in China proper are so thoroughly disgusted with Chinese culture, can they define their Chineseness as an exclusive commitment to wholesale Westernization? If their condemnation of things Chinese is total, does this mean that they have voluntarily forfeited their right to be included in a definition of Chineseness? For Chinese intellectuals living in China proper, can the meaning of being Chinese be sought in the limbo between a past they have either deliberately relegated to a fading memory or been coerced into rejecting or forgetting, a present they have angrily denounced, and an uncertain future, since they insist that the promise lies wholly in the alien unknown? The way these issues are formulated may appear relevant only for a tiny minority – the articulate and self-reflective intelligentsia – but the emotional intensity provoked by the debate has affected the Chinese populace in general.

The rise of Japan and the so-called Four Mini-Dragons (South Korea, Taiwan,

18 Yu-sheng Lin, *The Crisis of Chinese Consciousness: Radical Antitraditionalism in the May Fourth Era* (Madison: The University of Wisconsin Press, 1979).

Hong Kong, and Singapore) as the most dynamic region of sustained economic development since the Second World War raises challenging questions about tradition in modernity, the modernizing process in its different cultural forms. Does it suggest the necessity, indeed, the desirability of a total iconoclastic attack on traditional Chinese culture and its attendant comprehensive Westernization as a precondition for China's modernization?<sup>19</sup> From the perspective of economic organization, does this new capitalism, labeled as *guanxi* (network and connections) capitalism, contrasted with the classical capitalism of Western Europe, signal a new age – the age of the Pacific Rim?<sup>20</sup> Or, is it merely an epiphenomenon that can be explained in terms of existing European and American development models? Politically speaking, are we witnessing a process of democratization based more on consensus formation than on adversarial relationships, giving a wholly new shade of meaning to the concept of participatory democracy? Or, are we observing the continuous presence of the hierarchical authoritarian control of a political elite operating under the guise of majority rule?<sup>21</sup> Socially, do family cohesiveness, low crime rates, respect for education, and a high percentage of savings relative to that of other industrial societies indicate an ethos different from the individual-centered “habits of the

19 Tu Wei-ming, “A Confucian Perspective on the Rise of Industrial East Asia,” 1687th Stated Meeting Report, *Bulletin of the American Academy of Arts and Sciences* 42 (1) (October 1988): 32–50.

20 S. Gordon Redding, *The Spirit of Chinese Capitalism* (Berlin: de Gruyter, 1990).

21 Lucian Pye, *Asian Power and Politics: The Cultural Dimensions of Authority* (Cambridge, Mass.: Harvard University Press, 1985).

heart”?<sup>22</sup> Or, do they simply reflect an earlier stage of modern transformation, which will lead eventually to the anomie and alienation experienced in the West? Culturally, do these societies symbolize successful examples of advanced technology being combined with age-long ritual practices, or are they simply the passing phases of traditional societies?<sup>23</sup> In short, how does the rise of East Asia challenge our deep-rooted conceptions of economic growth, political development, social transformation, and cultural change?

These questions are significant for interpreting the meaning of being Chinese; they are potentially provocative to the overwhelming majority of Chinese intellectuals in mainland China who believe that Chineseness is incongruous with the modernizing process, defined purely in terms of science and democracy. If, indeed, the “Sinic world”<sup>24</sup> or the “post-Confucian”<sup>25</sup> region has succeeded in assuming a form of life definitely modern, distinctively East Asian – by implication Chinese as well – the sharp dichotomy between tradition and modernity must be rejected as untenable, as useless in analyzing developing countries as well as more highly industrialized or postindustrial societies. Any

22 Peter Berger and Hsin Huang Hsiao, eds., *In Search of an East Asian Development Model* (New Brunswick, N.J.: Transaction Books, 1988).

23 Tu Wei-ming, “The Rise of Industrial East Asia: The Role of Confucian Values,” *Copenhagen Papers in East and Southeast Asian Studies* (April 1989): 81–97.

24 Edwin O. Reischauer, “The Sinic World in Perspective,” *Foreign Affairs* 52 (2) (January 1974): 341–348.

25 Roderick MacFarquhar, “The Post-Confucian Challenge,” *The Economist* 9 (February 1980): 67–72.

attempt to measure the degree of modernization with a linear developmental scale is thought to be simpleminded. Although this point has been repeatedly argued by culturally sophisticated modernization theorists since the early 1970s,<sup>26</sup> the presence of an empirically verifiable phenomenon makes the argument even more convincing.<sup>27</sup>

Since traditional features of the human condition – ethnicity, mother tongue, ancestral home, gender, class, and religious faith – all seem to be relevant in understanding the lifeworlds of societies, both modern and developing, the need to search for roots, despite the pervasiveness of global consciousness, is a powerful impulse throughout the world today.<sup>28</sup> If there is an alternative path to capital formation, then democracy, technology, and even modernization may indeed assume different cultural forms. The most radical iconoclastic assertion, espoused by some of the articulate May Fourth intellectuals, that Chinese culture – and not just Confucianism but the ideographic language as well – will have to be abolished as a precondition

for China's modernization, is now regarded as completely outdated. Even the most ardent Westernizers in Beijing and Shanghai chose to see their ideas circulate in the Chinese print media. To Chinese intellectuals in industrial East Asia, the awareness that active participation in the economic, political, social, and cultural life of a thoroughly modernized community is not necessarily in conflict with being authentically Chinese implies the possibility that modernization may enhance rather than weaken Chineseness. Still, the meaning of being Chinese is itself undergoing a major transformation.

A recent economic phenomenon with far-reaching political and cultural implications is the great increase in intraregional trade in the Asia-Pacific region. Since the Four Dragons are providing 31 percent of all foreign investments in the countries of the Association of Southeast Asian Nations (ASEAN), the participation of "diaspora" Chinese becomes vitally important; they are now responsible for the largest capital transfer in the region, exceeding that of both Japan and the United States. Just as the public perception of the Chinese in the United States has changed from laundrymen to engineers and professionals, the image of the Chinese as economic beings is likely to be further magnified in Southeast Asia, changing perhaps from that of trader to that of financier. The Chinese constitute not only the largest peasantry in the world today but also the most mobile merchant class.<sup>29</sup>

Despite all these remarkable economic accomplishments in Asia and in the Pacific, the future is filled with uncertainties. As the United States reduces its

26 S. N. Eisenstadt, "Varieties of Political Development: The Theoretical Challenge," in S. N. Eisenstadt and Stein Rokkan, eds., *Building States and Nations: Models and Data Resources*, vol. I (Beverly Hills, Calif.: Sage Publications, 1973), 41–72; and S. N. Eisenstadt, "Post-Traditional Societies and the Continuity of Reconstruction of Tradition in Post-Traditional Societies," in S. N. Eisenstadt, ed., *Post-Traditional Societies* (New York: Norton, 1974), 1–27.

27 Peter L. Berger, "An East Asian Development Model?" in Peter Berger and Hsin Huang Hsiao, eds., *In Search of an East Asian Development Model*, 3–11.

28 Tu Wei-ming "The Search for Roots in East Asia: The Case of the Confucian Revival," in Martin E. Marty, ed., *Fundamentalisms Observed* (Chicago: The University of Chicago Press, forthcoming).

29 See Wang Gungwu, "The Culture of Chinese Merchants," unpublished lecture given at the University of Toronto, October 20, 1989.

budget deficit, it may not be the same catalyst for growth as it was in the 1980s when an American import spree fueled much of the economic expansion of the region. Also, with the advent of a unified European Community in 1992 and its growing preoccupation with Eastern Europe, not to mention the deterioration of the Soviet economy and the present Middle East crisis, the West may well turn its attention away from Asia and the Pacific. Although it is unlikely that a “fortress Europe” or a Western Hemisphere economic zone will quickly push the Asia-Pacific region toward a Japan-anchored trading bloc, the hazards of protectionism in North America are certainly not negligible.

Still, if the projection of a Pacific century is at all credible, the roles of Taiwan, Hong Kong, Singapore, and the Chinese communities in Southeast Asia ought not to be underestimated. Taiwan, for example, has the distinction of holding the largest foreign reserve in the world (over 70 billion dollars in 1990), surpassing Japan, the United States, and Germany. While this fact alone may not be particularly significant, the combination of government leadership, entrepreneurial ingenuity, and a strong work ethic has made Taiwan, despite its political isolation, an assiduous investor and an innovator in international trade. Taiwanese merchants (predominantly in small and medium industries) are noticeable worldwide; the Nationalist government has made a highly coordinated and strategically sophisticated effort to make Taiwan a valued partner in many joint ventures in a number of key states in North America.

If the Taiwan “economic miracle” has attracted the most attention with the American public, the fascinating and enduring feature of the Taiwan experience has been its conscientious effort

to chart a radically different course of development, deliberately to challenge the socialist experiment on the mainland. As a result, the perceptual gap between the two sides of the Taiwan Straits has been exceedingly wide; despite the rhetoric of unification, the two “countries” have vastly different economic structures, political systems, social conditions, and cultural orientations. The Taiwan independence movement has created perhaps the most controversial and explosive political issue on the island, but the democratization process initiated by the top Nationalist leadership under pressure in 1987 has undoubtedly caught the spirit of the moment. If Taiwan (the Republic of China) becomes truly democratic, the question of Taiwan’s Chineseness will inevitably become a matter of public debate. Much attention has recently been focused on what may be called the sedimentations of Taiwanese history. For the intelligentsia, especially those under forty who were born and raised in Taiwan, the recognition that there have been distinctive Dutch, Japanese, and American strata superimposed on the Chinese substratum since the eighteenth century – not to mention the upsurge of nativistic sentiments of the Polynesian aborigines – makes the claim of Taiwan’s Chineseness problematic.

Still, the very fact that more than a million Taiwanese residents travel each year to the mainland to sightsee, do business, carry on scholarly communication, and hold family reunions has created a sort of “mainland mania” in the island, compelling the Nationalist government to deal with the mainland question in ways scarcely imagined even a couple of years ago. In late December 1990 the president of Taiwan announced that its state-of-war “emergency” vis-à-vis the mainland will be terminated by

May 1991. This will certainly lead to other astounding activities. Taiwan's official ideological claim to be the true inheritor of Chinese culture has taken a strange turn. In response to the threat of the independence movement, the government deems it advantageous to underscore Taiwan's Chineseness, but the challenge from the mainland prompts it to acknowledge how far Taiwan has already departed from the Sinic world.

The tale of the two cities (Hong Kong and Singapore) is equally fascinating. All indications suggest that the average per capita income in Hong Kong in 1990 had already surpassed that of its colonial ruler, the United Kingdom. The latter, more than Hong Kong, seems to be the principal beneficiary in this two-way investment relationship. Hong Kong's free-market capitalism, ably guided by government-appointed local leaders, exemplifies the "loose-rein" political philosophy characteristic of traditional China. Even though its ruling style is noninterference, its approach to economic affairs is a far cry from *laissez-faire* as it is traditionally practiced. The role of Hong Kong in international finance and in the development of manufacturing and light industry appropriate to her specific geopolitical and cultural conditions provides an inspiring example for many other developing and developed societies. Lurking behind the scenes, of course, is the overwhelming presence of mainland refugees and their experiences of persecution, loss, escape, renewal, and uncertainty. An estimated 1.5 million Hong Kong residents demonstrated in support of the democracy movement in Beijing in May 1989; with a total population of 5 million, virtually every family was represented in these demonstrations. Hong Kong's concern for and involvement in the affairs of the

homeland cannot be overestimated. For the majority of Hong Kong residents, being Chinese as British subjects is, in human terms, arguably superior to being Chinese as citizens of the People's Republic of China.

The story of Singapore – which in less than two decades emerged from being an endangered entrepôt to become a major industrial center in the Asia-Pacific region in trade, high technology, petroleum, tourism, medicine, and finance – is no less dramatic. The linguistic situation alone offers a clue to the complexity of the human condition. As Ezra Vogel observes, among the 75 percent of the population who are Chinese (15 percent Malays and 7 percent Indians), at least six major groups "who found each others' dialects unintelligible" can be identified; the Chinese *lingua franca* is now Mandarin which is for Singaporeans a dialect learned in this generation and devoid of deep family-rooted ethnic significance.<sup>30</sup> Yet, Singapore as an independent state and a safe society with its own unique blend of cultural eclecticism has endured. Whether or not Singapore is practicing "capitalism with socialist characteristics,"<sup>31</sup> her success in providing adequate housing, transportation, education, security, and welfare for her citizens clearly indicates that, at least in the economic sphere, her leadership, both governmental and business, has charted a course of action congenial to the Singapore situation.

Because the omnipresence of governmental intervention has transformed Singapore into an administrative state,

30 Ezra Vogel, *The Spread of Industrialization in East Asia*, unpublished manuscript of the 1990 Reischauer Lectures, Harvard University, 73.

31 Statement by one Singapore leader as quoted in Vogel, *ibid.*, 79.



with a tight control of the press, mass media, and public discourse, a stigma attaches to Singapore in the prominent English-speaking newspapers, notably the *Far Eastern Economic Review* and the *Asian Wall Street Journal*. The fact that “the leading business entrepreneurs in Singapore are government bureaucrats”<sup>32</sup> and that there appears to be a one-party political system, raises serious doubts about the state’s commitment to democracy. Still, one has the impression that Singapore’s government is efficient and uncorrupt; that the society is fresh and clean; and that the people are healthy and hardworking. In contrast to Hong Kong, Singapore’s Chineseness is not pronounced; indeed, in a certain sense, it is artificially constructed. Despite the obvious fear that any emphasis on Chinese cultural identity will lead to racial disharmony,<sup>33</sup> Singapore is unmistakably a sanitized version of Chinese society. Vogel notes that “if Hong Kong entrepreneurs thought of Singapore as a bit dull and rigid, Singapore leaders thought of Hong Kong as too speculative, decadent, and undisciplined.”<sup>34</sup> In any case, both Hong Kong and Singapore have been instrumental in helping to spread the idea of a Pacific century. The Chinese communities in Malaysia, Thailand, Indonesia, and the Philippines are similarly participating in transforming these societies into newly industrial countries.

32 Ibid., 79.

33 For example, the concern over the government’s efforts to boost Mandarin has prompted Singapore’s former foreign minister, S. Rajaratnam, to raise the issue of possible racial disharmony as a result of emphasis on Chineseness. *Far Eastern Economic Review*, January 24, 1991, 19.

34 Vogel, *The Spread of Industrialization in East Asia*, 82.

The amazing aspect of all these scenarios is the glaring absence of mainland China. For thirty years (1949 – 1979), hostile external conditions and self-imposed isolation made the People’s Republic of China largely irrelevant to the rise of industrial East Asia. In the last decade, as the resumption of tourism, trade, and scholarly exchange thrust new responsibilities upon the Beijing regime, the Chinese intellectual community as well as the official establishment were appalled to discover that while the periphery of the Sinic world was proudly marching toward an Asia-Pacific century, the homeland remained mired in perpetual underdevelopment. Despite the insistence of the Beijing government to define China’s coming of age strictly in terms of the Four Modernizations – agriculture, industry, national defense, and science and technology – issues of political and social restructuring have been raised and not only by dissidents, but also by intellectuals in state organizations, research institutions, and universities. The “barracks mentality” is no longer tenable.

Although the phenomenon of Chinese culture disintegrating at the center and later being revived from the periphery is a recurring theme in Chinese history, it is unprecedented for the geopolitical center to remain entrenched while the periphery presents such powerful and persistent economic and cultural challenges. Either the center will bifurcate or, as is more likely, the periphery will come to set the economic and cultural agenda for the center, thereby undermining its political effectiveness.

Cultural China can be examined in terms of a continuous interaction among three symbolic universes.<sup>35</sup> The first

35 The term “cultural China” (*wenhua zhongguo*) is relatively new. A similar concept “over-

consists of mainland China, Taiwan, Hong Kong, and Singapore – that is, the societies populated predominantly by cultural and ethnic Chinese. The second consists of Chinese communities throughout the world, including a politically significant minority in Malaysia and a numerically negligible minority in the United States. These Chinese, estimated to number from twenty to thirty million, are often referred to by the political authorities in Beijing and Taipei as *huaqiao* (overseas Chinese).<sup>36</sup> More recently, however, they tend to define themselves as members of the Chinese “diaspora,” meaning those who have settled in scattered communities of Chinese far from their ancestral homeland. While the Han Chinese constitute an overwhelming majority in each of the four areas in the first symbolic universe, communities of the Chinese diaspora – with the exception of Malaysia already

---

seas China” (*haiwai zhonghua*) was used to designate Chinese communities outside mainland China. Since this term carries the political connotation of a Chinese-style commonwealth encompassing the mainland, Hong Kong, Taiwan, and Singapore, it has generated much controversy on both sides of the Taiwan Straits. I propose adopting the term “cultural China,” applied loosely, and invite the participation of all those trying to understand and bring understanding to China and Chinese culture, thus the idea of an intellectual discourse. See Tu Wei-ming, “‘Wenhua zhongguo’ chutan” (Probing “cultural China”), *The Nineties* (*Jiushi niandai*) 245 (June 1990): 60–61.

36 For scholarly treatment of the subject, see Wang Gungwu, *A Short History of the Nanyang Chinese* (Singapore: Eastern University Press, 1959); Edgar Wickberg, *The Chinese in Philippine Life, 1850–1898* (New Haven, Conn.: Yale University Press, 1965); G. William Skinner, *Chinese Society in Thailand* (Ithaca, N.Y.: Cornell University Press, 1957); and James L. Watson, *Emigration and Chinese Lineage: The Mans in Hong Kong and London* (Berkeley: University of California Press, 1975).

mentioned – rarely have a population exceeding 3 percent.

The third symbolic universe consists of individuals, such as scholars, teachers, journalists, industrialists, traders, entrepreneurs, and writers, who try to understand China intellectually and bring their conceptions of China to their own linguistic communities. For the last four decades the international discourse on cultural China has unquestionably been shaped by the third symbolic universe more than by the first two combined. Specifically, writings in English and in Japanese have had a greater impact on the intellectual discourse on cultural China than those written in Chinese. For example, Chinese newspapers abroad often quote sources from the *New York Times* and Japan’s *Asahi Shinbun* to enhance their credibility. The highly politicized Chinese media on both sides of the Taiwan Straits have yet to earn their reputation as reliable reporters and authoritative interpreters of events unfolding in their own domain. The situation, however, is rapidly changing. In cultural matters, the *New York Times* may be months out of date; the “River Elegy,” not to mention the so-called cultural fever,<sup>37</sup> did not catch the attention of Western journalists until months after it had engulfed the Chinese-speaking world. Japanese reporting also suffers from a lack of a systematic analysis of the cultural landscape. Still, foreign journalists continue to exert an unusually strong influence on the discourse of cultural China. Sinologists in North Amer-

37 Although the “River Elegy” had become a major topic of discussion in the Chinese mass media throughout the first and the second symbolic universes in the summer of 1988, with the exception of Fred Wakeman’s thoughtful piece in the *New York Review of Books*, references to the cultural phenomenon in the English-language publications began to appear in the spring of 1989.

ica, Japan, Europe, and even Australia have similarly exercised a great deal of power in determining the scholarly agenda for cultural China as a whole.

This tripartite division of cultural China is problematic. Hong Kong, Taiwan, and Singapore have much more in common with the Chinese diaspora than they do with mainland China. Despite Hong Kong's impending return to its homeland in 1997, an overwhelming majority of the working class as well as the intellectuals, if offered the opportunity, would not choose to identify themselves as citizens of the People's Republic of China. Hong Kong is, at least in spirit, part of the Chinese diaspora. Although the Republic of Singapore is establishing full diplomatic ties with the People's Republic of China, Singapore's leaders have had closer contact with the Nationalist government in Taipei than with the Communists in Beijing. Nevertheless, Taiwan, Hong Kong, and Singapore are grouped together with mainland China as the first symbolic universe because the life orientation of each of these societies is based in Chinese culture. If we define being Chinese in terms of full participation in the economic, political, and social life of a Chinese community or civilization, the first symbolic universe offers both the necessary and the sufficient condition.

Divergence in economic development, political system, and social organization notwithstanding, the four members of the first symbolic universe share a common ethnicity, language, history, and worldview. To be sure, ethnic awareness has been diluted by the admixture of a variety of races that constitute the generic Han people; linguistic cohesiveness is threatened by the presence of numerous mutually incomprehensible "dialects" (in the case of Singapore, the situation is further confounded by multilingual-

ism); historical consciousness has been undermined by varying interpretations of "Confucian China and its modern fate"<sup>38</sup> and, with increasing rapidity, worldviews have been affected by the importation of radically different belief systems. Still, if we view cultural China as being a psychological as well as an economic and a political interchange, then the nature of the interactions between mainland China and Hong Kong, Taiwan, and Singapore is sufficient to group these distinct nations together as integral parts of the first symbolic universe. For the last ten years, the cultural impact that Hong Kong has had on mainland China as a whole – and metropolitan Guangzhou (Canton) and the Shenzhen Special Economic Zone in particular – has been profound; the Hong Kong transformation of mainland China is likely to become even more pronounced in the 1990s. The effect on the modernization of China due to the recent participation of Taiwanese and Singaporeans – as scholars, teachers, advisors, traders, journalists, and tourists – indicates clearly that the potential for Taiwanization and Singaporization of selected geographic regions and social strata of the mainland may be realized in the coming decades.

This does not necessarily mean that this perceived convergence will eventually lead to a reintegrated China as a civilization-state. It is more likely that, as the peripheral regions of mainland China become "contaminated" or "polluted" by Hong Kong, Taiwan, and Singapore, relative economic prosperity and cultural richness will bring about a measure of political independence. Despite post-Tiananmen speculation about

38 Joseph R. Levenson, *Confucian China and Its Modern Fate: A Trilogy* (Berkeley: University of California Press, 1968).

military warlordism, the rise of economic and cultural regionalism seems inevitable. Whether by choice or by default, a significantly weakened center may turn out to be a blessing in disguise for the emergence of a truly functioning Chinese civilization-state. Of course, the destructive power of the center is such that the transformative potential of the periphery can be easily stifled. The unpredictability of the Beijing leadership and the vulnerability of the status quo in Hong Kong, Taiwan, and Singapore make the first symbolic universe fluid and a fruitful interaction among its members difficult. In spite of the so-called Middle Kingdom syndrome, a Chinese civilization-state with a variety of autonomous regions or even a loosely structured Chinese federation of different political entities remains a distinct possibility.<sup>39</sup>

Nevertheless, we are well advised to heed the observation of Lucian Pye who maintains that “China is not just another nation-state in the family of nations,” rather “China is a civilization pretending to be a state.”<sup>40</sup> Actually, “the miracle of China has been its astonishing unity.” In trying to find an analogy in Western terms, Pye characterizes China of today “as if the Europe of the Roman Empire and of Charlemagne had lasted until this day and were now trying to function as a single nation-state.”<sup>41</sup> We may not accept Pye’s assertion that “the overpowering obligation felt by Chinese rulers to preserve the unity of their civilization has meant that there could be no com-

39 Yan Jiaqi, as a political scientist, has been a consistent and articulate advocate of a federated China. See notes 3 and 19.

40 Lucian Pye, “China: Erratic State, Frustrated Society,” *Foreign Affairs* 69 (4) (Fall 1990): 58.

41 Ibid.

promise in Chinese cultural attitudes about power and authority,” but his general statement is well taken: “The fact that the Chinese state was founded on one of the world’s great civilizations has given inordinate strength and durability to its political culture.”<sup>42</sup> The beguiling phenomenon of China as a civilization-state requires further elucidation.

The idea of the modern state involving power relationships based on competing economic and social interests is anathema to the Chinese cultural elite as well as the Chinese ruling minority. To them, the state – intent on realizing the historical mission to liberate China from threats of imperialist encroachment and the lethargy and stagnation of the feudal past – symbolizes the guardian of a moral order rather than the outcome of a political process. The state’s legitimacy is derived from a holistic orthodoxy informed by Sinicized Marxism-Leninism, rather than from operating principles refined by actual political praxis and codified in a legal system. The state’s claims on its people are comprehensive and the people’s dependence on the state is total; the state exemplifies the civilizational norms for the general public and the leadership assumes ideological and moral authority. The civilization-state exercises both political power and moral influence.

It should be acknowledged, however, that for all her power and influence, China as a civilization-state is often negligible in the international discourse on global human concerns. The marginalization of the Middle Kingdom to the periphery is, by now, so much an accepted fact in the contemporary world that it is virtually taken for granted, even among those of us committed to Chinese studies in the West. The asymme-

42 Ibid.

try between the centrality of its magnetic pull in cultural China and the marginality of its significance to the “global village” as a whole makes the first symbolic universe a challenging issue for analysis and contemplation.<sup>43</sup>

The second symbolic universe, the Chinese diaspora, presents equally intriguing conceptual difficulties. Diaspora, which literally means the scattering of seeds, has been used to refer to Jews outside Palestine after the Babylonian exile or to Jews living in a Gentile world. Until the establishment of the modern nation-state of Israel, the saliency of faith in God and its attendant observance of law and ritual, rather than the state, characterized the distinctive features of the Jewish religious community.<sup>44</sup> In contrast to this, the state, or more precisely China as a civilization-state, features prominently in the Chinese diaspora. Because the Chinese diaspora has never lost its homeland, there is no functional equivalent to the cathartic yearning for Jerusalem. Actually the ubiquitous presence of the Chinese state – its awe-inspiring physical size, its long history, and the numerical weight of its population – continues to loom large in the psychocultural constructs of diaspora Chinese. For many, the state, either Nationalist or Communist, controls the symbolic resources necessary for their cultural identity. Although dual citizenship is no longer operative, both Beijing and Taipei expect the loyal support of their *huaqiao* (overseas Chinese). Few diaspora Chinese ever speculate about

the possibility of China disintegrating as a unified civilization-state. The advantage of being liberated from an obsessive concern for China’s well-being at the expense of their own livelihood is rarely entertained. The diaspora Chinese cherish the hope of returning to and being recognized by the homeland. While the original meaning of scattering seeds suggests taking root and perpetuating away from the homeland, many diaspora Chinese possess a sojourner mentality and lack a sense of permanence in their adopted country. Some return “home” to get married or send their children back for a Chinese education; they remain in touch with relatives and friends who keep them informed of the economic and political climate at home.<sup>45</sup>

The Chinese settlers who are scattered around the world come, historically, from a few well-defined areas along the southeast coast of mainland China – notably Guangdong and Fujian. For a specific group of settlers, the province itself was too extensive and diffuse to accommodate an emotional identification with their homeland. Until the recent waves of immigration to North America which began after 1949, the overwhelming majority of Chinese Americans identified themselves not as Cantonese – which was too cosmopolitan a term to evoke any real sense of rootedness – but as natives of subprovincial districts, such as Taishan, Zhongshan, or Panyu. Similar phenomena occurred in Europe and Southeast Asia. As a rule, mutual aid associations in Malaysia, Thailand, and Indonesia were organized according to county or village, rather than provincial affiliations. Secret societies that crossed local boundaries were either politically

43 I thank sociologist Ambrose King of the Chinese University of Hong Kong for illuminating discussions on this issue.

44 Ernest W. Nicholson, *God and His People: Covenant and Theology in the Old Testament* (Oxford: Clarendon Press, 1986).

45 For a typical example of these patterns, see Clarence E. Glick, *Sojourners and Settlers: Chinese Migrants in Hawaii* (Honolulu: University of Hawaii Press, 1980), 172 – 181.



oriented or economically motivated. It is understandable, then, that the Chinese diaspora was, for decades, so fragmented that there was little communication among groups within a host nation, let alone any transnational cooperation.

Nevertheless, despite apparent parochialism, the overseas Chinese have managed to adapt themselves to virtually all types of communities throughout the world. The impression that the overall cultural orientation of Chinese settlers has been shaped predominantly by the magnetic power of the homeland is simplistic. The reason that the overseas Chinese rarely consider themselves thoroughly assimilated in their adopted countries is much more complex. In the United States, racial discrimination against the Chinese was, until recently, blatant; the Chinatown mentality, as a response to the hostile environment, may be seen as a psychosocial defense and adaptation. The post-1949 immigrants from Taiwan and Hong Kong have developed entirely different patterns of assimilation. The arrival of "boat people" and refugees from the mainland has initiated yet another process where new style Chinatowns emerge in such unlikely places as the deep South and the Midwest in the United States.

The situation in Southeast Asia is radically different. The case of the Philippine Chinese, perhaps the smallest Chinese population of any major Southeast Asian country, merits special attention:

The history of the Chinese in the Philippine population is one of regular intermarriage with ethnic Filipinos and the generation, thereby, of a historically important body of Chinese *mestizos*. When waves of Chinese immigration have receded and the Chinese community was not replenished for a time, these *mestizos* have

flourished in business pursuits. And, as in Thailand, Chinese-Filipino intermarriage has produced, over long periods of time, much of the political, social and cultural leadership of the country.<sup>46</sup>

"In common with most of the rest of Southeast Asia," wrote Edgar Wickberg, "the Philippines has had little replenishment of its Chinese population since 1949." He further notes: "Other things being equal, then, we would expect such a population to be increasingly oriented towards the Philippines and Philippine culture and decreasingly interested in things Chinese."<sup>47</sup> This seems to be incongruent with the perceived phenomenon that the Chinese in the Philippines, unlike those in Thailand, have not yet been fully assimilated into the mainstream of Filipino society. In fact, their distinct Chineseness makes them vulnerable to nativistic assaults. Wickberg explains that, prior to 1975, "the Philippine policy of restricting certain occupations to citizens but making it difficult for Chinese to become citizens put the Chinese in an almost impossible situation."<sup>48</sup> It appears, therefore, that the push of local conditions as well as the pull of the homeland impels the Chinese to become unassimilable.

The story is complicated by the fact that Chinese-Filipino relations also figure prominently in determining the fate of the Chinese settlers in the Philippines. For example, in the 1950s and 1960s when the Nationalist government in Taiwan exerted profound influence in

46 Edgar Wickberg, "Some Comparative Perspectives on Contemporary Chinese Ethnicity in the Philippines," *Yazhou wenhua (Asian Culture)* (forthcoming), 5 (page numbers refer to his May 1989 manuscript).

47 *Ibid.*, 5 – 6.

48 *Ibid.*, 7.

the Philippines, “the Philippine government, on the whole, gave over to Taiwan the responsibility for defining the nature of the Chinese culture to be taught in the Philippine Chinese schools.”<sup>49</sup> As part of the united front between the Nationalist government and the Marcos regime to fight the spread of communism, Chinese schools were allowed to fly the flag of the Republic of China, to display pictures of Sun Yat-sen and Chiang Kai-shek, and to make use of textbooks from Taiwan. The recognition of Beijing in 1975 facilitated the Filipinization of the Chinese schools and prompted the Marcos government to grant full citizenship to about one-sixth of the entire Chinese population in the Philippines.

Perhaps the most encouraging sign was the approach of the newly established intellectual organization Kaisa Para Sa Kaunlaran, which advocated “the understanding and retention of one’s Chinese culture while fully identifying oneself with the Philippines and with Filipinos of non-Chinese backgrounds.”<sup>50</sup> Conceived in the 1970s by young university graduates of Chinese ancestry, the Kaisa vision intends to create a narrow ridge between cultural chauvinism and total assimilation. One of the most threatening issues confronting the Filipino-Chinese community is its public perception; although the Chinese are beneficiaries of the political and economic system, their contribution to social welfare is limited and their participation in the cultural life of the land is minimal. The resentment the local population feels toward the conspicuous consumption of the rich Chinese (for example, elaborate tombs in the style of miniature world-class hotels), which has often led to anti-Chinese riots

49 Ibid., 9.

50 Ibid., 21.

against the peddlers and retailers in Chinatowns in the past, remains a haunting memory.

The precarious nature of being Chinese in Southeast Asia is amply demonstrated by the institutionalized mechanism of desinification in Malaysia and Indonesia. For political reasons, the Malaysian and Indonesian governments consider Chineseness a potential threat to national security, not to mention national integration. Among the most tragic events in the second half of the twentieth century were the atrocities committed against the Chinese population in Indonesia in 1965, which were brought on by a perceived threat of Communist takeover. Between 250,000 and 750,000 people died in a matter of months, due, in part, to a coup d’état engineered by President Suharto. This Indonesian Chinese “holocaust” received little attention in the first symbolic universe of cultural China. The mainland was embroiled in its own holocaust, the Cultural Revolution; Taiwan condoned the heavy-handed attack on Communism; Hong Kong was too remote to be affected; and Singapore’s proximity to Indonesia – both geographically and politically – made it too vulnerable to offer a response. It was actually in the same period that growing anti-Chinese sentiment in Malaysia pushed Singapore to become an independent state.

The second symbolic universe, the Chinese diaspora, was too fragmented and isolated to even take notice of the tragedy. Malaysian Chinese, Thai Chinese, Philippine Chinese, and American Chinese were aware of what happened, but there was neither the infrastructure nor the resources to mount a transnational demonstration. In fact, the word *Chinese* qualified by *Malaysian*, *Thai*, *Philippine*, and *American* did not signify any underlying consciousness of ethnic

or cultural identity; these terms were used generically to designate communities that were culturally and racially similar, but which were otherwise totally unrelated. It is ironic that it was the third symbolic universe – consisting primarily of non-Chinese but who were committed, informed, and often sympathetic observers of things Chinese – that reacted most strongly to the holocaust and exposed it to the world at large.

Recent events have greatly improved the atmosphere for the Chinese in Southeast Asia, although the “Chinese question” continues to be a sensitive subject. Still, in Malaysia and Indonesia, being Chinese remains a stigma; things Chinese – especially symbols of Chinese high culture such as the written script – are viewed with suspicion. The economic success of the Chinese makes them hungry for cultural expression, and the host countries, while tolerating their economic well-being, are adamant about imposing cultural prohibitions. Signs of a Kaisa-like solution to the conflict between the political loyalty and the cultural identity of Chinese in Malaysia and Indonesia are yet to be found.

After having been ostracized from the diplomatic community of ASEAN for more than a decade, Taiwan is now returning as an investment giant. Records show that Taiwanese investments in the Philippines and Malaysia have taken the lead and now surpass those of Japan by a respectable margin. Taiwan’s presence in Indonesia is significant enough to have persuaded the Suharto government to relax its prohibition against Chinese schools, Chinese videocassettes, and publications in Chinese, which inspires a new vitality in Indonesian Chinese communities. Furthermore, operating with the full collaboration of merchants of Chinese origin in Bangkok, Taiwanese capital has also contributed to the eco-

nomie dynamism in Thailand. An obvious consequence for the second symbolic universe is the latent tension and visible conflict between Taiwan’s economic strength and the mainland’s political clout. The drama of the competition between the mainland and Taiwan is not confined to the ASEAN countries; the intensity is felt by Chinese communities in Tokyo, Paris, New York, San Francisco, Toronto, and Sydney. Although it is too early to tell whether a depoliticized cultural agenda will emerge as a result of this confrontation, it seems that Singapore may play a vital role in addressing economic and cultural issues and transcend the political animosity which exists on both sides of the Taiwan Straits.

Another example of the impact of the first symbolic universe on the second is the emigration of professionals from Hong Kong to North America and Australia. As 1997 draws near, Hong Kong emigrants with substantial capital and professional expertise are making their presence known in Chinese communities in Toronto, Vancouver, Los Angeles, San Francisco, New York, and Sydney. This seems symptomatic of a broader pattern: Chinese immigrants in these cities are also coming from mainland China, Taiwan, Singapore, Malaysia, Indonesia, the Philippines, and Vietnam. What we are witnessing, then, is a new era of the Chinese diaspora.

This phenomenon which historian Wang Gungwu, vice-chancellor of Hong Kong University, aptly depicts as a remigration of Chinese to North America, Europe, and Australia, is unprecedented and requires closer examination.<sup>51</sup>

<sup>51</sup> Remarks by Wang Gungwu at the conference on “The Meaning of Being Chinese,” East-West Center, October 24 – 26, 1990. See conference proceedings, Institute of Culture and Communication, East-West Center, Honolulu.

These financially secure Malaysian, Indonesian, Filipino, and Vietnamese Chinese have ostensibly emigrated from their adopted homelands for several generations in order to escape from policies which discriminated against their Chineseness. In order to combat the pressure to assimilate imposed by the new nation-states in Southeast Asia and to preserve a measure of Chineseness for their descendants, they have opted to immigrate to modern Western-style nations with strong democratic traditions. The irony of not returning to their ancestral homeland but going far away from China with the explicit intention of preserving their cultural identity seems perplexing, but as Wang Gungwu perceptively remarks, the transformation from a sojourner mentality to deliberate emigration is a new phenomenon.

As recently as the 1960s, the decision to renounce Chinese nationality (whether Nationalist or Communist) and to adopt local citizenship was, for many Chinese in the diaspora, a matter of great agony.<sup>52</sup> The massive exodus of the most brilliant Chinese intellectuals from the mainland during the last decade shows clearly that the civilization-state has lost much of its iron grip on the Chinese intelligentsia, and the Tiananmen brutality may have irreversibly severed the emotional attachment of the diaspora Chinese to the homeland. The meaning of being Chinese, a question that has haunted Chinese intellectuals for at least three generations, has taken on entirely new dimensions.

The term *cultural China*, coined in the last decade or so and often seen in intellectual journals outside mainland China, is itself an indication of the emergence of a “common awareness” (*gongshi*)

52 Ibid.

among Chinese intellectuals throughout the world. The presence of such an awareness prior to the opening up of mainland China in the late 1970s is made clear in the deliberate choice of *huaren* (people of Chinese origin) rather than *zhongguoren* (people of China, the state) to designate people of a variety of nationalities who are ethnically and culturally Chinese. *Huaren* is not geopolitically centered, for it indicates a common ancestry and a shared cultural background, while *zhongguoren* necessarily evokes obligations and loyalties of political affiliation and the myth of the Middle Kingdom. By emphasizing cultural roots, Chinese intellectuals in Taiwan, Hong Kong, and North America hoped to build a transnational network for understanding the meaning of being Chinese within a global context. For these intellectuals, the relevant political center that influenced their lives was the Nationalist government in Taiwan. Their efforts to depoliticize the cultural movement were an attempt to maintain a critical distance from the official anti-Communist line of the Guomindang (the Nationalist Party).

In the 1980s, with the advent of mainland China as an active participant in the discourse on cultural China, the symbol of *huaren* assumed a new significance: how could the overseas Chinese help the homeland to modernize? On the intellectual side, an unintended consequence of Deng Xiaoping's economic reforms was “cultural fever,” brought on by a revival of communication in the social sciences and humanities between scholars in mainland China and scholars abroad. The Tiananmen tragedy on June 4, 1989, symbolizes the near-total alienation of the Chinese intelligentsia from the ruling minority on the mainland. It is highly unlikely that the political regime that has brutally massacred peaceful demon-

strators and bystanders will ever be able to win back the hearts and minds of the intellectuals and those citizens who are committed to the dignity of China as a civilization.

The fate of the Chinese intelligentsia in the People's Republic of China inevitably elicits the horrifying question, How could the scholar, honored as a paradigm of the personality ideal in Chinese culture, have stooped so low for so long? The answer lies, in part, in the coexistence of political nationalism and cultural iconoclasm among the most articulate intellectual elite. The decline of China from being the Middle Kingdom for centuries to the "Sickman of East Asia" in just two generations time (beginning with the Opium War of 1839 and culminating in the collapse of the Hundred-Day reform in 1898), in conjunction with the disintegration of the Chinese political order, created such spiritual turmoil among the Chinese intelligentsia that the reconstruction of a political center became an overriding concern. Intent on creating the optimum conditions for China to recapture its position of wealth and power once again, the Westernized intellectuals launched a frontal attack on Confucian tradition: Confucianism was perceived to have nurtured a "national character" (*guominxing*) detrimental to China's modernization. The desire to increase China's chance of survival was therefore linked to an all-out attack on the very tradition which had shaped Chineseness throughout history.

This assertion – that we must totally reject that which has made us what we are – enabled the most forward-looking Chinese intellectuals to be receptive to foreign ideologies while still maintaining their nationalistic objectives. The May Fourth patriots experienced a keen sense of liberation when they confronted the national crisis by embracing vir-

tually all major Western philosophical currents of thought, including Dewey's pragmatism, Bergson's vitalism, Bakunin's anarchism, and Russell's empiricism. What was conspicuously absent was any persuasive form of fundamentalism or nativism that glorified Chinese culture for its own sake. However, beneath this intellectual commitment to alien Western values was a powerful surge of fundamentalistic and nativistic sentiment which was dangerously volatile among the Chinese populace throughout the country.

An unintentional and unfortunate consequence of this period of wholesale Westernization and anti-Confucianism was the marginalization of the intelligentsia from the center of the political arena. The thrust of their intellectual quest was the establishment of a political center; yet such a focus relegated them to the background. Furthermore, their demand for action was so overwhelming that the seeds of their own decline were embedded in the logic of the intellectual discourse. It is not surprising that Marxism-Leninism triumphed in the marketplace of ideologies; it met the requirements of both the cultural iconoclasts and political nationalists: it was Western to the core as the cultural iconoclasts had strongly recommended and its anti-imperialist stance was precisely what the political nationalists had demanded.

What the Chinese intelligentsia did not expect – and is still struggling to understand – is that the Party is not only the embodiment of socialist truth but also the bearer of the correct method for its eventual realization. The actual struggle undertaken by the masses (the peasants, the workers, and the soldiers) was too rooted in Chinese soil to benefit from the sophisticated intellectual consciousness framed in Western liber-



al democratic terms. The rise of Mao Zedong to the trinity of political leader, ideological teacher, and moral exemplar, though unprecedented in Chinese history, can be explained in terms of a fundamentalistic-nativistic challenge to the Westernization process as envisioned by the May Fourth intellectuals. In examining Mao Zedong's ideology we find, among other things, a combination of iconoclasm and nationalism; however, the iconoclasm is layered with numerous sediments of nativistic pathos, and the nationalism is imbued with fundamentalistic claims to China's uniqueness. Since his death, Chinese intellectuals may have radically changed their minds about Mao as the savior of the Chinese people, but for decades they were awed by his sagacity and, occasionally, charmed by his earthiness. The demonic power of destruction, which Mao unleashed repeatedly, stunned intellectuals to such a degree that they lost their ability even to describe it. Indeed, they have yet to develop adequate conceptual apparatuses to analyze that phenomenon, including their own roles as participants (willing or otherwise) and as victims.

Collective amnesia is so pervasive in China that the national memory has difficulty extending back even to the decade of the Cultural Revolution (1966 – 1976), let alone to the disaster of the Great Leap Forward (1958 – 1960) or the brutality of the Anti-Rightist campaign (1957 – 1958). Virtually all intellectuals of note were purged during the Anti-Rightist campaign that followed the short-lived domestic liberalization in the wake of Khrushchev's de-Stalinization in the Soviet Union. The Great Leap Forward – an ill-conceived utopian experiment intended to enable China to surpass the West in industrial productivity within fifteen years – in combination

with natural disasters led to massive starvation, killing an estimated 40 million people. Subsequently, neither the Party, nor the leadership, nor Mao was held accountable. In fact, Mao, disgusted with the inertia of the leadership and the bureaucratism of the Party, managed to rouse the Chinese youth to a crescendo of iconoclasm and nationalism by launching the Cultural Revolution in 1966. Yet, as some of the most perceptive minds in China have confessed, their faith in the truth of Marxism-Leninism, in the credibility of the Chinese Communist leadership, and in the legitimacy of the Party was not shaken until the mid-1980s.

In the spring of 1990, Chinese intellectuals worldwide developed a truly new, communal, critical self-consciousness in which the agenda of iconoclasm and nationalism was reversed; a search for cultural roots and a commitment to a form of depoliticized humanism became a strong voice in the discourse on cultural China.

China has witnessed much destructiveness and violence in her modern transformation. The agonizing question for us all in the three symbolic universes is raised with great poignancy by Stevan Harrell: "Why does a culture that condemns violence, that plays down the glory of military exploits, awards its highest prestige to literary, rather than martial figures, and seeks harmony over all other values, in fact display such frequency and variety of violent behavior, that is of the use of physical force against persons?"<sup>53</sup> Echoing Harrell's

53 Stevan Harrell, "Introduction," in Jonathan N. Lipman and Stevan Harrell, eds., *Violence in China: Essays in Culture and Counterculture* (Albany: State of New York University Press, 1990), 1. Harrell adds, "Whether Chinese culture is more violent than other cultures is dif-

puzzlement and frustration, Andrew Nathan, in a thought-provoking review essay, cites the condemnation of the authors of the aforementioned “River Elogy”: “What Confucian culture has given us over the past several thousand years is not a national spirit of enterprise, a system of laws, or a mechanism of cultural renewal, but a fearsome self-killing machine that, as it degenerated, constantly devoured its best and its brightest, its own vital elements.”<sup>54</sup> This is reminiscent of Lu Xun’s bitter satire against the Confucian legacy, which he mordantly denounced as cannibalistic ritualism.<sup>55</sup>

In retrospect, what the Chinese intelligentsia collectively experienced in the twentieth century is what Mark Elvin pointedly characterizes as the “double disavowal” of both Confucianism and Marxism.<sup>56</sup> The same indignation that Lu Xun’s generation felt about Confucian authoritarianism is now being expressed against Marxist totalitarian-

---

ficult to judge, but it is visibly not less violent than many.”

54 Quoted in Andrew J. Nathan, “A Culture of Cruelty,” *The New Republic* 203 (5) (July 30 – August 6, 1990): 32.

55 Lu Hsun, “Suigan lu sanshiba” (Random thoughts, no. 38), *Xin qingnian (The New Youth)* 5 (5) (November 15, 1918): 515 – 518. For an analysis of Lu Xun’s “complex consciousness,” see Leo Ou-fan Lee, *Voices from the Iron House: A Study of Lu Xun* (Bloomington: Indiana University Press, 1987).

56 Mark Elvin, “The Double Disavowal: The Attitudes of Radical Thinkers to the Chinese Tradition,” in Y. M. Shaw, ed., *China and Europe in the Twentieth Century* (Taipei: Institute of International Relations, 1986), 112 – 137. For his more recent view on the extinction of Confucianism, see Mark Elvin, “The Collapse of Scriptural Confucianism,” *Papers on Far Eastern History*, no. 41 (Canberra: The Australian National University, March 1990).

ism. Many intellectuals strongly believe that the collusion of the feudal past and the socialist present makes China a victim of a double betrayal. This, in a substantial way, explains the vehemence with which the authors of the “River Elogy” attacked the Confucian legacy and the enthusiasm they had for embracing the modern West. The matter, however, is complicated by the fact that the real challenge to the mainland Chinese intellectuals is not the modern West per se but the modern West mediated through industrial East Asia.

While Lu Xun’s generation, despite Spengler’s warning, never entertained the possibility of a path to modernity other than Westernization, the authors and producers of the “River Elogy” could not help but explore the courses of action most congenial to the Chinese situation. If Japan, South Korea, Taiwan, Hong Kong, Singapore, and Chinese communities throughout the world have shown not only the relevance of Confucian ethics to their modus operandi but also the dynamics of the Confucian tradition in shaping their forms of life, then the existential predicament of the mainland intellectual caught between a contemptible past and a brutal present is not indissoluble. Notwithstanding that “the inner strength of the Chinese intelligentsia has been sapped by the collusion of feudal Chinese traditionalism (the remnants of a politicized Confucian moralism) and the modern Western collectivism (the outmoded practice of Leninist dictatorship),”<sup>57</sup> the fruitful interaction between Confucian humanism and democratic liberalism in cultural China has already occurred. The authors and producers of the “River

57 Tu Wei-ming, “Iconoclasm, Holistic Vision, and Patient Watchfulness: A Personal Reflection on the Modern Chinese Intellectual Quest,” *Dædalus* 116 (2) (Spring 1987): 94.

Elegy,” some now as scholars in exile, have also begun to explore traditional symbolic resources (including those in Confucian humanism) in order to reformulate their strategy for China’s cultural reconstruction.<sup>58</sup>

The so-called Third Epoch of Confucian Humanism<sup>59</sup> may have been the wishful thinking of a small coterie of academicians, but the emergence of a new inclusive humanism with profound ethical-religious implications for the spiritual self-definition of humanity, the sanctity of the earth, and a form of religiousness based on immanent transcendence has already been placed on the agenda in cultural China. The real challenge to this new inclusive humanism is the narrowly conceived anthropocentricity informed by instrumental rationality and fueled by a Faustian drive to conquer and destroy. While the modern West has created virtually all major spheres of value for the twentieth century (science, technology, the free market, democratic institutions, metropolises, and mass communication, for example), the painful realization that it has also pushed humanity to the brink of self-destruction engenders much food for thought. The question of whether human beings are, in fact, a viable species is now being asked with a great sense of urgency.

It is ironic that, for the first half of the century, a major concern for the Chinese political leaders – notably Sun Yat-sen, Chiang Kai-shek, and Mao Zedong – was

58 Su Xiaokang, “Traditional Resources and Modernization: A Talk with Professor Tu Wei-ming in Paris,” *Minzhu Zhongguo (Democratic China)* 1 (2) (June 1990): 52–55.

59 Tu Wei-ming, “Toward a Third Epoch of Confucian Humanism: A Background Understanding,” in Irene Eber, ed., *Confucianism: Dynamics of Tradition* (New York: Macmillan, 1986), 3–21.

the very survival of the children of the Yellow Emperor. The fear, far from that of a population explosion, was actually the depletion of the Chinese race in the social Darwinian sense. With a view toward the future we need to ask, what form of life do the Chinese need to pursue that is not only commensurate with human flourishing but also sustainable in ecological and environmental terms?

What mainland China eventually will become remains an overriding concern for all intellectuals in cultural China. She may try to become a mercantilist state with a vengeance; she may continue to be mired in her inertia and inefficiency for years to come; or she may modernize according to a new holistic humanist vision. Saddled with a population burden approaching 1.2 billion, can this state succeed at any of these ambitions without first finding a viable way to liberate the energies of its people? Although realistically, those who are on the periphery (the second and third symbolic universes plus Taiwan, Hong Kong, and Singapore) are seemingly helpless in affecting any fundamental transformation of China proper, the center no longer has the ability, insight, or legitimate authority to dictate the agenda for cultural China. On the contrary, the transformative potential of the periphery is so great that it seems inevitable that it will significantly shape the intellectual discourse on cultural China for years to come. It is perhaps premature to announce that “the center is nothing, whereas the periphery is everything,”<sup>60</sup> but undeniably, the fruitful interaction among a va-

60 Professor Yu Ying-shih of Princeton, perhaps in jest, proposed this as the theme of his essay. If his schedule had permitted him to write a paper for the present volume, he would have underscored the intellectual and political significance of the periphery in cultural China.

riety of economic, political, social, and cultural forces at work along the periphery will activate the dynamics of cultural China.

The exodus of many of the most brilliant minds from the mainland, the emigration of Chinese professionals from Hong Kong, and the remigration of middle-class Chinese from Southeast Asia to North America and Australia suggest that it is neither shameful nor regrettable to voluntarily alienate oneself from a political regime that has become culturally insensitive, publicly unaccountable, and oppressive to basic human rights. The meaning of being Chinese is basically not a political question; it is a human concern pregnant with ethical-religious implications.

Is it possible to live a meaningful life as a Chinese individual if the dignity of one's humanity is lost? Does citizenship of a Chinese national state guarantee one's Chineseness? As a precondition for maintaining one's Chineseness, is it necessary to become a full participating citizen of one's adopted country? While the overseas Chinese (the second symbolic universe) may seem forever peripheral to the meaning of being Chinese, can they assume an effective role in creatively constructing a new vision of Chineseness that is more in tune with Chinese history and in sympathetic resonance with Chinese culture? Is it possible and even desirable for someone in the third symbolic universe who is not proficient in the Chinese language and who has no Chinese family ties by birth or marriage to acquire an understanding of Chinese culture such that he or she can greatly shape the intellectual discourse on cultural China and significantly contribute to the definition of being Chinese? An obvious no to the first two and a resounding yes to each of the remaining questions will give rich tex-

ture to the provocative inquiry into the meaning of being Chinese.

#### ACKNOWLEDGMENTS

I am indebted to David Chu and Rosanne Hall for searching criticisms of early versions of this paper.

*Cultural  
China:  
the periph-  
ery as the  
center*

# Amartya Sen

## *Indian traditions & the Western imagination*

This essay is concerned with Western images of Indian intellectual traditions and the interactions between those representations and a contemporary “internal” understanding of Indian culture.<sup>1</sup> I focus particularly on the elementary diversities that characterize Indian society and its intellectual traditions, as well as on the biases that result from paying inadequate attention to them. In an obvious way, this applies to seeing India as a “mainly Hindu” country (as Western newspapers often describe India, as do the newly powerful Hindu political parties within India); this “mainly Hindu” country is also the third-largest Muslim country in the world (with nearly 110 million Muslims).

Less conspicuously, the contrast applies also to Indian intellectual traditions. This home of endless spirituali-

ty has perhaps the largest atheistic and materialist literature of all the ancient civilizations. To be sure, this accounting of the *amount* of unorthodox writing may be a little misleading, since Indian traditions are characterized by some prolixity. For example, the Sanskrit epic *Mahabharata*, which is often compared with the *Iliad* and the *Odyssey*, is in fact seven times as long as the *Iliad* and the *Odyssey* put together. One of the more striking Bengali verses I remember from my childhood is a lamentation about the tragedy of death in a nineteenth-century poem: “Just consider how terrible the day of your death will be. / Others will go on speaking, and you will not be able to respond.” But even this extreme fondness for speech is associated with an enormous heterogeneity of programs and preoccupations. Irreducible diversity is perhaps the most important feature of Indian intellectual traditions.

The self-images (or “internal identities”) of Indians have been extremely af-

---

*Amartya Sen, recipient of the 1998 Nobel Prize in Economics, has been a Fellow of the American Academy since 1981. Sen is Lamont University Professor and Professor of Economics and Philosophy at Harvard University, positions he held at the time of this essay's appearance in the Spring 1997 issue of "Daedalus." He also served as Master of Trinity College in Cambridge, U.K., between 1998 and 2004.*

---

© 2005 by the American Academy of Arts & Sciences

1 This essay draws on an earlier article entitled “India and the West,” *The New Republic* (June 7, 1993). For helpful discussions, I am grateful to Akeel Bilgrami, Sugata Bose, Barun De, Jean Drèze, Ayesha Jalal, Dharma Kumar, V. K. Ramachandran, Tapan Raychaudhuri, Emma Rothschild, Lloyd Rudolph, Suzanne Rudolph, Ashutosh Varshney, Myron Weiner, and Nur Yalman.



fectured by colonialism over the past centuries and are much influenced – both collaterally and dialectically – by the impact of outside imagery (what we may call “external identity”). However, the direction of the influence of Western images on internal Indian identities is not altogether straightforward. In recent years, separatist resistance to Western cultural hegemony has led to the creation of significant intellectual movements in many postcolonial societies – not least in India. This has particularly drawn attention to the important fact that the self-identity of postcolonial societies is deeply affected by the power of the colonial cultures and their forms of thought and classification. Those who prefer to pursue a more “indigenous” approach often opt for a characterization of Indian culture and society that is rather self-consciously “distant” from Western traditions. There is much interest in “recovering” a distinctly Indian focus in Indian culture.

I would argue that this stance does not take adequate note of the dialectical aspects of the relationship between India and the West and, in particular, tends to disregard the fact that the external images of India in the West have often tended to *emphasize* (rather than downplay) the differences – real or imagined – between India and the West. Indeed, I propose that there are reasons why there has been a considerable Western inclination in the direction of “distancing” Indian culture from the mainstream of Western traditions. The contemporary reinterpretations of India (including the specifically “Hindu” renditions), which emphasize Indian particularism, join forces in this respect with the “external” imaging of India (in accentuating the distinctiveness of Indian culture). Indeed, it can be argued that there is much in common between James Mill’s impe-

rialist history of India and the Hindu nationalist picturing of India’s past, even though the former image is that of a grotesquely primitive culture whereas the latter representation is dazzlingly glorious.

The special characteristics of the Western approaches to India have encouraged a disposition to focus particularly on the religious and spiritual elements in Indian culture. There has also been a tendency to emphasize the contrast between what is taken to be “Western rationality” and the cultivation of what “Westerners” would see as “irrational” in Indian intellectual traditions. While Western critics may find “antirationalism” to be defective and crude, and Indian cultural separatists may find it cogent and penetrating (and perhaps even “rational” in some deeper sense), they nevertheless agree on the existence of a simple and sharp contrast between the two heritages. The issue that has to be scrutinized is whether such a bipolar contrast is at all present in that form.

I will discuss these questions and argue that focusing on India’s “specialness” misses, in important ways, crucial aspects of Indian culture and traditions. The deep-seated heterogeneity of Indian traditions is neglected in these homogenized interpretations (even though the interpretations themselves are of different kinds). My focus will be particularly on images of Indian *intellectual* traditions, rather than on its creative arts and other features of social life. After distinguishing between three of the dominant approaches in Western interpretations of Indian intellectual traditions, I will consider what may appear to be the overall consequence of these approaches in Western images of India and its impact on both external and internal identities.

A dissimilarity of perceptions has been an important characteristic of Western interpretations of India, and several different and competing conceptions of that large and complex culture have been influential in the West. The diverse interpretations of India in the West have tended to work to a considerable extent in the same direction (that of accentuating India's spirituality) and have reinforced each other in their effects on internal identities of Indians. But this is not because the distinct approaches to India are not fundamentally different; they certainly are very disparate. The similarity lies more in their *impact* – given the special circumstances and the dialectical processes – than in their content.

The analysis to be pursued here would undoubtedly invite comparison and contrast with Edward Said's justly famous analysis of "Orientalism." Said analyzes the construction of the "Orient" in Western imagination. As he puts it, "The Orient is an idea that has a history and a tradition of thought, imagery, and vocabulary that have given it reality and presence *in and for the West*."<sup>2</sup> This essay has a much narrower focus than Said's, viz. India, but there is clearly an overlap of subject matter since India is a part of the "Orient." The main difference is at the thematic level. Said focuses on uniformity and consistency in a particularly influential Western characterization of the Orient, whereas I shall be dealing with several contrasting and conflicting Western approaches to understanding India.

Said explains that his work "deals principally not with a correspondence between Orientalism and Orient, but with the internal consistency of Orient-

2 Edward W. Said, *Orientalism* (New York: Random House, 1978; Vintage Books, 1979), 5; italics added.

talism and its ideas about the Orient."<sup>3</sup> I would argue that unless one chooses to focus on the evolution of a specific conceptual tradition (as Said, in effect, does), "internal consistency" is precisely the thing that is terribly hard to find in the variety of Western conceptions of India. There are several fundamentally contrary ideas and images of India, and they have quite distinct roles in the Western understanding of the country and also in influencing self-perceptions of Indians.

Attempts from outside India to understand and interpret the country's traditions can be, I would argue, put into at least three distinct categories, which I shall call *exoticist* approaches, *magisterial* approaches, and *curatorial* approaches.<sup>4</sup> The first (*exoticist*) category concentrates on the wondrous aspects of India. The focus here is on what is different, what is strange in the country that, as Hegel put it, "has existed for millennia in the imagination of the Europeans."

The second (*magisterial*) category strongly relates to the exercise of imperial power and sees India as a subject territory from the point of view of its British governors. This outlook assimilates a sense of superiority and guardianhood needed to deal with a country that James Mill defined as "that great scene of British action." While a great many British observers did not fall into this category (and some non-British ones did), it is hard to dissociate this category from the task of governing the Raj.

The third (*curatorial*) category is the most catholic of the three and includes

3 Ibid., 5.

4 In the earlier article "India and the West" on which this essay draws, the third category was called "investigative" rather than "curatorial"; the latter is more specific and I believe somewhat more appropriate.

various attempts at noting, classifying, and exhibiting diverse aspects of Indian culture. Unlike the exoticist approaches, a curatorial approach does not look *only* for the strange (even though the “different” must have more “exhibit value”), and unlike the magisterial approaches, it is not weighed down by the impact of the ruler’s priorities (even though the magisterial connection would be hard to avoid *altogether* when the author is also a member of the ruling imperial elite, as they sometimes were). For these reasons, there is more freedom from preconceptions in this third category. On the other hand, the curatorial approaches have perspectives of their own, with a general interest in seeing the object – in this case, India – as very special and extraordinarily interesting.

Other categories can be proposed that are not covered by any of the three. Also, the established approaches can be reclassified according to some *other* organizing principle. I am not claiming any grand definitive status of this way of seeing the more prominent Western approaches to India. However, for the purpose of this essay, I believe this threefold categorization is useful.

I shall begin by considering the curatorial approaches. But first I must deal with a methodological issue, in particular, the prevalent doubts in contemporary social theory about the status of intellectual curiosity as a motivation for knowledge. In particular, there is much skepticism about the possibility of any approach to learning that is innocent of power. That skepticism is justified to some extent since the motivational issues underlying any investigation may well relate to power relations, even when that connection is not immediately visible.

Yet people seek knowledge for many different reasons, and curiosity about

unfamiliar things is certainly among the possible reasons. It need not be seen as a figment of the deluded scientist’s imagination, nor as a tactical excuse for some other, ulterior preoccupation. Nor does the pervasive relevance of different types of motivation have the effect of making all the different observational findings *equally* arbitrary. There are real lines to be drawn between inferences dominated by rigid preconceptions (for example, in the “magisterial” approaches, to be discussed presently) and those that are not so dominated.

There is an interesting methodological history here. The fact that knowledge is often associated with power is a recognition that had often received far too little attention in traditional social theories of knowledge. But in recent social studies, the remedying of that methodological neglect has been so comprehensive that we are now in some danger of ignoring other motivations altogether that may not link directly with the seeking of power. While it is true that any useful knowledge gives its possessor some power in one form or another, this may not be the most remarkable aspect of that knowledge, nor the primary reason for which this knowledge is sought. Indeed, the process of learning can accommodate considerable motivational variations without becoming a functionalist enterprise of some grosser kind. An epistemic methodology that sees the pursuit of knowledge as entirely congruent with the search for power is a great deal more cunning than wise. It can needlessly undermine the value of knowledge in satisfying curiosity and interest; it significantly weakens one of the profound characteristics of human beings.

The curatorial approach relates to systematic curiosity. People are interested in other cultures and different lands, and investigations of a country and its tradi-

tions have been vigorously pursued throughout human history. Indeed, the development of civilization would have been very different had this not been the case. The exact motivation for these investigations can vary, but the inquiries need not be hopelessly bound by some overarching motivational constraint (such as those associated with the exoticist or magisterial approaches). Rather, the pursuit may be driven primarily by intellectual interests and concerns. This is not to deny that the *effects* of these investigative pursuits may go well beyond the motivating interests and concerns, nor that there could be *mixed motivations* of various kinds, in which power relations play a collateral role. But to deny the role of curiosity and interest as powerful motivational features in their own right would be to miss something rather important. For the curatorial approaches, that connection is quite central.

A fine example of a curatorial approach to understanding India can be found in Alberuni's *Ta'rikh al-hind* (*The History of India*), written in Arabic in the early eleventh century.<sup>5</sup> Alberuni, who was born in Central Asia in A.D. 973, first came to India accompanying the marauding troops of Mahmud of Ghazni. He became very involved with India and mastered Sanskrit; studied Indian texts on mathematics, natural sciences, literature, philosophy, and religion; conversed with as many experts as he could find; and investigated social conventions and practices. His book on India presents a remarkable account of the intellectual traditions and social customs of early eleventh-century India.

Even though Alberuni's was almost certainly the most impressive of these

investigations, there are a great many examples of serious Arabic studies of Indian intellectual traditions around that time.<sup>6</sup> Brahmagupta's pioneering Sanskrit treatise on astronomy had first been translated into Arabic in the eighth century (Alberuni retranslated it three centuries later), and several works on medicine, science, and philosophy had an Arabic rendering by the ninth century. It was through the Arabs that the Indian decimal system and numerals reached Europe, as did Indian writings in mathematics, science, and literature.

In the concluding chapter of his book on India, Alberuni describes the motivation behind his work thus: "We think now that what we have related in this book will be sufficient for any one who wants to converse with [the Indians], and to discuss with them questions of religion, science, or literature, on the very basis of their own civilization."<sup>7</sup> He is particularly aware of the difficulties of achieving an understanding of a foreign land and people, and specifically warns the reader about it:

...in all manners and usages, [the Indians] differ from us to such a degree as to frighten their children with us, with our dress, and our ways and customs, and as to declare us to be devil's breed, and our

6 See Wilhelm Halbfass, *India and Europe: An Essay in Understanding* (New York: State University of New York Press, 1988), chap. 2.

7 *Alberuni's India*, pt. II, chap. LXXX, 246. The same Arabic word was commonly used for "Hindu" and "Indian" in Alberuni's time. While the English translator had chosen to use "Hindus" here, I have replaced it with "Indians" in view of the context (to wit, Alberuni's observations on the inhabitants of India). This is an issue of some interest in the context of the main theme of this essay, since the language used here in the English translation to refer to the inhabitants of India implicitly involves a circumscribed ascription.

5 See *Alberuni's India*, trans. E. C. Sachau, ed. A. T. Embree (New York: Norton, 1971).

doings as the very opposite of all that is good and proper. By the bye, we must confess, in order to be just, that a similar depreciation of foreigners not only prevails among us and [the Indians], but is common to all nations towards each other.<sup>8</sup>

While Arab scholarship on India provides plentiful examples of curatorial approaches in the external depiction of India, they are not, of course, unique in this respect. Chinese travelers Fa Hsien and Hsuan Tsang, who spent many years in India in the fifth and seventh centuries A.D. respectively, provided extensive accounts of what they saw. While they had gone to India for Buddhist studies, their reports cover a variety of Indian subjects, described with much care and interest.

Quite a few of the early European studies of India must also be put in this general category. A good example is the Italian Jesuit Roberto Nobili, who went to south India in the early seventeenth century, and whose remarkable scholarship in Sanskrit and Tamil permitted him to produce quite authoritative books on Indian intellectual discussions, in Latin as well as in Tamil. Another Jesuit, Father Pons from France, produced a grammar of Sanskrit in Latin in the early eighteenth century and also sent a collection of original manuscripts to Europe (happily for him, the Bombay customs authorities were not yet in existence then).

However, the real eruption of European interest in India took place a bit later, in direct response to British – rather than Italian or French – scholarship on India. A towering figure in this intellectual transmission is the redoubtable William Jones, the legal scholar and officer of the East India Company, who went to India in 1783 and by the follow-

ing year had established the Asiatic Society of Bengal with the active patronage of Warren Hastings. In collaboration with scholars such as Charles Wilkins and Thomas Colebrooke, Jones and the Asiatic Society did a remarkable job in translating a number of Indian classics – religious documents (such as the *Gita*) as well as legal treatises (particularly, *Manusmriti*) and literary works (such as Kalidasa's *Śakuntala*).

Jones was quite obsessed with India and declared his ambition “to know India better than any other European ever knew it.” His description of his selected fields of study included the following modest list:

... the Laws of the Hindus and the Mohamedans, Modern Politics and Geography of Hindustan, Best Mode of Governing Bengal, Arithmetic and Geometry, and Mixed Sciences of the Asiaticks, Medicine, Chemistry, Surgery, and Anatomy of the Indians, Natural Productions of India, Poetry, Rhetoric, and Morality of Asia, Music of the Eastern Nations, Trade, Manufacture, Agriculture, and Commerce of India.<sup>9</sup>

One can find many other examples of dedicated scholarship among British officers in the East India Company, and there can be little doubt that the Western perceptions of India were profoundly influenced by these investigations. Not surprisingly, the focus here is quite often on those things that are distinctive in India. The specialists on India pointed to the uncommon aspects of Indian culture and its intellectual traditions, which were obviously more interesting given the perspective and motivation of the

9 William Jones, “Objects of Enquiry During My Residence in Asia,” included in *The Collected Works of Sir William Jones*, 13 vols. (London: J. Stockdale, 1807; republished, New York: New York University Press, 1993).

8 Alberuni's *India*, pt. I, chap. I, 20.



observers.<sup>10</sup> As a result, the curatorial approaches could not escape being somewhat slanted in their focus. I shall come back to this issue later.

I turn now to the second category, the magisterial approaches. The task of ruling a foreign country is not an easy one when its subjects are seen as equals. In this context, it is quite remarkable that the early British administrators in India, even the controversial Warren Hastings, were as respectful of the Indian traditions as they clearly were. The empire was still in its infancy and was being gradually acquired, rather tentatively (if not in a fit of absentmindedness).

A good example of a magisterial approach to India is the classic book on India written by James Mill, published in 1817, on the strength of which he was appointed as an official of the East India Company. Mill's *History of British India* played a major role in introducing the British governors of India to a particular characterization of the country. Mill disputed and dismissed practically every claim ever made on behalf of Indian culture and its intellectual traditions, concluding that it was totally primitive and rude. This diagnosis fit well with Mill's general attitude, which supported the idea of bringing a rather barbaric nation under the benign and reformist administration of the British Empire. Consistent with his beliefs, Mill was an expansionist in dealing with the remaining independent states in the subcontinent. The obvious policy to pursue, he explained,

10 I have discussed the "positional" nature of objectivity, depending on the placing of the observer and analyst vis-à-vis the objects being studied, in "Positional Objectivity," *Philosophy and Public Affairs* (1993), and "On Interpreting India's Past," in Sugata Bose and Ayesha Jalal, eds., *Nationalism, Democracy and Development: State and Politics in India* (Delhi: Oxford University Press, 1997).

was "to make war on those states and subdue them."<sup>11</sup>

Mill chastised early British administrators (like William Jones) for having taken "Hindus to be a people of high civilization, while they have in reality made but a few of the earliest steps in the progress to civilization."<sup>12</sup> At the end of a comprehensive attack on all fronts, he came to the conclusion that the Indian civilization was on a par with other inferior ones known to Mill – "very nearly the same with that of the Chinese, the Persians, and the Arabians"; he also put in this category, for good measure, "subordinate nations, the Japanese, Cochinchinese, Siamese, Burmans, and even Malays and Tibetans."<sup>13</sup>

How well informed was Mill in dealing with his subject matter? Mill wrote his book without ever having visited India. He knew no Sanskrit, nor any Persian or Arabic, had practically no knowledge of any of the modern Indian languages, and thus his reading of Indian material was of necessity most limited. There is another feature of Mill that clearly influenced his investigations, to wit, his inclination to distrust anything stated by native scholars, since they appeared to him to be liars. "Our ancestors," says Mill, "though tough, were sincere; but under the glossing exterior of the Hindu, lies a general disposition to deceit and perfidy."<sup>14</sup>

Perhaps some examples of Mill's treatment of particular claims of

11 Quoted in Eric Stokes, *The English Utilitarians and India* (Oxford: Clarendon Press, 1959), 250.

12 James Mill, *The History of British India* (London, 1817; republished, Chicago: University of Chicago Press, 1975), 225–226.

13 *Ibid.*, 248.

14 *Ibid.*, 247.

achievement may be useful to illustrate the nature of his extremely influential approach. The invention of the decimal system with place values and the placed use of zero, now used everywhere, as well as the so-called Arabic numerals are generally known to be Indian developments. In fact, Alberuni had already mentioned this in his eleventh-century book on India,<sup>15</sup> and many European as well as Arab scholars had written on this subject.<sup>16</sup> Mill dismisses the claim altogether on the grounds that “the invention of numerical characters must have been very ancient” and “whether the signs used by the Hindus are so peculiar as to render it probable that they invented them, or whether it is still more probable that they borrowed them, are questions which, for the purpose of ascertaining their progress in civilization, are not worth resolving.”

Mill proceeds then to explain that the Arabic numerals “are really hieroglyphics” and that the claim on behalf of the Indians and the Arabs reflects the confounding of “the origin of cyphers or numerical characters” with “that of hieroglyphic writing.”<sup>17</sup> At one level Mill’s rather elementary error lies in not knowing what a decimal or a place-value system is, but his ignorant smugness cannot be understood except in terms of his implicit unwillingness to believe that a very sophisticated invention could have been managed by such primitive people.

Another interesting example concerns Mill’s reaction to Indian astronomy and

specifically the argument for a rotating earth and a model of gravitational attraction (proposed by Aryabhata, who was born in A.D. 476, and investigated by, among others, Varahamihira and Brahmagupta in the sixth and seventh centuries). These works were well known in the Arab world; as was mentioned earlier, Brahmagupta’s book was translated into Arabic in the eighth century and retranslated by Alberuni in the eleventh. William Jones had been told about these works in India, and he in turn reported that statement. Mill expresses total astonishment at Jones’s gullibility.<sup>18</sup> After ridiculing the absurdity of this attribution and commenting on the “pretensions and interests” of Jones’s Indian informants, Mill concludes that it was “extremely natural that Sir William Jones, whose pundits had become acquainted with the ideas of European philosophers respecting the system of the universe, should hear from them that those ideas were contained in their own books.”<sup>19</sup>

For purposes of comparison it is useful to examine Alberuni’s discussion of the same issue nearly eight hundred years earlier, concerning the postulation of a rotating earth and gravitational attraction in the still-earlier writings of Aryabhata and Brahmagupta:

18 Mill found in Jones’s beliefs about early Indian mathematics and astronomy “evidence of the fond credulity with which the state of society among the Hindus was for a time regarded,” and he was particularly amused that Jones had made these attributions “with an air of belief.” Mill, *The History of British India*, 223–224. On the substantive side, Mill amalgamates the distinct claims regarding 1) the principle of attraction, 2) the daily rotation of the earth, and 3) the movement of the earth around the sun. Aryabhata and Brahmagupta’s concern were mainly with the first two, on which specific assertions were made, unlike on the third.

15 Alberuni’s *India*, pt. I, chap. XVI, 174–175.

16 For a modern account of the complex history of this mathematical development, see George Ifrah, *From One to Zero* (New York: Viking, 1985).

17 Mill, *The History of British India*, 219–220.

19 Mill, *The History of British India*, 223–224.

Brahmagupta says in another place of the same book: "The followers of Aryabhata maintain that the earth is moving and heaven resting. People have tried to refute them by saying that, if such were the case, stones and trees would fall from the earth." But Brahmagupta does not agree with them, and says that that would not necessarily follow from their theory, apparently because he thought that all heavy things are attracted towards the center of the earth.<sup>20</sup>

Alberuni himself proceeded to dispute this model, raised a technical question about one of Brahmagupta's mathematical calculations, referred to a different book of his own arguing against the proposed view, and pointed out that the relative character of movements makes this issue less central than one might first think: "The rotation of the earth does in no way impair the value of astronomy, as all appearances of an astronomic character can quite as well be explained according to this theory as to the other."<sup>21</sup> Here, as elsewhere, while arguing against an opponent's views, Alberuni tries to present such views with great involvement and care. The contrast between Alberuni's curatorial approach and James Mill's magisterial pronouncements could not be sharper.

There are plenty of other examples of "magisterial" readings of India in Mill's history. This is of some practical importance, since the book was extremely influential in the British administration and widely praised. It was described by Macaulay as "on the whole the greatest historical work which has appeared in our language since that of Gibbon."<sup>22</sup>

20 *Alberuni's India*, pt. I, chap. XXVI, 276 – 277.

21 *Ibid.*, 277.

22 Quoted in John Clive's introduction to James Mill, *The History of British India* (repub-

Macaulay's own approach and inclinations echoed James Mill's:

I have no knowledge of either Sanskrit or Arabic . . . I am quite ready to take the Oriental learning at the valuation of the Orientalists themselves. I have never found one among them who could deny that a single shelf of a good European library was worth the whole native literature of India and Arabia.<sup>23</sup>

This view of the poverty of Indian intellectual traditions played a major part in educational reform in British India, as is readily seen from the 1835 "Minute on Indian Education," written by Macaulay himself (the quoted remark is actually taken from that document). The priorities in Indian education were determined, henceforth, by a different emphasis – by the need, as Macaulay argued, for a class of English-educated Indians who could "be interpreters between us and the millions whom we govern."<sup>24</sup>

The impact of the magisterial views of India was not confined only to Britain and India. Modern documents in the same tradition have been influential elsewhere, including in the United States. In a series of long conversations on India and China conducted by Harold Isaacs in 1958 with 181 Americans – academics, professionals in mass media, government officials, missionaries and church officials, and officials of foundations, voluntary social-service groups, and political organizations – Isaacs

---

lished, Chicago: University of Chicago Press, 1975), viii.

23 T. B. Macaulay, "Indian Education: Minute of the 2nd February, 1835"; reproduced in G. M. Young, ed., *Macaulay: Prose and Poetry* (Cambridge, Mass.: Harvard University Press, 1952), 722.

24 *Ibid.*, 729.

found that the two most widely read literary sources on India were Rudyard Kipling and Katherine Mayo, the author of the extremely derogatory *Mother India*.<sup>25</sup> Of these, Kipling's writings would be more readily recognized as having something of the "magisterial" approach to them. Lloyd Rudolph describes Mayo's *Mother India* thus:

First published in 1927, *Mother India* was written in the context of official and unofficial British efforts to generate support in America for British rule in India. It added contemporary and lurid detail to the image of Hindu India as irredeemably and hopelessly impoverished, degraded, depraved, and corrupt. Mayo's *Mother India* echoed not only the views of men like Alexander Duff, Charles Grant, and John Stuart Mill but also those of Theodore Roosevelt, who glorified in bearing the white man's burden in Asia and celebrated the accomplishments of imperialism.<sup>26</sup>

Mahatma Gandhi, while describing Mayo's book as "a drain inspector's report," had added that every Indian should read it and seemed to imply, as Ashis Nandy notes, that it is possible "to put her criticism to internal use" (as an over stern drain inspector's report certainly can be).<sup>27</sup> Gandhi him-

self was severely attacked in the book, but given his campaign against caste and untouchability, he might have actually welcomed even her exaggerations because of its usefully lurid portrayal of caste inequities. But while Gandhi may have been right to value external criticism as a way of inducing people to be self-critical, the impact of the "magisterial approach" certainly gives American perceptions of India a very clear slant.<sup>28</sup>

I turn now to the "exoticist" approaches to India. Interest in India has often been stimulated by the observation of exotic ideas and views there. Arrian's and Strabo's accounts of Alexander the Great's spirited conversations with various sages of northwest India may or may not be authentic, but ancient Greek literature is full of uncommon happenings and thoughts attributed to India.

Megasthenes's *Indika*, describing India in the early third century B.C., can claim to be the first outsider's book on India; it created much Greek interest, as can be seen from the plentiful references to it, for example, in the writings of Diodorus, Strabo, and Arrian. Megasthenes had ample opportunity to observe India since, as the envoy of Seleucus Nicator to the court of Chandragupta Maurya, he spent nearly a decade (between 302 and 291 B.C.) in Pataliputra (the site of modern Patna), the capital city of the Mauryan empire. But his superlatively admiring book is also so full of accounts of fantastic objects and achievements in India that it is hard to be sure what

25 See Harold Isaacs, *Scratches on Our Minds* (Cambridge, Mass.: MIT Press, 1958); republished as *Images of Asia: American Views of India and China* (New York: Capricorn Books, 1958). See also the discussion of this issue in the "Introduction" in Sulochana Glazer and Nathan Glazer, eds., *Conflicting Images: India and the United States* (Glen Dale, Md.: Riverdale, 1990).

26 Lloyd I. Rudolph, "Gandhi in the Mind of America," in Glazer and Glazer, eds., *Conflicting Images*, 166.

27 Ashis Nandy, *Traditions, Tyranny, and Utopias: Essays in the Politics of Awareness* (Delhi: Oxford University Press, 1987), 8.

28 On this, see Glazer and Glazer, eds., *Conflicting Images*. The influence of magisterial readings on American imaging of India has been somewhat countered in recent years by the political interest in Gandhi's life and ideas, a variety of sensitive writings on India (from Erik Erikson to John Kenneth Galbraith), and the Western success of several Indian novelists in English.

is imagined and what is really being observed.

There are various other accounts of exotic Indian travels by ancient Greeks. The biography of Apollonius of Tyana by Flavius Philostratus in the third century A.D. is a good example. In his search for what was out of the ordinary, Apollonius was, we are assured, richly rewarded in India: "I have seen men living upon the earth and not upon it; defended without walls, having nothing, and yet possessing all things."<sup>29</sup> How such contradictory things can be seen by the same person from the same observational position may not be obvious, but the bewitching charm of all this for the seeker of the exotic can hardly be doubted.

Exotic interests in India can be seen again and again, from its early history to the present day. From Alexander listening to the gymnosophists' lectures to contemporary devotees hearing the sermons of Maharishi Mahesh Yogi and Shri Rajneesh, there is a crowded lineage. Perhaps the most important example of intellectual exoticism related to India can be seen in the European philosophical discussions in the eighteenth and early nineteenth century, among the Romantics in particular.

Important figures in the Romantic movement, including the Schlegel brothers, Schelling, and others, were profoundly influenced by rather magnified readings of Indian culture. From Herder, the German philosopher and a critic of the rationalism of the European Enlightenment, we get the magnificent news that "the Hindus are the gentlest branch of humanity" and that "moderation and calm, a soft feeling and a si-

29 Quoted in John Drew, *India and the Romantic Imagination* (Delhi and New York: Oxford University Press, 1987), 95.

lent depth of the soul characterize their work and their pleasure, their morals and mythology, their arts."<sup>30</sup> Frederick Schlegel not only pioneered studies of Indo-European linguistics (later pursued particularly by Max Muller) but also brought India fully into his critique of the contemporary West. While in the West "man himself has almost become a machine" and "cannot sink any deeper," Schlegel recommended learning from the Orient, especially India. He also guaranteed that "the Persian and German languages and cultures, as well as the Greek and the old Roman, may all be traced back to the Indian."<sup>31</sup> To this list, Schopenhauer added the New Testament, informing us that, in contrast with the Old, the New Testament "must somehow be of Indian origin: this is attested by its completely Indian ethics, which transforms morals into asceticism, its pessimism, and its *avatar* (i.e., the person of Christ)."<sup>32</sup>

Not surprisingly, many of the early enthusiasts were soon disappointed in not finding in Indian thought what they had themselves put there, and many of them went into a phase of withdrawal and criticism. Some of the stalwarts, Schlegel in particular, recanted vigorously. Others, including Hegel, outlined fairly negative views of Indian traditions and presented loud denials of the claim of preeminence of Indian culture – a

30 J. G. Herder, *Auch eine Philosophie der Geschichte: Samtliche Werke*, translated by Halbfass, *India and Europe*, 70.

31 Translations by Halbfass, *India and Europe*, 74–75. Halbfass provides an extensive study of these European interpretations of Indian thought and the reactions and counterreactions to them.

32 A. Schopenhauer, *Parerga und Paralipomena*; translated by Halbfass, *India and Europe*, 112.



claim that was of distinctly European origin. When Samuel Coleridge asked: “What are / These potentates of inmost Ind?”<sup>33</sup> he was really asking a question about Europe, rather than about India.<sup>34</sup>

In addition to veridical weakness, the exoticist approach to India has an inescapable fragility and transience that can be seen again and again. A wonderful thing is imagined about India and sent into a high orbit, and then it is brought crashing down. All this need not be such a tragedy when the act of launching is done by (or with the active cooperation of) the putative star. Not many would weep, for example, for Maharishi Mahesh Yogi when the Beatles stopped lionizing him and left suddenly; in answer to the Maharishi’s question of why were they leaving, John Lennon said, “You are the cosmic one; you ought to know.”<sup>35</sup>

But it is a different matter altogether when both the boom and the bust are thrust upon the victim. One of the most discouraging episodes in literary reception occurred early in this century, when Ezra Pound, W. B. Yeats, and others led

33 See John H. Muirhead, *Coleridge as a Philosopher* (London: G. Allen & Unwin, Ltd., 1930), 283–284, and Drew, *India and the Romantic Imagination*, chap. 6.

34 The nature of exoticist reading has typically had a strongly “Hindu” character. This was, in some ways, present even in William Jones’s curatorial investigations (though he was himself a scholar in Arabic and Persian as well), but he was to some extent redressing the relative neglect of Sanskrit classics in the previous periods (even though the version of the *Upanishads* that Jones first read was the Persian translation prepared by the Moghul prince Dara Shikoh, Emperor Akbar’s great-grandson). The European Romantics, on the other hand, tended to identify India with variants of Hindu religious thought.

35 William Davis, *The Rich* (London: Sidgwick and Jackson, 1982), 99.

a chorus of adoration at the lyrical spirituality of Rabindranath Tagore’s poetry but followed it soon afterwards with a thorough disregard and firm denunciation. Tagore was a Bengali poet of tremendous creativity and range (even though his poetry does not translate easily – not even the spiritual ones that were so applauded) and also a great storyteller, novelist, and essayist; he remains a dominant literary figure in Bangladesh and India. The versatile and innovative writer that the Bengalis know well is not the sermonizing spiritual guru put together in London; nor did he fit any better the caricature of “Stupendranath Begorr” to be found in Bernard Shaw’s “A Glimpse of the Domesticity of Franklyn Barnabas.”

These different approaches have had very diverse impacts on the understanding of Indian intellectual traditions in the West. The exoticist and magisterial approaches have bemused and befuddled that understanding even as they have drawn attention to India in the West. The curatorial approaches have been less guilty of this, and indeed historically have played a major part in bringing out and drawing attention to the different aspects of Indian culture, including its nonmystical and nonexotic features. Nevertheless, given the nature of the curatorial enterprise, the focus inevitably leans towards that which is different in India, rather than what is similar to the West. In emphasizing the distinctiveness of India, even the curatorial approaches have sometimes contributed to the accentuation of contrasts rather than commonalities with Western traditions, though not in the rather extreme form found in the exoticist and magisterial approaches.

The magisterial approaches played quite a vigorous role in the running of

the British Empire. Even though the Raj is dead and gone, the impact of the associated images survives, not least in the United States (as discussed earlier). To some extent, the magisterial authors also reacted against the admiration of India that can be seen in the writings of curatorial observers of India. For example, both Mill and Macaulay were vigorously critical of the writings of authors such as William Jones, and there are some important dialectics here. The respectful curatorial approaches painted a picture of Indian intellectual traditions that was much too favorable for the imperial culture of the nineteenth century, and contributed to the vehemence of the magisterial denunciations of those traditions. By the time Mill and Macaulay were writing, the British Indian empire was well established as a lasting and extensive enterprise, and the “irresponsibility” of admiring the native intellectual traditions – permissible in the previous century for early servants of the East India Company – was hard to sustain as the favored reading of India in the consolidated empire.

Turning to the exoticist approaches, the outbursts of fascinated wonder bring India into Western awareness in big tides of bewildering attention. But then they ebb, leaving only a trickle of hardened exoticists holding forth. There may well be, after a while, another tide. In describing the rise and decline of Rabindranath Tagore in London’s literary circles, E. M. Forster remarked that London was a city of “boom and bust,” but that description applies more generally (that is, not confined only to literary circles in London) to the Western appreciation of exotic aspects of Eastern cultures.

The tides, while they last, can be hard work though. I remember feeling quite sad for a dejected racist whom I saw,

some years ago, near the Aldwych station in London, viewing with disgust a thousand posters pasted everywhere carrying pictures of the obese – and holy – physique of Guru Maharajji (then a great rage in London). Our dedicated racist was busy writing “fat wog” diligently under each of the pictures. In a short while that particular wog would be gone, but I do not doubt that the “disgusted of Aldwych” would scribble “lean wog” or “medium-sized wog” under other posters now.

It might be thought that since the exoticist approaches *give* credit where it may not be due and the magisterial approaches *withhold* credit where it may well be due, the two might neutralize each other nicely. But they work in very asymmetrical ways. Magisterial critiques tend to blast the rationalist and humanist aspects of India with the greatest force (this is as true of James Mill as of Katharine Mayo), whereas exoticist admirations tend to build up the mystical and extrarational aspects with particular care (this has been so from Apollonius of Tyana down to the Hare Krishna activists of today). The result of the two taken together is to wrest the understanding of Indian culture forcefully away from its rationalist aspects. Indian traditions in mathematics, logic, science, medicine, linguistics, or epistemology may be well known to the Western specialist, but they play little part in the general Western understanding of India.<sup>36</sup> Mysticism and exoticism, in contrast, have a more hallowed position in that understanding.

Western perceptions and characterizations of India have had considerable

36 On this issue, see Bimal Matilal, *Perceptions* (Oxford: Clarendon Press, 1986). See also Ronald Inden, *Imagining India* (Oxford: Blackwell, 1990).

influence on the self-perceptions of Indians themselves. This is clearly connected to India's colonial past and continued deference to what is valued in the West.<sup>37</sup> However, the relationship need not take the form of simple acceptance – it sometimes includes strategic responses to the variety of Western perceptions of India that suit the interests of internal imaging. We have to distinguish between some distinct aspects of the influence that Western images have had on Indian internal identities.

First, the European exoticists' interpretations and praise found in India a veritable army of appreciative listeners, who were particularly welcoming given the badly damaged self-confidence resulting from colonial domination. The admiring statements were quoted again and again, and the negative remarks by the same authors (Herder, Schlegel, Goethe, and others) were systematically overlooked.

In his *Discovery of India*, Jawaharlal Nehru comments on this phenomenon:

There is a tendency on the part of Indian writers, to which I have also partly succumbed, to give selected extracts and quotations from the writings of European scholars in praise of old Indian literature and philosophy. It would be equally easy, indeed much easier, to give other extracts giving an exactly opposite viewpoint.<sup>38</sup>

In the process of accepting the exoticist praise, the Indian interpretation of the

37 On this issue in general, and on the hold of "a predominantly third-person perspective" in self-perception, see Akeel Bilgrami, "What Is a Muslim? Fundamental Commitment and Cultural Identity," *Critical Inquiry* 18 (4) (1992).

38 Jawaharlal Nehru, *The Discovery of India* (Calcutta: Signet Press, 1946; centenary edition, Delhi: Oxford University Press, 1989), 158.

past has extensively focused on the *objects* of exoticist praise, concentrating more on the mystical and the antirationalist, for which many in the West have such admiration.<sup>39</sup>

Second, the process fit into the politics of elitist nationalism in colonial India and fed the craving for a strong intellectual ground to stand on to confront the imperial rulers. Partha Chatterjee discusses the emergence of this attitude very well:

... anticolonial nationalism creates its own domain of sovereignty within colonial society well before its political battle with the imperial power. It does this by dividing the world of social institutions and practices into two domains – the material and the spiritual. The material is the domain of the "outside," of the economy and of statecraft, of science and technology, a domain where the West had proved its superiority and the East had succumbed. In this domain, then, Western superiority had to be acknowl-

39 While the constitution of independent India has been self-consciously secular, the tendency to see India as a land of the Hindus remains quite strong. The confrontation between "secularists" and "communitarians" has been an important feature of contemporary India, and the identification of Indian culture in mainly Hindu terms plays a part in this. While it is certainly possible to be both secular and communitarian (as Rajeev Bhargava has noted in "Giving Secularism Its Due," *Economic and Political Weekly*, July 9, 1994), the contemporary divisions in India tend to make the religious and communal identities largely work against India's secular commitments (as Bhargava also notes). I have tried to scrutinize these issues in my paper "Secularism and Its Discontents," in Kaushik Basu and Sanjay Subrahmanyam, eds., *Unravelling the Nation: Sectarian Conflict and India's Secular Identity* (Delhi: Penguin, 1996). See also the other papers in that collection, and the essays included in Bose and Jalal, eds., *Nationalism, Democracy and Development: State and Politics in India*.

edged and its accomplishments carefully studied and replicated. The spiritual, on the other hand, is an “inner” domain bearing the “essential” marks of cultural identity. The greater one’s success in imitating Western skills in the material domain, therefore, the greater the need to preserve the distinctiveness of one’s spiritual culture. This formula is, I think, a fundamental feature of anticolonial nationalisms in Asia and Africa.<sup>40</sup>

There was indeed such an attempt to present what was perceived to be the “strong aspects” of Indian culture, distinguished from the domain, as Chatterjee puts it, “where the West had proved its superiority and the East had succumbed.”

Chatterjee’s analysis can be supplemented by taking further note of the dialectics of the relationship between Indian internal identity and its external images. The diagnosis of strength in that nonmaterialist domain was as much helped by the exoticist admiration for Indian spirituality as the acceptance of India’s weakness in the domain of science, technology, and mathematics was reinforced by the magisterial dismissals of India’s materialist and rationalist traditions. The emphases on internal identity that emerged in colonial India bear powerful marks of dialectical encounters with Western perceptions.

Third, as the focus has shifted in recent decades from elitist colonial history to the role of the nonelite, the concentration on the intellectual traditions of the elite has weakened. Here we run into one of the most exciting developments in historiography in India. There has been a significant shift of attention from the elite to the underdogs in the writing

40 Partha Chatterjee, *The Nation and Its Fragments* (Princeton, N.J.: Princeton University Press, 1993), 6.

of colonial history, focusing more on the rural masses and the exploited plebeians— a broad group often identified by the capacious term “subalterns.”<sup>41</sup> The move is entirely appropriate in its context (in fact, much overdue), and in understanding colonial history, this is a very important corrective.

While this shift in focus rejects the emphasis on elitist intellectual traditions in general (*both* of the materialist and the nonmaterialist kind), it is in many ways easier to relate the religious and spiritual traditions of the elite to the practices and beliefs of the nonelite. The cutting edge of science and mathematics is inevitably related to formal education and preparation. In this context, the immense backwardness of India in mass education (an inheritance from the British period but not much remedied yet) compounds the dissociation of elite science and mathematics from the lives of the nonelite. Acceptance of the achievements of Indian spirituality tends to look less “alienated” from the masses than the achievements in fields that demand more exacting formal education. Thus, the exoticists’ praise of India is more easily accepted by those who are particularly careful not to see India in elitist terms.

The fact remains, however, that illiteracy is a deprivation. The issue of interclass justice cannot be a matter only of recognizing the real role of the subalterns *in history* (for example, in anticolonial national movements), important

41 The most effective move in that direction came under the leadership of Ranajit Guha; see his introductory essay in *Subaltern Studies I: Writings on South Asian History and Society*, ed. Ranajit Guha (Delhi: Oxford University Press, 1982). See also the collection of “subaltern” essays edited by Ranajit Guha and Gayatri Chakravorty Spivak, *Selected Subaltern Studies* (New York: Oxford University Press, 1988).



though it is. It is also a matter of remedying the immense inequalities in educational and other opportunities that severely limit, even today, the *actual* lives of the subalterns.

Interestingly enough, even by the eleventh century, the seriousness of this loss was noted by Alberuni himself (one of the major curatorial authors whose work was referred to earlier). Alberuni spoke of the real deprivation of “those castes who are not allowed to occupy themselves with science.”<sup>42</sup> This substantive deprivation remains largely unremedied even today (except in particular regions such as Kerala), with half of the adult population of India (and nearly two-thirds of the adult women) still remaining illiterate.<sup>43</sup> In understanding the nature of Indian cultures and traditions, focusing mainly on the achievements – rather than deprivations – of the Indian subaltern can yield a deceptive contrast.

This shift in emphasis has also, to some extent, pushed the interpretation of India’s past away from those achievements that require considerable formal training. While this move makes sense in some contexts, a comparison of a self-consciously nonelitist history of India

with the typically classical understanding of the intellectual heritage of the West produces a false contrast between the respective intellectual traditions. In comparing Western thoughts and creations with those in India, the appropriate counterpoints of Aristotelian or Stoic or Euclidian analyses are not the traditional beliefs of the Indian rural masses or of the local wise men but the comparably analytical writings of, say, Kautilya or Nagarjuna or Aryabhata. “Socrates meets the Indian peasant” is not a good way to contrast the respective intellectual traditions.

The internal identities of Indians draw on different parts of India’s diverse traditions. The observational leanings of Western approaches have had quite a major impact – positively and negatively – on what contributes to the Indian self-image that emerged in the colonial period and survives today. The relationship has several dialectical aspects, connected to the sensitivity towards selective admirations and dismissals from the cosmopolitan West as well as to the mechanics of colonial confrontations.

The differences between the curatorial, magisterial, and exoticist approaches to Western understanding of Indian intellectual traditions lie, to a great extent, in the varying observational positions from which India has been examined and its overall images drawn. The dependence on perspective is not a special characteristic of the imaging of India alone. It is, in fact, a pervasive general feature in description and identification.<sup>44</sup> “What is India really like?” is

44 I have tried to discuss this general issue in “Description as Choice,” *Oxford Economic Papers* 32 (1980) (reprinted in *Choice, Welfare and Measurement* [Oxford: Blackwell; Cambridge, Mass.: MIT Press, 1982]), and in “Positional Objectivity,” *Philosophy and Public Affairs* (1993).

42 Alberuni’s *India*, chap. II, 32.

43 Indeed, in conceptualizing “the good life” even from the perspective of the deprived underdog, it would be a mistake to ignore altogether the intellectual achievements of the elite, since part of the deprivation of the exploited lies precisely in being denied participation in these achievements. While Marx might have exaggerated a little in his eloquence about “the idiocy of the village life,” there is nevertheless a substantial point here in identifying the nature of social deprivation. There is, in fact, no basic contradiction in choosing the subaltern perspective of history and taking systematic note of the scholarly accomplishments of the elite.



a good question for a foreign tourist's handbook precisely because the description there may sensibly be presented from the particular position of being a foreign tourist in India. But there are other positions, other contexts, other concerns.

The three approaches investigated here have produced quite distinct views of Indian intellectual history, but their overall impact has been to exaggerate the nonmaterial and arcane aspects of Indian traditions compared to its more rationalistic and analytical elements. While the curatorial approaches have been less guilty of this, their focus on what is really different in India has, to some extent, also contributed to it. But the bulk of the contribution has come from the *exoticist admiration* of India (particularly of its spiritual wonders) and the *magisterial dismissals* (particularly of its claims in mathematics, science, and analytical pursuits).

The nature of these slanted emphases has tended to undermine an adequately pluralist understanding of Indian intellectual traditions. While India has inherited a vast religious literature, a large wealth of mystical poetry, grand speculation on transcendental issues, and so on, there is also a huge – and often pioneering – literature, stretching over two and a half millennia, on mathematics, logic, epistemology, astronomy, physiology, linguistics, phonetics, economics, political science, and psychology, among other subjects concerned with the here and now.<sup>45</sup>

Even on religious subjects, the only world religion that is firmly agnostic

45 This contrast is discussed in my joint paper with Martha Nussbaum, "Internal Criticism and Indian Rationalist Traditions," in Michael Krausz, ed., *Relativism: Interpretation and Confrontation* (Notre Dame, Ind.: University of Notre Dame Press, 1989).

(Buddhism) is of Indian origin, and, furthermore, the atheistic schools of Carvaka and Lokayata have generated extensive arguments that have been seriously studied by Indian religious scholars themselves.<sup>46</sup> Heterodoxy runs throughout the early documents, and even the ancient epic *Ramayana*, which is often cited by contemporary Hindu activists as the holy book of the divine Rama's life, contains dissenting characters. For example, Rama is lectured to by a worldly pundit called Javali on the folly of his religious beliefs: "O Rama, be wise, there exists no world but this, that is certain! Enjoy that which is present and cast behind thee that which is unpleasant."<sup>47</sup>

What is in dispute here is not the recognition of mysticism and religious initiatives in India, which are certainly plentiful, but the overlooking of all the other intellectual activities that are also abundantly present. In fact, despite the grave sobriety of Indian religious preoccupations, it would not be erroneous to say that India is a country of fun and games in which chess was probably invented, badminton originated, polo emerged, and the ancient *Kamasutra* told people how to have joy in sex. Indeed, Georges Ifrah quotes a medieval Arab poet from Baghdad called al-Sabhadi, who said that there were "three things on which the Indian nation prided itself: its method of reckoning, the game of chess, and the book titled *Kalila wa Dimna* [a collection of legends and fa-

46 For example, the fourteenth-century book *Sarvadarsanasamgraha* (*Collection of All Philosophies*) by Madhava Acarya (himself a good Vaishnavite Hindu) devotes the first chapter of the book to a serious presentation of the arguments of the atheistic schools.

47 English translation from H. P. Shastri, *The Ramayana of Valmiki* (London: Shanti Sadan, 1959), 389.

bles].”<sup>48</sup> This is not altogether a different list from Voltaire’s catalog of the important things to come from India: “our numbers, our backgammon, our chess, our first principles of geometry, and the fables which have become our own.”<sup>49</sup> These selections would not fit the cultivated Western images of Indian historical traditions, which are typically taken to be pontifically serious and uncompromisingly spiritual.

Nor would it fit the way many Indians perceive themselves and their intellectual past, especially those who take a “separatist” position on the nature of Indian culture. I have tried to discuss how that disparity has come about and how it is sustained. I have also tried to speculate about how the selective alienation of India from a very substantial part of its past has been nourished by the asymmetrical relationship between India and the West. It is, oddly enough, the rationalist part of India’s tradition that has been affected most by this alienation. The impact of the West on internal identities in India has to be seen in fundamentally dialectical terms.

48 Ifrah, *From One to Zero*, 434.

49 Voltaire, *Les oeuvres complètes*, vol. 124; translated by Halbfass, *India and Europe*, 59.

# Robert Frost

## On Emerson

*On the evening of October 8, 1958 the American Academy of Arts and Sciences made its first award of the Emerson-Thoreau Medal to Robert Frost in recognition of his long and distinguished contribution to the creative arts. His response appears below, as revised by him from his words of that evening.*

*As befitted the eminence of the guest of honor and the importance of the occasion, the House of the Academy was filled. To accord with Mr. Frost's wishes, the atmosphere was informal and the ceremony simple – that of friends honoring a friend. Before dinner he chatted with some of his hosts, and after dinner he escaped for a short catnap. When he took his place before his enthusiastically applauding audience, it was as a member of the Academy, at home in its House, the man whom Amy Lowell called “a neighbor of neighbors.”*

*What he then said was essentially a conversation with friends. It followed the higher law of his own feeling rather than any formal*

---

*Robert Frost, one of America's leading twentieth-century poets and a four-time winner of the Pulitzer Prize, became a Fellow of the American Academy in 1931. His remarks, given on the night he accepted the American Academy's first Emerson-Thoreau Medal, appeared in the Fall 1959 issue of “Daedalus.” Frost died in 1963.*

---

© 2005 by the American Academy of Arts & Sciences

*rules for oral rhetoric, and his idiom was tuned for direct communication with responsive hearers. The printer has no font for the intonation of his voice, the rhythms of his speech, its change of pace, its eloquent pauses. All who were present carried away an indelible impression of the spirit, the personality, and the mind of the poet we were proud to salute. – K.B.M.*

All that admiration for me I am glad of. I am here out of admiration for Emerson and Thoreau. Naturally on this proud occasion I should like to make myself as much of an Emersonian as I can. Let me see if I can't go a long way. You may be interested to know that I have right here in my pocket a little first edition of Emerson's poetry. His very first was published in England, just as was mine. His book was given me on account of that connection by Fred Melcher, who takes so much pleasure in bringing books and things together like that.

I suppose I have always thought I'd like to name in verse some day my four greatest Americans: George Washington, the general and statesman; Thomas Jefferson, the political thinker; Abraham Lincoln, the martyr and savior; and fourth, Ralph Waldo Emerson, the poet. I take these names because they are going around the world. They are not just

local. Emerson's name has gone as a poetic philosopher or as a philosophical poet, my favorite kind of both.

I have friends it bothers when I am accused of being Emersonian, that is, a cheerful Monist, for whom evil does not exist, or if it does exist, needn't last forever. Emerson quotes Burns as speaking to the Devil as if he could mend his ways. A melancholy dualism is the only soundness. The question is: is soundness of the essence.

My own unsoundness has a strange history. My mother was a Presbyterian. We were here on my father's side for three hundred years but my mother was fresh a Presbyterian from Scotland. The smart thing when she was young was to be reading Emerson and Poe as it is today to be reading St. John Perse or T. S. Eliot. Reading Emerson turned her into a Unitarian. That was about the time I came into the world; so I suppose I started a sort of Presbyterian-Unitarian. I was transitional. Reading on into Emerson, that is into "Representative Men" until she got to Swedenborg, the mystic, made her a Swedenborgian. I was brought up in all three of these religions, I suppose. I don't know whether I was baptized in them all. But as you see it was pretty much under the auspices of Emerson. It was all very Emersonian. Phrases of his began to come to me early. In that essay on the mystic he makes Swedenborg say that in the highest heaven nothing is arrived at by dispute. Everybody votes in heaven but everybody votes the same way, as in Russia today. It is only in the second-highest heaven that things get parliamentary; we get the two-party system or the hydra-headed, as in France.

Some of my first thinking about my own language was certainly Emersonian. "Cut these sentences and they bleed," he says. I am not submissive

enough to want to be a follower, but he had me there. I never got over that. He came pretty near making me an antivocabularian with the passage in "Monadnock" about our ancient speech. He blended praise and dispraise of the country people of New Hampshire. As an abolitionist he was against their politics. Forty percent of them were states-rights Democrats in sympathy with the South. They were really pretty bad, my own relatives included.

The God who made New Hampshire  
Taunted the lofty land  
With little men; –

And if I may be further reminiscent parenthetically, my friend Amy Lowell hadn't much use for them either. "I have left New Hampshire," she told me. Why in the world? She couldn't stand the people. What's the matter with the people? "Read your own books and find out." They really differ from other New Englanders, or did in the days of Franklin Pierce.

But now to return to the speech that was his admiration and mine in a burst of poetry in "Monadnock":

Yet wouldst thou learn our ancient speech  
These the masters that can teach.  
Fourscore or a hundred words  
All their vocal muse affords.  
Yet they turn them in a fashion  
Past the statesman's art and passion.  
Rude poets of the tavern hearth  
Squandering your unquoted mirth,  
That keeps the ground and never soars,  
While Jake retorts and Reuben roars.  
Scoff of yeoman, strong and stark,  
Goes like bullet to the mark,  
And the solid curse and jeer  
Never balk the waiting ear.

Fourscore or a hundred is seven hundred less than my friend Ivor Richard's basic eight hundred. I used to climb

on board a load of shooks (boxes that haven't been set up) just for the pleasure I had in the driver's good use of his hundred-word limit. This at the risk of liking it so much as to lose myself in mere picturesqueness. I was always in favor of the solid curse as one of the most beautiful of figures. We were warned against it in school for its sameness. It depends for variety on the tones of saying it and the situations.

I had a talk with John Erskine, the first time I met him, on this subject of sentences that may look tiresomely alike, short and with short words, yet turn out as calling for all sorts of ways of being said aloud or in the mind's ear, Horatio. I took Emerson's prose and verse as my illustration. Writing is unboring to the extent that it is dramatic.

In a recent preface to show my aversion to being interrupted with notes in reading a poem, I find myself resorting to Emerson again. I wanted to be too carried away for that. There was much of "Brahma" that I didn't get to begin with but I got enough to make me sure I would be back there reading it again some day when I had read more and lived more; and sure enough, without help from dictionary or encyclopedia I can now understand every line in it but one or two. It is a long story of many experiences that let me into the secret of:

But thou, meek lover of the good!  
Find me, and turn thy back on heaven.

What baffled me was the Christianity in "meek lover of the good." I don't like obscurity and obfuscation, but I do like dark sayings I must leave the clearing of to time. And I don't want to be robbed of the pleasure of fathoming depths for myself. It was a moment for me when I saw how Shakespeare set bounds to science when he brought in the North Star, "whose worth's unknown although his

height be taken." Of untold worth: it brings home some that should and some that shouldn't come. Let the psychologist take notice how unsuccessful he has to be.

I owe more to Emerson than anyone else for troubled thoughts about freedom. I had the hurt to get over when I first heard us made fun of by foreigners as the land of the free and the home of the brave. Haven't we won freedom? Is there no such thing as freedom? Well, Emerson says God

Would take the sun out of the skies  
Ere freedom out of a man.

and there rings the freedom I choose.

Never mind how and where Emerson disabused me of my notion I may have been brought up to that the truth would make me free. My truth will bind you slave to me. He didn't want converts and followers. He was a Unitarian. I am on record as saying that freedom is nothing but departure – setting forth – leaving things behind, brave origination of the courage to be new. We may not want freedom. But let us not deceive ourselves about what we don't want. Freedom is one jump ahead of formal laws, as in planes and even automobiles right now. Let's see the law catch up with us very soon.

Emerson supplies the emancipating formula for giving an attachment up for an attraction, one nationality for another nationality, one love for another love. If you must break free,

Heartily know,  
When half-gods go  
The gods arrive.

I have seen it invoked in *Harper's Magazine* to excuse disloyalty to our democracy in a time like this. But I am not sure of the reward promised. There is such a thing as getting too transcended. There



are limits. Let's not talk socialism. I feel projected out from politics with lines like:

Musketaquit, a goblin strong,  
Of shards and flints makes jewels gay;  
They lose their grief who hear his song,  
And where he winds is the day of day.

So forth and brighter fares my stream, –  
Who drink it shall not thirst again;  
No darkness stains its equal gleam,  
And ages drop in it like rain.

Left to myself, I have gradually come to see what Emerson was meaning in "Give all to Love" was, Give all to Meaning. The freedom is ours to insist on meaning.

The kind of story Steinbeck likes to tell is about an old labor hero punch-drunk from fighting the police in many strikes, beloved by everybody at headquarters as the greatest living hater of tyranny. I take it that the production line was his grievance. The only way he could make it mean anything was to try to ruin it. He took arms and fists against it. No one could have given him that kind of freedom. He saw it as his to seize. He was no freedman; he was a free man. The one inalienable right is to go to destruction in your own way. What's worth living for is worth dying for. What's worth succeeding in is worth failing in.

If you have piled up a great rubbish heap of oily rags in the basement for your doctor's thesis and it won't seem to burst into flame spontaneously, come away quickly and without declaring rebellion. It will cost you only your Ph.D. union card and the respect of the union. But it will hardly be noticed even to your credit in the world. All you have to do is to amount to something anyway. The only reprehensible materiality is the materialism of getting lost in your material

so you can't find out yourself what it is all about.

A young fellow came to me to complain of the department of philosophy in his university. There wasn't a philosopher in it. "I can't stand it." He was really complaining of his situation. He wasn't where he could feel real. But I didn't tell him so I didn't go into that. I agreed with him that there wasn't a philosopher in his university – there was hardly ever more than one at a time in the world – and I advised him to quit. Light out for somewhere. He hated to be a quitter. I told him the Bible says, "Quit ye, like men." "Does it," he said. "Where would I go?" Why anywhere almost. Kamchatka, Madagascar, Brazil. I found him doing well in the educational department of Rio when I was sent on an errand down there by our government several years later. I had taken too much responsibility for him when I sent him glimmering like that. I wrote to him with troubled conscience and got no answer for two whole years. But the story has a happy ending. His departure was not suicidal. I had a postcard from him this Christmas to tell me he was on Robinson Crusoe's island Juan Fernandez on his way to Easter Island that it had always been a necessity for him some day to see. I would next hear from him in Chile where he was to be employed in helping restore two colleges. Two! And the colleges were universities!

No subversive myself, I think it very Emersonian of me that I am so sympathetic with subversives, rebels, runners out, runners out ahead, eccentrics, and radicals. I don't care how extreme their enthusiasm so long as it doesn't land them in the Russian camp. I always wanted one of them teaching in the next room to me so my work would be cut out for me warning the children taking my courses not to take his courses.

I am disposed to cheat myself and others in favor of any poet I am in love with. I hear people say the more they love anyone the more they see his faults. Nonsense. Love is blind and should be left so. But it hasn't been hidden in what I have said that I am not quite satisfied with the easy way Emerson takes disloyalty. He didn't know or ignored his Blackstone. It is one thing for the deserter and another for the deserted. Loyalty is that for the lack of which your gang will shoot you without benefit of trial by jury. And serves you right. Be as treacherous as you must be for your ideals, but don't expect to be kissed goodbye by the idol you go back on. We don't want to look too foolish, do we? And probably Emerson was too Platonic about evil. It was a mere *Tò μὴ οὔ* that could be disposed of like the butt of a cigarette. In a poem I have called the best Western poem yet he says:

Unit and universe are round.

Another poem could be made from that, to the effect that ideally in thought only is a circle round. In practice, in nature, the circle becomes an oval. As a circle it has one center – Good. As an oval it has two centers – Good and Evil. Thence Monism versus Dualism.

Emerson was a Unitarian because he was too rational to be superstitious and too little a storyteller and lover of stories to like gossip and pretty scandal. Nothing very religious can be done for people lacking in superstition. They usually end up abominable agnostics. It takes superstition and the prettiest scandal story of all to make a good Trinitarian. It is the first step in the descent of the spirit into the material-human at the risk of the spirit.

But if Emerson had left us nothing else he would be remembered longer than the Washington Monument for

the monument at Concord that he glorified with lines surpassing any other ever written about soldiers:

By the rude bridge that arched the flood  
Their flag to April breeze unfurled  
Here once the embattled farmers stood  
And fired the shot heard round the world.

Not even Thermopylae has been celebrated better. I am not a shiner but two things I never happen on unmoved: one, this poem on stone; and the other, the tall shaft seen from Lafayette Park across the White House in Washington.

# Umberto Eco

## *Innovation & repetition: between modern & postmodern aesthetics*

It is not by chance that modern aesthetics and modern theories of art (and I mean by “modern” those born with Mannerism, developed through Romanticism, and provocatively restated by the early twentieth-century avant-gardes) have frequently identified the artistic message with metaphor. Metaphor (the new and inventive one, not the worn-out catachresis) is a way to designate something by the name of something else, thus presenting that something in an unexpected way. The modern criterion for recognizing the artistic value was *novelty*, high information. The pleasurable repetition of an already known pattern was considered, by modern theories of art, typical of Crafts – not of Art – and of industry.

A good craftsman, as well as an industrial factory, produces many *tokens*, or occurrences, of the same *type* or model. One appreciates the type, and appreci-

ates the way the token meets the requirements of the type: but the modern aesthetics did not recognize such a procedure as an artistic one. That is why the Romantic aesthetics made such a careful distinction between “major” and “minor” arts, between arts and crafts. To make a parallel with sciences: crafts and industry were similar to the correct application of an already known law to a new case. Art, on the contrary (and by art I mean also literature, poetry, movies, and so on) corresponded rather to a “scientific revolution”: every work of modern art figures out a new law, imposes a *new paradigm*, a new way of looking at the world.

Modern aesthetics frequently forgot that the classical theory of art, from ancient Greece to the Middle Ages, was not so eager to stress a distinction between arts and crafts. The same term (*techne*, *ars*) was used to designate both the performance of a barber or a shipbuilder, the work of a painter or a poet. The classical aesthetics was not so anxious for innovation at any cost: on the contrary, it frequently appreciated as “beautiful” the good tokens of an everlasting type. Even in those cases in which modern sensitivity enjoys the “revolution” performed by a classical artist, his contemporary enjoyed the opposite aspect of

---

*Umberto Eco, best known in America for his best-selling novel “The Name of the Rose,” wrote this essay for the Fall 1985 issue of “Dædalus.” For more than two decades, he has been a professor of semiotics at the University of Bologna.*

---

© 2005 by the American Academy of Arts & Sciences

his work, that is, his respect for previous models.<sup>1</sup>

This is the reason why modern aesthetics was so severe apropos the industrial-like products of the mass media. A popular song, a TV commercial, a comic strip, a detective novel, a Western movie were seen as more or less successful tokens of a given model or type. As such they were judged as pleasurable but non-artistic. Furthermore, this excess of pleasurability, repetition, lack of innovation, was felt as a commercial trick (the product had to meet the expectations of its audience), not as the provocative proposal of a new (and difficult to accept) world vision. The products of mass media were equated with the products of industry insofar as they were produced *in series*, and the “serial” production was considered as alien to the artistic invention.

According to the modern aesthetics, the principal features of the mass-media products were repetition, iteration, obedience to a preestablished schema, and redundancy (as opposed to information).<sup>2</sup>

The device of *iteration* is typical, for instance, of television commercials: one distractedly watches the playing out of a sketch, then focuses one’s attention on the punch line that reappears at the end of the episode. It is precisely on this foreseen and awaited reappearance that our modest but irrefutable pleasure is based.

1 On the opposition between innovation and repetition, see my works *Opera aperta* (Milan: Bompiani, 1962) and *Apocalittici e Integrati* (Milan: Bompiani, 1964). Partial English translations appear in *The Role of the Reader* (Bloomington, Indiana: Indiana University Press, 1979).

2 I repeat here some of my old remarks in “The Myth of Superman” (1962), now in *The Role of the Reader*.

Likewise, the reading of a traditional detective story presumes the enjoyment of a scheme. The scheme is so important that the most famous authors have founded their fortune on its very immutability.

Furthermore, the writer plays upon a continuous series of connotations (for example, the characteristics of the detective and of his immediate “entourage”) to such an extent that their reappearance in each story is an essential condition of its reading pleasure. And so we have the by now historical “tics” of Sherlock Holmes, the punctilious vanity of Hercule Poirot, the pipe and the familiar fixes of Maigret, on up to the famous idiosyncracies of the most unabashed heroes of the hard-boiled novel. Vices, gestures, habits of the character portrayed permit us to recognize an old friend. These familiar features allow us to “enter into” the event. When our favorite author writes a story in which the usual characters do not appear, we are not even aware that the fundamental scheme of the story is still like the others: we read the book with a certain detachment and are immediately prone to judge it a “minor” one.

All this becomes very clear if we take a famous character such as Nero Wolfe, immortalized by Rex Stout. I shall review here the main characteristics of Nero Wolfe and his partners because it is important to ascertain how important they are for the reader of Stout’s books. Nero Wolfe, from Montenegro, a naturalized American from time immemorial, is outlandishly fat, so much so that his leather easy chair must be expressly designed for him. He is fearfully lazy. In fact, he never leaves the house and depends for his investigations on the smart and brilliant Archie Goodwin, with whom he indulges in a continuous sharp and tense polemic, tempered

somewhat by their mutual sense of humor. Nero Wolfe is an absolute glutton, and his cook, Fritz, is the vestal virgin in the pantry, devoted to the unending care of this highly cultivated palate and equally greedy stomach; but along with the pleasures of the table, Wolfe cultivates an all-absorbing and exclusive passion for orchids; he has a priceless collection in the greenhouse on the top floor of the villa where he lives. Quite possessed by gluttony and flowers, assailed by a series of accessory tics (love of scholarly literature, systematic misogyny, insatiable thirst for money), Nero Wolfe conducts his investigations, masterpieces of psychological penetration, sitting in his office, carefully weighing the information verbally furnished by Archie, studying the protagonists of each event who are obliged to visit him in his office, arguing with Inspector Cramer (attention: he always holds a methodically extinguished cigar in his mouth), quarreling with the odious Sergeant Purley Stebbins; and, finally, in a fixed setting from which he never veers, summoning the protagonists of the case to a meeting in his studio, usually in the evening. There, with skillful dialectical subterfuges, almost always before he himself knows the truth, he drives the guilty one into a public demonstration of hysteria by which he gives himself away.

The gamut is much more ample: Archie's almost canonical arrest under suspicion of reticence and false testimony; the legal diatribes about the conditions on which Wolfe will take on a client; the hiring of part-time agents like Saul Panzer or Orrie Carther; the painting in the studio behind which Wolfe or Archie can watch, through a peephole, the behavior and reactions of a subject put to the test in the office itself; the scenes with Wolfe and an insincere client.... Such is the "eternal" story that the

faithful reader enjoys in Stout's novels. To make it palatable, the author must invent every time a "new" crime and "new" secondary characters, but these details only serve to reconfirm the permanence of a fixed repertoire of *topoi*.

Not knowing who the guilty party is becomes an accessory element, almost a pretext. It is not a matter of discovering who committed the crime, but, rather, of following certain "topical" gestures of "topical" characters whose stock behavior we already love. The reader, little interested in the "new" psychological or economic motivations of the "new" crime, in fact enjoys those moments when Wolfe repeats his usual gestures, when he goes up for the *n*th time to take care of his orchids while the case itself is reaching its dramatic climax, when Inspector Cramer threateningly enters with one foot between the door and the wall, pushing aside Goodwin and warning Wolfe with a shake of his finger that this time things will not go so smoothly. The attraction of the book, the sense of repose, of psychological extension which it is capable of conferring, lies in the fact that, plopped in an easy chair or in the seat of a train compartment, the readers continuously recover, point by point, what they already know, and what they want to know again: that is why they have purchased the book. They derive pleasure from the nonstory (if indeed a story is a development of events which should bring us from the point of departure to a point of arrival where we would never have dreamed of arriving); the distraction consists in the refutation of a development of events, in a withdrawal from the tension of past-present-future to the focus on an *instant*, which is loved precisely because it is recurrent.

It seems that mechanisms of this kind proliferate more widely in the popular



narrative of today than in the eighteenth-century romantic *feuilleton*, where the event was founded upon a *development* and where the characters were required to march towards their death in the course of unexpected and “incredible” adventures.

If this were true, it would be because the *feuilleton*, founded on the triumph of information, represented the preferred fare of a society that lived in the midst of messages loaded with redundancy; the sense of tradition, the norms of social life, moral principles, the rules of proper comportment in the framework of a bourgeois society designed a system of foreseeable messages that the social system provided for its members, and which allowed life to flow smoothly without unexpected jolts. In this sphere, the “informative” shock of a short story by Poe or the *coup de théâtre* of Ponson du Terrail provided the enjoyment of the “rupture.” In a contemporary industrial society, instead, the social change, the continuous rise of new behavioral standards, the dissolution of tradition, require a narrative based upon redundancy. Redundant narrative structures would appear in this panorama as an indulgent invitation to repose, a chance of relaxing.

In fact, even the nineteenth-century novel was repetitive. Its fundamental patterns were always the same, and it was not so difficult, for a smart reader, to tell before the end of the story if Miss So-and-So was or was not the lost daughter of the duke of X. One can only say that the nineteenth-century *feuilleton* and contemporary mass media use *different* devices for making the expected appear unexpected. Archie Goodwin is explicitly expecting, with the readers, that Nero Wolfe will act in a certain way, while Eugène Sue pretended not to know in advance what her readers sus-

pected, namely, that Fleur-de-Marie was the daughter of Rodolphe of Gerolstein. The formal principle does not change.

Perhaps one of the first inexhaustible characters during the decline of the *feuilleton* and bridging the two centuries at the close of *la belle époque* is Fantomas. Each episode of Fantomas closes with a kind of “unsuccessful catharsis”; Juve and Fandor finally come to get their hands on the elusive one when he, with an unforeseeable move, foils the arrest. Another singular fact: Fantomas, responsible for blackmail and sensational kidnappings, at the beginning of each episode finds himself inexplicably poor and in need of money and, therefore, also of new “action.” In this way the cycle is kept going.

I would like to consider now the case of an historical period (our own) for which iteration and repetition seem to dominate the whole world of artistic creativity, and in which it is difficult to distinguish between the repetition of the media and the repetition of the so-called major arts. In this period one is facing the discussion of a new theory of art, one that I would label *postmodern aesthetics*, which is revisiting the very concepts of repetition and iteration under a different profile. Recently in Italy such a debate has flourished under the standard of a “new aesthetics of seriality.” I recommend my readers to take “seriality,” in this case, as a very wide category or, if one wants, as another term for repetitive art.

Seriality and repetition are largely inflated concepts. The philosophy of the history of art has accustomed us to some technical meanings of these terms that it would be well to eliminate: I shall not speak of repetition in the sense of Kierkegaard, nor of “*répétition différente*,” in the sense of Deleuze. In the history of

contemporary music, series and seriality have been understood in a sense more or less opposite what we are discussing here. The dodecaphonic “series” is the opposite of the repetitive seriality typical of all the media, because there a given succession of twelve sounds is used once and only once within a single composition.

If you open a current dictionary, you will find that for “repeat” the meaning is “to say something or do something the second time or again and again; iteration of the same word, act or idea.” For “series” the meaning is “a continued succession of similar things.” It is a matter of establishing what it means to say “again” or “the same or similar things.”

To serialize means, in some way, *to repeat*. Therefore, we shall have to define a first meaning of “to repeat” by which the term means to make a *replica* of the same *abstract type*. Two sheets of typewriter paper are both *replicas* of the same commercial *type*. In this sense one thing is the same as another when the former exhibits the same properties as the latter, at least under a certain description: two sheets of typing paper are the same from the point of view of our functional needs, even though they are not the same for a physicist interested in the molecular structure of the objects. From the point of view of industrial mass production, two “tokens” can be considered as “replicas” of the same “type” when for a normal person with normal requirements, in the absence of evident imperfection, it is irrelevant whether one chooses one instead of the other. Two copies of a film or of a book are replicas of the same type.

The repetitiveness and the seriality that interests us here look instead at something that at first glance does not appear the same as (equal to) something else.

Let us now see the case in which (1) something is offered as original and different (according to the requirements of modern aesthetics); (2) we are aware that this something is repeating something else that we already know; and (3) notwithstanding this – better, just because of it – we like it (and we buy it).

The first type of repetition is the *retake*. In this case one recycles the characters of a previous successful story in order to exploit them, by telling what happened to them after the end of their first adventure. The most famous example of retake is Dumas’s *Twenty Years Later*, the most recent ones are the “to be continued” versions of *Star Wars* or *Superman*. The retake is dependent on a commercial decision. There is no rule establishing whether the second episode of the story should reproduce, with only slight variations, the first one, or must be a totally different story concerning the same characters. The retake is not strictly condemned to repetition. An illustrious example of retake are the many different stories of the Arthurian cycle, telling again and again the vicissitudes of Lancelot or Perceval.

The *remake* consists in telling again a previous successful story. See the innumerable editions of *Dr. Jekyll* or of *Mutiny on the Bounty*. The history of arts and literature is full of pseudo-remakes that were able to tell at every time something different. The whole of Shakespeare is a remake of preceding stories. Therefore “interesting” remakes can escape repetition.

The *series* works upon a fixed situation and a restricted number of fixed pivotal characters, around whom the secondary and changing ones turn. The secondary characters must give the impression that

the new story is different from the preceding ones, while in fact the narrative scheme does not change. I have said something above on the scheme of the novels by Rex Stout.

To the same type belong the TV serials such as *All in the Family*, *Starsky and Hutch*, *Columbo*, etc. (I put together different TV genres that range from soap opera to situation comedy, and to the detective serial).

With a series one believes one is enjoying the novelty of the story (which is always the same) while in fact one is enjoying it because of the recurrence of a narrative scheme that remains constant. The series in this sense responds to the infantile need of hearing again always the same story, of being consoled by the "return of the Identical," superficially disguised.

The series consoles us (the consumers) because it rewards our ability to foresee: we are happy because we discover our own ability to guess what will happen. We are satisfied because we find again what we had expected, but we do not attribute this happy result to the obviousness of the narrative structure, but to our own presumed capacities to make forecasts. We do not think, "The author has constructed the story in a way that I could guess the end," but rather, "I was so smart to guess the end in spite of the efforts the author made to deceive me."

We find a variation of the series in the structure of the flashback: we see, for example, some comic-strip stories (such as Superman) in which the character is not followed along in a straight line during the course of his life, but is continually rediscovered at different moments of his life, obsessively revisited in order to discover there new opportunities for new narratives. It seems as if these moments of his life have fled from the narrator out of absentmindedness, but their

rediscovery does not change the psychological profile of the character, which is fixed already, once and for all. In topological terms this subtype of the series may be defined as a *loop*.

Usually the loop-series comes to be devised for commercial reasons: it is a matter of considering how to keep the series alive, of obviating the natural problem of the aging of the character. Instead of having characters put up with new adventures (that would imply their inexorable march toward death), they are made continually to relive their past. The loop solution produces paradoxes that were already the target of innumerable parodies. Characters have a little future but an enormous past, and in any case, nothing of their past will ever have to change the mythological present in which they have been presented to the reader from the beginning. Ten different lives would not suffice to make Little Orphan Annie undergo what she underwent in the first (and only) ten years of her life.

The spiral is another variation of the series. In the stories of Charlie Brown, apparently nothing happens, and any character is obsessively repeating his/her standard performance. And yet in every strip the character of Charlie Brown or Snoopy is enriched and deepened. This does not happen either with Nero Wolfe, or Starsky or Hutch: we are always interested in their new adventures, but we already know all we need to know about their psychology, their habits, their capacities, their ethical standpoints.

I would add finally that form of seriality that, in cinema and television, is motivated less by the narrative structure than by the nature of the actor himself: the mere presence of John Wayne, or of Jerry Lewis (when they are not directed by a great director, and even in these cases)

succeeds in making, always, the *same* film. The author tries to invent different stories, but the public recognizes (with satisfaction) always and ever the same story, under superficial disguises.

The *saga* differs from the series insofar as it concerns the story of a family and is interested in the “historical” lapse of time. It is genealogical. In the *saga*, the actors do age; the *saga* is a history of the aging of individuals, families, people, groups.

The *saga* can have a continuous lineage (the character is followed from birth to death; the same is then done for his son, his grandson, and so on, potentially forever), or it can be treelike (there is a patriarch, then the various narrative branches that concern not only his direct descendents, but also the collateral lines and the kin, each branch branching out infinitely). The most familiar (and recent) instance of *saga* is certainly *Dallas*.

The *saga* is a series in disguise. It differs from the series in that the characters change (they change also because the actors age). But in reality the *saga* repeats, in spite of its historicized form, celebrating in appearance the passage of time, the same story. As with ancient *sagas*, the deeds of the gallant ancestors are the same as the deeds of their descendents. In *Dallas*, grandfathers and grandsons undergo more or less the same ordeals: struggle for wealth and for power, life, death, defeat, victory, adultery, love, hate, envy, illusion, and delusion.

I mean by intertextual dialogue the phenomenon by which a given text echoes previous texts. Many forms of intertextuality are outside my present concerns. I am not interested, for example, in stylistic quotation, in those cases

in which a text quotes, in a more or less explicit way, a stylistic feature, a way of narrating typical of another author – either as a form of parody or in order to pay homage to a great and acknowledged master. There are imperceptible quotations, of which not even the author is aware, that are the normal effect of the game of artistic influence. There are also quotations of which the author is aware but that should remain ungraspable by the consumer; in these cases we are usually in the presence of a banal work and plagiarism.

What is more interesting is when the quotation is explicit and recognizable, as happens in literature or postmodern art, which blatantly and ironically play on intertextuality (novel on the techniques of the narrative, poetry on poetry, art on art).

There is a procedure typical of the postmodern narrative that has been much used recently in the field of mass communications: it concerns the ironic quotation of the commonplace (*topos*). Let us remember the killing of the Arab giant in *Raiders of the Lost Ark* and the staircase of Odessa in Woody Allen’s *Bananas*. What joins these two quotations? In both cases, the spectator, in order to enjoy the allusion, must know the original *topoi*. In the case of the giant, it is a situation typical of the genre; in the case of *Bananas* – on the contrary – the *topos* appears for the first and only time in a single work, and only after that quotation the *topos* becomes a shibboleth for movie critics and moviegoers.

In both cases the *topoi* are recorded by the “encyclopedia” of the spectator; they make up a part of the treasury of the collective imagination and as such they come to be called upon. What differentiates the two quotations is the fact that the *topos* in *Raiders* is quoted in order to contradict it (what we expect to

happen, based on our experience, will not), while in *Bananas* the topos is introduced only because of its incongruity (the staircase has nothing to do with the rest of the story).

The first case recalls the series of cartoons that was published years ago by *Mad*, under the heading “a film which we would like to see.” For example, the heroine, in the West, tied by bandits to the railroad tracks: the alternating shots show on one side the approaching train and on the other the furious cavalcade of rescuers trying to arrive ahead of the locomotive. In the end, the girl (contrary to all the expectations suggested by the topos evoked) is crushed by the train. Here we are faced with a comic ploy which exploits the presupposition (correct) that the public will recognize the original topos, will apply to the quotation the “normal” system of expectations (I mean the expectations that this piece of encyclopedical information is supposed to elicit), and will then enjoy the way in which its expectations are frustrated. At this point the ingenuous spectator, at first frustrated, overcomes his frustration and transforms himself into a critical spectator who appreciates the way in which he was tricked.

In the case of *Bananas*, however, we are at a different level: the spectator with whom the text establishes an implicit agreement (tongue-in-cheek) is not the ingenuous spectator (who can be struck at most by the apparition of an incongruous event) but the critical spectator who appreciates the ironic ploy of the quotation and enjoys its desired incongruity. However, in both cases we have a critical side effect: aware of the quotation, the spectator is brought to elaborate ironically on the nature of such a device and to acknowledge the fact that he has been invited to play upon his encyclopedic competence.

The game becomes complicated in the “retake” of *Raiders*, that is, in *Indiana Jones and the Temple of Doom*: here the hero encounters not one but two giant enemies. In the first case, we are expecting that, according to the classical schemes of the adventure film, the hero will be disarmed, and we laugh when we discover that instead the hero has a pistol and easily kills his adversary. In the second case, the director knows that the spectators (having already seen the preceding film), will expect the hero to be armed, and indeed, Indiana Jones quickly looks for his pistol. He does not find it, and the spectators laugh because the expectation created by the first film is this time frustrated.

The cases cited put into play an intertextual encyclopedia. We have texts that are quoted from other texts and the knowledge of the preceding ones – taken for granted – is supposed to be necessary for the enjoyment of the new one.

More interesting for the analysis of the new intertextuality in the media is the example of *E.T.*, in the scene where the creature from outer space (an invention of Spielberg) is led into a city during Halloween and encounters another personage, disguised as the gnome in *The Empire Strikes Back* (an invention of Lucas). *E.T.* is jolted and seeks to hurl himself upon the gnome in order to embrace him, as if he had met an old friend. Here the spectators must know many things: they must certainly know of the existence of another film (intertextual knowledge), but they must also know that both monsters were created by Rambaldi, that the directors of the two films are linked together for various reasons (not least because they are the two most successful directors of the decade); they must, in short, have not only a knowledge of the texts but also a knowledge of the world, circumstances



external to the texts. One notices, naturally, that knowledge of the texts and the world are only two possible chapters of encyclopedic knowledge, and that therefore, in a certain measure, the text always makes reference to the same cultural patrimony.

Such phenomena of “intertextual dialogue” were once typical of experimental art, and presupposed a Model Reader, culturally very sophisticated.<sup>3</sup> The fact that similar devices have now become more common in the media world leads us to see that the media are carrying on – and presupposing – the possession of pieces of information already conveyed by other media.

The text of *E.T.* “knows” that the public has learned, from newspapers or from television, everything about Rambaldi, Lucas, and Spielberg. The media seem, in this play of extratextual quotation, to make reference to the world, but in effect they are referring to the contents of other messages sent by other media. The game is played, so to speak, on a “broadened” intertextuality. Any difference between knowledge of the world (understood naively as a knowledge derived from an extratextual experience) and intertextual knowledge has practically vanished. Our reflections to come, then, must not only question the phenomenon of repetition within a single work or a series of works, but all the phenomena that make various strategies of repetition producible, understandable, and commercially possible. In other words, repetition and seriality in the media pose new problems for the sociology of culture.

Another form of intertextuality is the genre-embedding that today is very common in the mass media. For exam-

ple, every Broadway musical (in the theater or on film) is, as a rule, nothing other than the story of how a Broadway musical is put on. The Broadway genre seems to require (postulate) a vast intertextual knowledge: in fact, it creates and institutes the required competence and the presuppositions indispensable to its understanding. Every one of these films or plays tells how a Broadway musical is put on, and furnishes us, in effect, with all the information about the genre it belongs to. The spectacle gives the public the sensation of knowing ahead of time that which it does not yet know and will know only at the moment. We stand facing the case of a colossal pretention (or “passing over”). In this sense, the musical is a didactic work that takes account of the (idealized) rules of its own production.

Finally, we have the work that speaks of itself: not the work that speaks of a genre to which it belongs, but a work that speaks of its own structure, and of the way in which it was made. Critics and aestheticians were inclined to think that this device was an exclusive feature of the works of the avant-garde and alien to mass communications. Aesthetics knows this problem and indeed gave it a name long ago: it is the Hegelian problem of the Death of Art. But in these later times there have been cases of productions in the mass media capable of self-irony, and some of the examples mentioned above seem to me of great interest. Even here the line between “highbrow” arts and “lowbrow” arts seems to have become very thin.

Let us now try to review the phenomena listed above from the point of view of a “modern” conception of aesthetic value, according to which every work aesthetically “well done” is endowed with two characteristics:

<sup>3</sup> Cf., for the idea of “the model reader,” my *The Role of the Reader*.

(1) It must achieve a dialectic between order and novelty – in other words, between scheme and innovation;

(2) This dialectic must be perceived by the consumer, who must not only grasp the contents of the message, but also the way in which the message transmits those contents.

This being the case, nothing prevents the types of repetition listed above from achieving the conditions necessary to the realization of aesthetic value, and the history of the arts is ready to furnish us with satisfactory examples for each of the types in our classification.

*Retake.* Ariosto's *Orlando Furioso* is nothing else but a retake of Boiardo's *Orlando Innamorato*, and precisely because of the success of the first, which was in its turn a retake of the themes of the Breton cycle. Boiardo and Ariosto added a goodly amount of irony into material that was very "serious" and "taken seriously" by previous readers. But even the third *Superman* is ironical in regard to the first (mystical and very, very serious). It appears as the retake of an archetype inspired by the gospel, made by winking at the films of Frank Tashlin.

*Remake.* I have already suggested that Shakespeare remade a lot of very well-known stories of the previous centuries.

*Series.* Every text presupposes and constructs always a double Model Reader (let us say, a naive and a "smart" one). The former uses the work as semantic machinery and is the victim of the strategies of the author who will lead him little by little along a series of previsions and expectations; the latter evaluates the work as an aesthetic product and enjoys the strategies implemented in order to produce a model reader of the first level. This second-level reader is the one who enjoys the seriality of the series, not so much for the return of the same thing (that the ingenuous reader believed was

different) but for the strategy of the variations; in other words, he enjoys the way in which the same story is worked over to appear to be different.

This enjoyment of variations is obviously encouraged by the more sophisticated series. Indeed, we can classify the products of serial narratives along a continuum that takes into account the different gradations of the reading agreement between the text and the "smart" reader (as opposed to the naive one). It is evident that even the most banal narrative product allows the reader to become, by an autonomous decision, a critical reader, able to recognize the innovative strategies (if any). But there are serial works that establish an explicit agreement with the critical reader and thus, so to speak, challenge him to acknowledge the innovative aspects of the text.

Belonging to this category are the television films of Lieutenant Columbo. It is worth noticing that in this series the authors spell out from the beginning who the murderer is. The spectator is not so much invited to play the naive game of guessing (whodunit?) as (1) to enjoy Columbo's detection technique, appreciated as an encore to a well-known piece of bravura (and in this sense the pleasure provided by Columbo is not so different from the one provided by Nero Wolfe); and (2) to discover in what way the author will succeed in winning his bet, which consists in having Columbo do what he always does, but nevertheless in a way that is not banally repetitive. Every story of Nero Wolfe was written by Rex Stout, but every episode of *Columbo* is directed by a different person. The critical addressee is invited to pronounce a judgment on the best variation.

I use the term "variation" thinking of the classical musical variations. They,

too, were “serial products” that aimed very little at the naive addressee and that bet everything on an agreement with the critical one. The composer was fundamentally interested only in the applause of the critical listener, who was supposed to appreciate the fantasy displayed in his innovations on an old theme.

In this sense, seriality and repetition are not opposed to innovation. Nothing is more “serial” than a tie pattern, and yet nothing can be so personalized as a tie. The example may be elementary, but that does not make it banal. Between the elementary aesthetics of the tie and the recognized “high” artistic value of the Goldberg Variations, there is a graded continuum of repetitious strategies, aimed at the response of the “smart” addressee.

The problem is that there is not, on the one hand, an aesthetics of “high” art (original and not serial), and on the other a pure sociology of the serial. Rather, there is an aesthetics of serial forms that requires an historical and anthropological study of the ways in which, at different times and in different places, the dialectic between repetition and innovation has been instantiated. When we fail to find innovation in the serial, this is perhaps less a result of the structures of the text, than of our “horizon of expectations” and our cultural habits. We know very well that in certain examples of non-Western art, where we always see the same thing, the natives recognize infinitesimal variations and feel the shiver of innovation. Where we see innovation, at least in the serial forms of the Western past, the original addressees were not at all interested in that aspect and conversely enjoyed the recurrences of the scheme.

*Saga.* The entire *Human Comedy* by Balzac presents a very good example of a

treelike saga, as much as *Dallas* does. Balzac is more interesting than *Dallas* because every one of his novels increases our knowledge of the society of his time, while every program of *Dallas* tells us the same thing about American society – but both use the same narrative scheme.

*Intertextuality.* The notion of intertextuality itself has been elaborated within the framework of a reflection on “high” art. Notwithstanding, the examples given above have been taken up provocatively by the world of mass communication in order to show how even these forms of intertextual dialogue have by now been transferred to the field of popular production.

It is typical of what is called postmodern literature and art (but did it not already happen thus with the music of Stravinsky?) to quote by using (sometimes under various stylistic disguises) *quotation marks* so that the reader pays no attention to the content of the citation but instead to the way in which the excerpt from a first text is introduced into the fabric of a second one. Renato Barilli has observed that one of the risks of this procedure is the failure to make the quotation marks evident, so that what is cited is accepted by the naive reader as an original invention rather than as an ironic reference.<sup>4</sup>

We have so far put forward three examples of quotations of a previous topos: *Raiders of the Lost Ark*, *Bananas*, and *E.T.* Let us look closer at the third case: the spectator who knows nothing of the production of two films (in which one quotes from the other) cannot succeed in understanding why what happens does happen. By that gag, the movie focuses both upon movies and upon the media-universe. The understanding of

<sup>4</sup> “Dal leggibile all’illeggibile,” in Luigi Russo, ed., *Letteratura tra consumo e ricerca* (Bologna: Mulino, 1984).

this device is a condition for its aesthetic enjoyment. Thus, this episode can work only if one realizes that there are quotation marks somewhere. One can say that these marks can be perceived only on the basis of an extratextual knowledge. Nothing *in the film* helps the spectator to understand at what point there ought to be quotation marks. The film presupposes a previous world-knowledge on the part of the spectator. And if the spectator does not know? Too bad. The effect gets lost, but the film knows of other means to gain approval.

These imperceptible quotation marks, more than an aesthetic device, are a social artifice; they select the happy few (and the mass media usually hopes to produce millions of happy few . . .). To the naive spectator of the first level, the film has already given almost too much: that secret pleasure is reserved, for that time, for the critical spectator of the second level.

The case of *Raiders* is different. Here, if the critical spectator fails – does not recognize the quotation – there remain plenty of possibilities for the naive spectator, who at least can always enjoy the fact that the hero gets the best of his adversary. We are here confronted by a less subtle strategy than in the preceding example, a mode inclined to satisfy the urgent need of the producer, who in any case must sell his product to whomever he can. While it is difficult to imagine *Raiders* being seen and enjoyed by those spectators who do not grasp the interplay of quotations, it is always possible that this will happen, and the work is clearly open even to this possibility.

I do not feel like saying which, between the two texts cited, pursues the “more aesthetically noble” ends. It is enough for me (and perhaps for the moment I have already given myself much to think about) to point out a

(critically relevant) difference in the functioning and use of textual strategy.

We come now to the case of *Bananas*. On that staircase there descend, not only a baby carriage, but also a platoon of rabbis and I do not remember what else. What happens to the spectator who has not caught the quotation from the *Potemkin* mixed up with imprecise fancies about the *Fiddler on the Roof*? I believe that because of the orgiastic energy with which the scene – the staircase with its incongruous population – is presented, even the most naive spectator may grasp the symphonic turbulence of this Brueghel-like kermis. Even the most ingenious among the spectators “feels” a rhythm, an invention, and cannot help but fix his attention on the way it is put together.

At the extreme other end of the pole of the aesthetic interests, I would like to mention a work whose equivalent I have not succeeded in finding in the contemporary mass media; it is not only a masterpiece of intertextuality but also a paramount example of narrative metalanguage, which speaks of its own formation and of the rules of the narrative genre: I refer to *Tristram Shandy*.

It is impossible to read and enjoy Sterne’s antinovel novel without realizing that it is treating the novel form ironically. *Tristram Shandy* is so aware of its nature that it is impossible to find there a single ironical statement that does not make evident its own quotation marks. It brings to a high artistic resolution the rhetorical device called *pronuntiatio* (that is, the way of stressing imperceptibly the irony).

I believe that I have singled out a typology of “quotation marking” that must in some way be relevant to the ends of a phenomenology of aesthetic value, and of the pleasure that follows from it. I believe further that the strate-

gies for matching surprise and novelty with repetition, even if they are semiotic devices in themselves aesthetically neutral, can give place to different results on the aesthetic level.

Some conclusions follow:

Each of the types of repetition that we have examined is not limited to the mass media, but belongs by right to the entire history of artistic creativity: plagiarism, quotation, parody, the ironic retake, the intertextual joke, are typical of the entire artistic-literary tradition.

Much art has been and is repetitive. The concept of absolute originality is a contemporary one, born with Romanticism; classical art was in vast measure serial, and the “modern” avant-garde (at the beginning of this century) challenged the Romantic idea of “creation from nothingness,” with its techniques of collage, mustachios on the Mona Lisa, art about art, and so on.

The same type of repetitive procedure can produce either excellence or banality; it can put the addressees into conflict with themselves and with the intertextual tradition as a whole; thus it can provide them with easy consolations, projections, identifications: it can establish an agreement exclusively with the naive addressee, or exclusively with the smart one, or with both at different levels and along a continuum of solutions that cannot be reduced to a rudimentary typology.

Nevertheless, a typology of repetition does not furnish the criteria that can establish differences in aesthetic values.

Yet, since the various types of repetition are present in the whole of artistic and literary history, they can be taken into account in order to establish criteria of artistic value. An aesthetics of repetition requires as a premise a semiotics of the textual procedures of repetition.

I realize that all I have said until now still represents an attempt to reconsider the various forms of repetition in the media in terms of the “modern” dialectic between order and innovation. The fact, however, is that when one speaks today of the aesthetics of seriality, one alludes to something more radical, that is, to a notion of aesthetic value that wholly escapes the “modern” idea of art and literature.<sup>5</sup>

It has been observed that with the phenomenon of television serials we find a new concept of “the infinity of the text”; the text takes on the rhythms of that same dailiness in which it is produced, and that it mirrors. The problem is not one of recognizing that the serial text varies indefinitely upon a basic scheme (and in this sense it can be judged from the point of view of the “modern” aesthetics). The real problem is that what is of interest is not so much the single variations as “variability” as a formal principle, the fact that one can make variations to infinity. Variability to infinity has all the characteristics of repetition, and very little of innovation. But it is the “infinity” of the process that gives a new sense to the device of variation. What must be enjoyed – suggests the postmodern aesthetics – is the fact that a series of possible variations is potentially infinite. What becomes celebrated here is a sort of victory of life over art, with the paradoxical result that the era of electronics – instead of emphasizing the phenomena of shock, interruptions, novelty, and frustration of expectations – would produce a return to the continuum, the Cyclical, the Periodical, the Regular.

5 The “manifesto” of this new aesthetics of seriality is the special issue of the journal *Cinema & Cinema* 35–36 (1983): 20–24.



Omar Calabrese has thoroughly looked into this:<sup>6</sup> from the point of view of the “modern” dialectic between repetition and innovation, one can easily recognize how in the Columbo series, for example, on a basic scheme some of the best names in American cinema have worked in variations. Thus it would be difficult to speak, in such a case, of pure repetition: if the scheme of the detection and the psychology of the protagonist actor remains unchanged, the style of the narrative changes each time. This is no small thing, especially from the point of view of the “modern” aesthetics. But it is exactly on a different idea of style that Calabrese’s paper is centered. In these forms of repetition “we are not so much interested in what is repeated as we are in the way the components of the text come to be segmented and then how the segments come to be codified in order to establish a system of invariants: any component that does not belong to the system, can be defined as an *independent variable*.” In the most typical and apparently “degenerated” cases of seriality, the independent variables are not altogether the more visible, but the more microscopic, as in a homeopathic solution where the potion is all the more potent because by further “succussions” the original particles of the medicinal product have almost disappeared. This is what permits Calabrese to speak of the Columbo series as an “exercice de style” à la Queneau. We are, says Calabrese, facing a “neobaroque aesthetics” that is instantiated, not only by the “cultivated” products, but even, and above all, by those that are most degenerated. Apropos of *Dallas*, one can say that “the semantic opposition and the articulation of the elementary narrative structures

can migrate in combinations of the highest improbability around the various characters.”

Organized differentiations, polycentrism, regulated irregularity – such would be the fundamental aspects of this neobaroque aesthetic, the principal example of which is musical variations à la Bach. Since in the epoch of mass communications “the condition for listening . . . it is that for which all has already been said and already been written . . . as in the Kabuki theater, it may then be the most minuscule variant that will produce pleasure in the text, or that form of explicit repetition which is already known.”

What results from these reflections is clear. The focus of the theoretical inquiry is displaced. Before, mass mediologists tried to save the dignity of repetition by recognizing in it the possibility of a traditional dialectic between scheme and innovation (but it was still the innovation that accounted for the value, the way of rescuing the product from degradation and promoting it to a value). Now, the emphasis must be placed on the inseparable knot of scheme-variation, where the variation is no longer more appreciable than the scheme. The term *neobaroque* must not deceive: we are witnessing the birth of a new aesthetic sensibility much more archaic, and truly post-postmodern.

As Giovanna Grignaffini observes, “the neobaroque aesthetics has transformed a commercial constraint into a ‘formal principle.’” As a result, “any idea of unicity becomes destroyed to its very roots.”<sup>7</sup> As happened with Baroque music, and as (according to Walter Benjamin) happens in our era of “technological reproduction,” the messages of mass

6 “I replicanti,” *Cinema & Cinema* 35–36 (1983): 25–39.

7 “J.R. : vi presento il racconto,” *Cinema & Cinema* 35–36 (1983): 46–51.

media can and must be received and understood in a “state of inattention.”

It goes without saying that the authors I have quoted see very clearly how much commercial and “gastronomical” consolation there is in putting forward stories that always say the same thing and in a circular way always close in on themselves. But they do not only apply to such products a rigidly formalistic criterion, but also suggest that we ought to conceive of a new audience that feels perfectly comfortable with such a criterion. Only by presupposing such agreement can one speak of a new aesthetics of the serial. And only by such an agreement is the serial no longer the poor relative of the arts, but the form of art that can satisfy the new aesthetic sensibility, indeed, the post-postmodern Greek tragedy.

We would not be scandalized if such criteria were to be applied (as they have been applied) to abstract art. And in fact, here we are about to outline a new aesthetics of the “abstract” applied to the products of mass communication.

But this requires that the naive addressee of the first level will disappear, giving place only to the critical reader of the second level. In fact, there is no conceivable naive addressee of an abstract painting or sculpture. If there is one who – in front of them – asks, “But what does it mean?” this is not an addressee of either the first or second level; he is excluded from any artistic experience whatever. Of abstract works there is only a critical “reading”: what is formed is of no interest, only the way it is formed is interesting.

Can we expect the same for the serial products of television? What should we think about the birth of a new public that, indifferent to the stories told (which are in any case already known), only relishes the repetition and its own

microscopic variations? In spite of the fact that today the spectator still weeps in the face of the Texan families’ tribulations, ought we to expect in the near future a true and real genetic mutation?

If it should not happen this way, the radical proposal of the postmodern aesthetics would appear singularly snobby: as in a sort of neo-Orwellian world, the pleasures of the smart reading would be reserved for the members of the Party and the pleasures of the naive reading reserved for the proletarians. The entire industry of the serial would exist, as in the world of Mallarmé (made to end in a Book), with its only aim being to furnish neobaroque pleasure to the happy few, reserving pity and fear to the unhappy many who remain.

According to this hypothesis we should think of a universe of new consumers disinterested in what really happens to J.R., and bent on grasping the neobaroque pleasure provided by the form of his adventures. However, one could ask if such an outlook (even though warranting a *new* aesthetics) can be agreed to by an *old* semiotics.

Baroque music, as well as abstract art, is “a-semantic.” One can discuss, and I am the first to do so, whether it is possible to discriminate so straightforwardly between purely “syntactic” and “semantic” arts. But may we at least recognize that there are figurative arts and abstract arts? Baroque music and abstract painting are not figurative; television serials are.

Until what point shall we be able to enjoy as merely musical those variations that play upon “likenesses”? Can one escape from the fascination of the possible worlds that these “likenesses” outline?

Perhaps we are obliged to try a different hypothesis.

We can say then that the neobaroque series brings to its first level of fruition (impossible to eliminate) the pure and simple myth. Myth has nothing to do with art. It is a story, always the same. It may not be the story of Atreus and it may be that of J.R. Why not? Every epoch has its mythmakers, its own sense of the sacred. Let us take for granted such a “figurative” representation and such an “orgiastic” enjoyment of the myth. Let us take for granted the intense emotional participation, the pleasure of the reiteration of a single and constant truth, and the tears, and the laughter – and finally the *catharsis*. Then we can conceive of an audience also able to shift onto an aesthetic level and to judge the art of the variations on a mythical theme – in the same way as one succeeds in appreciating a “beautiful funeral” even when the deceased was a dear person.

Are we sure that the same thing did not happen even with the classical tragedy?

If we reread Aristotle’s *Poetics* we see that it was possible to describe the model of a Greek tragedy as a *serial one*. From the quotations of the Stagirite we realize that the tragedies of which he had knowledge were many more than have come down to us, and they all followed (by varying it) one fixed scheme. We can suppose that those that have been saved were those that corresponded better to the canons of the ancient aesthetic sensibility. But we could also suppose that the decimation came about on the basis of political-cultural criteria, and no one can forbid us from imagining that Sophocles may have survived by virtue of a political maneuver, by sacrificing better authors (but “better” according to what criteria?).

If there were many more tragedies than those we know, and if they all followed (with variations) a fixed scheme,

what would happen if today we were able to see them and read them all together? Would our evaluations of the originality of Sophocles or Aeschylus be different from what they are currently? Would we find in these authors variations on topical themes where today we see indistinctly a unique (and sublime) way of confronting the problems of the human condition? Perhaps where we see absolute invention, the Greeks would have seen only the “correct” variation on a single scheme, and sublime appeared to them, not the single work, but precisely the scheme. It is not by chance that, when speaking of the art of poetry, Aristotle dealt mainly with schemes before all else, and mentioned single works only for the sake of an example.

Since at this point I am playing what Peirce called “the play of musement” and I am multiplying the hypotheses – in order to find out, maybe later, a single fruitful idea – let us now reverse our experiment and look at a contemporary TV serial from the point of view of a future neoromantic aesthetics which, supposedly, has assumed again that “originality is beautiful.” Let us imagine a society in the year 3000 A.D., in which 90 percent of all our present cultural production had been destroyed and of all our television serials only *one* show of Lieutenant Columbo had survived.

How would we “read” this work? Would we be moved by such an original picture of a little man in the struggle with the powers of evil, with the forces of capital, with an opulent and racist society dominated by WASPs? Would we appreciate this efficient, concise, and intense representation of the urban landscape of an industrial America?

When – in a single piece of a series – something is simply *presupposed* by the audience, which knows the whole series,

would we speak perhaps of an art of synthesis of a sublime capacity of telling through essential allusions?

In other words, how would we read a “piece” of a series, if the whole of the series remained unknown to us?

Such a series of questions could continue indefinitely. I started to put them forth because I think that we still know very little about the role of repetition in the universe of art and in the universe of mass media.

*Innovation  
& repetition:  
between  
modern &  
postmodern  
aesthetics*

# Susan Sontag

## *An argument about beauty*

### 1

Responding at last, in April of 2002, to the scandal created by the revelation of innumerable cover-ups of sexually predatory priests, Pope John Paul II told the American cardinals summoned to the Vatican, “A great work of art may be blemished, but its beauty remains; and this is a truth which any intellectually honest critic will recognize.”

Is it too odd that the Pope likens the Catholic Church to a great – that is, beautiful – work of art? Perhaps not, since the inane comparison allows him to turn abhorrent misdeeds into something like the scratches in the print of a silent film or craquelure covering the surface of an Old Master painting, blemishes that we reflexively screen out or see past. The Pope likes venerable ideas. And beauty, as a term signifying (like

---

*Susan Sontag was elected a Fellow of the American Academy in 1993. Best known as a novelist and essayist – her books have been translated into thirty-two languages – she also composed stories and plays, wrote and directed movies, and worked as a theatre director in the United States and Europe. In 2000 she won the National Book Award for her novel “In America,” and in 2001 received the Jerusalem Prize for the body of her work. This piece appeared in the Fall 2002 issue of “Daedalus.” She died in 2004.*

---

© Susan Sontag

health) an indisputable excellence, has been a perennial resource in the issuing of peremptory evaluations.

Permanence, however, is not one of beauty’s more obvious attributes; and the contemplation of beauty, when it is expert, may be wreathed in pathos, the drama on which Shakespeare elaborates in many of the Sonnets. Traditional celebrations of beauty in Japan, like the annual rite of cherry-blossom viewing, are keenly elegiac; the most stirring beauty is the most evanescent. To make beauty in some sense imperishable required a lot of conceptual tinkering and transposing, but the idea was simply too alluring, too potent, to be squandered on the praise of superior embodiments. The aim was to multiply the notion, to allow for kinds of beauty, beauty with adjectives, arranged on a scale of ascending value and incorruptibility, with the metaphorized uses (‘intellectual beauty,’ ‘spiritual beauty’) taking precedence over what ordinary language extols as beautiful – a gladness to the senses.

The less ‘uplifting’ beauty of face and body remains the most commonly visited site of the beautiful. But one would hardly expect the Pope to invoke *that* sense of beauty while constructing an exculpatory account of several generations’ worth of the clergy’s sexual molestation of children and protection of



the molesters. More to the point – his point – is the ‘higher’ beauty of art. However much art may seem to be a matter of surface and reception by the senses, it has generally been accorded an honorary citizenship in the domain of ‘inner’ (as opposed to ‘outer’) beauty. Beauty, it seems, is immutable, at least when incarnated – fixed – in the form of art, because it is in art that beauty as an idea, an eternal idea, is best embodied. Beauty (should you choose to use the word that way) is deep, not superficial; hidden, sometimes, rather than obvious; consoling, not troubling; indestructible, as in art, rather than ephemeral, as in nature. Beauty, the stipulatively uplifting kind, perdures.

## 2

The best theory of beauty is its history. Thinking about the history of beauty means focusing on its deployment in the hands of specific communities.

Communities dedicated by their leaders to stemming what is perceived as a noxious tide of innovative views have no interest in modifying the bulwark provided by the use of beauty as unexceptionable commendation and consolation. It is not surprising that John Paul II, and the preserve-and- conserve institution for which he speaks, feels as comfortable with beauty as with the idea of the good.

It also seems inevitable that when, almost a century ago, the most prestigious communities concerned with the fine arts dedicated themselves to drastic projects of innovation, beauty would turn up on the front line of notions to be discredited. Beauty could not but appear a conservative standard to the makers and proclaimers of the new; Gertrude Stein said that to call a work of art beautiful means that it is dead. Beautiful has come

to mean ‘merely’ beautiful: there is no more vapid or philistine compliment.

Elsewhere, beauty still reigns, irrepressible. (How could it not?) When that notorious beauty-lover Oscar Wilde announced in *The Decay of Lying*, “Nobody of any real culture ever talks about the beauty of a sunset. Sunsets are quite old-fashioned,” sunsets reeled under the blow, then recovered. *Les beaux-arts*, when summoned to a similar call to be up-to-date, did not. The subtraction of beauty as a standard for art hardly signals a decline of the authority of beauty. Rather, it testifies to a decline in the belief that there is something called art.

## 3

Even when Beauty was an unquestioned criterion of value in the arts, it was defined laterally, by evoking some other quality that was supposed to be the essence or *sine qua non* of something that was beautiful. A definition of the beautiful was no more (or less) than a commendation of the beautiful. When, for example, Lessing equated beauty with harmony, he was offering another general idea of what is excellent or desirable.

In the absence of a definition in the strict sense, there was supposed to be an organ or capacity for registering beauty (that is, value) in the arts, called ‘taste,’ and a canon of works discerned by people of taste, seekers after more rarefied gratifications, adepts of connoisseurship. For in the arts – unlike life – beauty was not assumed to be necessarily apparent, evident, obvious.

The problem with taste was that, however much it resulted in periods of large agreement within communities of art lovers, it issued from private, immediate, and revocable responses to art. And the consensus, however firm, was never

more than local. To address this defect, Kant – a dedicated universalizer – proposed a distinctive faculty of ‘judgment’ with discernable principles of a general and abiding kind; the tastes legislated by this faculty of judgment, if properly reflected upon, should be the possession of all. But ‘judgment’ did not have its intended effect of shoring up ‘taste’ or making it, in a certain sense, more democratic. For one thing, taste-as-principled-judgment was hard to apply, since it had the most tenuous connection with the actual works of art deemed incontestably great or beautiful, unlike the pliable, empirical criterion of taste. And taste is now a far weaker, more assailable notion than it was in the late eighteenth century. *Whose taste?* Or, more insolently, *who sez?*

As the relativistic stance in cultural matters pressed harder on the old assessments, definitions of beauty – descriptions of its essence – became emptier. Beauty could no longer be something as positive as harmony. For Valéry, the nature of beauty is that it cannot be defined; beauty is precisely ‘the ineffable.’

The failure of the notion of beauty reflects the discrediting of the prestige of judgment itself, as something that could conceivably be impartial or objective, not always self-serving or self-referring. It also reflects the discrediting of binary discourses in the arts. Beauty defines itself as the antithesis of the ugly. Obviously, you can’t say something is beautiful if you’re not willing to say something is ugly. But there are more and more taboos about calling something, anything, ugly. (For an explanation, look first not at the rise of so-called political correctness, but at the evolving ideology of consumerism, then at the complicity between these two.) The point is to find what is beautiful in what has not hither-

to been regarded as beautiful (or: the beautiful in the ugly).

Similarly, there is more and more resistance to the idea of ‘good taste,’ that is, to the dichotomy good taste/bad taste, except for occasions that allow one to celebrate the defeat of snobbery and the triumph of what was once condescended to as bad taste. Today, good taste seems even more retrograde an idea than beauty. Austere, difficult ‘modernist’ art and literature have come to seem old-fashioned, a conspiracy of snobs. Innovation is relaxation now; today’s E-Z Art gives the green light to all. In the cultural climate favoring the more user-friendly art of recent years, the beautiful seems, if not obvious, then pretentious. Beauty continues to take a battering in what are called, absurdly, our culture wars.

#### 4

That beauty applied to some things and not to others, that it was a principle of *discrimination*, was once its strength and appeal. Beauty belonged to the family of notions that establish rank, and accorded well with social order unapologetic about station, class, hierarchy, and the right to exclude.

What had been a virtue of the concept became its liability. Beauty, which once seemed vulnerable because it was too general, loose, porous, was revealed as – on the contrary – excluding too much. Discrimination, once a positive faculty (meaning refined judgment, high standards, fastidiousness), turned negative: it meant prejudice, bigotry, blindness to the virtues of what was not identical with oneself.

The strongest, most successful move against beauty was in the arts: beauty, and the caring about beauty, was restrictive; as the current idiom has it, elitist. Our appreciations, it was felt, could be

so much more inclusive if we said that something, instead of being beautiful, was ‘interesting.’

Of course, when people said a work of art was interesting, this did not mean that they necessarily liked it – much less that they thought it beautiful. It usually meant no more than they thought they ought to like it. Or that they liked it, sort of, even though it wasn’t beautiful.

Or they might describe something as interesting to avoid the banality of calling it beautiful. Photography was the art where ‘the interesting’ first triumphed, and early on: the new, photographic way of seeing proposed everything as a potential subject for the camera. The beautiful could not have yielded such a range of subjects; and soon came to seem uncool to boot as a judgment. Of a photograph of a sunset, a beautiful sunset, anyone with minimal standards of verbal sophistication might well prefer to say, “Yes, the photograph is interesting.”

## 5

*What is interesting?* Mostly, what has not previously been thought beautiful (or good). The sick are interesting, as Nietzsche points out. The wicked, too. To name something as interesting implies challenging old orders of praise; such judgments aspire to be found insolent or at least ingenious. Connoisseurs of the interesting – whose antonym is the boring – appreciate clash, not harmony. Liberalism is boring, declares Carl Schmitt in *The Concept of the Political*, written in 1932 (the following year he joined the Nazi Party). A politics conducted according to liberal principles lacks drama, flavor, conflict, while strong autocratic politics – and war – are interesting.

Long use of ‘the interesting’ as a criterion of value has, inevitably, weakened its transgressive bite. What is left of the

old insolence lies mainly in its disdain for the consequences of actions and of judgments. As for the truthfulness of the ascription – that does not even enter the story. One calls something interesting precisely so as not to have to commit to a judgment of beauty (or of goodness). The interesting is now mainly a consumerist concept, bent on enlarging its domain: the more things that become interesting, the more the marketplace grows. The boring – understood as an absence, an emptiness – implies its antidote: the promiscuous, empty affirmations of the interesting. It is a peculiarly inconclusive way of experiencing reality.

In order to enrich this deprived take on our experiences, one would have to acknowledge a full notion of boredom: depression, rage (suppressed despair). Then one could work toward a full notion of the interesting. But that quality of experience – of feeling – one would probably no longer even *want* to call interesting.

## 6

Beauty can illustrate an ideal; a perfection. Or, because of its identification with women (more accurately, with Woman), it can trigger the usual ambivalence that stems from the age-old denigration of the feminine. Much of the discrediting of beauty needs to be understood as a result of the gender inflection. Misogyny, too, might underlie the urge to metaphorize beauty, thereby promoting it out of the realm of the ‘merely’ feminine, the unserious, the specious. For if women are worshiped because they are beautiful, they are condescended to for their preoccupation with making or keeping themselves beautiful. Beauty is theatrical, it is for being looked at and admired; and the word is as likely to suggest the beauty industry (beauty magazines, beauty parlors, beauty prod-

ucts) – the theatre of feminine frivolity – as the beauties of art and of nature. How else to explain the association of beauty – i.e., women – with mindlessness? To be concerned with one’s own beauty is to risk the charge of narcissism and frivolity. Consider all the beauty synonyms, starting with the ‘lovely,’ the merely ‘pretty,’ which cry out for a virile transposition.

“Handsome is as handsome does.” (But not: “Beautiful is as beautiful does.”) Though it applies no less than does ‘beautiful’ to appearance, ‘handsome’ – free of associations with the feminine – seems a more sober, less gushing way of commending. Beauty is not ordinarily associated with gravitas. Thus one might prefer to call the vehicle for delivering searing images of war and atrocity a ‘handsome book,’ as I did in the preface to a recent compilation of photographs by Don McCullin, lest calling it a ‘beautiful book’ (which it was) would seem an affront to its appalling subject.

## 7

It’s usually assumed that beauty is, almost tautologically, an ‘aesthetic’ category, which puts it, according to many, on a collision course with the ethical. But beauty, even beauty in the amoral mode, is never naked. And the ascription of beauty is never unmixed with moral values. Far from the aesthetic and the ethical being poles apart, as Kierkegaard and Tolstoy insisted, the aesthetic is itself a quasi-moral project. Arguments about beauty since Plato are stocked with questions about the proper relation to the beautiful (the irresistibly, enthrallingly beautiful), which is thought to flow from the nature of beauty itself.

The perennial tendency to make of beauty itself a binary concept, to split

it up into ‘inner’ and ‘outer,’ ‘higher’ and ‘lower’ beauty, is the usual way that judgments of the beautiful are colonized by moral judgments. From a Nietzschean (or Wildean) point of view, this may be improper, but it seems to me unavoidable. And the wisdom that becomes available over a deep, lifelong engagement with the aesthetic cannot, I venture to say, be duplicated by any other kind of seriousness. Indeed, the various definitions of beauty come at least as close to a plausible characterization of virtue, and of a fuller humanity, as the attempts to define goodness as such.

## 8

Beauty is part of the history of idealizing, which is itself part of the history of consolation. But beauty may not always console. The beauty of face and figure torments, subjugates; that beauty is imperious. The beauty that is human, and the beauty that is made (art) – both raise the fantasy of possession. Our model of the disinterested comes from the beauty of nature – a nature that is distant, overarching, unpossessable.

From a letter written by a German soldier standing guard in the Russian winter in late December of 1942: “The most beautiful Christmas I had ever seen, made entirely of disinterested emotions and stripped of all tawdry trimmings. I was all alone beneath an enormous starred sky, and I can remember a tear running down my frozen cheek, a tear neither of pain nor of joy but of emotion created by intense experience. . . .”<sup>1</sup>

Unlike beauty, often fragile and impermanent, the capacity to be overwhelmed by the beautiful is astonishingly sturdy

1 Quoted in Stephen G. Fritz, *Frontsoldaten: The German Soldier in World War II* (Lexington, Ky.: University Press of Kentucky, 1995), 130.

and survives amidst the harshest distractions. Even war, even the prospect of certain death, cannot expunge it.

## 9

The beauty of art is better, ‘higher,’ according to Hegel, than the beauty of nature because it is made by human beings and is the work of the spirit. But the discerning of beauty in nature is also the result of traditions of consciousness, and of culture – in Hegel’s language, of spirit.

The responses to beauty in art and to beauty in nature are interdependent. As Wilde pointed out, art does more than school us on how and what to appreciate in nature. (He was thinking of poetry and painting. Today the standards of beauty in nature are largely set by photography.) What is beautiful reminds us of nature as such – of what lies beyond the human and the made – and thereby stimulates and deepens our sense of the sheer spread and fullness of reality, inanimate as well as pulsing, that surrounds us all.

A happy by-product of this insight, if insight it is: beauty regains its solidity, its inevitability, as a judgment needed to make sense of a large portion of one’s energies, affinities, and admirations; and the usurping notions appear ludicrous.

Imagine saying, “That sunset is interesting.”



---

AMERICAN ACADEMY  
OF ARTS & SCIENCES

*President*

Patricia Meyer Spacks

*Executive Officer*

Leslie Cohen Berlowitz

*Vice President*

Louis W. Cabot

*Secretary*

Jerrold Meinwald

*Treasurer*

John S. Reed

*Editor*

Steven Marcus

*Librarian*

Robert C. Post

*Vice President, Midwest Center*

Geoffrey R. Stone

*Vice President, Western Center*

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 & Public Relations*

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*: Nanjing now – a Confucian temple housing an American fried-chicken franchise. See Tu Wei-ming on *Cultural China: the periphery as the center*, pages 145 – 167: “. . . the real challenge to the mainland Chinese intellectuals is not the modern West per se but the modern West mediated through industrial East Asia.” Image © Liu Liqun/Corbis.



KFC  
肯德基

肯德基

新鲜登场!

田园脆鸡堡

肯德基

KFC

肯德基

肯德基

公司  
店

coming up in Dædalus:

on body in mind      Antonio & Hanna Damasio, Jerry Fodor, Carol Gilligan, Gerald Edelman, Jorie Graham, Raymond Dolan, Arne Öhman, Mark Johnson, Jacques d'Amboise, and William E. Connolly

on the humanities      Patricia Meyer Spacks, Steven Marcus, Andrew Delbanco, Pauline Yu, Gerald Early, Anthony Grafton, Thomas Crow, Jack Balkin & Sanford Levinson, and Dagfinn Føllesdal & Michael L. Friedman

on aging      Henry J. Aaron, Paul Baltes, Linda Partridge, Dennis J. Selkoe, Caleb E. Finch, Sarah Harper, Chris Wilson, Jagadeesh Gokhale & Kent Smetters, Hillard Kaplan, and Lisa Berkman

on identity      Akeel Bilgrami, Wendy Doniger, Amartya Sen, Stephen Greenblatt, Kwame Anthony Appiah, Sydney Shoemaker, Joseph Koerner, Susan Greenfield, David A. Hollinger, Claudio Lomnitz, Carol Rovane, 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, and Mia Bloom

on sex      Joan Roughgarden, Terry Castle, Steven Marcus, Claudia Goldin, Brian Charlesworth, Elizabeth Benedict, Wendy Doniger, Lawrence Cohen, Anne Fausto-Sterling, Catharine MacKinnon, Tim Birkhead, Margo Jefferson, and others

on capitalism  
& democracy      Joyce Appleby, John C. Bogle, Lucian Bebchuk, Robert W. Fogel, Jerry Z. Muller, Peter Bernstein, and others

plus poetry by Charles Wright, Peg Boyers, Kevin Carrizo di Camillo, John Kinsella, Charles Simic &c.; fiction by Ree Davis, Adam Braver, Dorian Gossy &c.; and notes by Robert F. Nagel, Jeri Laber, Michael Cook, Norbert Schwarz, Joel F. Handler, William B. Quandt, William Galston, Richard Morris, Robert J. Sharer &c.

U.S. \$13; [www.amacad.org](http://www.amacad.org)