Approaches to Social Theory

Approaches to Social Theory

SIEGWART LINDENBERG
JAMES S. COLEMAN
STEFAN NOWAK

Editors

Based on the W.I. Thomas and Florian Znaniecki Memorial Conference on Social Theory

RUSSELL SAGE FOUNDATION
New York

The Russell Sage Foundation

The Russell Sage Foundation, one of the oldest of America's general purpose foundations, was established in 1907 by Mrs. Margaret Olivia Sage for "the improvement of social and living conditions in the United States." The Foundation seeks to fulfill this mandate by fostering the development and dissemination of knowledge about the political, social, and economic problems of America. It conducts research in the social sciences and public policy, and publishes books and pamphlets that derive from this research.

The Board of Trustees is responsible for oversight and the general policies of the Foundation, while administrative direction of the program and staff is vested in the President, assisted by the officers and staff. The President bears final responsibility for the decision to publish a manuscript as a Russell Sage Foundation book. In reaching a judgment on the competence, accuracy, and objectivity of each study, the President is advised by the staff and selected expert readers. The conclusions and interpretations in Russell Sage Foundation publications are those of the authors and not of the Foundation, its Trustees, or its staff. Publication by the Foundation, therefore, does not imply endorsement of the contents of the study.

Library of Congress Cataloging-in-Publication Data

Approaches to social theory.

Proceedings of the W.I. Thomas and Florian Znaniecki Memorial Conference on Social Theory, held Nov. 9–12, 1983 at the University of Chicago under the auspices of the Russell Sage Foundation.
Bibliography: p. 389
Includes index.
1. Sociology—Methodology—Congresses. I. Lindenberg, S. (Siegwart), 1941– . II. Coleman, James Samuel, 1926– . III. Nowak, Stefan. IV. Thomas, William Isaac, 1863–1947. V. Znaniecki, Florian, 1882–1958. VI. W.I. Thomas and Florian Znaniecki Memorial Conference on Social Theory (1983 : University of Chicago) VII. Russell Sage Foundation.

HM13.A66 1986 301′.01 85-62806
ISBN 0-87154-205-6

Cover and text design: Huguette Franco

10 9 8 7 6 5 4 3 2 1

This volume, and the conference on Social Theory that constituted its basis, are dedicated to W. I. Thomas and Florian Znaniecki. Their collaboration at University of Chicago, which led to *The Polish Peasant in Europe and America*, was an achievement both in theory-relevant research and in sociological study that crosses national boundaries.

Contents

Preliminaries

Introduction

In autumn 1982, Stefan Nowak and David Featherman planned, within the broader framework of the program of cooperation in the social sciences between the Polish Academy of Sciences and the American Council of Learned Societies and with encouragement from the International Research and Exchanges Board (IREX), a series of Polish-American conferences in sociology. The general methodological theme of the conferences was intended to be the qualitative-quantitative chasm in sociology, with the aim of helping to bridge that chasm, both in theory and in research. The first conference was to be on social theory, to be held in the United States in autumn 1983, and to be organized by James Coleman on the American side and Stefan Nowak on the Polish side. In the Spring of 1983, the prospective participants received an invitation from Coleman and Nowak which said in part:

> The Conference will be held at the University of Chicago beginning with dinner Wednesday, November 9, 1983, and continuing through Saturday, November 12, 1983. . . .
> Such a conference can be especially useful now, for sociological theory is in some disarray, yet there is promising work going on in a number of directions. The aim of the conference will be to bring together persons working in these different directions in Poland and the United States, increasing the level of discourse

> both across national boundaries and across sub-disciplinary boundaries. A particular focus of the conference will be on the complementary contributions of qualitative and quantitative approaches to social theory. The result should be of value for the general enterprise of theory-construction in sociology, through the stimulus it provides.
>
> Each session of the conference will focus on a particular direction of sociological theory, and will ordinarily consist of a paper by a Polish sociologist and one by an American sociologist, followed by comments from discussants and by general discussion.
>
> It is intended that the papers not be surveys of the field, nor that they be metatheoretical, about social theory. Rather, they should be original scholarly contributions. A given paper may exemplify a qualitative or formal approach, but it is hoped that in some papers there will be some combination of these two approaches. Whether a qualitative or formal approach is taken, it would be desirable to indicate how work from the other approach could supplement or complement the approach taken. It would also be desirable to specify which elements of the contribution could be used for understanding phenomena usually treated by the other approach. This might help show how the two approaches could be most usefully related, and would indicate how some bridges between these two approaches might be developed.
>
> The directions of theory listed in the sessions are intended to be reasonably comprehensive. It may be objected that they exhibit no single means of classifying theory, and that they have some degree of arbitrariness. To this we would, with some chagrin, assent. Our chagrin is not brought about by our own inability to lay out a coherent theoretical framework for the discipline, but by the disarray in which the discipline finds itself theoretically. For this reason, we have been eclectic, using some categories that reflect a substantive delimitation, others which reflect a general orientation to social theory. Thus the jarring quality of the collection of categories below is a measure of the extent of progress that one way or another, at one time or another, can be made.

In addition to the American participants, the conference schedule included the following intended Polish participants: Stefan Nowak, University of Warsaw (opening address); Antonina Kłoskowska, University of Warsaw (sociology of knowledge); Henryk Banaszak, University of Warsaw (network theory); Włodzimierz Wesołowski, University of Warsaw (structural theory); Klemens Szaniawski, University of Warsaw (purposive action theory); Janusz Ziołkowski, University of Poznan (ecological theory); Miłowit Kuninski, University of Krakow (interpretive sociology); Witold Morawski, University of Warsaw (organization theory); Aleksandra Jasinska-Kania, University of Warsaw (theory of social change); Andrzej Piotrowski, University of Łodz (sociolinguistics); Jacek

Szmatka, University of Krakow (social psychology); Piotr Sztompka, University of Krakow (social movements). Additional Polish discussants were Władyslaw Markiewicz, Edmund Mokrzycki, Jerzy Szacki, and Jerzy Wiatr.

But there was not to be a Polish-American Conference on Social Theory at the University of Chicago on November 9–12, 1983. Through no fault of their own, the Polish participants were not able to take part in the conference as planned, although they were able later on to present and to discuss their papers in Spring 1984 at the conference held in Grzegorzewice near Warsaw and organized by the Chair of Methodology of Sociological Investigations of the Institute of Sociology of Warsaw University. We hope the time is not far distant that a Polish-American Conference on Social Theory will be held, either in Poland or in the United States.

On November 9–12, 1983, the W. I. Thomas and Florian Znaniecki Memorial Conference on Social Theory was held at the University of Chicago. It was held under the auspices of the Russell Sage Foundation, not of IREX. It is the substance of this conference that fills the pages of this book.

The atmosphere of the conference was one of some intensity, perhaps due in part to the moment in the history of the discipline at which it took place. A proximate partial cause of the intensity was the number and interest of the *non*participants, who at every session filled the chairs surrounding the participants, sometimes crowding the library of Ida Noyes Hall to overflowing. This interest in turn arose in part from the identities of the participants, outstanding social theorists on the frontiers of the discipline. It arose in part from the evident seriousness of purpose and involvement of the participants themselves.

But then why were the participants, inured to the ways of conferences, so serious and involved? There were two reasons, we believe. First, they found themselves among their peers, all competing for center stage, for the attention of the discipline, and all ready to discover flaws in others' work. In such a gathering, one could hardly afford to do less than one's best. Second, there is a vacuum in sociological theory at present. This is evident by the extent and diversity of contenders for charting new directions for the discipline in theory. Some of those contenders and some of those directions may be found in this book. And the prize is not a trivial one: It is the capturing of a stream of energy and intellect of which the discipline of sociology consists, a stream now meandering but capable of great force when channeled. In what directions will the channels go? Virtually every part of this conference touches on this question, the panels as well as the papers and the discussions surrounding each paper.

One question remains. Why does such dialogue not take place at the meetings of the American Sociological Association or at another of the many conferences on sociological topics that take place each year?

We believe the answer has to do with fragmentation, with core and periphery, and with a misplaced egalitarianism. The annual meetings of the ASA (and of other sociological associations as well) are fragmented both by the many special topics within the field—traditional specializations as well as the new ones which arose out of the 1960s—and by the very format of the meetings. Most sessions at the annual meeting, and most funding of sociological work, are directed to these specialties, many on the periphery of sociology, and all, taken together, constituting a fragmentation of the discipline. The format of the meetings is also shaped by practical interests, which result in an egalitarian pattern in which each attendee presents a paper to which a scattering of like-minded others listens.

At the annual meetings of the ASA there are, to be sure, sessions on the substance of social theory which claim to address core issues in sociology. But seldom is there an occasion at which different approaches to social theory confront one another and defend themselves in the context of a whole range of approaches. Seldom is there confrontation at the core, a focusing on the central substance of sociology from a number of different approaches, each of which must defend itself in the presence of others.

We would not claim that this conference has brought together all approaches to sociological theory in fair and just representation. The conference naturally reflects not only the state of theoretical work in the field, but also the theoretical inclinations of its organizers. In particular, traces of the originally planned qualitative-quantitative focus of the Polish-American conference remain. In addition, there is a stronger rational-action representation than if the conference had been organized by others.

Yet despite the fact that not all the issues at the core of sociology were joined, and not all the contenders for control of that core were represented, the contrast is great between this confrontation—in the conference and in the book—containing a wide variety of approaches to social theory, all directed to core issues in the discipline, and most academic communications among sociologists, which are narrow, or on the periphery, or both. This conference and this book provide one of those rare occasions at which issues central to the discipline are joined.

The development of particular approaches to sociological theory resembles the evolutionary process on an archipelago of isolated islands. Each such theoretical orientation starts from its own system of assumptions about the nature of the social world or the phenomena in question, each shapes itself according to rules which would make it a typical Kuhn "paradigm," each develops cumulatively, guiding empirical research prescribed by the given paradigm.

The dividing lines between the various theoretical orientations usually constitute communication barriers between theorists on these different theoretical islands. In some more dramatic periods in the history of our discipline these communication gaps become the front lines of a war between the various theoretical paradigms. Both the occurrences of open conflicts and the usual mutual lack of interest are indications of the disarray in sociological theory. But the fact of this conference, and this book, indicates the vitality that can result from communication among the islands of the archipelago.

Although the Thomas-Znaniecki conference arose from a confluence of circumstances, the result was as if it had been designed with the above considerations in mind. We believe the conference marks a beginning toward sociology's rediscovery of a core, and we trust that there will be other such occasions, among sociologists from one nation, as this one was, bilateral, as the original conference was intended to be, and multilateral, involving sociologists from a number of countries.

The Substance of the Conference

If the claim is made that core issues in sociological theory can be illuminated in a conference like this one, it is reasonable to ask just what are those issues. The panels, the papers, the comments, and the general discussion following each paper give some answers to this question. In addition, it is useful in this introduction to focus attention on a few of the issues that arise from a comparison among papers.

We may push aside the underbrush of differences in method and theoretical assumptions and ask just what are the phenomena investigated by the papers. (We exclude here the panels, which were intended to be, and were, about social theory rather than social phenomena.) If we ask that question, then the papers appear to be concerned with behavior of three levels of social system.

Four of the papers focus their attention on the level of the society as a whole. Mancur Olson uses a particular principle of action to account for periods of economic growth and decline of whole societies, as well as the rise of caste systems of social stratification in certain societies. Olson's principle of action has to do with the condition under which potential groups will actually become organized, and the rate at which they will do so. Theda Skocpol and Ann Orloff attempt to fashion a theory of policy formation in nation-states which accounts for the earlier adoption of social-welfare policies in Britain than in the United States in the first decades of this century. In their theory, organized interest groups which are central to Olson's theory play no explicit role, though they lie in effect behind one (but only one) of Skocpol and Orloff's societal-level explanatory variables. Edward Laumann and David Knoke are also concerned with the formation of public policy at the nation-state level, in their case health-related policies, but over a much shorter period of time (1976–80), and confined to the United States. Their principal questions have to do with the structure formed by the groups interested in influencing public policy together with the particular policies that link them together. Gary Becker and Nigel Tomes are concerned with a societal-level question—the distribution of income and how it is affected by constraints on intergenerational transfers. Becker and Tomes's paper deals with the same phenomena of intergenerational mobility as do studies of status attainment, but differs in two ways: First, it is concerned with societal-level consequences rather than those at the individual level; and, second, its derivations flow from an assumed principle of maximization on the part of parents, in the investments they make in their children.

Another four papers are at the level of organizations or groups within whole societies. Joseph Ben-David examines how the institutional structure of nineteenth-century German universities facilitated the growth of a new discipline in the universities: experimental physiology. He examines not only the growth of a social organization but also the growth of a system of ideas, ideas of which this scientific discipline was composed. The goal of Anthony Oberschall's paper is similar to one of Ben-David's goals, to account for the emergence of a particular normative system. In Oberschall's case, the normative system is the set of local rules established in gold mining camps during the California Gold Rush. And there is compatibility between the explanatory mechanism that Ben-David uses—market competition among universities—and Oberschall's explanation, that those rules emerged which minimized transaction costs (i.e., the burden of social organization) among miners.

Robert Eccles and Harrison White's paper might be seen as focused on a societal-level system, for they see the economy as a matrix with multidivisional firms as columns, product areas or markets as rows, and the "profit center" within the firm as a cell in the matrix. But their attention is not on the matrix, i.e., the system, but on the firm, examining how the chief executive officers use pseudo-markets within the organization as a management tool. This conceptual structure contrasts sharply with that of Michael Hannan and John Freeman, who see birth and death of organizations rather than purposive action on the part of managers as the principal mechanism through which organizations evolve. Again, because this mechanism has implications for the distribution of types of organizations in the organizational population, Hannan and Freeman might well be concerned with problems at the level of the population distributions, but their concern is rather with specifying the mechanisms of selective retention which operate at the level of organizations.

Finally, three papers are at the level of individual behavior, or a system of interaction between two, or at most a few, individuals. In William Labov's work on sociolinguistics, the individual behavior or the interaction medium is speech behavior. Labov examines how this differs among social groups, in different social contexts and in different parts of a social network. David Heise's work concerns the language people use, but the orientation is entirely different from that of Labov. Following symbolic interactionist ideas, Heise treats words as carriers of affect intended by the speaker, while Labov uses them as tracers of some aspect of social organization. Both Heise and Labov treat language as a cultural product, but then their paths diverge. For Labov and other sociolinguists, the task is like that of an archaeologist, using language as the artifacts to discover the social and cultural content that lies hidden there. For Heise and other affect control theorists, the aim is, taking the shared meanings as given, to uncover the affective dynamics of interpersonal interaction.

Joseph Gusfield's work stems in part from the same symbolic-interactionist tradition on which Heise builds, though the methods they employ could hardly be more different. Both Heise and Gusfield are concerned with the meanings which persons vest in the language they use, but Gusfield is more concerned with the person-specific and situation-specific subjective meaning. Finally, the first part of the Laumann-Knoke paper, which discusses propositions to be found in network theory, is at the pairwise-interaction level, along with these three papers. It is, however, based on a behaviorist rather than a symbolic-interactionist orientation to action.

Taken together, the Heise paper, the Gusfield paper, and the first part of the Laumann-Knoke paper show merely the tip of the iceberg of theoretical work in social psychology that deals with microsocial relations.

What is most striking about these concentrations of attention at three different levels of social organization is the vastly different methods and theoretical assumptions. The preceding paragraphs point to some of these differences among the Labov, Heise, and Gusfield papers at the individual or pair-relation level. It is evident as well among the other papers. At the societal level, Laumann and Knoke use interview data and quantitative methods to study policy outcomes not greatly different from those for which Skocpol and Orloff use historical data and cross-national comparisons. Because Laumann and Knoke look within a society, rather than comparatively across societies, they cannot examine the factors that are central to Skocpol and Orloff's analysis, i.e., differences in the functioning of government. Their methods instead are suited to studying the detailed processes through which new policies come about within a given governmental and societal context. Olson's methods are similar to those of Skocpol and Orloff's: He uses cross-national comparisons to study similar societal-level outcomes. But he takes as problematic the very formation of interest and pressure groups which Skocpol and Orloff assume as given. They, in turn, take as problematic the structure of government, which Olson takes as given. It appears that each presents part of a theory, that part which is efficacious in accounting for the particular cross-national comparisons each is interested in explaining. Most different in theory and method from others at this level is the Becker-Tomes paper, which uses a formal model to show how assumptions of utility-maximizing behavior lead to different income distributions under different structural constraints.

Another contrast is provided by comparing the Hannan-Freeman paper, the Olson paper, the Eccles-White paper, the Ben-David paper, and the Oberschall paper in their assumptions about how organizations change. Hannan and Freeman assume that change occurs only by birth and death of organizations: Those that are unsuccessful are replaced over time, through a Darwinian process, by those with a higher survival value. Eccles and White, by contrast, suggest that changes in organizational forms occur through structural innovations made by chief executive officers in existing organizations (the multidivisional form, as initiated by General Motors) and copied by CEOs in other existing organizations, or in some cases through innovations initiated by management consultants (e.g., "investment portfolios" consisting of profit centers within a corporation). Clearly these theoretical positions are in direct conflict. Olson's mechanism of change is more compatible with that of Hannan and Freeman, yet different: While they examine factors making for differential ***death rates*** of organizations, he takes the ***birth*** process as problematic and focuses on one factor making for differential likelihood of birth. He sees organizations as dying only through some cataclysmic event, such as war.

Ben-David locates the engine of organizational change (where the organization in his case is the university, which in Germany changed radically during the nineteenth century) in the environment, showing how competition among universities facilitated the growth of new departments in the experimental sciences. By Ben-David's principle, organizations in highly competitive environments will change more rapidly than will those in less competitive environments. By Hannan and Freeman's principle the rate of change would also be more rapid in such environments, but by a quickening of the birth and death rates, not through internal change.

In Oberschall's paper, there is no explicit mechanism for change in the normative structure of mining camps: As with most theories involving a principle of minimization, there is no statement of the means by which the minimum is achieved. The principle is compatible with either the natural selection process of Hannan and Freeman or purposive design of the sort exemplified in the Eccles and White paper.

These are some of the comparisons and contrasts between papers in the conference and the volume. Do they give any sense of a central core to sociology? We believe they do, despite the very different "presentations of self" in the style and methods of the papers. First, and perhaps most important, the papers address real problems in the functioning of society. They are not merely discourses on social theory in general, but are attempts to contribute to theory, that is, to an understanding of the functioning of social systems.

Second, there is not simply a disparate scatter of problems. There are a few central problems addressed by these papers—in different ways and with different theoretical approaches. Some of the theoretical principles are in direct conflict; others are in no necessary conflict, but are not pushed far enough to exhibit compatibility or incompatibility.

The fact that there are a few central problems examined in different ways by these papers shows the potential vitality of the core of the discipline of sociology. The conflicts and incompatibilities among the theoretical explanations of these problems show merely that there is no dominant theoretical paradigm or framework in the discipline. It indicates that there is vitality in the various directions of search for appropriate theoretical underpinnings to the discipline. This is vitality that requires nurturance, in the form of continued attention to core problems of the discipline and clarity about just what theoretical assumptions lie behind particular explanations. It is vitality that depends both upon the existence of sociologists well prepared for work in the development and testing of sociological theory and upon their devoting time and attention to core problems of the discipline. The papers in this volume provide evidence that some sociologists are giving such attention; but they also provide evidence of the distance the discipline has to go to develop a strong theoretical base.

Siegwart Lindenberg
James S. Coleman
Stefan Nowak
Chicago, Illinois, January 1986

On W. I. Thomas and Florian Znaniecki

W. I. Thomas

William Issac Thomas (1863–1947), sociologist and social psychologist, was born in Virginia. Little is known about his early years, but he entered the University of Tennessee at the age of 17 and graduated in 1884. He remained at the university for the next four years as an instructor in modern and classical languages, thereby acquiring a linguistic facility that was to prove invaluable in his later work. After marrying Harriet Park in 1888, he spent a year studying in Germany and then joined the faculty of Oberlin College, where he taught English. In 1893, while on leave from Oberlin, he began graduate work in sociology at the University of Chicago and received his doctorate in 1896.

Between 1896 and 1910, when he attained his professorship at the University of Chicago, Thomas published a series of papers on the social psychology of sex (1907) and prepared the *Source Book for Social Origins* (1909).

Between 1908 and 1918 Thomas traveled extensively in Europe, with the support of the Helen Culver Fund for Race Psychology, and began to collect materials on Polish society and the migration of the peasants to America. In 1914 he began his collaboration with Florian Znaniecki.

Florian Znaniecki

Florian Witold Znaniecki (1882–1958) was born in Swiatniki, Poland. He did his undergraduate work at the University of Warsaw but was expelled shortly before receiving his degree—for leading a demonstration against the Russian administration. He pursued his graduate studies at the universities of Geneva, Zurich, Paris, and Cracow, receiving his doctorate from Cracow in 1909. He began his intellectual career as a poet and turned next to philosophy, a field in which he won an early distinction not only by original contributions but by translating Bergson's *Creative Evolution* into Polish. Finally, under the influence of W. I. Thomas, he turned away from philosophy and became a sociologist. Philosophy seemed to him to have become a discipline doomed to sterility, whereas sociology opened up new vistas for the future advancement of knowledge.

For political reasons Znaniecki was ineligible for an academic post in Poland, and therefore, after taking his doctorate, he found employment in an emigration bureau. It was here that Thomas, on one of his frequent trips to Europe, met Znaniecki.

Opening of the Conference

JAMES S. COLEMAN

I am pleased to have the occasion to welcome you to the W. I. Thomas and Florian Znaniecki Conference on Contemporary Social Theory. I am sad, however, to have the occasion to do so without a set of fourteen Polish participants who were intended to be here and who themselves intended to be here. Their papers are written, and some have been translated into English. They have, however, along with their papers, been unable to leave Poland to attend the conference.

Nevertheless the conference is being held—not with them, but as a memorial to Polish-American collaboration in sociology, as embodied in the work of W. I. Thomas and Florian Znaniecki. It is a conference with an unusual character, for there are few occasions when sociologists of very different persuasions, from phenomenology to small groups to macrosociology to rational choice, must contend with one another. This, I believe, is one of the two elements which attracted as many people as are here tonight. The other element is, I think, a testament to the people who are on the program tonight and later: It is a sterling cast.

The conference will be opened by a welcoming statement by Robert

McCormick Adams, who is provost of the University of Chicago. I asked him to be on the panel on current issues in social theory because he is a distinguished anthropologist. He said, "I don't have time to do anything like that. I'm provost of the university. All I have time for is to welcome the participants on behalf of the University of Chicago." But then he called me today and said, "You don't mind if I say a few substantive things, do you?" What could I say? We'll call his contribution "prologue."

Prologue

ROBERT McCORMICK ADAMS

Although it may be only an aspect of what Thomas Kuhn has suggested as the developmental path of any normal science, there is a worrisome drift in the social sciences away from an involvement with overarching issues and toward further specialization and progressively more detailed problem-solving. More disturbingly, a matter-of-fact acceptance of a limited and largely one-way relationship between producers of social science and their consumers, supporters, and observers is gradually taking root. Basic disciplinary premises and priorities are assumed to be fairly static rather than subject to active questioning and reshaping. Disciplinary boundaries that are best kept indefinite and in flux tend to become rigid categories of thought. All too rarely an opportunity comes along to override well-established boundaries and incrementalist solutions to problems and to concentrate instead on the explanatory stance of the social sciences as a whole. This is such an opportunity.

The possibility of a dialogue with Polish intellectuals that initially precipitated this conference touched a nerve in all of us that allows us to reach toward one another in an arrestingly different, creative way. Perhaps that is because the renaissance of Polish intellectual life—and not least the social-science component of it—has always been such an unexpected demonstration that outcomes are not determinate, that spirit can prevail over substance. What heartens us all is the seemingly irrepressible growth of a vigorous flower right between the millstones.

My own standpoint as I welcome you is that of a participant-observer with some direct hand on the guidance system of a large organization in a no-growth field, both of which are at risk. Perhaps that heightens my sense of the indeterminacy of things, of the frequency with which unexpected, even perverse outcomes seem to confront us in spite of our best efforts at prior planning and risk assessment. To try to run a university these days is to be driven toward a more and more pessimistic appraisal of the extent of discoverable regularity in human affairs. Those who confidently advocate the use of the social sciences as a guide to practical action need to give some attention to the many such appraisals that are to be heard on all sides. I think they provide some grounds for skepticism as to whether the results of most forms of social intervention are truly predictable. It would help if necessarily complementary avenues of investigation leading toward the discovery of regularities less often became tunnels lacking interconnections between one another.

In this respect Albert Hirschman has made a useful distinction between two different mechanisms that link social initiatives to their ultimately perverse effects. "Desertion," he writes, "is the only mechanism that economic theory recognizes while political science considers protest alone as worthy of attention. What we need instead is an approach transcending this disciplinary boundary permitting us to study the two mechanisms in their complementarity." The ironies involved in such a divergence are rather similar to those that James Duesenberry has noted between sociology and economics: Economics is all about why people make choices while sociology is all about why they don't have choices to make. Herein lies, I hope, the real subject matter of this conference. Can we begin to envision, if not yet necessarily to construct, a world of social theory that corresponds to the indeterminate world of our experience; to a world in which uncertainties preclude optimizing strategies; to a recognition that, as Kenneth Earl puts it, the past is relevant because it contains information which changes the image of the future; to the absence of any unambiguous boundaries in time or space on the interactions affecting our lives, both individually and collectively?

I do not want to be misunderstood as denying the existence of multiple spheres or levels of effectively autonomous thought and behavior and instead favoring the immersion of all social units into a featureless, indefinitely extending, entropic sea of social context. What the world of experience also makes plain is that social change cannot be understood if its agents are seen as mere products of their social structure. Cultural and institutional constraints assuredly often force choices in the last analysis. Similarly, institutions and cybernetic patterns could not even exist unless such constraints had some durability. But to focus on determinate elements of structure is to succumb to what Raymond Boudon has called "the chronic illness ravaging sociology and the other social sciences" unless at the same time we find a substantial place for preference-seeking, risk-, uncertainty- and opportunity-responsive, in broadest terms innovative or creative, courses of human action within the field of social analysis.

What otherwise threatens us, to be blunt, is a consistently oversocialized distortion of our human subject matter. Missing in the announced titles of the conference sessions, but I hope not in the papers and discussions composing them, is a concern for the intersections of the fields of personal and social preference and action in a flow of events that to some degree has to correspond to the impulses of profoundly intentional (as distinguished from necessarily rational or utilitarian) actors. What part of any general theory of causality will be assigned to the two respective sides of this interface? I imagine this will be touched on, although probably from mutually opposed perspectives, both of which are likely to treat the individual only as an abstraction, by the sessions devoted successively to structural theory and purposive action theory. Is the problem once again that the subject lies athwart an inconvenient disciplinary boundary, in this case with psychology?

How we bridge these fields seems, at least from the relatively distant perspective of an anthropologist, one of the most important opportunities we have to develop satisfying, as well as more useful, social theory. Perhaps the place to begin is with the recognition that decision-making is not an act, but a process. It is a process that should be ethnographically as well as analytically decomposable into component steps and influences. But I have already gone on longer than my primarily ceremonial function here can possibly justify. The importance of this conference is made manifest not only by the range of its subject matter but most centrally by the extraordinary group of participants who have assembled for it. You honor the University of Chicago by your presence and I can only say that we very warmly welcome you.

Finally, I would like to say one further word, speaking on behalf of the Russell Sage Foundation, which is sponsoring this conference. The welcome on behalf of the Russell Sage Foundation would normally have come from its president, Marshall Robinson, who had planned to be here but who is unfortunately ill. I am delighted to note that Marshall responded, I think essentially instantaneously, to Jim Coleman's request for funds when the cutoff of the Polish participation made funding from IREX impossible. Those of us who are involved with the Russell Sage Foundation, and I am a trustee of the foundation, like to think that the foundation will always be responsive to the needs and opportunities of basic research in the social sciences. I can only say that we will be if you make it so. From that point of view, on behalf of Marshall, let us hear from you. Thank you.

Current Issues in Social Theory

How Sociological Theory Lost Its Central Issue and What Can Be Done About It

SIEGWART LINDENBERG

Given the short time allotted to this panel, I will start with three concise theses and I will exaggerate to make my point quickly [cf. Lindenberg 1983a for a more detailed discussion].

The first thesis is: *There are virtually no current issues in sociological theory relevant to sociology as an empirical science right now.*

The second thesis is: *Unless something drastic happens, sociology departments in the United States and elsewhere will turn either into departments of social philosophy or into departments of solid social empiricism.* This process is already under way and I think eventually both kinds of departments will first become auxiliary service institutions and then vanish altogether, being absorbed by other disciplines and professional schools.

The third thesis is: *No call for the increased cooperation between theory and empirical research will have any effect.* In other words, the lack of relevance of sociological theory for social research has nothing to do with a lack of coordination between the two.

The Methodological Mistake of the Founding Fathers

How could it ever have come so far with such a promising social science? I think the basic mistake had been made way back when sociology was founded as a separate discipline and when it developed its own models of man. Economists, especially the model building economists, seemed to discount the importance of institutions and public opinion and to operate on a maxim that human action is determined by selfishness in the context of price and income constraints. By contrast, the founding maxim for sociology was that human action is socially determined by institutions and by public opinion, and the founding fathers believed that the heart of the trouble with economics was its model of man. This model did not express the social embeddedness of human action. As a result, sociologists developed their own *homo sociologicus,* in two different versions. Stressing the institutional side, there is what I have called SRSM—that is, Socialized, Role-playing, Sanctioned Man. Institutions are systems of role expectations, which in turn are internalized during a process of socialization. Imperfect socialization and situational strains are corrected by the fact that it is part of everybody's package of role expectations to sanction others who do not behave the way they should.

Stressing the public opinion side, there is OSAM—that is, Opinionated, Sensitive, Acting Man. People have opinions on everything and will act according to their opinions, but their opinions are sensitive to what others believe and do (social influence).

While both models of man highlight the social embeddedness of human action, neither of them is a good theory of action. Hence sociology is driven to develop in these two directions: empiricism that describes social phenomena and searches for sociological invariances in the data, rightfully ignoring "sociological theory," and social philosophical reflections that interpret the social world and expound and refine the insight that human action is socially embedded. In both areas, the competition from other disciplines and from specialized institutes and agencies is becoming stronger.

The charge of the founding fathers that economics (at that time) ignored the importance of institutions, norms, socialization, and social influence was neither irrelevant nor wrong. Their mistake was methodological, namely, the idea that the ills of economics were due to its "individualistic" model of man and that this model should be replaced by SRSM and OSAM. Another way of putting this is the following: The founding fathers and their followers failed to distinguish two kinds of primacies, the *analytical* primacy of society and the *theoretical* (or explanatory) primacy of the individual. By insisting that therefore the analytical primacy *and* the theoretical primacy lie with society, we are left without explanatory apparatus.

The central issue of sociological theory was the social embeddedness of human action, but as a consequence of the failure to distinguish analytical and theoretical primacies, many nonissues have come to prominence, such as "individual versus social," "micro versus macro sociology," "historical versus statistical analysis," "structural versus individualistic approaches." These are nonissues because the alleged opposites are artificially produced by this failure to acknowledge that in order to explain, we need, *among other things*, an adequate theory of action.

The Temporary Success of Sociology

If the sociological models of man are really as bad as I make them out to be, why was sociology so successful a social science for such a long time? I think there are three main reasons for this temporary success.

First, Adam Smith (in the context of the work of other Scottish moral philosophers) had established a broad social science that included prominently attention to institutions and social inequality and also focused on the formation of norms, conscience, and role-taking, integrated with the theory of rational choice. After Adam Smith, economics had become a very specialized science without much regard for institutions, social inequality, and social action. Against this neglect, sociology had a real and important message: the social embeddedness of human action.

Second, the *homo sociologicus* (in both forms) was a powerful tool for bringing the message across. But not just economics needed to heed the sociological insight. Other social sciences and folk theories were equally in need of being enlightened on this point. The success of sociology came from being an empirical "debunking" science with a central but neglected message. For example, it is believed that contracts are voluntary, mutual agreements between two unrelated partners. But let me show you that without extracontractual, institutionalized understandings, this agreement could never come about. It is believed that crime is determined by race or by skull size. Let me show you that it is socially determined. It did not really matter that sociologists did not come up with theories about how contracts are made, how they change and why, how categories of crime rates change and why, etc. Myth after myth was debunked and every time the result was the same: The sociological message proved more convincing than the previously held view.

Third, economists and other social scientists did not take to the streets to get people's opinions and to describe the ways they live. Sociologists, taking the anthropologists' example, did do that and amassed a wealth of information. In short, there were three good reasons why sociology should have been important and successful.

However, in the meantime things have changed. First, the sociological "perspective" has become part of the Western culture and thus also part of other social sciences. The debunking function is losing its target.

Second, institutional economics is moving into traditionally sociological areas, such as organizations, family, political institutions, social movements, education, crime. Some economists (e.g., the contractarians) are also moving into social philosophy, confronting sociological theorists who seek their refuge there.

This does not mean that all is well with economics. A hundred years of neglect of institutions is not overcome within a few years. But neoclassical economists did not systematically dismantle their own tools of explanation. They never believed that one can explain anything without a theory of action and they are not hampered by SRSM and OSAM. This gives them a competitive advantage even if they have not yet mastered network analysis and some other useful sociological tools.

Third, journalists and polling institutes have learned the sociological message and are taking to the streets to find out how people live and what their opinions are. Or course, sociologists still perform this task as well, but they meet with increasingly more competition, without the ability to lay a professional claim to this activity.

In sum, for almost a hundred years, potential competitors to sociology lost out because sociologists had the superior message. It is a triumph for sociologists that this message has virtually been accepted by just about everybody. In the past, each piece of empirical sociological research was a vindication of the sociological message (i.e., theory). That vindication linked theory and research and made one relevant for the other. Now that the message is taken for granted, there is nothing left to vindicate and "theory" has become irrelevant for empirical sociological research. With it, sociology has lost its central issue. Not surprisingly, many theorists seek refuge in social philosophy where relevance to empirical research is not at stake. The increasing attention paid to Habermas as a *sociological* theorist is a clear sign of this development.

What Can Be Done?

What we should do as sociologists, I submit, is to drop the traditional program that brought us SRSM and OSAM and the inability to develop theories that are lastingly relevant for empirical research.

Let us look at some minimal requirements of adequacy for a theory of action:

1. It must not require much information about each individual to which it is applied.
2. It must allow us to model institutional and social-structural conditions as defining goals and constraints of action (i.e., it must allow the analytical primacy of society).
3. It must allow psychological (including physiological) theories to influence its assumptions. For example, the information processing abilities of individuals must not be fixed by axiom.
4. It must allow us to introduce simplifying assumptions in such a way that they can be replaced with more complex assumptions as our knowledge increases (method of decreasing abstraction).
5. It must be well corroborated as a theory that explains behavior of human beings in the aggregate, inclusive of resourceful behavior.

The only theory to date that meets (or can be made to meet) these requirements is the theory of rational choice in various forms of elaboration. In the second half of the nineteenth century, economists used the capabilities described under 1 and 4 excessively, but it turns out that with relatively minor adjustments, this model of man could be made to meet all five requirements. By contrast, even the best psychological theories violate requirements 1, 2, and 4, and SRSM and OSAM violate 4 and 5 while allowing only a very limited selection of institutional and social-structural constraints and psychological correctives (partial violation of 2 and 3).

Once SRSM and OSAM are dropped, many issues that have been defined away by these models of man will come to the fore, while many prominent nonissues will vanish by themselves. I believe that the central issues in social theory will then be elaborations of the following: (a) the relation between utility and morality and (b) the relation between choice and structural imposition. These are only hints at what is involved [see Lindenberg 1982, 1983b, 1984], but traditionally these two classes of issues have been suppressed by setting SRSM and OSAM (morality and structural imposition) against *homo economicus* (utility and choice). By the same token it is clear that I do not suggest that theoretical sociologists become economists, for the suggested classes of issues are also not alive and well in economics.

Methodologically, new issues will evolve as well. Most prominently, I see the following problems emerge: (a) the problem of correspondence and (b) the problem of transformation. These problems arise from the *juxtaposition* of the theoretical primacy of the individual and the analytical primacy of society: (a) How can social conditions be modeled to provide both intermediate goals (i.e., production functions) and constraints for social action (the problem of correspondence), and (b) what is needed to explain the transformation of aggregate social action into collective social phenomena (problem of transformation)?

The Analysis of Diversity and the Diversity of Analysis

MORRIS JANOWITZ

For me, the house of sociology has many rooms. An interest in social institutions guides my work. I believe that the "institutional approach" to political sociology is close to the real world and at the same time supplies a basis for theoretical analysis of institution building in politics. Institutional analysis is especially useful in probing nation-states with democratic political institutions which are now experiencing great internal strain.

My approach to institutional analysis is based on a group of irreducible units and objects of analysis such as primary groups, communities, bureaucratic structures, and nation states. The units and objects are the ones postulated by W. I. Thomas, and we have not been able to improve extensively on his pattern of analysis. Of course, technology and economic and normative variables represent the substantive dimensions.

I shall not present to the participants of this conference my views of the basis on which a realistic institutional political sociology has been built and can be further elaborated. My views on these matters are contained in *The Last Half Century.*

The "new political economy" is the alternative approach which has generated considerable interest. Lindenberg's position is clearly an example of this alternative. He elaborated this position in more detail in a recent paper called "An Assessment of the New Political Economy: Its Potential for the Social Sciences and for Sociology in Particular" (1983a). This orientation connotes systematic and quantitative based economic rationality applied to a wide range of sociopolitical behavior. Lindenberg asserts that the redevelopment of such an orientation is the "best thing" that has happened to social science. No doubt the new political economy is an important intellectual achievement, but it could turn out that it is as much, or more than, a political movement as an intellectual movement. Likewise, in my view, the new political economy movement already often proceeds as if sociology was a narrow intellectual exercise or a dispensable subfield within contemporary economics.

Of course, it is unproductive to spend time discussing the merits of different boundaries of the various disciplines. Nevertheless, my basic orientation is that I am not at ease with this type of analysis. No doubt it is more accurate to assert that I do not fully understand the type of theoretical analysis it presents nor am I adequately informed of its substantive conclusions. I recognize the complexity in this type of theoretical and empirical effort. Nevertheless, the economic rationality approach of political economy just does not seem to me to be a realistic or broad enough approach for the sociological data we have available from quantitative studies—for example, voting behavior research.

However, one section of the Lindenberg paper is most attractive to me. Lindenberg states, correctly in my judgment, that there is an American type of sociological analysis. He stresses the difference between American and European sociological analysis. He states that "between 1920 and 1960, the face of modern sociology was developed in the United States. The context of this development was decidedly American." I believe that the format of American sociology was already launched by 1910. W. I. Thomas, the key figure in this enterprise, was by that time deeply involved in elaborating his integration of social theory and data. Lindenberg performs an important service in his conclusions; he seeks to elaborate comparative analysis. He goes on to assert that "while the classical European sociologists had been preoccupied with the causes and consequences of industrial and political revolutions, neoclassical American sociologists were drawn into the peculiarities of American nation-building." In short, Lindenberg joins the army of European intellectuals who found that the United States is peculiar.

I assume that these differences are important and not minor. I would be pleased if Lindenberg elaborated this crucial formulation. As so often happens, sociologists disagree on theoretical issues, but agree on substantive ones. Such an observation does not speak well of sociological theories. But in this case, I believe he is on the right track, that is, on a track we can both follow to some degree.

But the theoretical issues involved are important. For example, what does the difference between European and American sociology which Lindenberg asserts do to his efforts to create a more general sociology? I would assert that the difference he postulates limits his opportunity to create a more general sociology and this is, in my view, a good thing.

Lindenberg stresses the role and importance of social anthropology in dealing with the issues of immigration and ethnic groups. Thomas represents an early effort to link ethnic forms to the analysis of the nation-state. It is most fitting that James Coleman designated the name of W. I. Thomas in the headline of this conference. This is not to assert that the problems of incorporating social anthropology into macroanalysis have been solved. To the contrary, social anthropology is not easily articulated with macroanalysis of institutions and nation-states. Careful formulations can be and have been used to good ends in linking social anthropology with sociological analysis. The social anthropologists must be certain that they are aware of the impact of the state on the microunits of research and analysis.

In conclusion, macrosociology requires not only enriched monographic research, but carefully designed statistical trend analysis. These various observations I hope will prevent us at this conference from losing sight of the national differences Lindenberg has asserted—namely, the differences between European and American sociology. His formulations have the potential, in my opinion, to underline the levels of diversity which the macrosociologist of the Western industrialized nation-state needs to confront and to explain or, at least, to clarify.

The Hidden Issues: A New Agenda

IMMANUEL WALLERSTEIN

When James Coleman asked me to be on this panel, a panel on current issues in social theory, I asked him what he thought "issues" and "social theory" meant as a topic. He said he would be happy if we would discuss what ought to be the dominant themes and issues in sociology. "Current," as we use the term normally in the social sciences, is a term that refers to a decade or maybe 20 years, but not much more than that. I am more interested in a period of about 150 years, one running roughly from 1800 or 1850 somewhere up to about 1970 and another period, partially overlapping the previous one, starting about 1950 and running to about 2100. I would like to look at the first period not in terms of what was argued about but what wasn't. My topic for this panel is the issues that weren't issues.

The Great Consensus

It seems to me that a consensus emerged somewhere after the French Revolution. We all know that the reaction to the French Revolution in the Western intellectual world was to create two major streams of thought immediately: liberalism and conservatism. Liberalism was there before but it got a self-consciousness as the newly dominant theme. Something came along called conservatism, asserting itself as the other side of the coin. The other great division of the nineteenth century, as we all know, is, of course, Marxism and liberalism. I would like to underline not the differences, but what I see as the great *consensus* between Marxists and liberals in the nineteenth century (with those gloomy conservative inversions coming along as a sort of counterpoint but not really an opposition). I see five great points of consensus underneath all the debates and all the arguments.

The first point of consensus is that history is the story of human progress (and that, in turn, is the story of a better technology) and more human freedom. On that both Marxists and liberals were thoroughly in accord. Conservatives were not really in disagreement; they would merely call the historical description "regression" rather than "progress."

The second point of consensus that emerged from the French Revolution or the period following it is that the critical antinomy with which to analyze social phenomena is the antinomy of state and society, both terms that in some sense came into reality in the nineteenth century. They were implicitly terms that described roughly the same geographic boundaries. The great issue then became which of these two entities did in fact prevail over the other; which should prevail over the other; which one was more fundamental than the other. Views varied enormously on this, but the policy implications were similar: We should bring the state into relationship with society or society into relationship with the state.

The third point of consensus that I see between liberalism and Marxism is what I would call scientism, built on a long tradition already existent in Western thought. It said that, analogous to the physical world, the world of social reality can be described in the form of lawlike statements, and we can arrive at these statements through various kinds of empirical research from which we can abstract the truths, testing ever further our abstractions and getting to purer and purer abstractions. The more elegant, the more concise, the more pure our statements of these lawlike phenomena the better off we were. Above all we should eliminate time and history from our analyses.

The fourth point of consensus (I may stretch the degree of consensus, but not too much) is that the modern world was in fact the story of the rise of two new social categories: the bourgeoisie and the proletariat, otherwise known as the middle class and the working class. It was argued that these two social categories emerged in the modern world, filled it out, and that whatever older categories still existed were in some sense anomalies, survivals, or regressions. Given the inevitability of progress, the survival of older categories would be temporary.

The fifth point of consensus is Europocentrism. The modern world was basically an achievement of Western thought and Western man which then spread to the rest of the world. Further progress of the world was the progress of this diffusion of the wisdom that had by one means or another been achieved in the West. Of course one of the major intellectual questions then becomes: How was it that the West was able to make this great breakthrough?

The Great Change

Modern social science was the result of this consensus between liberalism and Marxism. I would say that modern social science is about 50 percent liberal and 50 percent Marxist, located in the interstices of these two major views of the world and building exactly upon those elements about which they agreed. The consensus prevailed roughly until after World War II and informed everything we did. And with regard to what we did, I agree with Lindenberg that we should cut our losses.

The consensus that informed everything we did began to be called into question as a result of the changing world: World War II, Stalinism, the rise of the Third World. It seems to me that we are living in the midst of a sea change in that 150- or 200-year-long consensus. All this discussion about the multiplicity of modes of sociology, social sciences in general, simply reflects both the discomfort and the defensiveness about the old consensus. I think that if all proportions are kept in mind, as the French would say, 1968 was as dramatic for the old consensus as 1789 was in many ways for the consensus of previous centuries.

What the sea change involves is questioning the five points of consensus. *The inevitability of progress:* Through the whole nineteenth century and early twentieth century, no one seemed to consider that progress might be possible but not inevitable. As a major point of view, that position simply was not there. It throws a whole different light on how we read history and it sees "transitions" or major social choices as great historical choices, which are not inevitable and in which the world can go in fundamentally different directions.

State and society: Both concepts are the inventions of the modern world, of what I would call the capitalist world-economy. Neither state nor society as we described them existed before that time and probably will not continue to exist in the future. It is therefore questionable to organize the categories of our analysis in terms of either state or society and to focus on our national paths (although Janowitz just reiterated his desire that in intellectual terms we reinforce that).

Scientism: The nineteenth century saw a great debate between those who were scientists and those who denounced science, who said in effect that within the social sciences, science is not possible, that we can only describe the particular. What was never offered as an alternative is that interpretation might be science, that the root of science is not the scraping away the details of the particular in order to move toward more and more pure, elegant, and concise statements of laws, but exactly the opposite. We start with our elegant abstractions and add elegant abstraction to elegant abstraction as we move closer and closer to an interpretation of a historical reality which is singular. The object of scientific activity is to end there. And why would we want to end in the interpretation of the singular? To act upon it morally.

The bourgeoisie and the proletariat: Looking again at our conception that the essential new characters of the modern world are the middle class and the working class, it will turn out that there is more social continuity than we thought from that premodern, feudal world to the modern capitalist world. The big break in social categories and social actors is still to come.

Finally, *Europocentrism* has clearly been challenged in the last few years. But let me point to a peculiarity of the history of the world. Up until the existence of the modern world we had a large number of multiple, rather autonomous cultural worlds that lived side by side in a complicated way. One of the consequences of the capitalist world-economy and its life has been that it became the first social system to cover the entire globe. We therefore have a very fundamental question about the relationship of the peculiar entities called civilizations to the fact that we exist within a singular world-system and that what may replace it (if something does replace it) may well be equally a singular world-system.

As we debate these new issues, which are new issues only in the sense they were not debated for 150 years, we may arrive at some new consensus over the next 30 or 40 years which will hide other issues. I don't believe in the end of history. But I do believe that history equals theory, that it is not a separate activity.

General Discussion

Jack Goldstone: Lindenberg reminds us again of the demise of sociological theory and predicts its disappearance. Failure of classical mechanics to explain the atom didn't mean the end of physics. It seems that as long as phenomena exist that require explanation, sociology is not going to disappear regardless of how many consensuses break down, regardless of how many theories turn out to be failures. I'd like to know why you think that inequality, social change, and revolution will cease to be problems that will draw sociological analysis.

Siegwart Lindenberg: I never said that the phenomena will go away but that unless the traditional sociological program is changed drastically, sociology will be absorbed by other social sciences who cover those very phenomena, notably historiography and economics. There are parts of economics which I think will not survive for the same reasons—namely, macroeconomics—but that's a different point.

Immanuel Wallerstein: I would agree that sociology as such is going to disappear. But I also think that anthropology as such, history as such, and economics as such are going to disappear. I foresee the likelihood of a melting of all of these into something new, after which there is a redivision of the pie along considerably new lines over the next 30 or 40 years. In the same way botany and zoology, which were university departments everywhere, have been merged into biology, and biology has resubdivided in many new ways.

Edward Shils: But the Aristotelian groupings are still basically around. There is some change in content but there is much more stability here.

Randall Collins: Lindenberg was implicitly invoking the fact that so many different kinds of conditions and specialties in sociology bring chaos, leading to collapse. At the opposite pole, Wallerstein invoked a field which really has proceeded by induction from a very wide kind of historical research, and he says rather confidently we can actually use this as a frame for all kinds of intellectual issues. Without falling into Wallerstein's camp, I am more optimistic than pessimistic, not so much because of things in the future but by looking at what has happened intellectually in the last 20 years. If we take sociology field by field, we have gone through great development in many specific areas. The whole area of historical comparative sociology has unquestionably gone through its golden age. Although we still do not have a theory of revolution we know a lot more about it and have more sophisticated theoretical positions than we had before. The same is true at the other end of the scale. We know a lot more in empirical detail and in some kind of inductive theory about extremely micro phenomena than we have ever known before. Sociology is much more powerful now than it was 20 years ago, and it was more powerful 20 years ago than it was 30 years before that. Lindenberg has proposed a particular way of trying to bring it all together, and I think it's all to the good to try to make that effort. But in fact there is much reason for optimism.

Siegwart Lindenberg: There has been some progress but there is an irony to this progress. As I see it, there has been a lot of progress in the last 20 years because of the revival of Marxism. What was positive about the revival of Marxism was that Marx, at least in his very concrete studies if not in his purely theoretical work, used a rather good theory of action, a rational choice kind of theory. That slipped into sociology against Durkheim, against Parsons, against the role-playing folklore, through the revival of Marx. That in itself was an enormous improvement. But the irony is that this was never seen as the particular part that made the improvement. This of course has made it very difficult to see what direction to take. Many sociologists still believe they must invoke Marx when all they want to do is to introduce interests rather than value systems and role-playing.

Unidentified Speaker: What about Max Weber?

Siegwart Lindenberg: Like Marx, Weber paid attention to the role of interests. Besides that, he is a gold mine of ideas and everyone digs in his works for nuggets. Yet, Weber will not save sociology from the fate I described. His classification of action types is *not* an action theory and *not* a solution to the vexing problem of the interaction between utility and morality. His theory of bureaucracy cannot compete with Downs, Niskanen, Tullock, or Wintrobe. His sociological-historical contributions to the study of religion, cities, capitalism, politics, and music were all milestone contributions at the time, but between then and now historians have learned enough from him to make these studies honorably obsolete. Besides, for Weber sociology was mainly auxiliary for historiography, so he would not be surprised about this development.

Theda Skocpol: I would like to hear from each of the panelists one problem they think is worth theorizing about. That at least would give me a better sense of what turns us on rather than what locks us in.

Siegwart Lindenberg: For me, the most exciting issue is the one that was lost when sociology and economics developed in such a way that they took opposite bases as given, as I pointed out in my presentation: structural imposition and morality for sociology and choice and utility for economics. Dealing with the interrelation of structural imposition and choice, of morality and utility, is I think vital for theoretical sociology and therefore exciting for me. And for that issue, the Scottish moral philosophers, especially Hume and Smith, are much more important than the classical sociologists.

Morris Janowitz: Not because I like it but out of an obligation to the society, I study civil-military relations. They are the crux of the matter in contemporary society.

Immanuel Wallerstein: My substitute for society as the unit of analysis is a historical system. My major question is how it functions in its historical life, particularly its trend patterns and cyclical patterns. Why do certain kinds of phenomena repeat themselves with temporal cyclicity and why do historical systems develop structural contradictions that are eventually insurmountable—and what happens when they do? For example, the capitalist economy has clear cyclical rhythms of expansion and contraction over time. Why this regularity, what brought the system into existence, what kinds of ways are there for going out of existence, and what alternatives are historically posed?

Peter Blau: Now you say you want to explain regularities and that's exactly what I want to do, too. But you are contradicting what you said in your presentation, namely, that you are not interested in general principles, but in what happened in this particular case and that particular historical situation.

Immanuel Wallerstein: Fair enough. Since Aristotle was invoked against me a moment ago, I must say that I am a great devotee of the Aristotelian *via media.* On the one side are the people who look for universal laws and on the other are those who say there are no structures. I am in between. I do believe there are structural regularities within the framework of specific historical systems. I do not believe there are, in any significant level of importance, statements that cross over all of space and time, as descriptive and meaningful statements on human behavior. But within the framework of reasonably large-scale social entities, which I call historical systems, there are rules that govern how they operate.

Unidentified Speaker: One of the good things I've gotten out of Wallerstein's point, although I have differences with him on another level, is his notion that there are highly arbitrary disciplinary divisions within the social sciences. The kind of problems that Janowitz is concerned with, and those that I am concerned with (levels of trade union organization), could be dealt with in other fields, particularly politics and history.

Arthur Mann: I speak not as a sociologist. I teach American history. Two of the three speakers said that sociological theory had come to the end of the line. I understand why Lindenberg is saying that sociological theory is coming to the end of the line. So I have a question for you, Mr. Wallerstein, that historians will ask. Why, in your view, were people wrong for 150 years?

Immanuel Wallerstein: Because people are always wrong for 150 years. The truths of the time are historically rooted and they reflect both the social realities of the time and the state of knowledge of the time. Truths last as long as the historical reality. Therefore, we will come up with new ways of looking at the world which will equally only last as long as that historical reality lasts.

Arthur Mann: So your theory is from wrong to wrong.

Immanuel Wallerstein: No, from the best attempt that we can make at any given time to interpret the reality in which we live to the best attempt at the next time.

Arthur Mann: I insist that you said: "People have been wrong for the past 150 years because it was inevitable." And you did say that sociology is at the end of the line. Basing theories on demonstrable fact, checking facts so that some people turn out to be more right than others, how do you fit this into your view?

Immanuel Wallerstein: The first point I made was a plea for a concept of progress as a possibility but not an inevitability. Therefore, I would apply that equally well to the world of knowledge. Progress in knowledge is possible but it's not inevitable. We don't necessarily build on the shoulders of giants. But we are now aware of things that were said previously and we have insights that people in previous eras did not have. They will be distorted, and better ones may come along in the future.

Theda Skocpol: How do you know it's progress?

Immanuel Wallerstein: That is a totally different question. At that point we have to get, I'm afraid, out of the context of the variables with which we have traditionally dealt in social science, which means getting into very fundamental value questions. We will know what is progress in terms of the values we bring to the judgment of history. Those values themselves do not come out of the air nor out of natural law. They in turn reflect the social reality of the world in which we live.

The Emergence of Sociology as a Discipline

Three Sociological Traditions: On Creating the Future While Creating the Past

RANDALL COLLINS

Robert Bierstedt, in his lucidly written *American Sociological Theory* (1981), comments that Durkheim, Weber, and Marx were known by American sociologists before 1940 but not especially adulated. They were just three names among many who made up the history of the field, and not among the most important or interesting. We see the same thing in Pitirim Sorokin's *Contemporary Sociological Theories*, published in 1928. Here Durkheim, Marx, and Weber get a few pages, but relatively superficial ones, and far less space than that devoted to Le Play, Pareto, or Otto Ammon.

Today, the reputations of Marx, Durkheim, and Weber are at their height. We have created their importance by what we have been doing in the last few decades. They have become core traditions of sociology, not merely because of what they were originally, but by virtue of what we have been able to make out of them. I say this in opposition to a certain school of thought which maintains that sociologists should treat their theoretical forebears merely as intellectual history, figures to be appreciated in their own contexts and not, as they would say, anachronistically brought up into their relevance for the present.

I think it is a sign of our intellectual growth as a discipline that sociology can now be conceived as the development of certain classic theoretical traditions. Émile Durkheim is now a more important figure than Adolphe Coste (although Sorokin would have had it reversed), and Weber stands above Lapouge or Lester Ward. We have, in short, created our intellectual past because we have found certain ideas that do constitute breakthroughs, upon which we have been able to build. In this sense, we may speak of creating the past as an offshoot of our creating the intellectual future.

I do not wish to imply that all is completely sunny today. For one thing, there are at least three major core traditions in sociology, not necessarily in agreement with each other. We have a long way to go to arrive at consensus on sociological truths, but at least three cumulating traditions is a sign of progress over dozens of noncumulating ones.

There are, I think, two major flaws in this conception. One is that Marx and Weber are actually the same intellectual tradition. Weber, in an important sense, is one of the successors of the Marxian tradition. Most of the developments of sociological Marxism in the last thirty years have consisted of revisions in a Weberian direction. This is even true for the idealist-philosophical wing of Marxism which has been so prominent in recent years, which fits closely enough with the idealistic and historicist side of Max Weber. There is a deeper reason for this connection: Both Marx and Weber are part of the historically oriented German intellectual tradition, which reaches back to Kant and Hegel and to the great scholars of the nineteenth century who founded history as an academic discipline. Similarly, the tradition that we label with Durkheim's name is the French and English tradition of positivism, evolutionism, and comparative anthropology, with its rather direct lineage through Saint-Simon, Comte, Mill, Spencer, Durkheim, Mauss, Lévi-Strauss, as well as other followers, including many Americans.

This reduces the list to two core traditions. A third, which I think has been unfairly neglected, is obvious enough in the United States, for it is our own native tradition in social psychology—or possibly one should say microinteractionism. It has no really towering figure upon whose name it can be hung. George Herbert Mead possibly is the most eminent candidate, but there are also Cooley and W. I. Thomas. I want to suggest that this American tradition is broader than symbolic interactionism, though, because recent contributions to it include such theorists as Goffman and Garfinkel, who take it in quite different directions. In fact, we have another illustration of how we create the past as we create the future, as the endeavors of Garfinkel and his school have now roped Husserl, Heidegger and Wittgenstein into our intellectual ancestry.

Now that we have these three traditions—or fictions, if you like—conveniently in hand, I will embark on a quick tour of their accomplishments as well as their weaknesses. I begin with the Marx-Weber tradition. Its greatest strength, I think, is that it takes stratification as its central topic. Stratification is theoretically crucial because it stands at the crossroads of social institutions. Its explanation leads directly into a theory of politics, economics, and social change. Moreover, the debates among Weberian and Marxian views have been fruitful in producing empirical research and genuine theoretical progress. I would say, for instance, that we know a fair amount about the general determinants of the varieties of class cultures; something about the causes of varying distributions of wealth; something about the structure and the power struggles of formal organizations; as well as the conditions of revolutions, uprisings, and social movements. Moreover, this has often been put in the necessary context of historical comparisons, which has now led us to formulate theories at the level of the world system itself rather than some artificially selected segment of it. And important new areas have opened up, such as the breaching of an ideological screen that used to exist—until about fifteen years ago—that kept us from seeing theoretically that there was such a thing as sex stratification. Among other things this is creating a theoretical upheaval in our understanding of the family.

But not all is so rosy with this Marx-Weber-stratification-historical-change tradition. I will mention just a few of its concomitant flaws. On the empirical side, I think we let statistical techniques too much dictate our concerns about stratification. Undoubtedly the most popular research topic in the last twenty years has been the study of social mobility. (Some of its practitioners like to refer to it as the status attainment tradition, but that has a certain ideological ring that I would prefer to avoid). It has to be admitted that there is no general theory of social mobility. We have been largely concerned with measuring the amount of mobility which has occurred in recent decades in the United States and describing some of its channels—insofar as these can be inferred from conventional surveys of individual traits. It is not merely that the structural side has been left out almost entirely, but that the really fundamental requisite of general theory has been slighted: I mean the method of systematic comparison which in this case involves comparing structural features of different societies to establish the determinants of varying amounts and processes of mobility. It is a little absurd to think that a series of individual attributes which apparently governed the progress of American school children in the 1960s constitutes a generalizable theory of social mobility.

There is an opposite weakness that afflicts the Marx-Weber tradition, which is its tendency to let historicist philosophy overwhelm all matters of theoretical substance. This was a deep-seated trait of nineteenth-century German intellectual life, having been formulated in different ways by both Hegel and Dilthey. Weber was even more deeply bitten than Marx, since Weber was an economist of the German historical school, which rejected general economic theory in favor of analyses of specific historical contexts.

Marx on this point was ambivalent. Although it is typical now to claim Marx as authority for a complete historical relativism, and indeed for a kind of Diltheyesque concentration of meaningful human constructions —what was aptly called *Geisteswissenschaft*—Marx nevertheless has a respectable lineage of concern for objective explanatory laws.

Today, the philosophical-idealist side of the German tradition has been bidding hard to eclipse any scientific aims in theory, leaning on the names of both Marx and Weber to do so. From the kind of statements now being made, especially in Germany, there is a tendency to turn the last 150 years into a series of footnotes on Hegel.

Although the historicist-idealist wing has a certain ideological popularity within the politics of the discipline, it seems to me not likely to make a significant contribution. It confines itself largely to metatheory and methodological polemic, and lacks the vitality to come out with new ideas.

The second major tradition has more or less the opposite strengths and weaknesses. The Durkheimian tradition has the merit, above all in Durkheim's own writings, of stressing explanatory generalizations. It also formulates the essential method for producing these: the method of systematic concomitant variation, which Durkheim stated should take especially the form of historical comparisons.

Unfortunately, this scientific aim was vitiated by Durkheim as well as his predecessors and successors, by mixing it up with other elements borrowed from biology. One is the argument from analogy between society and an organism, which has severely derailed the search for causal explanations. It has made it easy to substitute functionalist interpretations of existing institutions for a comparative analysis of the conditions that produce them. Along with the facile unilinear evolutionism, this has given the Durkheimian tradition a reputation for political conservatism, or at least support of the *status quo*. The impression is bolstered by the fact that the Durkheimian tradition does its best to ignore stratification, inequality, domination, and conflict, except as residual categories subsumed under pathological or transitional conditions. These are sufficient reasons, no doubt, why the Durkheimian tradition has tended to be in bad repute for the last fifteen years.

Nevertheless, I would argue that this tradition has made solid progress, if we strip away its conservatizing and normatizing tendencies. The proper antidote is precisely a dose of the German tradition, with its historical realism, its emphasis on conflict, and its recognition that a "society" is not an organism but an artificial construct abstracted from the larger world system in which it is embedded.

The strength of the Durkheimian tradition is not on the macro level, but on the micro level. Its strength is a theory of how social solidarity is produced by varying degrees and densities of social interaction. This is especially Durkheim's theory of rituals, set forth in his work on religion, which was a starting point for so much social anthropology. In recent generations this has been brought back into sociology proper: partly by anthropologists like Lloyd Warner and, in our own day, Mary Douglas, who have migrated from their tribes back to the mainland. Significantly, it reenters sociology as an analysis of stratification—of the ritual differences among class cultures and of the interactional conditions that produce them.

This can be seen by the way the analysis of ritual interaction and its corresponding forms of language and consciousness have been formulated by Basil Bernstein, Erving Goffman, Rose Coser, Pierre Bourdieu, and me. This puts the analysis of solidarity on the right level: instead of assuming that the entire "society" is the unit, we look empirically at the groups which actually do have given degrees of solidarity; these turn out to be social classes, or even smaller segments Weber would have termed status groups. I have suggested that the Durkheimian theory of rituals is in fact an essential link in any conflict theory of stratification and domination. It gives us a theory of the "means of emotional production," analogous to the "means of intellectual production" which the Marxian tradition started to delineate.

The third tradition, microinteractionism, extends from Cooley and Mead to Goffman and Garfinkel. Its main weakness, I would say, is too much of the metatheoretical gushiness of romanticist idealism. It claims to represent real-life, human consciousness against the impersonalities of scientific analysis. Unfortunately its claim to freshness and spontaneity wears thin after the same programmatic claim is repeated for years without producing much in the way of new ideas.

But this judgment is too harsh. The microinteractionist tradition has made some important gains; among the most important is the way

Garfinkel and the ethnomethodologists have revealed how everyday-life cognition depends upon a nonrational, taken-for-granted substratum. My own hypothesis is that this dark underpinning of society is in fact accessible, as the study of emotions, as well as of the so-called speech actions part of language.

This leads to a way that the microinteractionist tradition, which has long been so hostile to the macro and positivist traditions in sociology, can contribute something vital to them. I have suggested that macrosociology can be linked theoretically onto a firm microinteractional foundation, via a network of interaction ritual chains. These chains create not only varying degrees and links of solidarity, but also rev up and slow down the emotional energies which motivate people: sometimes to maintain routines of power and property, sometimes to move individually or collectively into new social arrangements. Interaction rituals, seen as nodes of social networks, operate not only as the "means of emotional production" (and "means of intellectual production," too) which provide the inner dynamics of the macrostructure of domination and conflict; they also link to the empirical foundation upon which the microinteractionists have placed proper stress.

To conclude: I am often amused when people tell me that sociology seems in a rut. If one looks at nothing but metatheoretical pronouncements, then surely things never seem to change. Perhaps one can't see theoretical development unless one has substantive theoretical interests of one's own. If one knows where to look beneath the clouds of dust that make up normal intellectual controversy, it is evident that the twentieth century has accomplished a lot. A real sociological science has slowly been coming into being.

The Development of Scholasticism

ARTHUR L. STINCHCOMBE

Sociological Prestige and the Distance Between Natural Humans and Humans as Sociological Objects

My general argument is that the development of sociology as a discipline led us systematically away from the study of humans acting in society. The higher the prestige of a piece of sociological work, the fewer people in it are sweaty, laughing, ugly or pretty, dull at parties, or have warts on their noses. Field work is the lowest status in methodology, because surprising humans keep popping out and bewildering us by doing things we do not understand; much better to have people answering closed-ended questions so that they fall neatly into cross-classifications to be analyzed by loglinear methods. Similarly, the highest prestige theory for many years, that of Talcott Parsons, started off with its people and their actions reduced to a cross-classification of five pattern variables (or per-

---

Note: This paper was prepared for the Thomas and Znaniecki conference on contemporary social theory. An extended version appeared in *Acta Sociologica* 27 (1983).

haps different numbers at different times). No room for laughter or embarrassing burps in the middle of a lecture here.

If we range theories from the prolix fashion of Herbert Blumer—who knows how people will define the situation and consequently what they will do—to the lean and spare rational actors models that allow us to use mathematical maximization methods to specify at least one feature of the behavior exactly (e.g., what the net profit will be), it is the theories that are most divorced from blood, sweat, and tears that have highest prestige. And we find that the highest prestige books on substance are those that treat other books. The most learned person in a department of sociology, the Reinhard Bendix or Hans Gerth or Edward Shils, will be found, in general, to be the one who gets his or her facts from books rather than from people.

The Dynamics of Disciplines

As a social structure, I take a discipline to be a social system with the following features:

1. It distributes prestige to cultural objects, such as scientific papers, by a set of standards which are maintained by debate and consensus rather than by an authoritative administrative or legal order.
2. It penetrates the administrative systems in which scholars or other producers of culture work, especially universities, and shapes them so that money and power are made dependent on the prestige derived from cultural production.
3. It lays claim to a jurisdiction of cultural production within which its standards, formed in this consensual-debate way, will be taken as authoritative.

My argument is that the dilemma any such social structure creates is that it has to render the process of cultural production regular and limited in order to be able to form the consensus on which the operation of a discipline works. Cultural messiness would be illustrated by a loglinear analysis of a cross-tabulation, with a list of nineteen answers at the bottom which didn't fit the classification criteria.

We always have to forget, in our devotion to the best maximum likelihood estimates of loglinear parameters, that the 25 percent who did not answer or who gave unclassifiable answers are equal in the sight of God to those who got into our tables, in spite of our best efforts to construct analyses that exclude them. Similarly, an action that might appear in full feather in a novel, such as Alyosha's dismay at the fact that the saintly Father Zossima's body stunk after he died in Dostoyevski's *The Brothers Karamazov,* has to be reduced to expressive rather than instrumental, particularistic rather than universalistic, and so on, before we know what to do with it in our structure of sociological theorizing (Parsons 1951, pp. 58–67).

I am not objecting to the abstraction as a scientific procedure. [Nor do I think that the fact that divorce is a matter of rage, tears, passion, and love for children should keep us from abstracting certain parts of it out, to deal with in the courts.] A successful abstraction is a great achievement. But if one abstracts from the other 25 percent of the population just because they do not answer survey questions, rather than because it is scientifically justifiable to treat politics as consisting of the other 75 percent then this distorts the scientific value of the achievement. If one excludes Alyosha's experience from sociology because "expressive and particularistic" does not seem to catch the essence of his dismay at Father Zossima's body stinking, one misses the main test of the conceptual scheme.

What we want, for example, is not a maximum likelihood estimate of the relationship between alienation from politics and political behavior for the population who answer survey questions, but any old nonoptimal kind of estimate that takes into account that the people who refuse to answer survey questions are quite a lot more alienated than those who agree to answer (e.g., see Stinchcombe, Jones, & Sheatsley 1981, pp. 374–75).

The Damage Caused by Disciplinary Structures

The two subdisciplines of sociology which would be predicted to have the highest prestige according to the above analysis are methodology, in the sense of the study of statistics and of abstractions such as a mobility table, and theory, in either mathematical treatments of oversimplified people or analyses of historically important books in the development of the discipline. The subdisciplines that we would expect to be lowest in prestige are those in which people come to us in raw form, without having been culturally processed, such as the substantive study of the family, or of criminals and crime, or whatnot. But, after all, the purpose of the methodology and of the theory is to understand social behavior like that of families or criminals. The result of making methods and theory into high prestige subdisciplines is that they tend to close themselves off from sources of information that would tell them when they are wrong.

To take a very simple, but it seems to me telling, indicator of this, I believe that one could not find the percentage of nonresponse in the original survey from which the social mobility table came in any of Leo Goodman's articles analyzing such tables—it seems quite possible that disciplined lower-class Danes do answer survey questions while English lower classes are more resistant and rebellious to survey interviews, so that the whole set of ingeniously estimated differences between the Danish and English mobility tables are fictional (e.g., Goodman 1969, pp. 34–37). The hint that this is a possibility does not appear in Goodman's analysis of the table.

How Do Facts Get into Sociology in Spite of the Discipline?

While no one today would publish a book as chock full of human beings and human action as, say, Thomas and Znaniecki's ***The Polish Peasant*** (1918), clearly there is a lot of empirical commitment in sociology. If it doesn't come from sociology's structure as a discipline, where does it come from?

The sources of empirical commitment are various. Let me mention briefly and elaborate on a few of these sources: (1) applied interests, either of a "practical" or of an ideological sort, (2) general intellectualism, (3) reflections on one's own life, (4) sectarian commitments in sociology, and (5) teaching. I will argue that each of these is likely to produce an impulse in individual sociologists or groups of sociologists which makes them psychologically committed to explaining, reporting, or otherwise dealing with raw facts that are not preshaped to our disciplinary taste.

1. Applied interests: One can have either narrowly practical or broadly ideological purposes which commit one to understanding a given empirical reality. For example, governments want to know what effect ethnic segregation has on educational attainment and what effects of various public policies with respect to ethnic segregation would be on that attainment. A sociologist hired to find out about such matters, or with a consulting practice providing expert testimony on such questions for plaintiffs or defendants in such cases, has good reason to understand the human complexity of the scientific area.

The low prestige of such activities is sufficiently communicated by calling them "applied" or "ideological."

2. General intellectualism: A good many sociologists are folks who like to read books, whose eyes stray at the breakfast table to anything in sight that has print on it. A sociologist led to read novels in order to be an

intellectual does eventually realize that Alyosha's dismay at the stink of Father Zossima's body is social production and that we do not really understand it.

The commitment to facts as reported in books read by general intellectuals, then, is a source which pulls against the pull of sociological prestige. Of course, a fact about the current situation in El Salvador is unlikely to appear in the *American Journal of Sociology* until it is cleaned up and made into theory or method. But knowing about it does produce a penumbra of a sense of social reality in a large segment of the discipline that keeps the disciplinary dynamic from having its full force.

The low prestige of such activities is embedded in calling them "journalism," or calling the person "a *Commentary* sociologist."

3. Many of us, I suppose, came into sociology because we had difficulties of one sort or another managing social relations. I suppose that a lot of the troubles were of the general form that one's prestige ranking in the intellectual status system of secondary school was much higher than one's rank in the social-class dominated informal social life of the school: Our own hidden injuries of class sustain our interest in the reality of stratification. Some of the most interesting recent work in organization theory has been about university organization (especially Cohen, March, & Olsen (1972) on the garbage can model of decision-making), which is surely part of most of our daily lives. The drift of many of our best minds into the sociology of science suggests the same sort of fascination with what shapes our own experiences.

There does not seem to be a general pejorative for such activities, though I have heard "narcissistic" and "self-indulgent."

4. Sectarian commitments: By sectarian commitments, I mean loyalty to such social movements in sociology as symbolic interactionism, ethnomethodology, psychiatry, or occasionally historical sociology. The mechanism of commitment in such sociological sects to empirical materials seems to have to do with a passionate concern with a methodological principle which entails incorporating undisciplined social facts. Although George Herbert Mead did no field work, as far as I know, and Herbert Blumer published very little of his, their abstract methodological arguments implied that one could not get at the essence of social life without observing interactions in which the people themselves define the purposes and meaning attached to the actions.

Similarly, there is not much way that one can have a commitment to the ordinary psychiatric interview as a basic source of data without getting all sorts of messy facts about sex and misery and delusion, facts that fit uneasily into received social theory (What pattern variables describe sex? What kinds of sex do the nonrespondents have?).

The low prestige of such activities is sufficiently communicated by my own pejorative, "sectarian."

5. Teaching: Teaching undergraduates or Masters of Business Administration students is a personal war over who gets to define what is interesting. MBA students do not want a cleaned up social mobility table, but the inside dope on how to get ahead. Insofar as they sometimes win this war, at least in the limited environment of the classroom, they force the teacher to think about how people get ahead—thinking about how to get ahead undermines the commitment to loglinear modeling of the table.

Similarly, our students are interested in marriage because they are going to make one or are going to decide against it. An occasional teacher will discover, in trying to turn topics of love and sex into a more or less orderly lecture, that there is a lot of social interaction connected to sex that we understand very little about. How do men (it is usually only men) deal with the intimacy of masturbation when at sea on a submarine with very little privacy? How do people cope when they find out they have married someone who wants a different amount of sex than they do? Such topics confine a scholar to the peripheral journals of the discipline as a structure, but students can keep him or her worrying about such subjects.

The pejorative for the activity of talking about what students are interested in in the classroom is "catering to the students" or occasionally, by an unpopular teacher about a popular one, "seductive."

Conclusion

As a discipline, sociology as a social structure leads toward scholasticism. But fortunately scholars are not allowed to construct monasteries, so there is a constant stream of empirical pollutants that threaten the scholastic structure. There are enough general intellectuals, enough people who deal with problems in their own lives through intellectualization, and enough people who take what they say about sex and ambition in the classroom seriously in their scholarly life to supply the materials for a constant tension within the sociological community. The thing that keeps our scholastic structure from being perfect and eternal is that we keep having our attention called to social facts that we cannot yet manage. This threatens the discipline, yet keeps it alive. The disorganized flow of empirical social reality creates problems difficult enough to make it worthwhile to have a discipline trying to tame the flow into theoretically and methodologically unimpeachable sociology.

I should add that I got two kinds of criticism on my major argument, one by Howard Becker and the other by Stephen Warner. Both are substantial enough to warrant being reported here.

Howard Becker's point is that yes, there *was* a discipline of sociology when he was a boy and when I was a boy, but there is no discipline anymore. There are many subdisciplines in sociology, with separate journals. Even the *AJS* and *ASR,* which used to be the core of the discipline, have changing criteria, depending on the subdiscipline that dominates them at the moment. When a scholar fails to get an article into these journals, he or she sends it to a journal of the subdiscipline: The people who are really interested in what the article has to say are the subdiscipline's readers and referees.

Becker urged a couple of tests of this. One of them was to ask the folks who have been awarded the Sorokin prize or the MacIver prize how many of them followed the methodology being advocated by the head of the methodology section of the ASA at the time they were awarded the prize. If one did that examination, Becker believes, one would find that they never used that prestigious methodology. A rather similar thing happens on the theoretical side. For example, *The Presentation of Self in Everyday Life* sells better than Talcott Parsons. Thus, Becker's counter-argument is that the assumption on which I based my argument, namely, the unity of sociology as a discipline, is fundamentally flawed.

The second criticism, by Stephen Warner, is in some ways more fundamental. He argues that when one looks closely there really isn't theory or methodology. Instead, to call something "theoretical" is simply another way of saying that it is prestigious. A really first-rate work on, say, the growth of science in the seventeenth century will get published under a title that includes "theory" because it grips our imagination and seems to illuminate our substantive understanding. Thus, whatever is prestigious is what we call "theory." And the very best kind of methodological analysis operates much the same way. Thus, in Warner's view, theory and methods get to the center of our discipline because they deserve it. Retrospectively we call them "theory" or "methods" to indicate that we think they are deserving.

Two Traditions: Substantive Analysis and Empirical Research

EDWARD SHILS

I was so filled with admiration by Professor Collins's paper that I want to go him one better. I see only two traditions in sociology. First, the tradition of empirical study, the study of the data, either using officially gathered statistics or creating data where they didn't exist before by direct observation, in the way field anthropologists or the old-fashioned participant observers used to do or the way it is done now by surveys. Second, the tradition of substantive analysis.

The empirical tradition is an illustrious one, full of disconnectednesses. I recommend to you, for a much better account of very important aspects of this, a wonderful essay by the late Professor Lazarsfeld reprinted in his collection of papers edited by Professors Merton, Coleman, and Rossi. I think it is the most original article written on the history of sociology since Talcott Parsons's chapters on the growth of sociological theory. It begins in the seventeenth century with political arithmetic in England. In Germany, as Lazarsfeld points out, something grew out of efforts to see the regularities of God's will manifested in demographic phenomena. This was at the time when data scarcely existed, which was generally true until the early nineteenth century when censuses began to be taken.

Surveys were done by going out and looking in the horse's mouth so to speak or in a pickpocket's pocket. There were surveys of what they used to call in America "the debtor, dependent and dangerous classes" and in France "les classes dangereuses." Even Karl Marx, that recently acquired father of sociology, prepared to make an *enquête ouvrière* and drew up a questionnaire. Questionnaires were in the air at that time. A lot of information was gathered by interrogation of individuals as to how much money they had, how many rooms they had, how many children slept in each bed, and so on. That was a very important strand of sociology. The great monument of Charles Booth's *Life and Labor of the People of London* is a certain culmination of that strand and became a model, which then came to America in the Pittsburgh and Springfield surveys. If you read in the beginnings of Dickens's *American Notes,* you will see Dickens refer to various surveys carried on by retired army officers on the number of deaf children in Boston and things like that. There was a lot of home-cooked sociology going on.

The other tradition was a substantive tradition. I won't call it analysis; I won't call it theory. It is the concern about society, an intent to understand the dynamics of a society and the special properties which distinguish societies from each other.

Durkheim represented a coming together of these two traditions. The two traditions came together also in Chicago. W. I. Thomas and Park were in Chicago. Especially Park, who knew the surveys and the classical ecological studies, was concerned with a picture of the working of a society. He improvised all sorts of techniques, mostly participant observation, just looking on the hoof, the way a newspaper journalist would do. At one time there were journalists who went about, talked to people in the streets, looked at the police blotter, went to houses and asked people questions, and so on. Journalism was an important part of sociology. The Chicago ecological approach in many respects continued some of the things done in the nineteenth century, such as the search for spatial distributions of various phenomena and the concentration of criminality. It was a spot mapping technique, a cartographic technique.

Durkheim also used the same technique; however, he used on a national and international basis what Park and Burgess used for these small precincts or census tracts in Chicago.

Then an extraordinary development came, oddly enough, out of market research. But before I go on, I would like to ask you, Robert: Paul Lazarsfeld wrote under a false name, Elias Smith; did he have in mind Vladimir Elias Berg who was one of the forerunners of market research and advertising analysis in general?

Robert Merton: "Much more pedestrian than that. Elias stood for 'Alias.' Lazarsfeld used this name because he would be writing all eight articles in a single issue of the journal."

So Lazarsfeld was like John Wayne on the Alamo, running from one place to the other. Back to the story. The new development was coming out of market research and to some extent out of psychology and scaling and analysis of characteristics of individuals, which owed much to Wundt and German psychology. Paul Lazarsfeld, with extraordinary ingenuity and daring, picked up those things and put them all together. With him and Robert Merton you had another coming together of the two traditions of substantive analysis and empirical research. It is not to be flattering our distinguished guest here this evening to say that he especially represented both the substantive and the empirical, whereas Paul Lazarsfeld was more interested in the empirical; although he was full of ingenious ideas in substantive matters, they were more disconnected.

Now let me turn away from that and concentrate on the substantive tradition, which is mainly what Professor Collins was speaking about. The substantive tradition is a rope composed of sorts of heterogeneous but overlapping strands. In order to describe these strands, the terminology of Tönnies is good enough as a start, with one thing added: *Gemeinschaft, Gesellschaft* and an idea of a cultural or moral order. Those who have chosen to read Durkheim remember him well enough to know that these were the things on his mind. It is not by accident that Max Weber's great treatise on sociology *Wirtschaft und Gesellschaft* was meant to be just one part of a textbook in economics. One major chapter is called "The Sociological Categories of Economic Activity" and fundamental phenomena like the market and the division of labor constitute an important part of that. That whole set of phenomena which is supposed to be characteristic of *Gesellschaft*—the society of organic solidarity, the disintegrated society, a society of individuation, a society of separation, a breaking up of the larger society—is a common theme in all of the writers. Durkheim, Max Weber, W. I. Thomas, even Sorokin—all of them came out of that tradition.

The reactionaries and French patriots used to abuse Durkheim for bringing Germanity into France and they were right. Durkheim came out of the tradition of the study of the folk spirit, the study of folklore, and the German ethnography called *Völkerkunde.* Although the fundamental ideas of Weber and Durkheim are really different, they belong to variants of the same tradition. Durkheim read a great number of German ethnological books, books on the theory of evolution of society, evolutionary theory of sociology, and the data he used for documentation came in part from Germans and in part from British and American ethnographers.

He applied his conception of collective conscience, the forms of the affirmation and threatening of collective conscience. That was on a direct line from the German conception of the folk psychology. He was coming more or less from the same tradition as Tönnies. In addition, both had studied Hobbes. In fact Tönnies even discovered some unpublished manuscripts of Hobbes.

That tradition then came into the United States, in what became in its time the dominant sociology, although it only occasionally referred to Durkheim. Thomas knew the German folk psychology literature, the *Völkerkunde,* and the ethnography. One of his main books before *The Polish Peasant* was *Systematic Source Book of Social Origins*, which shows that Thomas read a prodigious amount from the same sources that Durkheim was reading from. In 1915, Robert Park wrote an essay called "The City, a Spatial Pattern of Moral Order." This essay shows all the different and intertwined strands that make up the "dominant" sociology. It brings together the ecological, the quantitative description, the market, the division of labor, and then the system of the collective spirit entering into social formations.

General Discussion

Arthur Mann: I do have a question. If one were to look at the founding fathers of sociology in the United States, how important was "Do Good," amelioration, solving social problems?

Edward Shils: Fundamentally important. The answer is obvious.

Arthur Mann: Why then does it not become one of the major traditions in sociology as a discipline? It was not mentioned. I heard three traditions from one person, two traditions from another.

Edward Shils: Because it was pervasive. It didn't have to be mentioned.

Arthur Mann: To what extent is it still true of sociologists that they think they can find a handle on the world so to reform it?

John Freeman: If you look at the discipline of sociology as an enterprise, you ask what drives it? Why are students attracted to this discipline? And the answer is that sociology has something to say about things that students care about. I think we are in a cultural period right now where people care less about social issues than they did a few years ago, but I have no doubt that the issues are not going away and that people's concern with them will return. The material basis of the discipline in terms of grants, the willingness of universities to allocate slots to sociology departments, and so forth, is directly related to the degree to which the discipline addresses these issues. In short, I think that "do-goodism" is essential for sociology.

Arthur Stinchcombe: One should not exaggerate the difference between sociology and the rest of intellectual life. For example, take the concern with rescuing people from the Chilean government. You will find many sociologists among those concerned, but you will also find many literary people and linguists and all kinds of folks. When we speak of the problem of intellectuals in Poland, we are not just talking about sociologists but about people who come out of a general tradition of being politically in favor of intellectual freedom and many other things that are embedded in a humanist tradition. In other words, it is not only we who are "do-gooders." By and large you can predict a person's degree of cultural leftism by how many books he reads a year. We partake in this "dangerous" character of the general intellectual community.

Ronald Burt: I would like to know why Stinchcombe was so pessimistic about sociology in his presentation.

Arthur Stinchcombe: I wouldn't say that I was not optimistic. What I was trying to get at was this: What fundamentally gives us vigor is understanding social reality. We must get social reality into the discipline. On the other hand, I am firmly in favor of then transforming it into culturally valuable products that might be published in prestigious journals.

Jack Goldstone: My contact with people in other fields leads me to be optimistic about sociology. Political scientists and historians are paying attention to work done by sociologists in the theory of family organization. Political scientists are paying attention to structural theories of revolution and social change. Business schools think enough of sociologists trained in organizational theory to hire a fair number of them. There may not be a widespread sense that there is a core of sociological theory but the vitaliy of our discipline comes from the fact that people in other fields are paying attention to what many of us do.

Edward Laumann: I think there is a problem with sociology, a problem of numbers. There are some 12,000 sociologists who belong to the ASA. About 90 percent of sociologists do not write anything once they have their degree. Lawyers, on the other hand, have 500,000 troops. Their scale of operations is quite considerable. This is the problem of numbers.

Second, it seems that we are at the domestic putting-out stage of operation. Most of us, particularly the so-called theorists, tend to pick a crafted type of activity where we can do the job ourselves, where we can interview a few people. The federal level of funding for sociological inquiries, about $3 million a year, would buy us a foot of a Polaris submarine. We do not have the concentration of intelligent people with the technology and the theoretical guts and nerve to collect information that we need in order to generate theories with power. So my thesis is that we lack the critical mass and that much of sociology is just a sort of teaching enterprise to undergraduates.

Theda Skocpol: Everyone is discussing the central dilemmas that we are facing now, but no one has discussed what are the substantive problems a sociologist might care about. Sociology needs theory, it needs methods, and it certainly needs great people to talk about. But above all what it needs is problems to investigate and people talking to each other about how to understand those problems.

Edward Shils: It's true, in a vital field, that people are busy; not so much talking about problems but doing important things. But I don't think I do Professor Skocpol an injustice by saying that one of the parts of her intervention was to suggest that we ought to draw up some kind of an agenda for sociology. Well, we shouldn't. Everybody should do what he thinks important provided he can get somebody to pay for it if it takes money to do it.

I don't think that programs ever help very much. When I was younger, people were always making lists of problems for other people to solve. The Social Science Research Council in the 1920s and 1930s thought that one of its functions was to work out an agenda. It fell on barren soil, not because sociologists were morons but because people want to do what they want to do. I think one of the cats in T. S. Eliot's book of practical cats will do only what he is going to do and that is what a real sociologist, a real physicist, a real mathematician does. If someone else gives him a problem, he transforms it. There is a certain unity because there is an elite which studies what it wants to study and there are others who run after the leaders. If you have an appearance of consensus that way, there is also dialogue among the important discoverers in the field and I view that as a very desirable thing.

Randall Collins: Let me respond to Skocpol's question not by talking about what I think is important personally, but by pointing to work that is represented in this room.

We have had some debate here between two positions. These positions are so powerful because they are working programs the adherents of which have made progress and feel that something extensive can be done with them. For example, Oberschall has taken the rational choice model and is going very far with it. Lindenberg gave us more programmatic cheerleading for it without the specifics of what he is doing, but clearly there is a big working program there. On the other side, and this may be some of the dialectic today, is the network analysis with its resistance to a certain kind of individualistic reductionism. Again, this is an area that is relatively new; it hardly existed fifteen years ago.

There are quite a few other things. There must be at least three or four people in this room who have produced fairly extensive ways of attempting to resolve the macro-micro dispute not by arguing about reductionism or the opposite but by actually producing theoretical systems connecting the levels. There is also a complete program fully under way for developing a theory about the world system. It is easy to look among us and see people who have made tremendous strides in political sociology, the ongoing development of theory of revolution and other forms of politics. There is a much more powerful form of analysis developing in microsociology; wait until we get to the papers by David Heise and William Labov. There is also the vigorous field of the population ecology of organizations. Not that all these things are congruent, but they are more than programs. They are fighting over who is going to control what territory, and I think that's a good situation to be in.

Joseph Ben-David: We seem to be going back and forth between an optimistic and a pessimistic appraisal of sociological theory. But perhaps the state of sociological theory is not independent of the state of society. Societies are created by people and they are created with certain visions, with certain purposes which change through history. There are great periods in history when people believe they have a sense that their society has a meaning and there are periods in history when people are confused because society doesn't seem to be going anywhere.

Now these facts influence sociological theory very seriously. Durkheim and Weber both wrote great sociology because they both lived in great historical periods. Durkheim was the sociologist of the Third Republic and Weber was the sociologist of the German empire at a time when there was a chance of its becoming a liberal society. The great age of American sociology was, in my opinion, the postwar period and had a lot to do with the work of Talcott Parsons. But his work was influenced by the sense that the United States had become a society which plays a tremendous role in world history. If we do not have such theories today, it is because we do not have such societies.

James Coleman: There is a different perspective which has not been discussed tonight. A large number of people in society today are appalled at the growth, importance, and dominance of sociology. In the field of criminology, a large number of people ask, Why are sociologists so powerful? In the field of education, a large number of people ask, Why and how is it that sociologists are so powerful? In the field of the military, the same thing is true. I could go on for five more fields. Obviously we have a discipline which has grown extraordinarily in importance in society. But there is another question, to which this conference begins to provide some answers: While we have a great deal to say to the rest of society, do we have things to say to one another which we can mutually understand, if not agree upon? Is there a *central core* to the discipline? I for one am not willing to say that sociology is a set of applied fields. It is growing enormously in strength because of its applications, but for that very reason it is extremely important that we pay attention now to the central core.

Robert Merton: It can be argued that this has been partly derived (in an almost Spencerian sense) from not only the absolute growth in numbers but also from the changing rate of growth in numbers. I would provisionally interpret the malaise as an inevitable response to this enormous differentiation of interests, concerns, subject matters and primary foci of attention. That malaise derives from a sense that the discipline is growing apart. The differentiation has produced malaise owing to the concern that we cannot easily communicate with one another, that there is not that intellectual core which enables specialists in seemingly unconnected spheres to recognize what they have in common, namely, some sort of sociological perspective.

Since we are each advancing diagnoses, short-term or longer-term, I will add another to the lot. I think that a conference of this sort, annually reproduced, is one of the few ways in which we can reduce that cumulating sense of malaise. A conference provides occasion for discovering, in a rather direct and intimate way, differences in orientation which do not involve cognitive contradictions. I do not, of course, deny that cognitive contradictions occur. However, as often as not, the malaise is less a matter of cognitive conflict than derivative of strong and exclusive commitment to one's own theoretical and substantive foci of attention. Who among us can conceive of working away at our lasts privately, believing that these are really of no consequence? Obviously, we are saying: I think this is important. That is what I took to be part of Shils's message. And this sense of importance, this claim for nothing more than the attention of peers, introduces potential structural conflict in a discipline that is growing relatively rapidly and becoming differentiated at a very rapid pace. The malaise arises in part, I think, from a sense that what one is doing is not recognized widely enough; that awareness of the work is confined to a small subset of people in a special field.

A diagnosis along these lines suggests that pessimism results from the growing pains of a rapidly differentiating discipline in which the differentiation has multiplied so fast that we haven't the human resources to develop each sphere of inquiry in sufficient degree. The sociological enterprise requires a sense of tolerance rather than of battle, consolidating a mutual awareness of various theoretical orientations with a reasonable confidence that their mutual theoretical connections will be progressively defined.

One last observation. Leo Goodman is in my direct line of vision. Those of us who have known Goodman's work over the decades know that he represents a superb specimen of someone who has made a personal commitment to a particular line of inquiry; specialized as it is, it is not merely private. It is a commitment which has made its presence known and felt in large parts of the collective sociological enterprise. In each of the developing specialties, problems do emerge (and here I am connecting with what Theda Skocpol apparently had in mind). The problems derive from the antecedent work. One does not set out a schedule of problems; one discovers problems as part of the ongoing process of focused inquiry.

Sociology of Knowledge

Academic Market, Ideology, and the Growth of Scientific Knowledge: Physiology in Mid–Nineteenth-Century Germany

JOSEPH BEN-DAVID

The Background and the Problem

Sociology of knowledge was very popular in continental European sociology during the 1920s and 1930s, went into almost complete oblivion during and after World War II, to reappear again with added force about 1970 (Curtis & Petras 1970; Fuhrman 1980; Hamilton 1974; Merton 1973; Remmling 1967; Stark 1958). This recent interest in the sociology of knowledge has been distinguished from prewar sociology of knowledge by three characteristics: (1) the field is now as popular in Britain as on the continent of Europe and has many adherents also in the United States; (2) it also claims as its domain science, a branch of knowledge which was not considered amenable to sociological interpretation by the central Mannheimian school of the prewar period (although there were even then attempts to include science); and (3) sociology of science has strong support among philosophers and historians of science (Barnes 1977; Bloor 1976; Hesse 1980; Mulkay 1979; Mulkay & Miliĉ 1980). What has remained unchanged is that the field claims to replace universalistic methodologies and epistemologies of science and cogni-

Note: This research was supported by the Spencer Foundation.

tion in general with relativistic criteria of truth varying from social context to social context. Therefore, the research program has also remained similar and consists of attempts to reinterpret important cultural events, such as scientific discoveries, as the outcomes of the social motivations and interests of the discoverers acting in particular historical and/or organizational situations rather than as events in a coherent chain of attempts to find logically satisfactory explanations of puzzling phenomena of nature.

This approach has not been universally accepted. There is an ongoing debate among sociologists of science, paralleled by one among philosophers, between those in favor of and those against the attempts at a reduction of scientific knowledge to the interplay of social forces.[1]

The purpose of this paper is to examine whether the terms in which these debates are conducted in sociology are adequate to the present state of research. It will be suggested that they are not and that the issues have to be reconceptualized in order to bring them in line with the results of empirical studies.

To begin, I shall state some propositions formulated in the early 1970s, shortly after the recent debate had started (Ben-David 1971). First, scientists, as well as other thinkers, go about finding answers to their problems through examining the relevant information and by analyzing the problems with the aid of whatever conceptual tools are available to them in the disciplinary and/or professional traditions they have inherited.

In addition, every scientist may, and probably will, be influenced by a variety of other conditions, such as personal or collective interests, political currents, religious beliefs, and knowledge derived from other intellectual traditions than those of his field of research. These may help or hinder the finding of satisfactory solutions to the problems under investigation by drawing his attention to, or distracting it from, data and ideas important for the solution or by creating emotional attachments helpful or harmful in finding a solution. Therefore, such influences may be important in the explanation of the work of any particular individual or group or the explanation of any particular discovery.

When, however, the purpose is not the explanation of one-time events but the development of a disciplinary or subdisciplinary tradition, the

[1] There were emotional elements in this difference of opinion. But many also found the very concern with the demarcation of science constricting. This reaction was, however, not entirely undesirable. No one was deterred from asking exciting new questions about the demarcationists' no-man's land, namely, the relationship between science and other fields; and the fact that there was no empirical evidence to show that scientific knowledge was indeed shaped by social conditions was eventually taken as a challenge, leading to an increasing flow of empirical research from about the mid-1970s on.

importance of these changing historical conditions will tend to be eliminated. All subsequent users of a contribution to a tradition will approach it from the point of view of its usefulness for the solution of the substantive problems of the field. Like the original contributors, they will also be influenced by a variety of conditions not related to the tradition in question. Since, however, the considerations internal to the field are constant, while the unrelated "external" influences vary from person to person, the effects of these external influences on the development of the field will probably be eliminated in the long run. This process of selection need not be very prolonged, since it is speeded up by the decentralization of the scientific effort which ensures some instant randomization of external influences and by the socialization and reward system in science which tends to select for success researchers who have the intellectual perspicacity and emotional discipline to keep out of their work influences not contributing to the solution of scientific problems they are working on.

Therefore, it was concluded that these external influences will not be "systematic," that is, "regular and predictable. Occasional influences may provide the theme for historical investigations but not for a sociology of science" (Ben-David 1971, p. 8). Furthermore, it was argued that this was true not only of natural scientists but also of social scientists and political and moral philosophers. These, of course, will ask questions relating to the current state of society, which is the subject matter of their enquiries, but will still deal with those problems with conceptual tools acquired in the course of their training in intellectual traditions which are the creation of people from many ages and societies. These traditions are and can only be held together by an internal logic and interest in long-lasting specific problems, not by the changing variety of social conditions, beliefs, and other influences impinging on the people who created, and contributed to, them in different ages and societies.

These views were supported by empirical evidence, such as a comparison of the success of field theories, probably influenced by romantic *Naturphilosophie,* in physics, and the rejection of these ideas in biology—in which they had originally much greater influence—because in biology they did not prove useful for empirical research.

Two Views of Physiology in Mid–Nineteenth-Century Germany

With a single exception (Mulkay & Miliĉ 1980) these arguments were not paid attention to by the protagonists of a sociology of scientific knowledge. Nevertheless, I still believe them as essentially correct, although in need of some change of emphasis. In the following I shall try to show their usefulness through the analysis of a historical case, that of physiology in mid–nineteenth-century Germany.

Timothy Lenoir, a historian of science interested in the relationship between philosophical and political ideas and biological theory, and well versed in sociology of science, has recently investigated the rise of the "organic physics" school, comprising Brücke, Helmholtz, C. Ludwig, and, in particular, the career of E. Dubois-Reymond, who was the main ideologist of this group (Lenoir 1983). The subject matter of his work partially overlapped the work that Avraham Zloczower and I did about twenty years ago and can usefully serve as an illustration of the change in the sociological perspective during the intervening period (Ben-David 1960; Ben-David & Zloczower 1961; Zloczower 1981).

When Zloczower and I looked at the steep rise of German scientific research in the middle of the nineteenth century, we addressed the following questions: (1) What were the conditions in the German states which motivated so many people of outstanding ability to take up research as a vocation and those in power to encourage and support their endeavor? (2) How did it happen that in spite of the initial predominance of idealistic and romantic views on science, and an explicit prejudice against experimental research at universities, experimental fields, namely, chemistry and physiology first, followed later by physics, gained within a short time so much the upper hand, that eventually—beginning in the 1850s—German scholars attempted to turn even philosophy into an experimental science (an effort which eventually gave rise to experimental psychology)? (See Ben-David & Collins 1966.)

We found the answer to these questions in the existence of a—by the standards of those times—large (about twenty-four universities), decentralized academic system, in which the universities (or rather the Ministries of the various Länder financing the universities—Turner 1970) fiercely competed with each other for academic fame. This created a sellers' market for able and successful researchers working at the forefront of science. The opportunities were particularly favorable in the experimental sciences, in which success was judged by relatively objective standards, academic recognition was worldwide, and the opportunity to start a new specialty—and thus to realize the dream of every competitor in a market for obtaining at least a temporary monopoly—was greater than in the humanities. These conditions attracted talented people and gave them the bargaining power needed to obtain new academic positions, good laboratories, and to overcome the prejudice against experimental fields.

Physiology was a particularly good illustration of the way the system worked. Originally a subsidiary branch of anatomy, it became a separate and autonomous field due to the successful bargaining of Carl Ludwig with the University of Zurich when offered a chair at that university. The precedent of Zurich was followed by all other universities. The extremely productive and inspiring Ludwig went on eventually to Leipzig to establish at that wealthy university a center for research and training of unprecedented size and importance.

The view of science implied in this model is that of a scientific community doing its best when provided with sufficient funds, able to work in an institution which safeguards its autonomy and to apply purely intellectual criteria in the selection of research problems and evaluation of results (Merton 1973a, 1973b; Polanyi 1951). To this model should be added competition, to keep both scientists and university administrators on their toes.

Therefore, the research strategy followed by us was as follows: first to examine the structure of the university system and then to concentrate on the way this structure affected careers and research, in particular the growth of new specialties and disciplines.

Lenoir accepts this account as part of the story and adds important new information to it, but his view of science and his basic theoretical model are different. If I may characterize the approach that Zloczower and I take as looking primarily at the internal social structure of science and considering the interface with the environment only from the point of view of the conditions needed to secure the boundaries and sustenance of the internal structure, then one can characterize Lenoir's approach as concentrating primarily on the interface and regarding the internal system as much more unstable and punctured by many more interpenetrations of people and ideas than we assumed. Even more important is that Zloczower and I were interested only in the growth of research and academic recognition of new fields of enquiry, while Lenoir wants to give a sociological explanation of how and why certain scientific ideas arose and/or became adopted by scientists in a given field.

His different perspective suggests to him questions not asked by us. He asks: Who were those attracted to physiology? What were their purposes prior to entering this field? What extrascientific resources did they bring with them in order to forge ahead in the field? How did they manipulate the internal environment in order to realize their preconceived goals? How did they balance, at various times, their scientific work with their political and status interests? How did they find a way to relate their research interests to their social ambitions?

Thus, we obtain a new, different picture: At its center is a group of very able young men, from upwardly mobile middle-class backgrounds, fired with ambition to "occupy a respected position in society," who chose science as a vocation in which one could do something important for society and as an avenue of upward mobility. All but one of them (Carl Ludwig) were students of Johannes Müller, whose laboratory was the foremost in physiology during the 1830s (Cranefield 1957).

The most interesting question addressed by Lenoir is: Why did this group of people choose a radical biophysical program, "organic physics," for their research which, as Cranefield had shown, was not quite practicable at that time. Lenoir stresses that there was in this an element of revolt against Müller and an attempt at establishing themselves as a distinct group (a kind of "product differentiation"). Above all, he shows that the serious study of physics which all the members of this group undertook (Helmholtz eventually became a physicist, and Dubois-Reymond was also suggested for a chair in physics at one time) was related to their participation in a scientific-political movement in Berlin headed by Gustav Magnus.

Magnus was an experimental physicist of bourgeois background who fought against the monopoly of the theoretical physicists at the University of Berlin. He succeeded in obtaining a full professorship, which meant official recognition of experimental physics, in 1845. The discrimination against experimental physics—a field which he believed was of greatest importance for industry—was for him part of the same system of aristocratic prejudice and absolutist authoritarianism which also discriminated against the "working bourgeoisie" and kept Germany dismembered into a number of small states, the existence of which was in the interest only of the members of the royal or princely families ruling those states. Cultivating and teaching physics, in particular experimental physics, was for him a mission of spreading enlightenment and knowledge useful for economic growth.

The organic physicists, alongside a number of junior officers from the army engineering corps, inventors (like Siemens, Halske, and Leonhardt), and several university students joined Magnus in the Berlin Physical Association established in 1845. This was done partly out of their wish to perfect themselves in their different professional endeavors through the application of physics and partly because they shared with Magnus the prevailing liberal views on political, economic, and academic reform. Thus, organic physics was not simply an endeavor to advance physiological knowledge but was also part of a world view, related to a sociopolitical ideology.

To consider their scientific views as part of the liberal world view must have been an encouraging and exhilarating feeling during the pre-1848 period when liberalism was on the rise and seemed to be on the verge of political success. Identification with the liberal movement and advocacy of academic reform could also appear, at that time, as useful for the advancement of the views and careers of the organic physicists. It is quite likely that these circumstances were instrumental in welding these young physiologists into a cohesive group, with a boldly declared purpose and program of reconstituting physiology on purely physical-chemical foundations.

Following the failure of the Revolution, their mood changed, and for a while they felt that their hopes for scientific and professional advancement would have to be given up together with their hopes for political reform. Fortunately, their fate did not depend on support from their political friends or on ideological climate. They could still resort to the internal system of science and get ahead according to the politically neutral criterion of scientific success and recognition concentrating on research and publication. Eventually, all of them attained central positions in world science and in the German academic system. They could and did wait within the safe precincts of the university for the turning of the political tide. When this came about, and the semiliberal political mood of the 1860s and 1870s, shared now by the highest circles, officially favored experimental science, the ambitious physiologists were not averse to returning to the public arena, adding intellectual luster to the economic and military successes of the new German empire and benefiting from favors granted to them on the assumption that research was useful for industry, the military, and the country as a whole.

An Attempt at Sociological Interpretation

Several remarks can be made on this comparison of these two studies of German physiology: (1) They do not contradict each other; (2) nevertheless, they convey significantly different views of physiology and physiologists; (3) it would be a pity to relegate the kind of study done by Lenoir to some place outside sociology (as I suggested twelve years ago; see Ben-David 1971); but (4) it is far from obvious how to incorporate it into sociology.

I shall attempt to deal with this last problem (which is the core of this paper) by making some more detailed comparisons. First, I shall discuss the question apparently asked in both the earlier and later studies: Why was experimental physiology—at that time a field having no practical application—so attractive to many able young people and thus destined to develop rapidly?

For Zloczower and me this question was inseparable from the parallel question about chemistry, which was the other experimental field that underwent a spectacular process of growth at about the same time. The explanation we suggested was that these two fields were particularly favored by the conditions then prevailing in the competitive academic system. In Germany both fields came under the sway of speculative *Naturphilosophie* and other philosophical systems during the first three decades of the nineteenth century. These theories proved to be relatively sterile for experimental research (although theoretically some of the ideas proved valuable—Lenoir 1982), and as a result German physiology and chemistry were backward in comparison with their development in France or in the school of Berzelius in Sweden, which rejected speculation and adopted an experimental-quantitative approach. While in Germany research was stagnating, in France and Sweden it advanced impressively. What happened as a consequence was predictable by the market model: Sooner or later someone was bound to import the new approach from abroad and, once imported and having proved successful in one place—philosophical opposition notwithstanding—it would spread rapidly throughout the system. And this is precisely what happened. The work of Stephen Turner (1971) and now the present paper by Lenoir have added important details to the picture, such as who made the decisions about appointments and how and under what conditions those decisions were made, but these have only confirmed the usefulness of the academic market model.[2]

Lenoir, who looks at physiology from the viewpoint of its place in the context of German society as a whole, pays no attention to its parallel with chemistry. What he sees, and to which we did not pay attention, is the link of this field with experimental physics, and beyond this with technology and industry in general. Physiology, perceived as "organic physics," had a particular attraction to young people like Dubois-Reymond, Helmholtz, Ludwig, and Brücke, because this approach based physiology on a general principle of reducing the mysteries of life to measurable and experimentally testable physical and chemical processes

[2]By the way, this model also fits the fact that the development of experimental physics lagged considerably behind that of chemistry and physiology (a point not sufficiently considered before). German physicists, unlike biologists and chemists, had never accepted *Naturphilosophie* (although their disdain for experimental physics was sometimes, probably wrongly, attributed to the influence of this philosophy). Theoretical physics was, and continued to be, an internationally respected and successful field, and there was nothing ostensibly better that could be offered by experimentalists such as Gustav Magnus. Unlike chemistry and physiology, in physics the experimentalists did not have a competitive advantage.

which, they believed, was a world view bound to emerge victorious not only in science but also in society as a whole. Their youthful scientific aspirations were reinforced by a vision of playing a leading role in the liberal reform of their society.

At first blush it may appear that the same question revealed two different stories and resulted in two different answers. However, there are, in fact, two different questions here. One question is: Why were so many able and energetic people attracted to physiology? The second question is: Why did they, or at least some of them, commit themselves in the early and mid-1840s to a dogmatically conceived and aggressively declared research program of "organic physics"? The first question is common to both investigations and the answers are not mutually exclusive. One can be attracted to a field for several reasons, and several people can be attracted for different reasons. The opportunities for success in research offered by experimental fields could attract people to either chemistry or physiology; some people might have been attracted to them also by ideology and social ambition. This latter kind of motivation might have played a part in the case of Dubois-Reymond and perhaps two or three others. But, I maintain, this was not a necessary or even important cause for the young people's interest in the field; the important and sufficient cause was the existence of favorable conditions in the academic market in chemistry and physiology and the intellectual opportunities available in both fields.

On the other hand, the adoption of the organic physics research program in a demonstrative and somewhat dogmatic manner can be explained only as a result of the ideological attractiveness of the particular approach and cannot be dealt with by the academic market model. This is consistent with the fact that after 1848, when they had to rely for success exclusively on the academic market, all but one (Dubois-Reymond) of the organic physicists either abandoned physiology or disregarded the program in their actual research (Cranefield 1957).

Another issue which illustrates the relationship between the approaches of these two sets of investigations is the importance attributed in Lenoir's paper to the class background and affiliation of the physiologists—an issue which was quite disregarded by Zloczower and me. It appears paradoxical that this should have proved a variable of importance to a historian, while it was found to be of no significance by a pair of sociologists (especially in view of the fact that one of the pair was a Marxian scholar).

This discrepancy is, again, the result of a difference between two points of view. Lenoir is interested in the place of science in the class system and politics of the German states, especially Prussia, in the 1840s.

From this point of view, experimental science, like industry, appears as a new avenue for the middle classes toward improving their status in society. Exploring the social background and the career choices of those who entered a scientific profession because, or partly because, they regarded it as an avenue to success suited to their circumstances and tastes is an excellent way of viewing physiology in the broadest social context, as well as coming to grips with the individual biographies of some key scientific figures of the day. Placing science in this context can show how politics can be linked to the furthering of the interest of scientific fields and the advancement of some scientists.

As against this, the earlier work that Zloczower and I did derived from a different interest: How did the phenomenon called science arise, spread, and grow in what can be loosely referred to as the Western world and later also in the entire world? We were well aware that politics, religion, and ideology had at times favored and at other times opposed science; that science had been very frequently used for political, ideological, and, of course, economic purposes; that quite a few scientists had welcomed, or even initiated, these uses; and that some (many fewer) scientists had bent their scientific views to political and other nonscientific uses. But the phenomenon we were interested in was how, despite all this, scientific research on an ever-broadening range of problems retained a coherence, continuity, and constancy of purpose to produce improved knowledge through rational enquiry and systematic observation and experimentation.

If one could view science as a ship trying to keep on course amidst storms and dangerous reefs, requiring many detours, then one could say that Lenoir is interested in the winds, currents, and reefs without which one could not explain the actual course taken by the ship, while we were interested in the steering mechanism of the ship which determines its long-term direction. Of course, there are also favorable winds, currents, and sea-lanes which propel and guide the ship on its intended course, and then the two viewpoints converge. The adoption of the physicalist program described in this paper was such a case of temporary convergence.

For us, the fascinating aspect of German physiology, and German science in general, during the nineteenth century was that, whereas previously the support and flourishing of science had depended almost entirely on favorable political and ideological conditions (on favorable "winds and currents"), in the nineteenth century there emerged a relatively autonomous and competitive academic system, capable of advancing research rapidly, with no need for justifying claims for the support of science by specific nonscientific purposes and values. Sticking to the nautical simile, it was something like the transition from sailing vessels to steamships.

Therefore, the decisive issue was that looking at the development of German academic science, or at specific fields of science, from the foundation of the University of Berlin in 1809 to about the time of World War I, one could see a logic that was reasonably consistent with the academic market model and was relatively independent from the numerous changes of the political and class systems. From this point of view, the class-consciousness and liberal involvement of the most important group of physiologists in the 1840s was an insignificant phenomenon; the most important thing was that the system led to rapid growth and innovation between about the 1830s and the 1870s (when growth was deflected to new specialties), irrespective of the revolution and its suppression in the 1840s.

What these comparisons show is that the same social events and phenomena could be viewed as parts of two different social systems: The group of young "organic physicists" in the 1840s could be located in the social system of prerevolutionary Berlin; at the same time they could also be located in the nineteenth-century German academic system. Furthermore, within each of these systems the group in question can be at the center of the picture, as in the study done by Lenoir, or part of a panoramic view of a larger scene, as in the investigations done by Zloczower and me. Or, it is possible to try to do both, to focus first on the group picture and then to look at the panoramic picture, locating in it the group by, so to speak, magnifying the spot where they are.

All of these ways of looking at the events are sociological. Whether the question is about a one-time event, such as the emergence of the organic physics program, or about a class and chain of events, such as the growth of experimental sciences in Germany during a period of time, makes the analysis neither more nor less sociological, as long as the answer to the question locates the event in an appropriate social system.

This is not to say that one should be content with case studies of one-time events. It is important to search for parallels and precedents. The confluence between a social movement, its political ideology and scientific world view, which served as an explanation of the emergence of the organic physics program, may have similarities with the confluence between the Puritan ethos and Baconian research strategy; or with the congruence between latitudinarian Anglicanism and a corpuscularian view of physical nature in seventeenth century England (Jacob & Jacob 1980; Merton 1973b; Webster 1975).

Another category of events into which the sociology of organic physics could be fitted may be a seemingly recurring relationship between reductionist approaches in science and moderately reformist ideologies in politics, such as have been witnessed in England beginning in the 1650s and ending at the turn of the century; in pre-Revolutionary and post-Thermidor France in the eighteenth century; or in the Western democracies from the end of World War II to the late 1960s (Ben-David 1983).

Conclusions for the Sociology of Scientific Knowledge

One clear conclusion from the present discussion is that the sociology of scientific knowledge represented by interpretations such as those of Lenoir, J. R. Jacob, and M. C. Jacob have no implications for epistemology. The scientific world views and programs discussed in these interpretations might have been embraced with enthusiasm because of ideological reasons. But the scientific success of those adopting them as manifested in the incorporation of their discoveries in the disciplinary tradition was determined by their usefulness for the solution of specific substantive problems. The relationship between science, ideology, and social structure established in these studies is functional and does not postulate a commitment to any relativistic or other epistemology. This substantiates the assertion made in the beginning of the paper that the continuing debate about the alleged relativistic epistemological implications of the sociology of scientific knowledge has no relationship to the findings of empirical research.

However, there may be another issue arising from recent research which ought to be formulated and addressed. Placing the concerns in the local and temporary framework, such as the prerevolutionary social system of Berlin or Paris, or the social system of a "school," makes it possible to reconstruct a kind of local and temporary knowledge state in science—namely, a picture of this state as it looks from the point of view of people deeply involved not only in their research, but perhaps also in politics, personal conflicts, religious commitments, career concerns, etc.

The integration of these local and temporary knowledge states in a history showing the long-term development of scientific traditions is a task still to be accomplished. This paper is an attempt to make such an integration.

It seems, however, that there is a tendency among some sociologists and philosophers of science to deny the legitimacy of such attempts on the ground that they are based on an erroneous reconstruction of social and intellectual history in which developments are viewed from their end point backward, rather than in the sequence in which they actually occurred, when the outcome, at each point, was still unpredictable.

This is an argument which needs to be heeded. But while one has to be aware of the pitfalls of reconstructing the past from the point of view of the present, one must equally be aware of the pitfalls of trying to view it as a kind of biological evolution, which is completely blind as far as the future is concerned, and the outcome of which is determined by some kind of natural selection and survival of the fittest.[3]

The point is that what is being selected out in the social world are social and cultural systems, such as human roles, intellectual contributions, institutions, and cultural traditions, which are not analogous to the biological species, the object of evolutionary selection. Social systems have an ability to modify themselves as well as their environment in quite fundamental ways and are designed to act for the attainment of long-range purposes. Taking into account these long-range purposes of social systems, and paying attention to the social devices—usually norms—by which the long-time (relative) stability of these systems is secured, is not a falsification of the past but the only way of its realistic reconstruction.

Concentrating exclusively on the local and temporary social systems in which scientific activity takes place to the exclusion of the longer-term, normatively regulated academic or other systems, will not disclose a "naturalistic" unbiased view of the social conditions of science, or any other tradition, but a mutilated one. The temporary-local and the long-term normative systems are equally needed in making sense of any social or cultural development.

[3] For an example of such—probably unintentional—evolutionary imagery, see Bourdieu (1981): "The structure of the scientific field at any given moment is defined by the state of power distribution between the protagonists in the struggle (agents or institutions), i.e., by the structure of the distribution of the specific capital, the result of previous struggles which is objectified in institutions and disposition and commands the strategies and objective chances of the different agents or institutions in the present struggles" (p. 267). Elsewhere in the same paper he speaks about science not being "an exception to the fundamental laws of all fields—in particular the law of interest, which is capable of introducing ruthless violence into the most 'disinterested' scientific struggles" (p. 273).

Comment

RANDALL COLLINS

If we go back to the 1920s, we find philosophies of knowledge which for the most part exempted science from the realm of explanation. The first people who pushed into a sociology of science were Marxists in the 1930s: J. D. Bernal in England, Bernard Stern on this side of the Atlantic, and others who had rather strong ideological concerns about why science should not be exempt from being seen as part of a social system and hence molded to social purposes. This line of analysis largely disappeared for some time, but not without leaving a residue, and not without eliciting counterthoughts. Most important, it created a context within which the sociology of science per se was generated, to a large extent the work of Robert Merton. He developed a sociology of science as a social institution in itself, an attempt to map out its culture and its dynamics. This may have been a somewhat heroic and lonely operation for a few decades until about the 1960s when three important contributions were made.

The first and clearly the most famous of them was Thomas Kuhn's book on the structure of scientific revolutions. Kuhn tried to demonstrate that science has a paradigm upheld by a social group that attempts to disregard countervailing evidence as long as it possibly can. In this sense, Kuhn put science back in the same realm as other forms of thought in human activity.

The second contribution (methodologically probably the most influential one) was the publication of two books by Derek de Solla Price. Price attempted to generate what he called the "science of science" by taking the literature of science as something that one could quantitatively measure and attempt to make lawlike generalizations about. A rather prospering industry measuring scientific literature developed from this basis: citation analysis, co-citation chains, and the like.

The third important factor is the work of Joseph Ben-David. I think it is particularly important because it is clearly the most sociological of all these endeavors. Ben-David is concerned with science as a nested and overlapping set of organizations. Scientific disciplines are seen as communities, which fits rather nicely with Price's invisible colleges and perhaps with Kuhn's paradigm groups. But for Ben-David these are organizations nested within other organizations; hence the structure of, say, the

German university system (or, by generalization, the structure of any such interorganizational networks) is absolutely crucial as a setting for what happens to the internal network.

Since the 1960s there has been great growth in all these areas, perhaps most in the work influenced by Kuhn, the demarcationism debate about science and nonscience, and efforts to examine particular fields historically. We've also had extremely close empirical analyses of what actually goes on in scientific laboratories, heavily influenced by the ethnomethodologists, both methodologically and conceptually. This produces what might be called an epistemologically microradical position: What is this person in the laboratory actually doing as compared with the interpretations we make of it? But we also find studies that ignore everything we've learned about the internal structure of science and simply try to correlate ideas of scientists with their social positions in the society outside, like the old Marxist view in the 1930s. The debate continues between internalists and externalists.

I would like to argue that if we look at a nested set of organizations as Ben-David is doing, we don't have pure ideas internally versus social structure externally. Rather, we have an external organizational setting: a larger society with its economic and political structure; then an intervening organizational structure like a university or laboratory which allows some autonomy from the outside; and, finally, within that various groups of specialized intellectuals. These groups are what Mullins calls "theory groups," groups with a discernible history of development of their social networks and typical intellectual sequences of program statements and ideologies. From an intellectual viewpoint, those ideologies and program statements are the idea content of sciences.

My own position on the demarcationism issue is that indeed there *is* something distinctive about science as compared with other social activities. But it is distinctive not because science is not social, but because it is a highly nested set of networks of a particular sort. This gives us the possibility of saying something about the social determinants of scientific ideas without having to reduce them to some gross external form of social structure. Several people—Bourdieu, Karen Knorr, myself, and others—use a theory that scientific ideas are resources or, as Bourdieu calls them, cultural capital that one brings to any given situation of the intellectual field. The scientific actors are attempting to negotiate the best possible situation for themselves using those ideas as capital. Looking at science from this organizational perspective, we have the potential to advance a good deal further in explaining the actual ideas, the content of science, as they develop.

General Discussion

Harrison White: In describing the fierce competition of the new German university system, Ben-David spoke of the fact that one of the things going for the experimentalists was, after all, that they were doing lab work, and he seemed to take it for granted that lab work was more communicable. I found that a fascinating puzzle because in my own experience, lab work is one of the least communicable things in the world. It is not at all easy to replicate experimental work or lab work. I would argue that the advantage of experimentalists is a paradox. You have to self-consciously interdigitate with other people in lab work and come to a commonly agreed position, while theorists can take it for granted that they can read and understand each other.

Joseph Ben-David: I would prefer answering all questions at the end.

Robert Merton: I'm not sure, Joseph, that I have a question; rather, I have some remarks. There are, of course, theoretical questions at issue in the sociology of science, sometimes deep ones. You have identified some of them in your paper as Randall Collins has. I begin by observing that a close local knowledge of the sociology of science and scientific knowledge notes that some theoretical developments have come thick and fast while others have come slowly. One example of a slow development: One calls to mind the fourfold table which the young Stephen Cole developed in the late 1960s and put into further print in the Jonathan and Stephen Cole monograph, *Social Stratification in Science* (1973, p. 3). Across the top are the two categories, internal to the institution of science and external to it; down the side are the two categories of types of influence: social and cognitive (or intellectual). That fourfold table immediately alerts us to the idea that "science" has its own social structure as well as its cognitive structure (scientific knowledge). That insight developed quite slowly over a period of years. Before then, the historiography of science was largely devoted to the development of scientific knowledge—"the history of ideas"—and to "external" social and cultural influences on science. It took some time before scholars recognized that science had a social structure internal to it, a structure that variously affects its cognitive development.

I relate this observation to Randall Collins's remarks. Back in the 1930s, when Blackett, Needham, Huxley, Hogben, and the rest of that noble band of scientists initiated widespread public interest in the social contexts of science, it did not at all occur to them that science has its own internal social structure. Their observations fit into the category of inquiry given over to the external social (principally economic) structure. Interestingly enough, they ignored the other, the fourth box in the fourfold table, the intellectual currents external to science which, in principle, can variously influence the character and course of scientific knowledge.

What can be described as the great debates during the past decade or so in the sociology of scientific knowledge are fundamentally epistemological. My position is congruent with that of Ben-David's, namely, that the sociology of knowledge is epistemologically neutral. The interplay between ideas in the philosophy of science and in the sociology of science is reinstituting a major, direct, and fundamental concern with epistemological questions. I happen to believe that these constitute the Achilles' heel of sociological inquiry into the workings and development of science. This, of course, is a subject requiring and deserving much discussion.

A third set of basic questions properly agitating the rapidly growing specialty of the sociology of science relates to what has come to be known as the problem of the demarcation of science (from other spheres of culture). It is a question implied by Marx and especially by Engels but, to my knowledge, not examined by them in an interesting way. Thomas Kuhn's reintroduction of the generic question about what distinguishes what we call science—in an older terminology, positive science—from other cultural products recapitulates that development in the sociology of knowledge. If this were a conference devoted wholly to this one field, we could profitably pursue that matter at length.

A final remark. This question of the demarcation of science from other kinds of cognitive activity and outcomes is a crucially important theoretical issue. It is a question which Mary Douglas, there across the table, confronts head on. And since I believe that we fundamentally disagree on this, it would be interesting for us to go at it for a time—in, of course, a friendly way. My own position is that science and technology, each in its own way, do indeed separate out from other spheres of culture and cognitive activity, that they do differ in describable, distinctive, and consequential ways. This matter comes to a head in arriving at a position concerning scientific, and technological, progress, the position that one can speak of progress in these domains in a sense in which one cannot speak of it in other departments of culture and society.

It has been a matter of notoriety for a very long time that experimental science has in practice been commonly built on a logical fallacy. Everyone "knows" that. I recall having put that notorious claim into print some forty years ago as then a self-evident fact. As I put it then, "the paradigm of 'proof through prediction' is, of course, logically fallacious," the fallacy of affirming the consequent and inferring the truth of the antecedent. Now how can it be that science has built its imposing cognitive edifice on a long-recognized logical fallacy? Perhaps by establishing necessary *and* sufficient conditions for a predicted outcome and thus avoiding the fallacy? Yet one knows that in principle it cannot be demonstrated that the twin types of conditions have been established. The answer is, I think, that in the actual practice of science, when many predictions are derived from the same theory and are found to obtain empirically, these are taken as the kind of confirmatory or corroborative evidence that is then socially located at the core of what is taken to be canonical knowledge.

In scientific practice, precision takes on special importance. Those predictions which are comparatively precise and precisely confirmed by observation make for acceptance of the underlying theory from which they have been derived. Such precision also reduces the number of rival theories from which the same observations, and others as well, can be logically derived. Logical principle then gives way to the kind of expedience that has always marked empirically oriented science. The conditions making for the batch of precise and preferably diverse confirmed predictions are taken to be necessary and sufficient. The competition among theories is for the time lessened.

I should like to recommend further consideration of this question of the demarcation of scientific from other forms of knowledge at future meetings of this conference on social theory.

Mary Douglas: I come from that no-man's land which you mentioned where no one likes to tread and I have one or two things to say to Mr. Ben-David. Since I completely agree with you that science must be a self-demarcating form of discourse, would you not say that *all* forms of discourse must be demarcating and be bounded and dependent upon a community of people who are in the discourse? Is it unique to science that it's self-demarcating? I would think not.

What I got from your discussion was three points: whether the internal structure of science is competitive and decentralized; whether, externally, there is a market for its products; and that there are probably particularly intricate forms of nesting in science compared with other kinds of discourse. I wonder whether these three are not closely connected. That the self-nesting comes from the market demand; and the market demand itself is an aspect of the competitiveness that is able to be established within the scientific community.

Finally, I always knew that Merton thought we disagreed, but he is so polite that I have never been able to get him to say what he thought we disagreed about. We don't disagree.

Arthur Mann: I speak as a historian, not as a sociologist. But I want to carry something away from these deliberations. If I understood you rightly, Professor Ben-David, you said that in looking at science in the past, sociologists have to know the players and their game and to link both players and their game to what was going on in their time in their local circumstances. Now if that is the way you get knowledge about knowledge, about science, my question is this: Where is the difference between that method and the historical method?

Joseph Ben-David: I agree with so much that has been said, I don't have anything to debate with Randall Collins or with Bob Merton. Now there was an interesting question by Harrison White about a puzzle. How does laboratory work, which is so difficult to understand and no one really repeats it, lead to fairly consensual evaluation? There is a manuscript, relevant to this question, by Susan Cozzens, from NSF. She compared the 1966 paper that Collins and I wrote, on the beginnings of experimental psychology, which happens to be an often quoted paper, with a similarly often quoted paper in chemistry. What emerges is devastating but instructive. The way the chemists quoted their paper was consistent. They related to very specific points with which they agreed and disagreed, which they modified, etc. The way our paper was quoted, only God could have made sense of it. It was completely random and usually irrelevant to the main point we wished to make. It seems that the communicative advantage of the experimental method, and of course the same applies to mathematical work, is that you have to specify very precisely what you are investigating, what exactly your question is, and you have to do this operationally. This limits the range of possible interpretations.

Concerning the questions of Mary Douglas, I wish we had a whole morning, as Robert Merton suggested. They are important questions. I certainly agree that there are other self-demarcating systems and that it is very important to see the parallels between science and other cognitive systems (something that I and some of my colleagues did not pay sufficient attention to until lately). For example, there are wonderful papers by Robin Horton on African witchcraft which I should have considered much more carefully. Still, as both Douglas and Horton have pointed out, science demarcates itself in a different way than other cognitive systems or "forms of discourse."

The last question was by Arthur Mann. Who said there was any difference between history and sociology? But there is sometimes a division of labor between them. Historians usually do more descriptive work. They concentrate on clarifying the facts which need to be clarified and make sense of local-temporal systems. On the other hand, sociologists construct conceptual frameworks which can be very useful for historiography. Research in history and sociology is complementary and frequently overlapping.

Network Theory

Social Network Theory

EDWARD O. LAUMANN
AND DAVID KNOKE

Our task here is threefold. First, we shall quickly adumbrate the particular sort of social structural analysis with which we and our associates have been concerned over the past ten years or so. We will then step back to indicate some of the central theoretical issues that have been plaguing us as we set about applying this theoretical perspective to empirical problems. To avoid confusion, we shall restrict the term "network analysis" to the emerging set of analytic techniques developed to examine network data—i.e., data typically in the form of a square matrix of rows and corresponding columns referring to a set of actors, with entries in the cells indicating the presence or absence of some relational tie between every possible pair of actors (see Marsden & Laumann 1984). Finally, we shall take a particular problem, issue linkage and the communication structure among influential corporate actors in a national policy domain, to illustrate the intricate interplay of theory and substance in motivating the sorts of questions structural theory and network analysis can and should address. We shall try to show how these questions depart in crucial and promising ways from those conventionally raised in policy analysis.

Principles of Structural Analysis

First, what do we mean by social structural analysis?[1] Structure and various descriptive terms related to structure—such as hierarchy, dominance, structural differentiation, structural change, power, and market or class structure—are probably the most popular concepts in the sociological lexicon. Despite the many differences in nuance that various authors associate with the term "structure," the root meaning refers to a persisting order or pattern of relationships among some units of sociological analysis, be they natural persons, corporate actors (Coleman 1973), collections of actors (Laumann 1973), or even behavioral patterns (cf. Nadel 1957, pp. 1–19; Mayhew 1971; Blau 1974, 1975). The apparent consensus in the usage of the term masks the unfortunate fact that there is little agreement on the concepts and the methodology in terms of which one measures or, perhaps more modestly, describes given "social structures."

In our approach to analyzing a social system, we first identify the individual *actor* (be it a natural person, a corporate actor, or a set of actors) *in* a particular kind of *social position* in that social system (cf. Parsons 1951). It is important to bear in mind the distinction between *incumbent* and *social position* inasmuch as an incumbent of a given social position may simultaneously occupy many other positions.

A *social structure* will be defined as a persisting pattern of social relationships among social positions (see Laumann 1966). A *social relationship* is any direct or indirect linkage between incumbents of different social positions that involves mutual, but not necessarily symmetric, orientations of a positive, neutral, or negative affectual character and/or may involve the exchange of goods, services, commands, or information (see Homans 1950; Parsons 1951; Blau 1964; Marsden & Laumann 1977; Heinz & Laumann 1982; Knoke & Laumann 1983). The unit of structural analysis is, then, the specific relationship obtaining between any pair of actors, as defined earlier. The absence of the specified relationship between a pair is as theoretically important as its presence (cf. Lorrain & White 1971; White, Boorman, & Breiger 1976).

We adopt three postulates in our approach to structural analysis. These

[1] See Laumann and Pappi (1976, chaps. 1 and 13) for an extended discussion of this question. The following is drawn from their discussion on pp. 5–9.

principles are stated as assumptions that do not readily lend themselves to direct empirical tests. Rather, we see them as basic tenets of structural analysis that underlie an extensive array of substantive applications—from small group interactions to the world system of nations.

> Postulate I (relationship-specific structures) asserts that there exists a multiplicity of social structures in any complex social system that arises out of the many possible types of social relationships linking positions to one another.

Positions that are "close together" because they have close links with respect to a given type of social relationship (e.g., friendship, business ties, neighborhood, or information exchange on community affairs) can be "far apart" when a different relationship is considered. In short, relationship-specific structures arrange social positions in quite different patterns of association, and there is no inherent theoretical reason to suppose that these different patterns are to be regarded as more or less adequate approximations of *the* "true" underlying social structure. On the contrary, each structure may be thought of as reflecting its own logic of social and functional constraint. Note that this postulate does not preclude the formulation of a theory of structural priority, which would hold, as in certain Marxian formulations, that certain types of relationship structures are more fundamental (in the sense of being formative) than others.

> Postulate II (distance-generating mechanism) asserts that, for any given relationship-specific structure, there exists a principle of systematic bias in channeling the formation of (or making more likely the) relationships between certain kinds of positions and the avoidance of such relationships between others.

In other words, we assume that relationships among social positions are usually not formed on a strictly random or chance basis, but rather in accord with some principle of the differentiation among positions. When discussing the distance-generating mechanism for social structures of intimate association, for example, Laumann (1973, p. 5) argued: "Similarities in status, attitudes, beliefs, and behavior facilitate the formation of intimate (or consensual) relationships among incumbents of social positions." And, conversely: "The more dissimilar two positions are in status, attitudes, beliefs and behavior of their incumbents, the less likely the formation of intimate (or consensual) relationships and, consequently, the 'farther away' they are from one another in the structure." He was then able to interpret the relative proximities among various ethnoreligious groups in Detroit (Laumann 1973, chap. 3) and among various jointly defined religious and occupational groups in Jülich, West Germany (Laumann & Pappi 1976, chap. 3), by pointing to their differentiation in relative social prestige, socioeconomic status, and commitment to religiously linked values and activities.

Whatever the distance-generating mechanism among social positions, be it subjectively perceived prestige differences, inequalities in the possession of resources, differentiated mutual interdependencies, or what have you, it is what will be meant by the *principle of organization* of the social structure under analysis.

> Finally, Postulate III (structural contradictions) holds that, given two or more relationship-specific structures predicated on different principles of organization, among the possible features of such a complex social system are both structural contradictions and structural crystallizations.

A structural contradiction exists when the relative proximities of social positions in a given relationship-specific social structure are negatively correlated or simply unrelated to those of another structure. To the extent that several functionally significant social structures result in similar arrangements of positions (i.e., to the extent that proximities of positions in various social structures are positively correlated), one may speak of the social system in question as being *structurally crystallized.*

Structural crystallization may have serious *negative* implications for the overall functional integration of the system, insofar as all the various relationship structures tend to place the same positions either close together or far apart. A structurally crystallized system tends to be characterized by groupings of similarly circumstanced positions with respect to their patterns of forming various relationships, with the groupings themselves located at some distance from one another. We can readily see how the postulate of structural contradictions may be linked to the classic discussions in political sociology of cross-cutting versus reinforcing societal cleavages and their divergent impacts on voting decisions (cf. Lazarsfeld et al. 1944; Lipset 1963a, 1963b; Sheingold 1973).

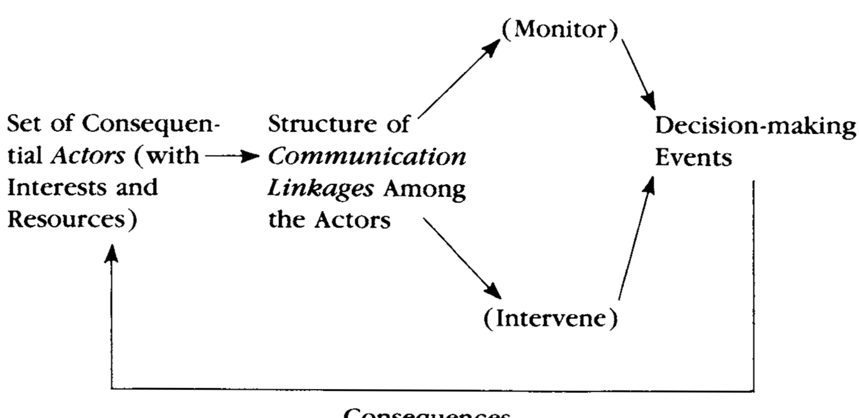
We shall argue that the structural differentiation of large-scale complex social systems has two fundamental implications for the integrative problems that such systems confront. First, structural differentiation is the basis of the objective differentiation of interests—i.e., claims for scarce goods, services, and facilities that component actors make on the larger social system, and for the differentiation of means (or relative power) by which they assert these claims with greater or lesser effect.

Second, structural differentiation is also likely to lead to the differentiation of evaluative standards (values) that are used by various system elements to specify and establish the priorities among competing ends or goals that the system as a whole should collectively seek to achieve.

Structural Analysis in Practice

We would now like to apply this discussion of social structural theory and network analysis to a detailed empirical illustration of issue linkage and communication structure in national policy domains. We refer the reader to other papers which describe the rationale and methods by which we identified and interviewed the *population* of consequential corporate actors, including such entities as the American Medical Association and the Director's Office of the National Institutes of Health, active in formulating national health policy (Laumann, Marsden, & Prensky 1983; Knoke & Laumann 1982; Laumann & Knoke 1983; Knoke & Laumann 1983; Laumann, Knoke, & Kim 1985).

The orienting framework motivating the entire analysis is that there is a *set of consequential actors,* possessing profiles of variable *interests* in a range of national health issues and relevant mobilizable *resources,* embedded in a particular *structure of communication linkages.* This flow of specialized communication among the actors enables them to monitor and to communicate their intentions in relevant decision-making events that, in turn, have consequences for their interests. Diagrammatically, we represent the model as follows:



Using this general framework, we shall first discuss the more general substantive problem of issue linkage. To illustrate the approach, we shall then present some results from an analysis of the issue structure of national health policy in the 1970s. After that, we shall briefly characterize the structure in which the same set of actors are tied to one another by routine communication linkages and then examine the extent to which the structure of issue linkage "explains" the structure of routine communication linkages. To complete the analysis according to our diagrammatic model, we should, of course, also discuss how an organization's location in the issue linkage and communication structures jointly determine its activation on decision-making events related to national health policy. While this analysis has been completed with quite satisfactory results, we cannot report it in detail here because of space limitations (see Laumann, Knoke, and Kim, 1985).

Issue Linkage in Policy Domains

At least three general features of issues can be identified that link issues to one another. First, every issue possesses an inherent *substantive or conceptual content* that may link it to certain other issues on grounds of logical implications that would be generally recognized and accepted by most observers. For example, an issue regarding the regulation of experimentation on human subjects is logically entailed in the more inclusive issue of the organization of biomedical research, should it arise as an issue at all. Second, every issue eventually comes to occupy an *institutional locus of relevant decision-makers* (with its attendant audience) in which it is typically debated and resolved. Thus, the organization of biomedical research is the mandated responsibility principally (but not exclusively) of the National Institutes of Health. And third, the resolution of a particular issue generates a potential or actual flow of harms and/or benefits to an *identifiable set of politically active or inactive individuals and corporate actors,* who are likely to be or become concerned about this flow. In this case linkage between two issues arises because the recipients of the harms or benefits perceive that the ways in which both issues are resolved will have similar or contradictory beneficial or disadvantageous effects on them. It is precisely the nature of this socially organized concatenation of diverse other interests around the focal issue of concern among the interested parties that provides the critical driving mechanism linking issues into diversely organized complexes of issues in the policy domain. In other words, issue linkage arises from the ways in which the "natural" constituencies of the several issues are intersected.

In the remainder of this paper, we demonstrate the value of analyzing issue linkages from the third perspective. We shall examine in some detail the issue structure arising from the overlapping subsets of corporate actors concerned about the political issues of the national health policy domain in the 1970s.

An initial list of issues with policy import during the 1970s was compiled from various popular press and academic sources and subsequently consolidated by absorbing minor or narrowly focused topics into broader categories. This preliminary issue list was sent to a panel of about a dozen academic and journalistic specialists who were asked to nominate additional issues that had been overlooked. The final set of 56 health issues took their suggestions into account.

The set of issues was then presented to informants who, taken together, represented the population of corporate actors relevant to health issues. The interviewer presented a booklet displaying the list of domain issues, along with the instruction to place a checkmark in front of each issue in which the informant's organization had an interest, then to indicate the level of interest on a five-point scale ranging from "minor" to "moderate" to "major interest."

Table 1 displays the full set of domain issues presented to informants, clustered on the basis of interests expressed by informants, according to a hierarchical clustering algorithm described below. These clusters are presented in descending order of mean organizational interest (using a 0–5 scale). The second column in Table 1 presents the percentage of all 133 organizations expressing moderate-to-major interest in each issue. The 56 health issues agglomerated into 16 distinctive clusters, with the funding of medical research, the containment of health care costs, and disease prevention-control issues attracting the greatest organizational attention. Note that no set of issues attracted more than "moderate" average concern from the entire set of organizations, while the bulk of issues were rated only slightly higher than "of minor interest." Intense interest tended to be concentrated within small, specialized subsets of organizations, a fact having important structural implications.

Now imagine a matrix in which the 133 respondent organizations are listed in the rows and the columns refer to the 56 issues. The entries in row *i* are the 0–5 scores of organization *i* for each of the 56 issues in which it could express a level of interest. The structure of this organizations-by-issues matrix can be examined from two fundamentally different points of view—namely, from the point of view of the linkages among the issues (i.e., correlations between the columns) or from that of the organizations' profiles of interest (i.e., correlations between the rows). Thus, the matrix provides an operationalization of the third specification of issue linkages—that is, the commonality of issues interests across all the corporate actors at the core of the policy domain. To the extent, however, that the actors themselves operate with some notion of the

Table 1 Issue Clusters in the Health Domain

Cluster	Issues	Mean	%
1.	Medical Funding Issues	3.15	67.2
	Medicare/Medicaid reimbursements for health personnel	2.89	62.4
	Medicare/Medicaid funding	3.29	69.2
	National health insurance	3.28	69.9
2.	Health Costs Issues	2.56	54.9
	Certificates of need	2.14	44.4
	Health maintenance organizations (HMOs)	2.49	55.6
	Health planning	3.12	67.7
	Hospital cost containment	2.79	57.9
	Professional standards review organizations (PSROs)	2.26	48.9
3.	Prevention and Control Issues	2.54	58.7
	Disease control programs	2.08	47.4
	Preventive care	2.99	69.9
4.	Biomedical Research Issues	2.21	46.1
	Biomedical research programs, e.g., targeted disease funding	2.85	61.9
	National Institutes of Health funding	2.83	56.4
	DNA research	1.29	27.8
	Human experimentation	1.85	38.3
5.	Elderly Care Issues	2.04	44.8
	Home health programs	2.13	46.6
	Nursing homes	1.95	42.9
6.	Health Professionals Education Issues	2.00	42.9
	Allied health professional training	2.20	46.6
	Health Education	2.65	56.4
	Federal aid to nursing schools	1.41	27.8
	Nurse training	1.52	31.6
	Health professional certification	2.08	45.1

Table 1 (Continued)

Cluster	*Issues*	*Mean*	%
	Federal aid to medical schools	1.85	40.6
	Physician training	2.27	51.9
7.	Mental Health and Drug Abuse Issues	1.97	43.1
	Community mental health centers	1.79	37.6
	Mental health	2.23	49.6
	Drug/alcohol abuse programs	1.89	42.1
8.	Environmental and Occupational Health	1.82	39.9
	Environmental health	1.80	41.4
	Occupational health and safety	1.84	38.3
9.	Public Health Issues	1.77	37.5
	Community health centers	1.82	39.1
	Rural health care	2.03	44.4
	Migrant health programs	1.17	22.6
	Public health services	2.25	48.9
	Health professional redistribution programs	1.99	43.6
	National Health Service Corps	1.71	35.3
	Emergency health care programs	1.38	27.8
	Federal aid to hospitals	1.81	38.3
10.	Professional Abuse Issues	1.65	33.5
	Medical malpractice insurance	1.61	34.6
	Medicare/Medicaid fraud	1.68	32.3
11.	Disability Issue	1.59	35.3
	Developmental disabilities programs	1.59	35.3
12.	Miscellaneous Issues	1.57	33.9
	Indian health care	1.20	25.6
	Medical records privacy	1.93	42.1
	Abortion	1.24	24.8
	Maternal and child health insurance	1.89	42.9

Table 1 (Continued)

Cluster	Issues	Mean	%
13.	Drug and Device Regulation Issues	1.56	33.0
	Drug development regulation	1.83	39.1
	Drug industry regulation	1.68	36.1
	Drug labeling regulation, e.g., expiration date labeling, safety brochure inserts	1.82	39.8
	Generic versus brand-name drugs	1.46	30.8
	Medical advertising	1.05	20.3
	Medical device regulation	1.53	31.6
14.	Kidney Dialysis Issue	1.37	27.1
	Kidney dialysis	1.37	27.1
15.	Veterans and Unions Issues	1.11	21.1
	Hospital housestaff unionization	0.92	17.3
	Veterans health care	1.30	24.8
16.	Food Additives Issues	0.84	14.1
	Cyclamates	0.68	10.5
	Saccharin	0.90	15.0
	Food additive regulation	1.13	21.1
	Laetrile	0.63	9.8

substantive logic of linkage among issues and are themselves oriented to specific institutional loci for influencing decision-making, the other two bases of issue linkage must inevitably be reflected in the constituency-driven structure of issue linkages.

From the vantage point of the linkage among issues, we can determine the extent to which the levels of interest in issue j covary with the levels of interest in issue k across the set of consequential organizations. The greater the commonality of concern two issues elicit across the set of organizations—whatever its basis in underlying substantive, institutional, or sociopolitical similarities—the higher the correlation. A negative correlation between two issues, on the other hand, suggests that intense levels of interest in issue j are associated with an absence of interest in issue *k*, and vice versa. In general, we should expect positive correlations among the issues because they are all drawn from the same policy domain.

The range of the 1,540 correlations among the 56 issues was between +.829 and −.284, but only 6.2 percent of the correlations were negative in sign. Nearly all of these negatively signed correlations were observed between the four issues concerning biomedical research and the other 52 issues in the domain—indicating the highly distinctive character of the interested parties concerned with these issues in contrast to the rest of the policy domain. The fact that nearly all the correlations among the remaining issues are positive suggests that the issues confronting the policy domains enjoy a degree of mutual coherence and interrelationship in the concerns of the active participants.

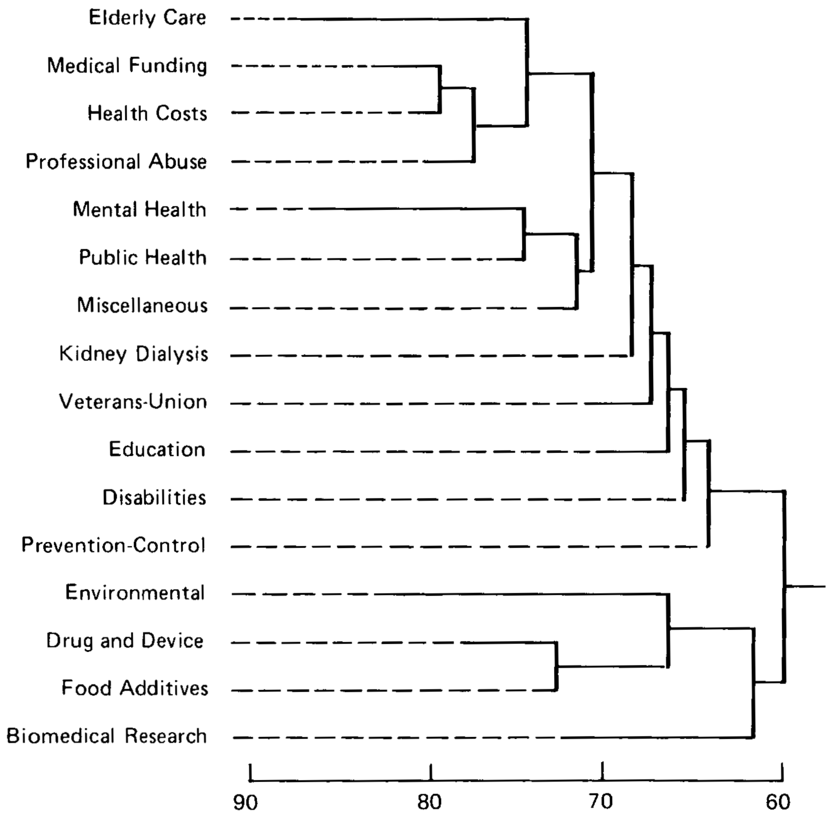
This matrix was then submitted to hierarchical cluster analysis to locate distinct subsets of issues based on similar patterns of organizational interests. (See Prensky 1985, chap. 3, for an extended discussion of these procedures and a related data analysis.) Hierarchical clustering produces nested subsets with increasing internal heterogeneity as fewer but larger clusters are identified (see Bailey 1974 for a useful overview of clustering methods). We used an agglomerative approach that locates pairs of highly similar issues (based on the patterns of interests expressed by all organizations), then successively adds other issues or clusters of issues to these core clusters, based on the average similarity to the variables already in the initial clusters.

Figure 1 is a dendrogram or tree diagram that reveals how the basic issue clusters would merge if further hierarchical clustering was carried out.[2] Although hierarchical clustering suggests which sets of issues are adjacent to one another, in the sense of attracting similar organizational audiences, it fails to convey the overall pattern of proximity and distance among types of issues. We next turn to a more adequate depiction of the overall issue linkage structure.

The correlation coefficient for a pair of issues can be treated as a relative measure of their similarity. When the full matrix of issue-by-issue correlation coefficients is submitted to a multidimensional scaling routine (Kruskal & Wish 1978), the result is a spatial diagram showing the location of domain issues in an issue space. Sets of issues with high similarity of attracted interest will be located close to each other, while

[2]The numerical values for the branches represent a rescaling of the average correlation coefficients within sets of merged clusters, with an $r = 1.00$ scaled to 100 and an $r = 0.0$ scaled to .50.

Figure 1 Dendrogram of Health Domain Issue Clusters

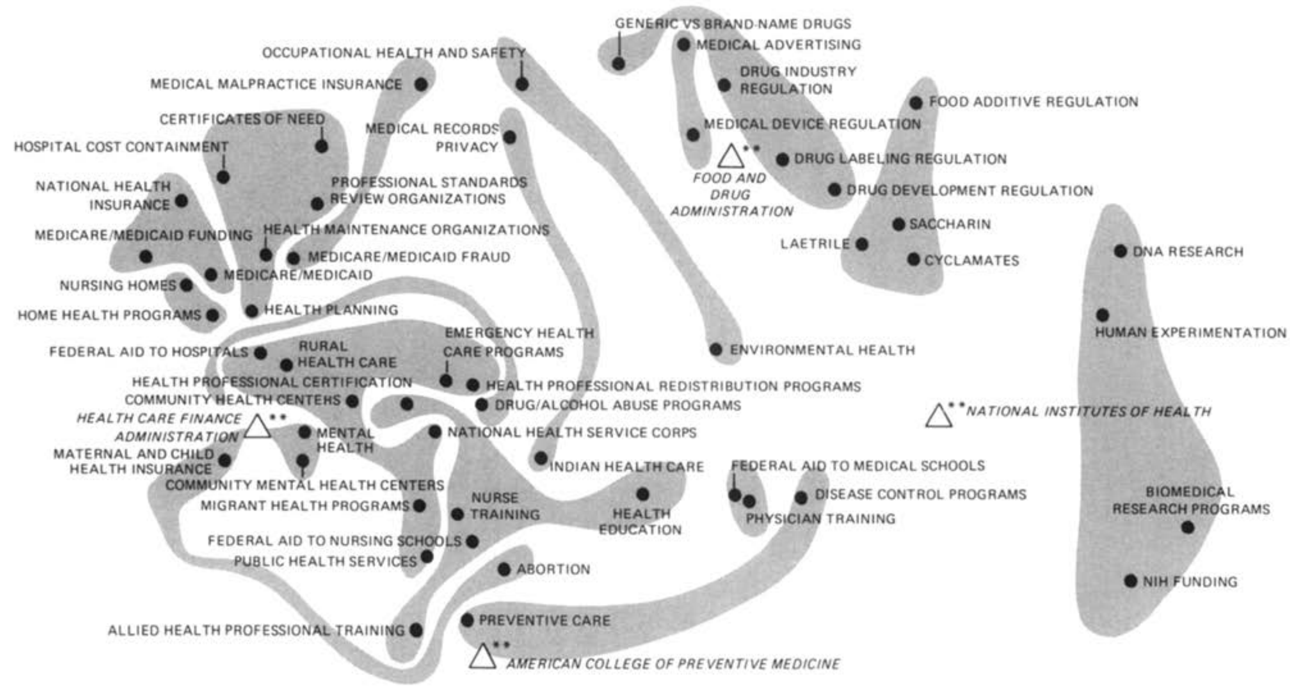


issues appealing to divergent organizational interests will be plotted far apart. Figure 2 displays the two-dimensional solution.[3]

The most prominent feature of the picture is the circle or rim pattern in the distribution of issues. That is, no issues occupy the center of the issue space. For the most part, the subsets of issues composing each issue cluster (determined from the hierarchical clustering analysis) are located adjacent to one another, as shown by the closed lines drawn

[3]Multidimensional scalings were performed with the ALSCAL program of Young in the SAS package. The Kruskal stress values were .25 for the two-dimensional solution and .17 for the three-dimensional solution. Although the three-dimensional analysis provided a better fit, the two-dimensional pattern is easier to present and does not differ in its gross patterns from the more complex plot.

*Figure 2 Two-dimensional Picture of the Issue Structure of the National Health Policy Domain for the 1970s**



*Issues belonging to the same issue cluster in the hierarchical clustering analysis (displayed in Table 1) are indicated by shaded areas.

**Triangle denotes a marker organization to be used in aligning Figures 2 and 3.

around them. (In the few cases where interpenetrating clusters occur, the subsets would be separated when a third dimension is fitted.) The circular configuration occurs because there are no central issues which manage to attract the interest of all domain actors in such a fashion that the central issue attains comparable covariation with all the other domain issues.

Figure 2 reveals that the greatest issue distance on the horizontal axis occurs between such financial issues as Medicare/Medicaid funding and health care costs on the left side of the figure and biomedical research issues at the extreme right side. The vertical dimension distinguishes issue clusters associated with governmental regulation (drug and device, food additives, and environmental-occupational health) at the top end from professional manpower and public health matters (professional education, disease prevention-control, and some public health issues) at the bottom of the figure. In fact, it might be useful to conceive of the elliptical arrangement of the health issues as "hung" onto four institutional foci, each located in one of the four quadrants of the space, where the organizational dimension of the issues-by-organizations relationship is highlighted. In the upper right-hand quadrant, one finds the issues of particular relevance to the Food and Drug Administration; while in the lower right-hand quadrant, roughly halfway between the extreme outlyers of biomedical research and physician training might be located the National Institutes of Health. In the middle of the left-hand side one can locate the Health Care Finance Administration and the various congressional subcommittees concerned with funding the health care system, while in the lower left-hand quadrant is located the American College of Preventive Medicine. Such an interpretation gives some credence to the notion that a key principle of organization of the issue space is the institutional locus in terms of which policy options are selected and implemented.

A central contention of our analysis is that issues in close proximity to one another attract a common *public* of interested organizations whose specific interests and mobilized activities to influence significant outcomes for themselves help to define in fundamental ways the very nature of the debate about how specific policy options should be defined and selected (see Knoke & Laumann 1983 for an extended discussion of issue publics). That is, terms of the debate are determined in good measure by the character of the organizations which constitute the relevant mobilized public. To illustrate the argument, the reader's attention is directed to the location of the abortion issue in the lower left-hand quadrant of the issue space where a variety of issues having to do with the provision of public health services, medical care for the disadvantaged, and mental health issues are also found. There is no a priori or inherent substantive logic one could use to deduce that "abortion"—a major ethical and moral issue of the human condition—should "obviously" belong in such an arena. Its location arises because the interested parties happened to link the general issue of abortion to efforts to establish a prohibition to funding abortions at community health clinics and thus mobilized the concerns of the substantial "welfare lobby" who are concerned with these issues more generally.[4]

The Structure of Organizational Interests: Purposive Action Among Organizations

Let us now turn to an examination of the matrix from the perspective not of issue linkage, but of the organizations' profiles of interest. Here we are interested in seeing how *organizations* come to occupy close or distant proximities from one another as defined by their relative similarities or dissimilarities in their profiles of interest across the 56 issues. It is useful to think of a given organization as having a portfolio of investment decisions to make on how it proposes to spend its scarce resources in influencing policy choices of interest to itself. Some organizations have a broad mandate to be concerned with the overall character of national health policy and must therefore monitor a broad and rather unselective range of concerns across the diverse issues up for resolution on the national agenda. Others, in contrast, may have a much more narrowly focused set of concerns.

We propose to treat each organization's profile of expressed interests in the full set of 56 health issues as a key variable in the analysis—namely, its most general purposive orientation or goal objectives within the na-

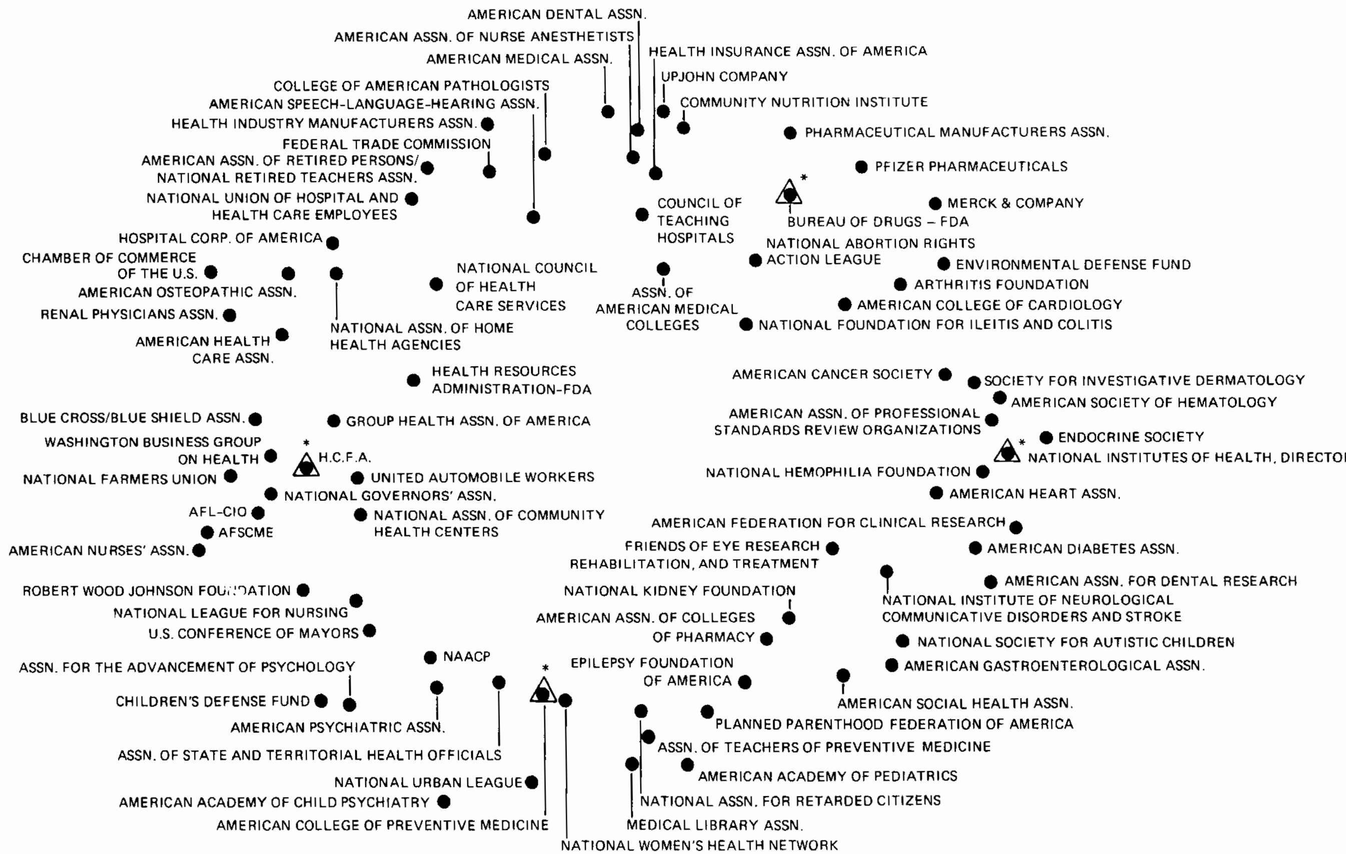
[4]While we have no direct evidence on the location of the issue "the legal definition of death as cessation of brain functioning" because we did not include it in our list of issues, there is little doubt that the issue of brain death would have been located in the lower right-hand quadrant in the arena of biomedical research and NIH funding issues. In an excellent case study, Leslie Ann Rado (1982) identified the principal interested parties on this issue as being drawn almost exclusively from high-status medical specialty groups and academia. The nature and course of this controversy in the 1970s thus followed a very different path from that of abortion, which, at least philosophically and substantively speaking, has to do with very similar questions about the meaning of human life and death. In short, the "affected constituency" principle of issue linkage appears to provide a more compelling account of the markedly different politics associated with these two issues than the principle of inherent substantive logic linking the issues would have suggested.

tional health policy domain. By restricting attention to interests in issues regardless of intensity of outcome preferences, organizations having divergent policy outcome preferences but with shared foci of concern will be located in close proximity. Of course, many health domain participants—such as the U.S. Chamber of Commerce and the AFL-CIO—have heavy investments in issues located in other policy domains (e.g., education, civil rights, the domestic economy). To the extent that such generalist organizations allocate substantial efforts to other arenas, we may err in exaggerating their capacities to pursue objectives within the health policy domain. However, for analytic as well as empirical simplicity, we disregard any issue interests of an organization that may lie outside the domain under study.

We shall confine attention to proximities as a function of organizational similarities and dissimilarities in interest profiles. When the matrix of 8,778 organization-by-organization correlation coefficients is submitted to a multidimensional scaling routine, the result is a spatial diagram showing the locations of domain organizations ordered according to the organizations' differential levels of joint concerns and apathies about 56 health issues.[5] For simplicity of presentation, we display the two-dimensional solution in Figure 3, which clearly arrays the organizations in a circular configuration. The key structural implication to draw from such a pattern of organizational proximities is that no actors are located in the central position, equidistant from all other organizations. A central position would be occupied only if one or more organizations expressed substantial concerns across the full range of issues, although not necessarily about every issue on the list. In short, the policy domain is organized in a peripheral pattern according to specialized organizational interests that shade gradually from one issue concern to the next in an ordered way. The Office of Management and the Budget, the American Medical Association, and the Secretary of Health and Human Services, all of whom might have been expected to have strongly generalized interests running the gamut of national health policy concerns, are located in the upper rim of the figure, reflecting, as we shall see, their predominant concerns for containing the costs of health care to the

[5]As before, higher dimensionality provides a better fit to the data but is less easy to display and interpret than the two-dimensional configuration. The Kruskal stress for the three-dimensional solution was .19 and rose to .28 for the two-dimensional solution. While the stress—a measure of goodness of fit—for the two-dimensional solution is quite high, it is apparent from examining the three-dimensional solution that the overall structure is one of a sphere or hypersphere with an empty center.

Figure 3 ***Two-dimensional Picture of the Structure of Organizational Proximities According to Their Levels of Similarities in Profiles of Issue Interests***



*Triangle denotes a marker organization to be used in aligning Figures 2 and 3.

government. As we all have suspected for some time, there is no idealized rational actor who is coordinating the whole enterprise from the Olympian heights of the center. Rather, the more appropriate metaphor is of bit players saying their lines who are recruited for a number of one-act plays written by different playwrights.

But one can detect some order in the cacophony. The horizontal axis broadly distinguishes between the organizations of full-time medical specialties and their consorts, the drug companies, on the right-hand side from the lay consumer (potential and actual patient) groups on the left. The vertical axis roughly sorts the lay and technically specialized organizations concerned with hospital and medical care, rather broadly conceived and generally "establishment," at the top from mental health, welfare, and community and public health organizations—the less "established" and reputable—at the bottom.

A provocative exercise for the interested reader is to superimpose the issue space depicted in Figure 2 on the organization space of Figure 3. To aid the reader in doing this, we have marked four organizations, Food and Drug Administration (Drugs), the National Institutes of Health (Director's Office), the American College of Preventive Medicine, and the Health Care Financing Administration (HCFA), with triangles in Figure 3. Align these with the corresponding four triangles located among the issues relevant to these organizations in Figure 2. One can then note a rough correspondence between the two configurations which together jointly define the circular arrangement of issue publics, consisting of continually changing constituencies of organizations concerned with particular issues as one proceeds around the two circles.[6]

The Communication Structure in the National Health Policy Domain

A crucial assumption motivating the entire analysis is that the flow of information in a policy domain plays the critical role in orchestrating the activities of the various participants and in determining the outcomes of various policy deliberations (cf. Deutsch 1966; Allison 1971; Pool 1983). We therefore expect the communication structure of the policy domain to be related to the organizational purposes of the participant actors, as measured above. It is to the examination of the health domain's routine

[6]See Knoke and Laumann (1983) for an extended and more formal treatment of the notion of issue publics in two national policy domains, health and energy.

communication structure and its relation to its issue structure that we now turn.[7]

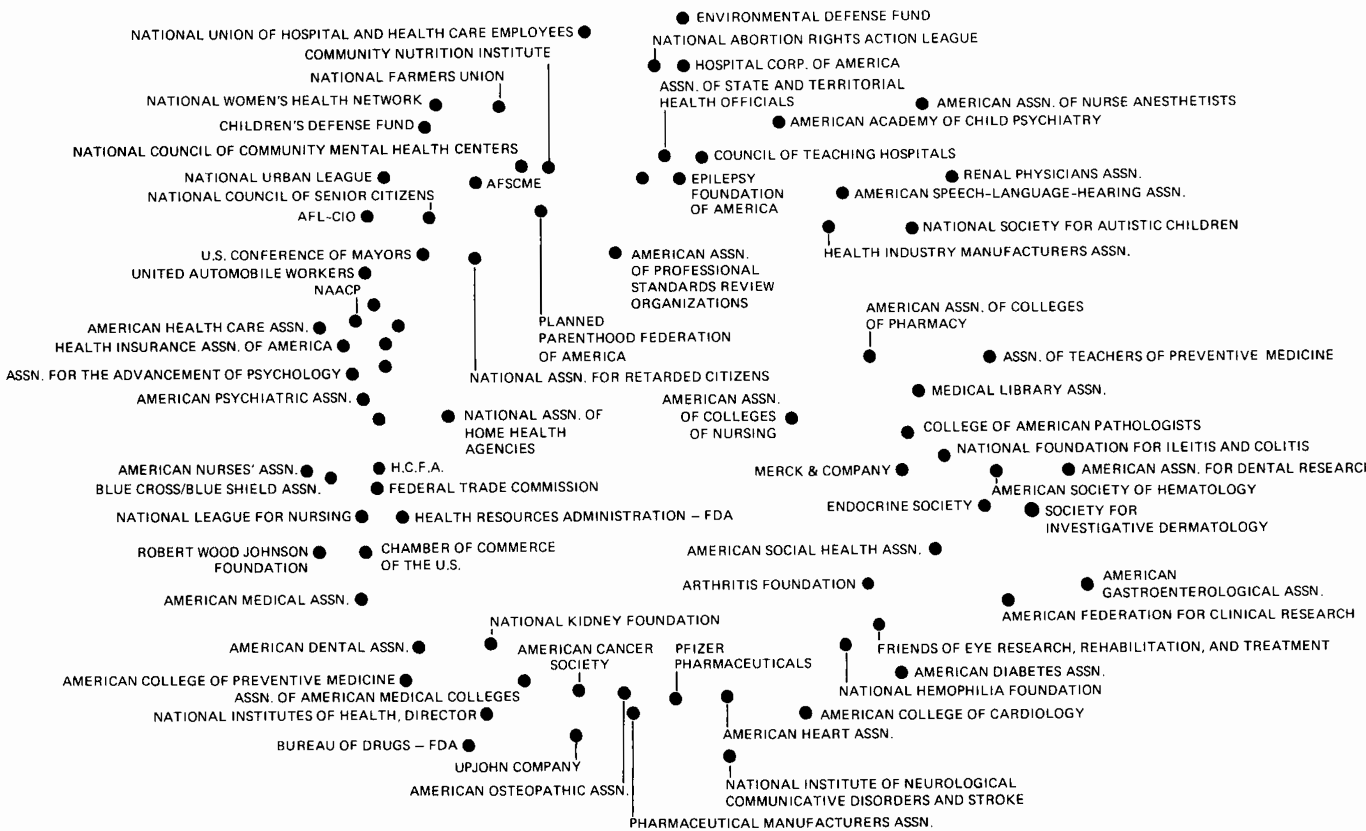
Each organization in the policy domain was asked to name all the organizations in the domain with which it "regularly and routinely discusses national health policy matters." We then asked whether the respondent organization usually initiated the discussion, or whether the other organization did, or whether discussions were initiated as often by one as the other. A rectangular binary matrix (respondent organizations in the rows and targets in the columns) was constructed.[8] Of the 22,005 potential linkages in this matrix (135 interviewed organizations × 164 target organizations), 7,157 communication ties were observed, yielding a density of 32.5 percent. Such a density is apparently typical for interor-

[7]There are many alternative ways in which information exchange in a policy domain could be organized. The most likely structure is what we and our associates call the information-brokerage model. As Prensky (1985) observes: ". . . the uneven distribution of information will cause organizations to form links with other organizations that can more easily gather such information. Also, organizations will form links so that they can send policy information to other organizations that they wish to have that information. . . . [S]ome organizations are specialists, with interests in only selected areas of the domain. An area of the domain is a collection of issues that are viewed by the organizations as being similar. For example, the domain members might view funding issues as one area and research issues as another. Such specialists would not maintain information flows with other specialist organizations in areas of no interest to them. Other organizations are generalists, interested in many areas of the domain. They communicate with specialists in their many areas of interest. Indeed, the generalists serve as brokers of information (Laumann et al. 1980), gathering information from specialists in a particular area of the domain, processing it for its relevance for other areas in which the generalist is interested, and then sending it to specialist organizations in those other interested areas. Thus, specialist organizations are linked to specialists in other interest areas by two-step information flows through generalist organizations."

The implications of such a model are several. First, the patterns of communication among organizations occupying a structurally equivalent interest position (i.e., those assigned to the same issue cluster in our hierarchical clustering analysis above) will be more similar than those among organizations in different issue clusters. Second, greater overlap in interest areas between two issue clusters (i.e., adjacent issue clusters in the rim arrangement depicted in Figure 3) will motivate a higher density of inter-cluster communication. Finally, there is a two-step information flow through generalist organizations, between specialist positions that have no areas of interest overlap. Prensky (1983) reports findings generally supportive of these propositions. We shall describe below a more rough-and-ready examination of the data pointing to similar conclusions.

[8]See Knoke and Laumann (1983) for more detail on the methodological treatment of the data.

Figure 4 Two-dimensional Picture of the Structure of Routine Communication Ties Among Consequential Organizations in National Health Policy



ganizational communication networks (e.g., Galaskiewicz 1979; Knoke & Laumann 1983). The mean number of initiated discussion contacts is 53.0, while the mean number of received discussion contacts is 26.6, necessarily lower because of the rectangular structure of the matrix. No organization had fewer than three contacts, indicating overall that the policy domain network is fully connected (in digraph terms), probably in two, or at most three, steps.

As we expected, the flow of policy information is highly unequal among domain actors. Governmental actors generally rank at the top in communication activity, with the Department of Health and Human Services, the American Medical Association (the only private organization being so central in communication activity), and several congressional subcommittees sending and receiving information from 75 percent or more of the domain participants. At the other end of the spectrum, the least connected actors tend to be small, highly specialized organizations or associations with marginal interests in domain issues, such as the National Urban League.

The data matrix was subjected (after transformation; see Knoke & Laumann 1983) to the same multidimensional scaling procedure used for the matrix of interests. Figure 4 depicts the familiar circular configuration, now constructed on the basis of relative similarities among the organizations in their communication profiles. The interested reader might again like to try his or her hand at superimposing Figure 4 on Figure 3, in order to gain some sense of the degree to which the embeddedness of organizations in the issue space corresponds to their embeddedness in the routine communication structure. For a rough but more rigorous indicator of that correspondence, we calculated the product-moment correlation between the interpoint distances in Figure 3 and those in Figure 4, which turned out to be .33, a moderate association. The correlation of the interpoint distances for the three-dimensional solution is .35, a very modest improvement. These correlations are substantively quite large, especially in light of the fact that there are certainly other important factors determining the patterns of organizational communication and issue linkage to be considered.[9]

[9] Think, for example, of enormous differences among organizations in their relative monitoring capacities and other resources relevant to participating in the communication structure. Regardless of the scale and scope of organizational interests in health policy, organizations with very limited resources obviously can maintain a far less elaborate network of communication links than those who control substantial resources to participate in the "Washington scene." General interest organizations like the AFL/CIO and the U.S. Chamber of Commerce have

Conclusion

The foregoing analyses demonstrate a substantial correspondence between the structural configurations of actors concerned with various issues in national health policy and their reported regular-routine communication linkages. Further analysis, not reported in this paper, discloses that the location of organizations in the issue and communication structures have substantial implications for policy domain behavior, including the rank-order reputations of influence among national health policy organizations (Laumann & Knoke 1983), the patterns of organizational activation on specific decision-making events over time (Laumann, Knoke, & Kim 1985), and the organizations' relative successes in achieving their preferred outcomes in specific policy controversies.

This empirical illustration provides just one instance of many potentially insightful applications of the structural method. Our discussion of alternative ways policy domain issues may be linked together demonstrates that a single question may be addressed from a variety of substantive angles. The value of the structural approach depends greatly upon the analyst's skill and sensitivity in framing the inquiry to measure relationships in ways that will illuminate both system behavior and the consequences for individual actors.

At present, researchers can use the tools of structural analysis to construct fairly robust descriptions of empirical social systems, a necessary first step in their efforts to build more powerful, general analytic explanations of such phenomena. Armed with such finely honed, subtly nuanced structural descriptions, we can launch the ambitious task of formulating more penetrating theories of social structure and collective action.

huge staffs devoted to a broad range of policy domains outside of health affairs. Their ties to other organizations, as well as staff strength, permit them to supplement their personnel and efforts on health matters whenever it is deemed necessary. In contrast, a one- or two-person staff can hardly undertake all the activities that might be warranted to further its interests.

Comment

RONALD S. BURT

I serve two purposes here. One is to give you a brief and sympathetic exegesis of the paper. The second is to call your attention to three points indicating how the paper's argument is to my mind wrong-headed—not wrong in fact, but wrong in perspective.

Laumann and Knoke offer four interesting ideas or results. (1) They distinguish ways in which issues are linked through social structure. Issues can be perceived to be substantively similar and so be linked in content. Issues can involve the same participants and so be linked through the people active in them. Issues can have implications for the same constituency and so be linked by their consequences. (2) There is more than a single pattern of interest in issues evident among the leading organizations in American health policy in the 1970s; there is structural differentiation. To the extent that separate issues draw the interest of separate organizations you observe the circle pattern that Laumann and Knoke report with Figure 2. (3) Organizations are differentiated by their interests just as issues are differentiated. The dimensionality of the issue by organization interest matrix is the same by rows or columns, so in Figure 3 you again see the circular distribution of organizations around a rim. (4) Finally, the spatial map of the interorganization communication network (Figure 4) looks very much like the space of organization interest similarities (Figure 3). The correspondence is weak, but nonnegligible, suggesting that structurally proximate organizations have similar interests. The paper ends with the exhortation that the method illustrated for describing social structure makes it possible to accumulate a literature of descriptions from which we can begin the more ambitious task of formulating structural theory.

Given the exegesis, consider three points in critique of the paper. In general, empirically driven theory is difficult in network analysis. Laumann and Knoke quite rightly point to some current problems: excessive formalism, a profusion of symbols, arcane propositions, the difficulty of defining system boundaries, the micro-macro linkage. There are even more serious obstacles to network theory emerging from empirical research. Network content is one obstacle pertinent to this discussion. Network analysis is most successful when applied where at least one kind of relationship is clearly defined: market relations of buying and selling in organization theory, population flows between places and ages in demography, people moving between occupational statuses in mobility studies. Network analysis does not prosper when based on ambiguous relations such as the discussion relations studied here. Distinctions between the presence and absence of relations and distinctions between kinds of relations become unreliable. The stochastic regressors created when network measures of distance and structure are based on such data can be expected to obscure correlations that would be strong and obvious if the data were reliable. Observed correlations will be weak. The weak association that Laumann and Knoke observe between interest similarities and structural proximity is either a result of the association being weak in fact or a result of inconsistencies across respondents identifying interorganizational relations in which national health policy matters were "regularly and routinely discussed." Providing a hazy view of social structure and a bias toward negligible correlations, empirical research describing data on ambiguous relationships is a weak foundation for theoretical statements.

Second, amplifying the first point, network theory's greatest strength makes it especially sensitive to data problems. Network theory builds its explanations from patterns of relations. It captures causal factors in the social structural bedrock of society, bypassing the spuriously significant attributes of people temporarily occupying particular positions in social structure. In contrast, sociology typically builds its explanations from the attributes associated with positions in society: role labels such as mother, father, male, female, worker, manager, old, young. Linton long ago cautioned anthropologists against describing social systems in terms of easily observed statuses without describing the patterns of relations giving each status meaning. Such wisdom is difficult to incorporate in popular research methods. Preserving the distinction between attributes and relationships requires more accurate measurement than is typical in sociology. A person's sex, a person's race, or a corporation's size is more easily observed and recorded than the strength and content of a relationship. It is easier to obtain and record data on the distribution of males, females, blacks, whites, large and small corporations in a study population than to obtain and record data on relations between each kind of actor. Phrasing the issue more succinctly, the marginals in a network are more reliable than its structure. Given the comparatively fine grain measurement that network theory presumes and the suspect reliability of much network data, it seems unlikely that precise, cumulative network theory will emerge from empirical research.[1]

[1] Drawn from a national probability sampling frame, the network data available in the 1985 General Social Survey will go far toward improving the rigor of network theory based on discussion relations and the reliability of data on such relations.

Third, although some decry the formalism of network theory, the fact that network theory lends itself so naturally to formal models is one of its strengths. At minimum, rigorous attention to the aggregation of relations in network concepts clarifies operationalization for empirical research. Laumann and Knoke illustrate the difficulty of relying too heavily on empirical research in concept development. Their analysis is most pertinent to debate in network theory over the mechanisms by which social structure affects social perceptions. Under cohesion, socializing communication is the mechanism. People talk to one another about topics of mutual interest and come to shared understandings, expressing similar opinions on the topics. Under structural equivalence, communication is expanded to include competition as a socializing mechanism. Envy and relative deprivation driven by the desire to perform well in one's roles create homogeneity in the behavior and opinions of people who occupy similar positions in a social system. A detailed comparison of the cohesion and structural equivalence mechanisms is available in the second section of *Toward a Structural Theory of Action.* Laumann and Knoke describe their results in terms of cohesion—similar interests and support for public policies arise from communication between organizations. Appropriately, proximate organizations in Figures 2 and 3 have similar profiles of interests across the fifty-six issues studied. But when the same operationalization is applied to the network of interorganizational discussion relations, it measures structural equivalence, not cohesion. Proximate organizations in Figure 4 have similar patterns of relations with others. They may or may not actually communicate with one another. In other words, the social proximity of organizations demonstrated to predict interest similarity is their structural equivalence, not their tendency to communicate with one another. As have recent studies of structural equivalence in innovation diffusion and elite judgments, Laumann and Knoke's empirical results are evidence of response similarity arising from structural equivalence rather than cohesion, an important theoretical distinction easily lost in empirically driven theory.

Let me reemphasize that none of my comments detracts from the intriguing findings that Laumann and Knoke report here briefly and in detail in their book. Constrained to be uncomfortably brief, I merely remind you of some difficulties in adopting their research strategy as the most effective way to develop social structural theory.

General Discussion

Michael Hechter: This question is equally addressed to Ed Laumann and to Ron Burt. I wonder whether network theory or structural theory in general can ever successfully account for change—that is, dynamics—and if it can, I want to know what the mechanism is. So I want to know how we answer questions like, "Where do structures come from?" "How do they change?" "Where do the relationships come from?" "How do they change?"

Other papers in this conference deal with social dynamics. Ecological theories posit selection mechanisms that quickly get translated into change, and rational choice theories posit some kind of change in relative prices as a motor of structural change. We can and will argue about the adequacy of these and other approaches, but at least they specify a dynamic mechanism.

Edward Laumann: Talcott Parsons wrote a book in 1960 called *Structure and Process in Modern Society,* in which he argued that structure and process are in some ways two sides of the same coin. I always agreed with Parsons's dictum that you can't talk about social change until you can talk about the structure that is changing. My enterprise here is essentially to say that different policy products are likely to come out of different structures, depending on their respective organizational principles.

Systematic structural observations are necessary to be able to say, for example, that this was the arrangement in the health policy domain in the 1970s. We are currently looking at the health structure as of the Reagan period, and can thus begin looking at what changed. Not to use this sort of systematic data analysis at two time points would require an impressionistic and possibly misleading reconstruction of the Carter period.

I can tell you all kinds of "dynamic" stories that arise out of understanding how those issues got organized earlier into specific social arrangements. I could give you a theory of the history of the emergence of organized modern medicine that helps us understand why we are where we are in the 1970s. For example, the whole decision-making about NIH. This was constructed in the 1950s around a set of medical specialist groups that had their own "disease of the month" support crowd. It resulted in the development of highly specialized medicine with no effectively organized constituency for the general practitioners who provide routine care to the vast bulk of "medical consumers." So there is a biomedical research community with a captive set of institutes that act on their behalf and a congressional constituency as well that is impervious to the demands of welfare-oriented groups or lay groups demanding routine medical care. In addition, these specialist orientations got mapped into the way medical schools are organized. This produced an arrangement that some are now trying to undo.

In 1965 Medicare was adopted, and its effect was almost like the Treaty of Vienna. It settled the issue of universal national health insurance so that it is not likely to happen in this century, for the very simple reason that Medicare bought off the most important sets of user groups that would have supported universal health insurance.

So I am not so pessimistic as to say that the structural model that I am concerned with has nothing to do with dynamics. But we can't have a theory of dynamics without having gained an understanding of the basic structural arrangements that frame the dynamic process.

David Rubenstein: Laumann made reference to multiple structures and to dominant structures. My impression is that there are all kinds of structures out there: friendship structures, power and authority and status structures, and kinship structures. My question would be whether you see this heterogeneity of structure to be a problem, and if you think that network theory must move toward finding dominant structures in order to develop something approaching a general theory.

Ronald Burt: Yes, it is possible for a general theory. The key notions—access to relations, brokerage of relations, demand for relations—are the vehicles for doing this across different contents. The result is general propositions that work in affect relations, economic relations, political relations.

Structural Theory

The California Gold Rush: Social Structure and Transaction Costs

ANTHONY OBERSCHALL

This paper will provide an example of how to account for social structure by means of transaction costs. Transaction costs refer to the costs of interaction and of exchange itself, such as the costs of collecting information on interaction partners and on the commodity or action that is exchanged, the costs of negotiating an agreement or contract and of monitoring its implementation, and the actual enforcement costs of the agreement. Transaction costs exist because human beings' rationality is bounded, not least by the time and effort of collecting and processing information; because some people are opportunists who violate trust and take advantage of vulnerable interaction partners, if they can get away with it; and because real-life situations are complex and uncertain, making it unlikely that all contingencies can be anticipated and explicitly agreed upon (Williamson 1981; Ben-Porath 1980; De Alessi 1980).

The underlying idea of the transaction costs approach is that different modes of transacting, i.e., different institutional arrangements for interdependent actions and exchanges, are best understood from the point of view of minimizing all costs, including the very important transaction costs. The approach brings social structure and institutions back into economics and allows a growing overlap and integration among the social sciences. It seeks to explain the boundaries of a group or social unit, their internal structure, the types of norms that prevail in them, the likelihood of nonconformity. In a frequently commented on passage in the *Division of Labor,* Durkheim wrote that

> when men unite by means of contract, it is because they need each other . . . as a consequence of the division of labor. But if they are to cooperate harmoniously, it is not sufficient for them to form relationships, nor to experience a state of mutual dependence. . . . The conditions of cooperation must be agreed upon for the entire duration of their relationship. The rights and obligations of each must be defined. . . . Were it not so, new conflicts and disputes would occur at each moment.

The transaction cost approach seeks to explain just what the "conditions for cooperation" are that have been "agreed upon" and the properties of the institutional arrangement which define the "rights and obligations of each." It is a useful tool for social theory, and in particular for anyone who seeks to account for variations in social structure and institutions.

I will apply transaction cost analysis to property rights and governance institutions that emerged in the California gold rush following 1848. California's interior was a wilderness, sparsely inhabited by a few ranchers and some Indians. Civil government hadn't yet been constituted following the war with Mexico, and the military administration was effectively limited to the coast. The rivers, streams, and canyons rich with gold were for the most part located in empty land unencumbered by property claims, even of Indians. Protecting life, limb, and property and devising a means of allocating mining claims were up to the miners themselves.

The simple technology of placer mining required no more than a shovel, pickaxe, pan, and simple tools and could be mastered in a few days simply by observing more experienced miners. To operate equipment such as the cradle or rocker used for separating gold from slush, a team of two or three miners was necessary, and was more efficient than each individually using a pan. But beyond teams of two or three partners, there were no economies of scale in joint production. Given the hazards and hardships of the journey to California, whether by land or sea, and then from the coast to the interior, and the physical stamina needed for mining and camping under pioneer conditions, it is hardly surprising that miners were almost all men in their young and middle adult years. They came from forty different countries and from every state in the Union (Hamilton 1978, p. 1470).

For purposes of travel, the future miners would form "companies" at their locality of origin. For instance, from Massachusetts alone, 124 companies set out for California in 1849, each company from a specific town. Eighty-three companies organized in France. These companies, however, disbanded in California (Hamilton 1978). The goal of each and every miner was to strike it rich and return home.

The 1848 rush was limited to a few thousand inhabitants who were already in California and Mexico, and included soldiers of the U.S. garrison and crews of ships who deserted in large numbers. The situation changed in 1849 and subsequent years when tens of thousands kept arriving. By the end of 1852, the population of California was estimated at 264,000 and was concentrated in the mining areas (Umbeck 1981).

At first, with plenty of mining land in relation to miners, there was no incentive for creating a permanent governance structure. This is how Shinn (1884) describes early mining prior to crowding effects:

> If a few prospectors found a small deposit of mineral they began to work it with but one idea in mind, that of obtaining as much as possible before anyone else discovered their new camp. When this happened, the newcomers suggested some definite size for claims, and by "mutual consent" and desire for compromise, they soon arrived at some understanding. In this primitive, informal way, many a group of men worked together for several weeks or months, in a mountain gorge. . . . In several cases on record the first workers laid out unusually large claims to which later groups of miners demurred, and the question was put to vote; thus in one instance reducing each of the eight original claims to one-fifth of its former size, but without any objection from the owners. . . . [p. 34]

But such informal arrangements wouldn't do when hundreds converged on the same area in the rush:

> March 27, 1850, five New England prospectors . . . discovered the famous "Kennebee Hill" placer deposits, and within thirteen days there were eight thousand miners in the new camp, saloons, hotels, stores, and all sorts of human parasites . . . of this mushroom city. Fourteen days from the discovery a meeting was called "to organize and govern this camp." . . . In 1852, the united district polled more than twelve hundred votes. The laws adopted were enforced and the place was noted for its orderly appearance. [p. 26]

Just how was a mining camp or district organized? When crowding effects led to disputes and fights, miners called an assembly and constituted themselves into a territorial mining district, with visible, physical boundaries; chose criteria for district membership with specific entrance and exit rules; allocated land and water rights among members; and built a governance structure for enforcing and changing these rights, as well as orderly camp life. In physical terms, Shinn (1884) describes a typical camp in the following way:

> So narrow and deep are most of the wild mountain gorges in which the miners toiled . . . that the typical camp . . . was a long line of men extended for several miles up and down the ravine, and returning at night across each others' claims to the little collection of tents and cabins, saloons, hotels . . . that was the "camp." . . . [p. 18]

In the first decade of placer mining in California, such a means of self-government was undertaken spontaneously in about 500 mining districts (Umbeck 1981). What were these rights and governance institutions?

District membership included all those present at the original constituent assembly. The land was divided into equal lots (e.g., 100 by 200 feet along the riverbed and bank), marked by visible stakes on which the owner's name was posted. Each member was given one lot only in the district and had the right to own only one lot. Exceptions occurred: Sometimes the original discoverers were given an extra lot, and one additional lot might be purchased by sale. Ownership right had to be maintained by occupancy on the lot for a specific number of days per week or month, varying with the season. Lots could be sold to newcomers by sale, and purchase would automatically make one a district member. Abandoned claims were assigned to, or could be claimed by, newcomers on a first come, first served basis. Lots distant from the riverbanks were larger in size than the "wet and gulch" claims, but also measured out equally. These lots were less productive of gold per unit area and required more labor for extracting it because water was distant (Umbeck 1981; Shinn 1884).

A low-cost governance structure was also agreed upon at the constituent assembly, usually consisting of a recorder or clerk, a handful of elected executive officers who acted as mediators, arbitrators, and, if need be, judges, with a right to appoint a jury to try cases. A schedule of compensation for serving in these, mostly temporary positions was agreed upon. Only miners in the district could serve on these bodies. Within this broad framework, there were considerable differences in detail, for instance on plot size, occupancy rules, and the like. Sometimes, one Grand Enforcer—called an "alcalde" following the Mexican local government model—was provided all executive and judicial powers, but could be deposed any time by a majority vote.

The governance system worked remarkably well. Most disputes were settled by arbitration, with offenders compensating victims in gold or cash. In cases of homicide, theft, and assault, offenders would be banished from the district, fined, or hanged from a tree after being tried in a miner's court, depending on the gravity and circumstances of the crime. But there was little violence in the mining camps themselves, compared with the lawlessness of San Francisco and other California towns, and the fighting that inevitably accompanied the gambling and drinking in the hastily built mining towns where miners repaired for supplies, "rest," and recreation. Bancroft (1888b, p. 131) cites the following figures for 1855: Of 538 recorded homicides in California, only 12 originated from fights about mining claims.

It is by no means obvious why the miners on every occasion should choose the particular property rights and governance institutions just described, in preference to some other arrangements. A striking feature of these institutions is their egalitarianism and individualism: one man, one vote; and equal sized and shaped plot to each; private property rights in land; no absentee ownership. Actually matters are not quite this simple, because foreigners, especially of non-European ethnicity, were at times discriminated against. Bancroft described instances of Mexicans and Hispanic-Americans evicted from desirable claims by newcomers to a camp, with the tacit approval of other miners. Chinese were not allowed in some camps, and French miners were also at times expelled or excluded (Bancroft 1888a, pp. 403–7). But with these notable exceptions, egalitarianism prevailed.

Let us now apply transaction cost analysis to the emergence of property rights and governance institutions in mining camps. Transaction costs are the costs of making an agreement, monitoring its terms, and enforcing it. In the gold camp situations, the miners had to define the boundaries of the group itself within which agreements were to be made; to agree on how to divide the land or the gold; finally, to decide on what governance structure to create for enforcing property rights, peaceful interpersonal relations, and collective defense against outsiders.

Four matters stand out. First, the miners were strangers to each other and of very diverse backgrounds; and no one knew ahead of time how many would converge on the same area. Second, the miners had roughly the same capacity for violence—the gun was an equalizer of this capacity, and all miners were armed—and had roughly the same capacity for production, since they used the same technology and were men of the same age cohorts. Third, gold was a nonrenewable resource and a windfall for all of them. They were not interested in permanent settlements and governance institutions, as would be the case for managing a renewable resource such as crops. Each miner had an interest in maximum mobility opportunity over the entire gold region, with flexible entrance and exit from a particular mining camp. Fourth, given the large number of miners converging on California and the concentration of gold in riverbeds and banks, and the technology of placer production, gold camps were densely packed with miners literally rubbing elbows against neighbors. How do these four features bear on transaction costs?

Any particular group claiming property rights had to be in a position to enforce them against outsiders and had to have a mechanism of making would-be freeriders in the group participate in common defense and governance. Freeriding is an obstacle to sharing the costs of governance and is a component of transaction costs, i.e., the cost of policing the agreement on governance within the group. If miners were within sight and voice relay range of each other, there existed a simple mechanism of convening the group when circumstances demanded and of deterring freeriders who might be tempted to work their claim while others were attending to the camp's governance. A camp consisting of several hundred miners in many small, physically proximate, autonomous miners' groups along a two to three mile stretch of canyon has lower transaction costs than a more populous, spatially extensive miners' camp vulnerable to freeriding and possessing a formal, costly means of communication. The pattern and boundaries of the mining camps is thus explained in terms of minimizing transaction costs.

One can also explain the occupancy rules against absentee ownership and prolonged absence from the camp. These rules assured that a large part of the group would be present in the camp at any particular time, enabling the group to protect itself against outsiders, and ensuring cheap, interpersonal communication within the group.

Note also that miners did not enforce a principle of selection or exclusion for group membership, beyond the simple fairness rule of first to arrive, first to join, until all lots are filled, with the already mentioned exception of foreigners' exclusion in some camps. Any screening based on prior criminality, to take an obvious possibility, would be impossible to implement since miners were strangers to each other. Given the high density of camp settlement, crime detection was not much of a problem, as would be the case in low-density, anonymous environments. The social control of criminality could be handled cheaply, swiftly, and effectively without prior selection.

What of the egalitarian aspects of the governance institutions and of the property rights to land and water? The one man, one vote and one equal-sized plot principle stands out prominently by its fairness and ease of administration compared with all other possibilities and principles. Consider the cost of reaching an agreement among a large number of strangers, with but a temporary interest in maintaining cooperative relationships with each other, on any other principles. Each would have an interest in putting forward rules or exceptions that favored him personally—for example, number of dependents or "need," or resources expended on getting from home to California—and each would have an interest in deceiving the others by lying, whatever the criterion selected. The time it would take to reach an agreement, and to monitor its enforcement, would be long, if indeed an agreement could be reached at all. No such problem with the egalitarian principle, which is easy to justify and to enforce. Umbeck (1981) makes a somewhat different argument based on each miner's equal capacity for gold production and equal capacity to defend property rights by means of violence. I believe that the prominence of the fairness norm and its low transaction cost is a more convincing explanation for its adoption.

Finally, why were the camp norms not only egalitarian, but individualist, in the sense that the land rather than the gold was divided into equal shares, as might well have occurred if miners shared not just governance, but production as well? Umbeck (1981) provides an explanation based on transaction costs. The cost of reaching some agreement on dividing resources within a constituted group is the same, whether land is divided or the gold is. The costs of monitoring the agreement differ and vary with the size of the miners' group in the two instances. In the land allotment scheme, there is a once-and-for-all initial cost of measuring out lots, assigning them to miners, and posting stakes. Each miner hereafter has an incentive to keep trespassers out, which he could easily do just by being physically present, or his neighbors would if he was temporarily absent. In the gold division scheme, some means of continuously monitoring amount of work time put into group production would have to be devised, and surveillance against theft, that is, keeping some of the gold rather than surrendering it into the common pool for division. With a small group of miners working with each other and next to each other and living in the same tent or cabin, monitoring and surveillance is built into routine group interaction, with no additional cost. As size of group increases and as physical distance between members increases, monitoring, bookkeeping, and surveillance costs increase, as does the incentive for dishonesty. Thus, in all but the smallest groups, the gold sharing scheme has higher transaction costs than the land allotment scheme.

To the extent that miners in a district had been earlier socialized uniformly to the same standards of honesty and hard work, and had forged attachments and social bonds with each other with some expectation of permanence, the monitoring, metering, and enforcement costs of the gold sharing scheme would be less burdensome. It also provided insurance against risk, not possible under land allotment. But miners were a collection of strangers, of very diverse and unknown backgrounds and moral qualities, with no expectation of permanence beyond the life of the gold camp, terminated by mass exodus once the gold had been mined or a richer area discovered.

And lastly, the land allotment scheme did make it possible for small, voluntary partnerships that pooled land and divided gold to form in any case, which would spread risk somewhat, make teamwork for rocker and cradle possible, and enable partners to maintain occupancy while one of them was away resupplying or prospecting. Such small partnerships of two to four miners were in fact very common in the California gold rush.

Are there other theories that account for the property rights and governance institutions of mining camps? The best known and original historian of the mining camps, Charles Howard Shinn, writing in 1885, attributed them to some sort of innate "unconscious socialism" joined to an equally elemental belief that all men had an equal right in the pursuit of fortune. Bancroft (1888a), the historian of California, thought in terms of national character and an Anglo-Saxon disposition for democratic institutions.

Apart from the easy objections to the antiquated conceptions and terminology there is the fact that New England towns from which many miners came were highly stratified, rank-conscious communities where land was certainly not divided into equal lots, not even at the time of initial settlement. Further, gold miners, including many California veterans, readily submitted to a hierarchical and bureaucratic form of colonial government in the British Columbia gold rush of the late 1850s and early 1860s. There they conformed to the authority of gold commissioners, paid for gold mining licenses not required in the United States, and kept to rules set down by a distant governor appointed by a distant Queen. Trimble marveled at this (1914):

> In the gold commissioner of British Columbia . . . were centered the powers of the American mining camp and of a British magistrate. . . . A marked feature of life in Kootenai (a British Columbia mining district) was the submission of the miners to the lawful authorities. Here were a thousand or two rough miners, all collected from the American territories at the time when Montana was going through vigilante throes; government was represented by a lone magistrate with two or three constables unsupported by any possibility of aid. . . . And yet the testimony of the British officials is unanimous to the orderliness of the miners. [pp. 203, 58–59]

It appears that in a situation where government and property rights already existed, the miners simply conformed to the rules, whatever their dispositions, and thus "saved" transaction costs which were borne by government.

Opposed to any argument that egalitarianism arose from sentiments of we-feeling and solidarity are descriptions by both Shinn and Trimble of instances where discoverers of a gold bed tried to keep this information secret as long as possible and assigned to themselves large, irregular claims until overwhelmed by newcomers and forced to change (Shinn 1884, p. 17). Also, when physical conditions for mining required more capital-intensive mining techniques than in California placer mining, with a division of labor based on hired wage labor, as occurred in the very rich Cariboo fields of British Columbia, the miners became internally stratified. "Companies" of miners formed to put up money for timber and pumps. These "companies" would then hire other miners as wage laborers, which many miners ended up doing for a long time when they ran out of funds (Trimble 1914, pp. 54–55). Although miners' camps and towns were surely marked by considerable fellow feeling, and there are many instances of generous help and assistance to the ill and needy, no deep commitment to egalitarianism prevented unequal institutions, when circumstances and technology made them attractive. In yet other situations typical of California during the dry season, labor-intensive teamwork was needed to divert low streams and rivers by digging ditches, channels, and dams. This enabled miners to get at the dry riverbeds rich in gold. "Companies" based on team labor temporarily formed for this purpose repeatedly. Once river diversion had succeeded, each miner would revert to individual digging up of his prior, allotted claim in the dry riverbed. Thus, teamwork and sharing was situational and opportunistic, not a commitment to principles.

I have argued that the property rights and governance institutions of the California gold rush mining camps were such as to minimize transaction costs. In particular, the size and boundaries of the camp, the entrance and exit rules to the group, the fairness norm on land allotment and the one man, one vote rule, the prohibition against absentee ownership and the occupancy rule are explained with reference to lower costs of reaching an agreement, of enforcing the agreement, and of communication among the miners compared with alternative norms and governance institutions. The transaction costs approach promises to be an important tool for explaining how social structures come about, whether or not it happens as a result of deliberate creation or an evolutionary process.

Comment

PETER M. BLAU

Oberschall starts his paper by emphasizing that recent developments in economics have overcome the limitations of the assumption of unbounded rationality of neoclassical economic theory. Specifically, by taking transaction costs into account, it is possible for economics—and by implication for sociology—to maintain the model of rational choice without ignoring or neglecting the external social conditions that restrain human action and that explain why it is not based on purely rational self-interest. The body of Oberschall's paper uses a discussion of the California gold rush in the years around 1850 to illustrate this claim.

He shows that the miners had formed companies already for the purpose of travel, and later, as their numbers in California grew from a few thousand to tens of thousands and then to more than two hundred thousand, developed governance institutions. Although the miners were initially interested only in maximizing rewards—in striking it rich and returning home—the situation forced them to enter into cooperative arrangements and develop institutions for dividing the land, establishing fixed boundaries, and settling disputes. To enforce the emerging institutionalized norms, an assembly with legislative powers had to be created, and executive officers and judges had to be elected or appointed. In short, acting initially purely in terms of rational self-interest, miners produced a rudimentary political system, which was designed to protect all their interests by imposing some limits on the purely rational choices of any one individual. Oberschall suggests that these changes, and even the specific nature of the institutions that developed—for instance, the egalitarianism and individualism—can be explained in terms of transaction costs. His implication is that institutional and structural limits on rational choice can generally be accounted for by an analysis of transaction costs. It is this conclusion with which I take issue.

Let me hasten to add that I do not disagree with Oberschall's domain assumption that rational choice, broadly defined, governs most if not all human decisions, not merely strictly economic ones; nor do I disagree with his claim that taking transaction costs into account helps explain many external restraints limiting rational choice, notably those resulting from the common culture, the institutionalized rights and duties, and the prevailing normative standards. But shared cultural orientations, political institutions, and legal or informal norms are not the only structural constraints on social life, indeed, not even the major one. The core of a community's or society's social structure rests on its population composition and refers to people's distribution among different positions along various lines and their social relations. You cannot occupy a position or relate to someone in it unless there are such positions that people can and do occupy. The major structural constraints on people's preferences and rational choices are the result of the distribution of positions in a collectivity, since this distribution governs the opportunity structure. Only by taking into consideration how these distributions affect chances and options do we avoid psychological reductionism—methodological individualism, if you prefer—and give full due to the emergent properties of social structure. Let me illustrate this point with two examples.

As many of you undoubtedly know, I am not completely disidentified with the status attainment model. However, this model has been criticized, and quite correctly so, for dealing only with the resources, costs, and qualifications of individuals. It completely ignores the fact that achieving a status, which is measured in this model by occupation, also requires that there are jobs; therefore, the market structure has to be taken into account. This criticism is the basis of the dual and multiple market theories. It is also the basis of Boudon's analysis of why increasing levels and equality of education have not produced greater equality in socioeconomic status. In other words, the status attainment model deals only with the demand for jobs and not with the supply of jobs in the labor market.

The second example comes from my recent work. The sociological objective is to study the structure of positions and the structure of relations in a collectivity. The theory seeks to show how the social structure exerts external constraints and defines opportunities by examining how the structure of positions influences people's choices of associates. I make two assumptions: The first is that people have "ingroup" preferences. By this I mean that most people prefer associating with others in their group or socioeconomic stratum to associating with people in different groups or strata. Of course people cannot realize this preference in all respects because not all groups completely overlap. The second assumption is that population structures influence the opportunities for social relations. By population structure I mean nothing but the degree of differentiation of a population in various dimensions, which are simply the distributions of the population among social positions. In the case of groups (nominal social positions), like ethnic or religious groups, differentiation is a form of heterogeneity; in the case of status differences (ranked positions), like income, education, or socioeconomic status, differentiation is a form of inequality.

The first assumption implies, for example, that people marry disproportionately often somebody in their own group or somebody similar in income and in education to themselves, and the data confirm this implication. Inequality means that the average difference in status is great and therefore inequality will push people to marry someone not as close to them as they would ideally want. The case of heterogeneity is similar. The second assumption implies that the opportunity to marry somebody in your own group is less if there is much heterogeneity than if there is little, and the opportunity to marry somebody close to you in income or in education is less if there is much inequality. Hence, the prediction from the theory is that as the degree of inequality in a society or community increases, the difference in income or in education between spouses increases. That is what the findings confirm. In other words, the structure of opportunities limits the influences of individual preferences. This illustrates Durkheim's external constraints. I claim that this cannot be taken into account by the rational choice model, whether or not transaction costs are analyzed, and I consider the study of such structural influences the core of sociology.

General Discussion

Siegwart Lindenberg: I think there is a stubborn misunderstanding concerning rational choice and I see that again in Peter Blau's comments. The misunderstanding is that anyone claims that rational choice theory is a structural theory. What Oberschall and I and other rational choice theorists claim is that there will be no adequate structural theory *without* rational choice theory. The examples given by Blau only confirm this. His own theory would be even better if he would make the rational choice part of it explicit, as I tried to show in an article some years ago.

Once this is understood, more important things will arrest our attention: An explanation of behavior does not tell what the collective consequences of the concert of behaviors will be. For example, in Olson's theory of collective action, what the microeconomic model really does is to explain apathy. It will not explain underproduction of collective goods. For that are needed things that have nothing to do with a rational choice model. For example, if class action in court cases is institutionally possible then one individual can produce a collective good for millions of others through a court decision. If there is no such institutional arrangement, then this particular way does not work.

Only if I explain collective consequences, can I, in turn, explain the constraints for rational choice. But this sword cuts both ways. I cannot explain the collective consequences unless I have an adequate theory of action, something sociology is in dire need of acquiring. If we agree on that, we would reduce our transaction costs by 90 percent and start digging for gold, to stay in Oberschall's story.

James Coleman: Peter Blau made a point which seems to involve a confusion between two things. One is the structure through which a social product is generated from individuals and the other is the behavioral model that's involved. For example, microeconomic theory has two elements. First, its behavioral model is a rational choice model, and, second, the structure through which a social product, price, and distribution of goods is generated, is a private goods market. But these two elements are separable. For example, another structure through which a social product is generated from individuals is an election mechanism. That social product is a collective decision. But that does not imply that the behavioral model cannot be one of rational choice. Similarly, an organizational structure is different from a market. In that sense it seems to me that sociology could take one aspect of microeconomic theory, namely, a rational choice behavioral model, and add to the market a variety of other structures through which a social product is generated. This is a point at which I disagree with Becker's orientation. He maintains for sociological activities not only the behavioral assumptions of rational choice but also the market structure at some points where the social product is generated by a structure other than a market.

Peter Blau: I said at the beginning that I agree with most of Oberschall's assumptions on rational choice. And if you put it the way Lindenberg did, that one has to consider rational choice and the structural aspects, then there is no disagreement. But this approach is not what methodological individualism has meant in sociology. Homans has stated it most extremely. For him there is no such thing as sociological theory, there is only psychological theory and ultimately everything has to be reduced to that. This is what methodological individualism has meant in sociology and I strongly disagree with it.

Siegwart Lindenberg: It would be useful and prevent deep misunderstandings if we would reserve the term "methodological individualism" for the position I described (with which Peter Blau could agree) and use the term "reductionism" for the position Homans has sometimes taken.

Randall Collins: I think there is an advantage in starting from a network analysis on the level that Ed Laumann was arguing for, in the sense that one can derive precisely those preferences and resources from the position of each individual in the network at any given moment. And the lineup of resources and interests that comes out of each situation leads naturally, by a mechanism perhaps not too different from rational choice, to yet further network situations, and so forth. Let me fill in a bit about how I think that happens, where the interest and resource aspects come in. In everyday life, any organization includes conversational interactions. People come into every conversation with some kind of currency that they have already picked up from a previous situation. You can call it "cultural capital" if you like, following Bourdieu, and in fact you can see this currency circulating through a series of interactions. People negotiate situations for optimal domination, building up further contacts into their personal networks, which in turn will bring more situations in which this cultural capital can be used for domination.

Allen Grimshaw: It seems to me that if individuals make rational choices, it's a bit difficult to understand the nature of the world in which we are now living. We are in substantial danger of being blown away. If we are going to talk about rational choice being a useful explanatory model, at the very least we have to realize that there is institutionalized irrationality and that it has very substantial effects on outcomes for all of us.

Michael Hechter: Rational choice says nothing about collective rationality. There are many collective dilemmas that result from individuals pursuing their own selfish interests.

Edward Laumann: I would like to come back to Oberschall's presentation. How would you locate the following arrangement? We have a substantial amount of potential gold to be found in this area. Let's say that one of the people who come into this valley to find it is an organizer. He puts together a group and says, "Listen, if you follow me, we can organize this situation so that we can claim more land to search for gold. We, as a collective actor, can mobilize the resources to coerce the others. I'll assign some of you to be policemen and we'll cover your claim." How would you deal with that in your rational choice models?

Anthony Oberschall: This mechanism was used to form vigilante committees to do away with individual crime, not with outside intruders. The process that I described of agreeing on a form of government and the handling of property rights was repeated hundreds of times between 1848 and the early 1860s. There wasn't on any one of these occasions a major departure from what I have described to you. It is surprising that they always chose this form.

But looking at it from a transaction cost point of view, it is not such a mystery. How could an entrepreneur convince others to participate in his scheme? There was high uncertainty as to whether gold would be found and, if so, how much, for any particular area. Miners had to be mobile. Neither capital investment nor family tied them to a particular spot. Given this uncertainty, what would make it worth their while to invest much effort setting up an organization and driving others away? What would have made some of them willing to be policemen only, while others may have been striking it rich? On what basis would they have trusted the entrepreneur that they get a fair share doing police work rather than digging for gold?

Once the gold in the riverbed was gone, it could only be mined with sophisticated equipment. In this situation, uncertainty was much lower and high capital investment tied the miners to a particular area. This is where the entrepreneur came in. Once such a corporation, often financed with eastern capital, bought up claims they just used existing corporate forms for mining. The various state supreme courts and the federal Supreme Court did recognize these, what they called the "miners' codes." The common law then was written into state law and recognized by the court so that a more capitalized enterprise could safely come in without legal risks.

Immanuel Wallerstein: Let's take the mining example. Yes, if you hold everything constant, which is what Tony Oberschall did, they will behave that way. But we start with saying that it's a windfall. "Windfall" sounds like it comes from God, but in fact there has been research that demonstrates that the moments at which mining discoveries are made are not chance. There are systematic patterns. I can predict at what point in historical time mining discoveries will be made and therefore that a situation will be created in which lots of people will go into an area which was relatively underpopulated previously and will suddenly begin to mine. Then one has to ask why there was not immediately a state structure established which would provide the protection mechanisms at a cost that would not be burdensome for the individuals, which is one of your premises. I can begin to say things about why that should happen, why in fact other agencies deliberately abstain from intruding state protection services in such areas at early points in mining procedures, and why at later points in time, when larger scale discoveries are made, all of a sudden in comes state protection services, because the mining operations tend to get concentrated.

Theda Skocpol: In the mining example, a moment is picked out in which something close to a deliberate decision-making process probably did occur, in a situation where participants had equal resources and little information about one another and no pre-existing structural ties. But there were probably some Indians on that land who weren't consulted or given that chance to have an equal claim in this situation. Would you deny that we have to take into account a series of things about the institutional structure, if any, that exists among the participants before the windfall arrives and the notions they have about who is included in the community and who might rank higher or lower in that community?

Donald Levine: If Oberschall's point is that windfall situations are universally accompanied by access norms of nonexclusivity and of distribution norms of equality, that's clearly counterfactual. My own experience in Africa is full of examples where windfalls of dead animals or gifts of sheep or chickens are immediately responded to by identifying the different parts of the animal in a hierarchical manner; and it seemed the higher status people get those parts, and the lower status people get the lower parts of the animals, and other people don't get any or only get what is left over when everyone else is done. That seems to be a common kind of pattern among cultures in the world.

Anthony Oberschall: There is virtually no time left to respond in the framework of this session. Still, let me at least respond to one, namely, Blau's question of how to deal with the supply side rather than just the demand side. I am studying the New Christian Right in North Carolina. The supply side of this movement consists of moral entrepreneurs; they are very conservative in their religious and political orientation. These moral entrepreneurs, some of them on TV, others in churches and associated in various ways in a movement, can be studied and theoretically analyzed and I am doing that. Maybe I will not satisfy you when I am ready to produce the results, but I am working on it.

Purposive Action Theory

Human Capital and the Rise and Fall of Families

GARY S. BECKER AND NIGEL TOMES

Introduction

Although discussions of inequality among families and discussions of inequality between generations of the same family have been almost entirely separated, they are analytically closely related. In particular, regression away from the mean in the relation between the incomes of parents and children would be associated with large and growing inequality of income over time, while regression toward the mean would be associated with a smaller and more stable degree of inequality. These statements are obvious in a simple stochastic model of the relation between parents and children:

$$I_{t+1} = a + bI_t + \varepsilon_{t+1}, \tag{1}$$

where I_t is the income of parents, I_{t+1} is the income of children, a and b are constants, and the stochastic forces affecting the income of children (ε_{t+1}) are assumed to be independent of the income of parents.

Inequality in income would continue to grow over time if b were greater than unity, but would approach a constant level if b were smaller than unity in absolute value. Clearly, the size of b also measures whether children of richer parents tend to be less rich than their parents, and whether children of poorer parents tend to be better off than their parents. Even in rigid and caste-dominated societies, many of the elite and underprivileged families would change places over generations unless inequality continued to grow over time ($b > 1$).

The degree of regression toward or away from the mean in the achievements of children compared with those of their parents is a measure of the degree of equality of opportunity in a society. The purpose of this paper is to analyze the determinants of unequal opportunities, sometimes called "intergenerational mobility" or, as in the title of our paper, "the rise and fall of families." We use these terms interchangeably.

The many empirical studies of mobility by sociologists have lacked a framework or model to interpret their findings. We try to remedy this defect and fill a more general lacuna in the literature by developing a systematic model that relies on utility-maximizing behavior by all participants, equilibrium in different markets, and stochastic forces with unequal incidence among participants.

An analysis that is adequate to cope with the many aspects of the rise and fall of families must incorporate, among other things, concern by parents for children, as expressed in altruism toward children, assortative mating in marriage markets, the demand for children, the treatment by parents of exceptionally able or handicapped children, and expectations about events in the next or even later generations. Although these and other aspects of behavior are incorporated into a consistent framework based on maximizing behavior (see also Becker 1974, 1981; Becker & Tomes 1976, 1979, 1986; and Tomes 1981), we do not pretend to handle them all in a fully satisfactory manner. However, our approach indicates how a more complete analysis can be developed in the future.

Earnings and Human Capital

Perfect Capital Markets

Some children have an advantage because they are born into families with greater ability, with greater emphasis on childhood learning, and with other favorable cultural and genetic attributes. Both biology and culture are transmitted from parents to children: one encoded in DNA and the other in a family's culture. We do not need to separate cultural from genetic endowments, and will not try to specify the exact mechanism of cultural transmission. We will follow our previous paper (Becker & Tomes 1979; see also, e.g., Bevan 1979) in assuming as a first approximation that both are transmitted by a stochastic linear equation:

$$E_t^i = \alpha_t + hE_{t-1}^i + v_t^i, \tag{2}$$

where E_t^i is the endowment (or vector of endowments) of the ith family in the tth generation, h is the degree (or vector of degrees) of "inheritability" of these endowments, and v_t^i measures unsystematic components or luck in the transmission process.

The assumption that endowments are only partially inherited, that h is less than unity and greater than zero, is a plausible generalization to cultural endowments of what is known about the inheritance of genetic traits. This assumption implies that endowments regress to the mean: Children with well endowed parents tend also to have above average endowments, but smaller relative to the mean than their parents', whereas children with poorly endowed parents tend also to have below average endowments, but larger relative to the mean than their parents'.

The term α_t can be interpreted as the social endowment common to all members of a given cohort in the same society. If the social endowment were constant over time, and if $h < 1$, the average endowment would eventually equal $1/(1 - h)$ times the social endowment (i.e., lim $\bar{E}_t = \alpha/[1 - h]$). However, we do not assume that α is constant because, for example, governments may invest in the social endowment, but we do assume that parents cannot invest in their children's endowment.

Practically all formal models of the distribution of income that consider wages and abilities assume that abilities automatically translate into earnings, mediated sometimes by demands for different kinds of abilities (see, e.g., Mandelbrot 1962; Roy 1950; Tinbergen 1970; Bevan & Stiglitz 1979). This is useful in understanding certain gross features of the distribution of earnings, such as its skewness, but is hardly satisfactory for analyzing the effect of parents on their children's earnings. Parents not only pass on some of their endowments to children, but also influence the adult earnings of their children by investments in their skills, health, learning, motivation, and many other characteristics. These expenditures are determined not only by the abilities of children, but also by the incomes, preferences, and fertility of parents, public expenditures on education and other human capital of children, and still other variables. Since earnings are practically the sole income for most persons, parents influence the economic welfare of their children primarily by influencing their potential earnings.

To analyze these influences in a simple way, assume two periods of life, childhood and adulthood, and that adult earnings depend on human capital (H) and market luck (l):

$$Y_t = \gamma(T_t, f_t)H_t + l_t, \tag{3}$$

where γ, the earnings of one unit of human capital, is determined by equilibrium in factor markets, and depends positively on technological knowledge (T) and negatively on the ratio of the amount of human capital to nonhuman capital in the economy (f). Since we are concerned with differences among families, we can usually avoid paying close attention to the exact value of γ because that is common to all families. Therefore, we assume that the measurement of H is chosen so that $\gamma = 1$.

Although human capital takes many forms, including skills and abilities, personality, appearance, and reputation, we further simplify by assuming that it is homogeneous and the same "stuff" in different families. Since much research demonstrates that investments during childhood are crucial to later development (see, e.g., Bloom 1976), we assume that the total amount of human capital accumulated, including on-the-job training, is proportional to the amount accumulated during childhood. Then adult human capital and expected earnings would be determined by endowments inherited from parents and by parental (x) and public expenditures (s) on his or her development:

$$H_t = \psi(x_{t-1}, s_{t-1}, E_t), \text{ with } \psi_j > 0, j = x, s, E. \tag{4}$$

Ability, early learning, and other aspects of a family's cultural and genetic "infrastructure" usually raise the marginal effect of family and public expenditures on the production of human capital; that is,

$$\frac{\partial^2 H_t}{\partial j_{t-1} \partial E_t} = \psi_{jE} > 0, \qquad j = x, s. \tag{5}$$

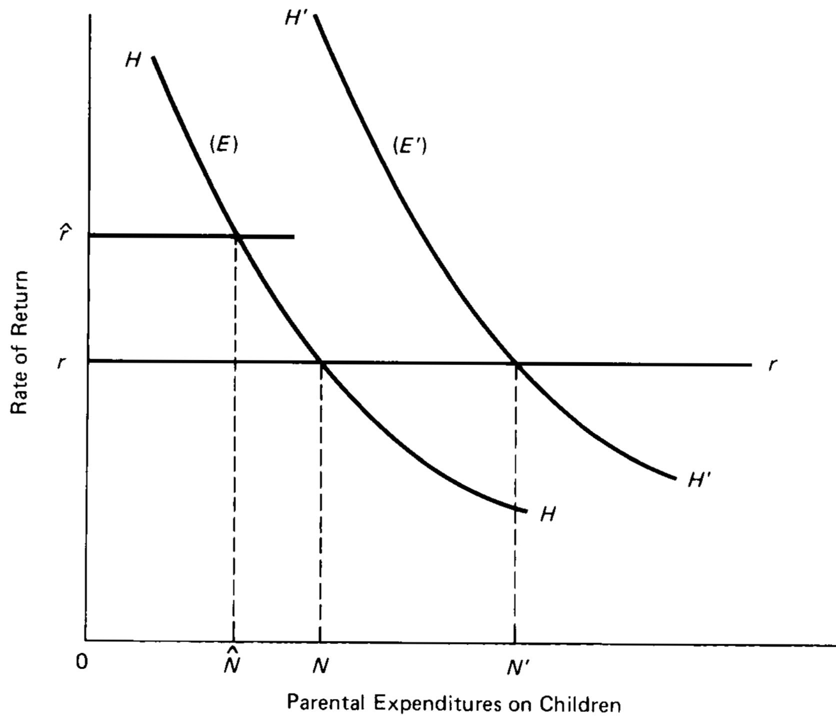
The marginal rate of return on parental expenditures (r_m) is defined by the equation

$$\frac{\partial Y_t}{\partial x_{t-1}} = \frac{\partial H_t}{\partial x_{t-1}} = \psi_x = 1 + r_m(x_{t-1}, s_{t-1}, E_t), \tag{6}$$

where $\partial r_m/\partial E > 0$ by inequality (5).

Although the human capital of different persons may be close substitutes in production, each person forms a separate human capital "market" because rates of return depend on the amount invested in him as well as on aggregate stocks of human capital. Marginal rates of return eventually decline as more is invested in a person because investment costs eventually rise as his foregone earnings rise, and benefits decline increasingly rapidly as his remaining working life falls (see the more

Figure 1



extended discussion in Becker 1975). The decline in marginal returns can be stated as

$$\psi_{xx} = \frac{\partial r_m}{\partial x} < 0. \tag{7}$$

The negatively inclined curves *HH* and *H'H'* in Figure 1 give the effect of parental expenditures on marginal rates of return for given endowments and public expenditures. *H'H'* is above *HH* because the endowment *E'* exceeds *E*.

Since nonhuman capital or assets can usually be purchased and sold in relatively efficient markets, presumably returns on assets are less sensitive than returns on human capital to the amount owned by any person. Little is known about the effect of abilities, other endowments, and wealth on returns from different assets, although some theory suggests a positive relation (see Ehrlich & Ben-Zion 1976). Our analysis requires only the reasonable assumption that returns on assets are much less sensitive to endowments and accumulations by any person than are returns on human capital (a similar assumption was made in Becker 1967, 1975). A simple special case of this assumption is that the rate of return on assets is the same to all persons.

Since much of the endowed luck of children (v_t) is revealed to parents prior to most of their investment in children, we assume that rates of return on these investments are fully known to parents (as long as the social environment (α_t) and public expenditures (s_{t-1}) are known). Parents must decide how to allocate their total "bequest" to children between human capital and assets. We assume initially that parents can borrow at the asset interest rate to finance expenditures on children and that this debt can become the obligation of children when they are adults.[1]

We also assume that parents try to maximize the welfare of children when no reduction in their own consumption or leisure is entailed. Then parents would borrow whatever is necessary to maximize the net income (earnings minus debt) of their children, which requires that expenditures on the human capital of children equate the marginal rate of return to the interest rate:

$$r_m = r_t, \text{ or } \hat{x}_{t-1} = g(E_t, s_{t-1}, r_t), \tag{8}$$

$$\text{with } g_E > 0 \text{ (by eq. (6)), } g_r < 0, \text{ and with } g_s < 0 \tag{9}$$

if public and private expenditures are substitutes. Parents can separate investments in children (an example of the separation theorem) from their own resources and altruism toward children because borrowed funds can be made the children's obligation.

The optimal investment is given in Figure 1 by the intersection of the horizontal "supply curve of funds," *rr*, with a negatively inclined demand curve (*HH* or *H′H′*). This figure clearly shows that better endowed children accumulate more human capital; those with the endowment E accumulate $0N$ units, while those with $E' > E$ accumulate $0N' > 0N$. Therefore, better endowed children would have higher expected earnings because equation (3) converts human capital into expected adult earnings. Since more is spent on better endowed children, the total effect of endowments on earnings, and the inequality and skewness in earnings

[1] Note, however, that the debt "bequeathed" to children is less than parental expenditures on rearing children when altruistic parents have the utility-maximizing number of children (see Becker & Barro 1986).

relative to that in endowments, is raised by the positive covariance between endowments and expenditures.

Clearly, an increase in the rate of interest reduces the investment in human capital, and hence earnings: Compare $0N$ and $0N'$ in Figure 1. The effect of an increase in public expenditures is less clear. If public expenditures were perfect substitutes dollar for dollar for private expenditures, the production of human capital would be determined by their sum ($x + s$) and by E; an increase in public expenditures would then induce an equal decrease in private (parental) expenditures, and the accumulation of human capital would be unchanged. Even then, since private expenditures cannot be negative, a sufficiently large increase in public expenditures would raise the accumulation of human capital.

Note that the human capital and earnings of children would not depend on their parents' assets and earnings because poor parents can borrow what is needed to finance the optimal investment in their children. However, the income of children would depend on parents because gifts and bequests of assets and debt would be sensitive to the earnings and wealth of parents. Indeed, wealthy parents would tend to self-finance the whole accumulation of human capital and add a sizable gift of assets as well.

Although the earnings and human capital of children would not be directly related to parental earnings and wealth, they would be indirectly related through the inheritability of endowments. The greater the degree of inheritability, the more closely related would be the human capital and earnings of parents and children. To derive the relation between the earnings of parents and children, substitute the optimal level of x given by equation (9) into the earnings-generating equation (3) to get

$$\begin{aligned} Y_t &= \psi[g(E_t, s_{t-1}, r_t), s_{t-1}, E_t] + l_t \\ &= \phi(E_t, s_{t-1}, r_t) + l_t, \end{aligned}$$

where (10)

$$\phi_E = \psi_g g_E + \psi_E = \frac{\partial Y}{\partial x}\frac{\partial x}{\partial E} + \frac{\partial Y}{\partial E} > 0.$$

Since this equation relates E to Y, l, g, and r, E_t can be replaced by E_{t-1} from (2), and thereby relate Y_t to Y_{t-1}, l_t, v_t, l_{t-1}, and other variables:

$$Y_t = F(Y_{t-1}, l_{t-1}, v_t, h, s_{t-1}, s_{t-2}, r_t, r_{t-1}, \alpha_t) + l_t. \quad (11)$$

The earnings of parents and children are linearly related when the effect of E_t on Y_t is independent of the level of endowments ($\phi_{EE} = 0$):

$$Y_t = c_t + \alpha_t\phi_E + bY_{t-1} + l_t^*,$$

where (12)

$$l_t^* = l_t - bl_{t-1} + \phi_E v_t, \text{ and } c_t = c(s_{t-1}, s_{t-2}, r_t, r_{t-1}).$$

Although the direct relation between the earnings of parents and children would be linear when $\phi_{EE} = 0$, c_t may differ among families if government expenditures (s_{t-1}, s_{t-2}) differed, and l_t^* would be negatively related to the market luck and earnings of parents.

Holding constant adult and children luck (l^*), earnings of children regress to the mean at the rate of $1 - b$. However, in an OLS regression of the actual earnings of children on the actual earnings of parents (Y_t on Y_{t-1}), the coefficient is biased downward by the generational "transitory" earnings of parents (l_{t-1}). If c_t is the same for all families, the expected value of the regression coefficient would equal

$$E(b_{t,t-1}) = b\left(1 - \frac{\sigma_l^2}{\sigma_y^2}\right), \tag{13}$$

where σ_l^2 and σ_y^2 are the variances of l_t and Y_t. This coefficient is closer to the degree of inheritability, the less important is generational transitory earnings in the total inequality in earnings.

Discrimination against families of particular races, religions, castes, or other characteristics transmitted across generations tends to reduce public expenditures (s_{t-1}), to reduce the effect of E_t on Y_t, and to increase the relative importance of stochastic forces (l) by reducing access to information about opportunities. Equation (12) indicates that such discrimination does not change the coefficient of parents' earnings, but raises the variability in the residual. *Even when earnings of parents are held constant,* minorities have lower earnings than others. Moreover, since the transitory component of earnings would be more important, the regression coefficient and the fraction of inequality explained by OLS regressions of children's on parents' earnings would be smaller among minority families. Note, however, that the total inequality in earnings may not be larger among minorities, and may be smaller, because the systematic component (determined by ϕ_E) would be smaller.

Imperfect Access to Capital

Access to capital markets to finance investments in children separates the transmission of earnings from the generosity and resources of parents. Economists have argued for a long time, however, that human capital is poor collateral to lenders. Children can "default" on the debt inherited from parents by working less energetically or by entering occupations with lower earnings and higher psychic income. Such "moral hazard" from the private nature of information about work effort and employment opportunities can greatly affect the earnings realized from human capital. Moreover, most societies are reluctant to enforce the collection of parental debts left to children, perhaps because the minority of parents who do not care much about the welfare of their children would raise their own consumption by leaving large debts to children.

To bring out sharply the effect of imperfect access to debt contracted for children, we assume that all parental investments in children must be financed either by selling assets or by reducing their own consumption. Consider parents without assets[2] who have to finance the efficient investment in human capital (say, $0N$ in Figure 1) partly by reducing their own consumption because they cannot contract debt for their children. Since the marginal utility of their consumption would then be raised relative to the marginal utility of resources invested in children, they would be induced to cut back their expenditure on children. Consequently, both the amount invested in children and parental consumption would be reduced by limitations on the debt contractable for children.

Further reductions in parental consumption are encouraged as less is invested because the marginal return from expenditures on children is raised. If parents maximize expected utility, the first order condition requires that the expected marginal utility from expenditures on children equals the marginal utility of parental consumption (assumed to be known) discounted by the known marginal return on investments in children. If parents are altruistic, with utility that is a separable and additive function of their own consumption and the utility of children, the first order condition is

$$\varepsilon_{t-1} \frac{\partial V_t}{\partial Y_t}(1 + r_m) = u'(Z_{t-1}), \tag{14}$$

[2]Even parents who accumulate assets over their lifetime may lack assets while investing in children.

where V_t depends on the utility of children, Z_{t-1} is parents' consumption, ε_{t-1} are expectations conditional on the information available to parents, and r_m is the marginal rate of return on their foregone consumption $(1 + r_m = -dY_t/dZ_{t-1})$.

Therefore, expenditures on children by parents without assets depends not only on endowments of children and public expenditures, as in equation (8), but also on earnings of parents (Y_{t-1}), their generosity toward children, and uncertainty (ε_{t-1}) about the luck of children and later descendants, as in

$$\hat{x}_{t-1} = g^*(E_t, s_{t-1}, Y_{t-1}, \varepsilon_{t-1}, V_t), \text{ with } g^*_Y > 0. \qquad (15)$$

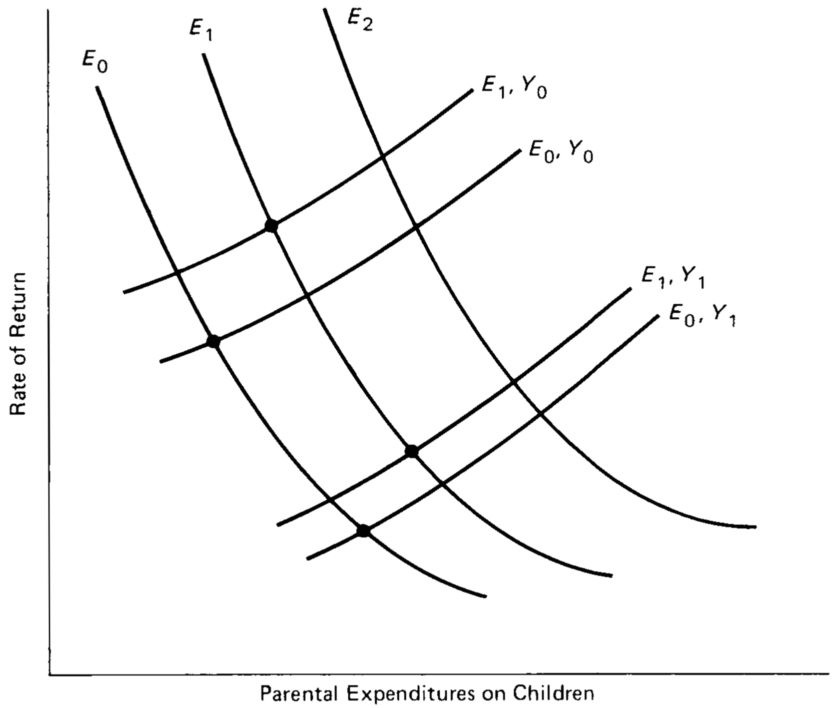
If public and private expenditures were perfect substitutes, g^* would depend on the sum of s_{t-1} and Y_{t-1} because an increase in public expenditures would then be equivalent to an equal increase in parental earnings. The effect of children's endowments on investments is now ambiguous ($\hat{g}^*_E \gtreqless 0$) because an increase in their endowments raises the resources of children as well as the productivity of investments in their human capital. Expenditures on children are discouraged when they are expected to be richer because parents must reduce their own consumption when they spend more on children.

The demand curves for expenditures in Figure 2 are similar to those in Figure 1 and are higher in families with better endowed children. The cost of funds to a family is no longer constant because increased expenditures on children lower the consumption of parents, and thereby raise their subjective discount rate (the shadow cost of funds). Moreover, supply curves are no longer the same for all families because the subjective cost of given expenditures is smaller to parents with higher earnings. Expenditures on children in each family are determined by the intersection of the supply and demand curve for that family. An increase in parental earnings shifts the supply curve to the right and induces greater expenditures on children. The distribution of intersection points determines the distribution of investments and rates of return, and hence, as shown in Becker (1967, 1975), the inequality and skewness in the distribution of earnings.

Substituting equation (15) into the earnings-generating equation (3) yields

$$\begin{aligned} Y_t &= \psi\{g^*(E_t, Y_{t-1}, r^*_{t-1}), E_t\} + l_t \\ &= \phi^*(E_t, Y_{t-1}, r^*_{t-1}) + l_t, \end{aligned} \qquad (16)$$

Figure 2



where r^*_{t-1} includes s_{t-1}, ε_{t-1}, and V_t. Earnings of children now depend directly on the earnings of parents as well as indirectly through the transmission of endowments. Some (e.g., Bowles 1972; Meade 1976; and Atkinson 1983) argue for a direct effect because "contacts" of parents are said to raise the opportunities of children; others argue for a direct effect because parents receive utility from the human capital of children. Fortunately, the effect of parental earnings on access to capital can be distinguished analytically from its effect on "contacts" or "utility."

The indirect effect of parents' earnings on the earnings of children operates through the transmission of endowments and can be found by substituting E_{t-1} for E_t, and then using equation (16) for E_{t-1}:

$$Y_t = F(Y_{t-2}, Y_{t-1}, l_{t-1}, v_t, h, \alpha_t, r^*_{t-1}, r^*_{t-2}) + l_t. \qquad (17)$$

Earnings of grandparents and grandchildren are directly linked because of the constraints on financing investments in children: Earnings of parents are not sufficient to describe the effects on children of both the resources and endowments of parents.

If Y_t were approximately linearly related to E_t and Y_{t-1}, then

$$Y_t \approxeq c'_t + (\beta + h)Y_{t-1} - \beta h Y_{t-2} + l^*_t, \text{ with } \beta = \phi^*_y.^3 \quad (18)$$

Equation (18) shows that an increase in the earnings of grandparents *lowers* the earnings of grandchildren *when parents' earnings and grandchildren's luck are held constant.* Therefore, restrictions on borrowing to finance investments in children introduces a negative relation between the earnings of grandparents and grandchildren and raises the effect of parents on children. The coefficient of parents' earnings exceeds the degree of inheritability by the marginal propensity to invest in children (β).

As in equation (13), OLS estimates of the coefficient of Y_{t-1} are biased downward by the transitory component of generational earnings. OLS estimates of the simple relation between Y_t and Y_{t-1} would tend toward

$$\beta \leq b_{t,t-1} = \frac{b_{t,t-1\cdot t-2}}{1 + h\beta} \leq \min\,(1, \beta + h, b_{t,t-1\cdot t-2}),^4 \quad (19)$$

where $b_{t,t-1\cdot t-2}$ is the partial regression coefficient between Y_t and Y_{t-1}. Therefore, both partial and simple regression coefficients between the earnings of parents and children provide *upper limits* of the effect of capital market constraints on investments in children. The biases in OLS estimates can be overcome by using instruments for the earnings of parents, such as the earnings of uncles or great-grandparents (see Goldberger 1979; and Behrman & Taubman 1983).

[3] A similar equation is derived in Becker and Tomes (1979, eq. [25]); however, the coefficient called β there refers to the propensity to bequeath all capital, *including debt,* to children, not the propensity to invest in the human capital of children by parents who cannot leave debt.

[4] Equation (18) implies that

$$\begin{aligned} b_{t,t-1} &\approxeq \beta + h - h[b_{(\beta y_{t-2} + l_{t-1})\cdot y_{t-1}}] \\ &\approxeq \beta + h - \frac{h\sigma^2_l}{\sigma^2_y} - h\beta b_{t-1,t}. \end{aligned}$$

If the economy is in long-run equilibrium (see Becker & Tomes 1979), then $b_{t,t-1} = b_{t-1,t}$, $\sigma^2_{y_{t-1}} = \sigma^2_{y_t}$, and the equality in (19) follows. The relation between $b_{t,t-1}$ and the right-hand side of (19) is derived in Becker and Tomes (1979 app. E).

Rich families can usually far more readily self-finance investments in children than can poor and middle level families. Since our analysis implies that richer children have better than average endowments, presumably either the variability in transitory earnings substantially exceeds the variability in the effect of endowments on earnings, or endowments regress significantly to the mean so that the endowments of richer children are much below those of their parents, or both. If returns on assets are not highly sensitive to earnings and endowments, equilibrium *marginal* rates of return on investments in children would be lower in richer families than in constrained poor and middle level families, even though endowments and average rates of return are higher in richer families. Equilibrium marginal rates would tend to decline, but perhaps not monotonically, as parental earnings rise. Eventually, marginal rates on human capital would equal the rate of return on assets, and then marginal rates would be relatively constant as their earnings rose.

If marginal rates were lower in richer families, a small redistribution of human capital away from these families and toward children from less advantaged families would *raise* the average marginal rate of return and hence would raise efficiency, even though endowments and the average productivity of investments in children were greater in richer families (see also Becker 1967, 1975). The usual conflict between "equity" (measured by inequality) and efficiency is absent because a redistribution of investments toward less advantaged children has a similar effect to an improvement in the efficiency of capital markets.

Larger public expenditures on the human capital of children would not raise their total capital if the efficient amount had been invested in children and if public and private expenditures were perfect substitutes: Parental expenditures would be reduced dollar for dollar of the increased public expenditures. However, if parental expenditures are constrained by limited access to capital, increased public expenditures would raise the human capital of children by raising family resources (assuming taxes to finance these public expenditures are imposed on other families). The increase is shared between parents and children in a ratio determined by the marginal propensity to invest in children (β).

Some have argued that Head Start and other programs to raise the achievements of poor children are doomed to failure because of the relatively low abilities and other endowments of the children participating (see, e.g., Jensen 1969). Although our analysis does imply that endowments and the *average* productivity of expenditures on human capital are lower in poor families, it also implies that *marginal* rates of return may well be ***higher*** in poor families. This implication appears to suggest that compensatory programs for the poor, if administered reasonably well, would raise the achievements of participants.

Such a conclusion, however, neglects "compensatory" responses by parents. To further equity toward other members, parents redistribute their time and expenditures away from participants to siblings and themselves, which would offset the effects of these programs on participants. Parental compensatory responses apparently offset the effect of public health programs, food supplements to poorer pregnant women, and social security programs, as well as perhaps some Head Start programs (see the discussion in Becker 1981, pp. 125–26, 251–53). If public and private expenditures on the human capital of children were perfect substitutes, the effect of public expenditures on earnings would be determined by the propensity of parents to invest in children (β). The evidence discussed in Becker and Tomes (1984) implies that compensatory public expenditures are largely offset by reduced parental expenditures because the propensity to invest appears to be less than .25.

Endowments do not change much during the prime adult years of the life cycle and can be said to be a "fixed effect" or fully "inherited" (h = 1). If a person's total human capital were approximately constant during these years—as determined by parental investments augmented by on-the-job and other adult training—the regression coefficient between earnings at different ages would equal $1 - \sigma_l^2/\sigma_y^2$, where σ_l^2/σ_y^2 is the share of transitory earnings in the inequality in annual earnings. Since constraints on borrowing to invest in children appear to have only a moderate effect ($\beta \leqq .25$), earnings would be more closely related over the life cycle than across generations because endowments are quite imperfectly transmitted to children (they are not a "fixed effect" across generations).

Summary and Conclusions

This paper develops a model of the transmission of earnings and assets from parents to children and later descendants. The model is based on utility maximizing by parents concerned about the welfare of their children. The degree of intergenerational mobility, or the rise and fall of families, is determined by the interaction of utility-maximizing behavior with investment and consumption opportunities in different generations, including different kinds of luck.

Cultural and genetic endowments are assumed to be automatically transmitted from parents to children, with the relation between the endowments of parents and children determined by the degree of "inheritability." The intergenerational mobility of earnings is shown to be closely related to the inheritability of endowments, even when earnings depend mainly on the human capital invested in children. Indeed, if all parents can readily borrow to finance the optimal investments in children, the degree of intergenerational mobility in earnings essentially would equal the inheritability of endowments.

However, poor families may have difficulty financing investments because loans to supplement their limited resources are not readily available when human capital is the collateral. Such capital market restrictions lower investments in children from poorer families. Intergenerational mobility in earnings would then depend not only on the inheritability of endowments, but also on the willingness of poorer families to self-finance investments in their children.

Unlike human capital, assets are traded in markets with relatively similar rates of return to well-endowed and less well-endowed persons. Therefore, assets can act as a buffer to offset regression to the mean in the endowments, and hence in the earnings, of children. In particular, successful families would bequeath assets to children to offset the downward regression in the earnings of their children. As a consequence, assets would regress more slowly to the mean than earnings and might even regress away from the mean when earnings regressed rapidly to the mean.

Comment

RUSSELL HARDIN

Once you grant Gary Becker his assumptions, you generally have to agree with his conclusions. So I will quibble with one of his assumptions: that the number of children in each generation is the same as the number of parents. I will take it on in a way that I think is fully consistent with what Becker wants to do. I will assume that the decision as to how many children to have is a matter of parental choice of investments in the next generation and therefore of course in further generations.

The major element in Becker's account is the role of wealth in deciding *how much to invest* in one's children. I will assume that wealth is also an important part of the decision of parents when they go about deciding *how many* children to have. This decision is subject to the same rational analysis that Becker uses. The only charge I have against what he has done is that he has not gone far enough. In particular, that part of the parental decision on investment in children that he assumed away, namely, how many children to have, plays a major role in explaining the circulation of elites and the circulation of the impoverished or the poor. I think it explains what one might call the disappearance of great families and by implication their replacement by others who previously were not great.

Wealth plays a very important role in addition to those that Becker has it playing. It provides a cushion for risk-taking. The well-off can risk having fewer children because they do not need to depend on their children's survival for their own. The poor need their children for their later lives.

As the sociologist James Scott has argued, a peasant who lives on the margin of subsistence will invest in seed varieties whose yield is low but stable and will plant in several small plots scattered all over rather than in one big, efficient plot for the purpose of ensuring against the worst outcome, which is failure of the crop and starvation. In order to ensure survival, the peasant avoids the opportunity to become well off. That's not because peasants are more risk-averse than the rest of us but because the costs that they might incur if the risk goes wrong are much worse than what middle-class and well-off people may suffer from their risks.

Peasants probably view their children the way they do seed varieties. They prefer low variability and security to high variability and significant risk. Security means support in their old age. They want to maximize prospects for such support. They have many children, and the only investment they can make in them, other than keeping them fed, is the investment in the apprenticelike working conditions that the children grow up in to become essentially identical to their parents in the next generation.

The well-to-do are in a different position. They can set about deliberately producing essentially high-quality children. They can invest in one only. The only problem with that is the statistics of small numbers. If your family line has only one scion in each succeeding generation, the half-life of your family will be short. With a single accident, illness, or genetic failure, the well-to-do parents will not be grandparents and the family will come to an end. Focusing all investment on one or two children, though, suppresses regression toward the mean for those children who do come along. The well-to-do children do well with that kind of investment behind them.

One would have to have more than the knowledge that great families disappeared from the rolls to justify Becker's claim that what we see in the data is regression toward the mean. The appearance of new great families would still remain to be explained if my account of why the old ones disappear, rather than regress, is correct. One would conclude that circulation among the underprivileged exceeds circulation among the elite, as Becker notes, for the simple reason (not Becker's) that the elite do not circulate but exit. They don't go down, they just go out, leaving vacancies to be filled by children of those who had many children in a generation.

The supposition that the well-off have few children and therefore exterminate themselves is not original. It's an old worry among nineteenth-century French demographers who saw the upstart Germans surpassing the demographically stable French because the French had family sizes of two to three children rather than the typical German family size of half a dozen and more.

There have been past efforts to explain extinction of well-off families as a function of their wealth, but to my knowledge no one has done the Becker-type explanation that I have just presented. The other efforts usually suppose that wealth is somehow hazardous, that the wealthy die because they are wealthy. But the rational explanation which is entirely within Becker's own framework of explanation is a cogent way of understanding such trends. While all I have done is to recommend how one might go about adding to Becker's kind of analysis, I think that the result of including this additional element of choice, how many children to have, makes an enormous difference in one's account of what actually happens. If the initial choice of investment in children should include this other factor, we might get a more complete account of the historical issues that have interested the sociologists Becker cites as motivating his own analysis.

General Discussion

Gary Becker: I agree that fertility is important; indeed, our complete paper analyzes the effects of fertility on the relation between the earnings and assets of parents and children. We show that a negative relation between fertility and parents' income tends to reduce the degree of regression to the mean in both earnings and assets and that a decline in the average level of fertility raises the regression to the mean in earnings and lowers it in assets.

My reading of the historical evidence [see Becker 1981, chap. 5] is that until the nineteenth century, income and number of children were positively related, particularly after correcting for the negative relation between income and child mortality. The relation between income and fertility changed during the nineteenth century in Western countries to a negative relation, a change that should have reduced the degree of intergenerational mobility.

Howard Aldrich: I have a question that has to do with the distribution of wealth. I am curious whether the question of transmitting wealth would be relevant for most of the population in the period you are dealing with. For nineteenth-century urban areas, about 20 percent of the population have 80 percent of the wealth. This meant that for most of the population the question of how wealth affected their ability to transmit would be irrelevant. The best current study we have is the Projector and Wise 1962 Federal Reserve Board study which showed that about half of the population had zero wealth if their assets were sold off to take care of their liabilities. Only about 6 percent of the population had 40 percent of the wealth. This means that you would have a large section of the population which is zero on your independent variables. There isn't any wealth to transmit. That leads me to the second question. You said for some people there is no wealth, but they are saving human capital. I have been trying to think what the poor can save, what human capital the poor can save and I can't imagine what that is.

Gary Becker: Note that our model *implies* rather than assumes that the distribution of tangible wealth is highly skewed, with a small fraction of the population owning much of the tangible wealth. So the facts cited by Aldrich are fully consistent with our analysis.

When we say that investments in children are limited by parents' wealth, earnings or human capital are included as part of wealth. The poor can and do invest in many kinds of human capital, including expenditures on the food and care that affect the death rate and other dimensions of the children's health. They also invest by providing for schooling and other training. These are investments by parents because they forego earnings from children when they are in school or at other training programs.

Michael Hechter: I wonder if you could tell us briefly why you made a decision to include altruism in the parents' utility function.

Gary Becker: Altruism plays a crucial role in the organization, production, and distribution of household resources [see the extended discussion in Becker 1981, chap. 8], although several other assumptions about the attitude of parents toward children could have served equally well for most issues considered in our paper. Clearly, however, an assumption that parents have strong feelings, attachments, and love for their children is not inconsistent with common views about parents' attitudes toward children.

Victor Vanberg: For its explanatory power, the rational choice approach is dependent on the empirical content of its basic behavioral assumptions. I see the danger that including altruism in the utility function eliminates any empirical content. The traditional strategy of the rational choice approach is to assume something like self-interested behavior and to make a distinction between the motivation and the results of behavior. If the results of behavior don't look like the outcome of self-interest, we would not therefore abandon the assumption of self-interest but have a closer look at the constraints. Don't you see the problem of losing empirical content?

Gary Becker: I do not believe that is a weakness of our analysis. Why should altruism be excluded from preferences or from rational choice models more generally? Certainly, there are no major analytical difficulties that justify its exclusion. Moreover, altruism has several striking and testable implications about behavior [see Becker 1981, chap. 8] that could be investigated empirically. As long as preferences are assumed to be *stable*, one is protected against arbitrary shifts in assumptions to "explain" particular bodies of evidence.

Siegwart Lindenberg: I would like to point to yet another problem. In many countries, conscious policy-making attempts to influence the correlations you are studying. Take a simple example. The relation between social class and the health of children's teeth is purposefully influenced by programs that bring dentists into schools. Basically, parents' investment decisions are neutralized by such policies. The beta weights of your model would thus have to be interpreted against the background of public policy that is aimed at making parental investment decisions irrelevant, especially at the lower end.

Gary Becker: We agree that public policies can significantly affect the degree of mobility. For example, large subsidies to education may have contributed to the weak relation between the earnings of parents and children in the United States and Scandinavia (see the discussion in our larger paper). However, we claim that the net effect of many policies are not clear because of reaction by the families affected. In particular, the apparent effect of Head Start programs on participating children are rather weak probably because of "compensatory" responses by their parents.

Ecological Theory

The Ecology of Organizations: Structural Inertia and Organizational Change

MICHAEL T. HANNAN AND
JOHN FREEMAN

Human ecology seeks to explain the forms of human social systems and their development (Hawley 1968). It directs attention primarily to the ways in which environmental conditions interact with internal processes of development to shape the forms of social organizations. In so doing, it attempts to understand the forces that control the diversity of forms of social organization.

In sociology, ecological theory has been applied primarily in the study of communities (territorially based social systems). However, ecological reasoning has also played an important role in many kinds of macrosociology (although the ideas are seldom called "ecological"). This paper considers issues that arise in one such application—the ecology of organizations.

Note: The work reported here was supported by National Science Foundation grant SES-8109382. Glenn Carroll, Jeffrey Pfeffer, Susan Olzak, and Arthur Stinchcombe made helpful comments on an earlier draft. An extended version of this paper appears in the *American Sociological Review* 49 (1984):149–64.

The conventional wisdom in organizational theory holds that the diversity of organizations in society reflects primarily the history of adaptations by individual organizations. Earlier (Hannan & Freeman 1977), we challenged this view and argued that adaptations of organizational structures to environments occurs principally at the *population* level, with forms of organization replacing one another as conditions change. This initial statement of population ecology theory rested on a number of simplifying assumptions. A major one was the premise that individual organizations are subject to strong inertial forces; that is, they seldom succeed in making radical changes in strategy and structure in the face of environmental threats.

How strong are inertial forces on organizational structure? This question is substantively interesting in its own right. It is also strategically important, because the claim that adaptation theories of organizational change should be supplemented by population ecology theories depends partly on these inertial forces being strong.

Many popularized discussions of evolution suggest that selection processes invariably favor adaptable forms of life. In fact the theory of evolution makes no such claim, as we made clear earlier (Hannan & Freeman 1977; Freeman & Hannan 1983). This paper goes beyond our earlier theory in acknowledging that organizations change in some ways frequently and occasionally manage to make radical changes in strategy and structure. Nevertheless, we argue that selection processes tend to favor organizations whose structures are difficult to change. That is, we claim that high levels of structural inertia in organizational populations can be explained as an outcome of an ecological-evolutionary process.

Background

Our earlier formulation of an ecological theory of organizational change pointed to a variety of constraints on structural change in organizations:

> for wide classes of organizations there are very strong inertial pressures on structure arising from both internal arrangements (for example, internal politics) and from the environment (for example, public legitimation of organizational activity). To claim otherwise is to ignore the most obvious feature of organizational life. Failing churches do not become retail stores, nor do firms turn themselves into churches. [Hannan & Freeman 1977, p. 957]

Some of the factors that generate structural inertia are internal to organizations: These include sunk costs in plant, equipment, personnel, the dynamics of political coalitions, and the tendency for precedents to become normative standards. Others, including legal and other barriers to entry and exit from realms of activity, investments in networks of exchange relations with other organizations, and the threat of losing institutional support as a consequence of attempting radical change in strategy and structure, are external.

We continue to believe that inertial pressures on most features of organizational structure are quite strong—much stronger than the mainstream view acknowledges. Moreover, the assumption that organizations rarely make fundamental changes successfully has proven to be a useful strategic simplification. It has allowed a rich and evocative set of ecological theories and models to be applied to the problem changes in organizational form over time (see, e.g., Brittain & Freeman 1980; Carroll 1983; Carroll & Delacroix 1982; Freeman 1982; Freeman & Hannan 1983; Freeman, Carroll, & Hannan 1983).

However, the claim that organizational structures change rarely is the subject of dispute. March (1981, p. 563) begins his review of research on organizational change by asserting that "organizations are continually changing, routinely, easily, and responsively, but change within organizations cannot be arbitrarily controlled . . . What most reports on implementation indicate . . . is not that organizations are rigid and inflexible, but that they are impressively imaginative."

The contemporary literature contains at least three broad points of view on organizational change. Population ecology theory holds that most of the variability in organizational structures comes about through the creation of new organizations and organizational forms and the replacement of old ones (Hannan & Freeman 1977; Freeman & Hannan 1983; McKelvey 1982). A second view, which might be called rational adaptation theory, proposes that organizational variability reflects designed changes in strategy and structure of individual organizations in response to environmental changes, threats, and opportunities. There are numerous variants of this perspective which differ widely on other dimensions. Contingency theories emphasize structural changes that match organizational structures to technology-environment pairs (Thompson 1967; Lawrence & Lorsch 1967). Resource dependence theories emphasize structural changes that neutralize sources of environmental uncertainty (Pfeffer & Salancik 1978). An institutionally oriented version of this perspective holds that organizational structures are rationally adapted to prevailing normatively endorsed modes of organizing (Meyer & Rowan 1977; DiMaggio & Powell 1983). Marxist theories of organization typically assert that organizational structures are rational solutions for capitalist owners to the problem of maintaining control over labor (Edwards 1979; Burawoy 1979). The third broad perspective, which might be called random transformation theory, claims that organizations change their structures mainly in response to endogenous processes but that such changes are only loosely coupled with the desires of organizational leaders and with the demands and threats of environments (March & Olsen 1976; March 1981; Weick 1976).

The selection and adaptation perspectives are so different that it is hard to believe that they are talking about the same things. If in fact they are not, it is useful to clarify the conditions under which the two perspectives apply.

Transformation and Replacement

All accepted theories of biotic evolution share the assumption that innovation, the creation of new strategies and structures, is random with respect to adaptive value. Innovations are not produced because they are useful; they are just produced. If an innovation turns out to have adaptive value, it will be retained and spread through the population with high probability. In this sense, evolution is blind. How can this view be reconciled with the fact that human actors devote so much attention to predicting the future and to developing strategies for coping with expected events? Can social change, like biotic evolution, be blind?

Almost all evolutionary theories in social science claim that social evolution has foresight, that it is Lamarckian rather than Darwinian in the sense that human actors learn by experience and incorporate learning into their behavioral repertoires (see, for example, Nelson & Winter 1982). To the extent that learning about the past helps future adaptation, social change is indeed Lamarckian—it transforms rather than selects. In other words, major change processes occur *within* behavioral units.

Even when actors strive to cope with their environments, action may be random with respect to adaptation as long as the environments are highly uncertain or the connections between means and ends are not well understood. It is the *match* between action and environmental outcomes that must be random on the average for selection models to apply. In a world of high uncertainty, adaptive efforts by individuals may turn out to be essentially random with respect to future value.

The realism of Darwinian mechanisms in organizational populations also turns on the degree to which change in organizational structures can be controlled by those ostensibly in command. Suppose that individuals learn to anticipate the future and adapt strategies accordingly and that organizations simply mirror the intentions of rational leaders. Then organizational adaptations would be largely nonrandom with respect to future states of the environment. On the other hand, if March and others are right, organizational change is largely uncontrolled. Then organizations staffed by highly rational planners may behave essentially randomly with respect to adaptation. In other words organizational outcomes may be decoupled from individual intentions; organizations may have lives of their own. In this case it is not enough to ask whether individual humans learn and plan rationally for an uncertain future. One must ask whether organizations as collective actors display the same capacities.

The applicability of Darwinian arguments to changes in organizational populations thus depends partly on the tightness of coupling between individual intentions and organizational outcomes. At least two well-known situations generate loose coupling: diversity of interests among members and uncertainty about means-ends connections. When members of an organization have diverse interests, organizational outcomes depend heavily on internal politics, on the balance of power among the constituencies. In such situations outcomes cannot easily be matched rationally to changing environments.

When the connections between means and ends are obscure or uncertain, carefully designed adaptations may have completely unexpected consequences. Moreover, short-run consequences may often differ greatly from long-run consequences. In such cases, it does not seem realistic to assume a high degree of congruence between designs and outcomes.

In earlier papers we emphasized another facet of the problem. The environment of each organization consists mainly of other organizations. Each of them tries in some degree to control its environment, to predict environmental variations, and to take appropriate action. Consequently the environment for each actor is highly unpredictable. The paradoxical situation is that Lamarckian flexibility, which allows each organization to change strategy and structure as events unfold, makes the consequences of actions and strategies all that much more uncertain.

Learning and adjusting structure enhances the chance of survival only if the speed of response is commensurate with the temporal patterns of relevant environments. Indeed, the worst of all possible worlds is to change structure continually only to find each time upon reorganization that the environment has already shifted to some new configuration that demands yet a different structure. Learning and structural inertia must be considered in a dynamic context. Can organizations learn about their environments and change strategies and structures as quickly as their environments change? If the answer is negative, replacement or selection arguments are potentially applicable.

To claim that organizational structures are subject to strong inertial forces is not the same as claiming that organizations never change.

Rather, it means that organizations respond relatively slowly to the occurrence of threats and opportunities in their environments. Therefore, structural inertia must be defined in relative and dynamic terms. It refers to comparisons of the typical rates of change of the processes identified above. In particular, structures of organizations have high inertia when the speed of reorganization is much lower than the rate at which environmental conditions change. Thus, the concept of inertia, like fitness, refers to a correspondence between the behavioral capabilities of a class of organizations and their environments.

We continue to argue that it is useful in analyzing patterns of long-term change in organizational forms to emphasize Darwinian mechanisms rather than Lamarckian ones. The fact that members of organizations plan rationally for change and that organizations often develop structures designed to plan and implement change does not undercut the value of this view as long as organizations are political coalitions and as long as environments of organizations change.

Reproducibility, Inertia, and Selection

The creation of a permanent organization as a solution to a problem of collective action is costly compared with other alternatives. Why do individuals and other social actors agree to commit scarce resources to such expensive solutions to problems of collective action? A number of answers to this question have been put forth (see Scott 1981, pp. 135–63, for an insightful review). The new institutional economics argues that organizations arise to fill the gaps created by market failure (Arrow 1974). Williamson's influential analysis (1975) proposes that organizations are more efficient than markets in situations in which economic transactions must be completed in the face of opportunism, uncertainty, and small-numbers bargaining. Although sociologists tend to deny that organizations arise mainly in response to market failures, they tend to agree that organizations have special efficiency properties, but emphasize their efficiency and effectiveness for coordinating complex tasks (Blau & Scott 1962; Thompson 1967).

Although these efficiency arguments are plausible, it is not obvious that they are correct. From the perspective of the performance of a *single* complex collective action, it is not obvious that a permanent organization has any technical advantage.

We emphasize different kinds of competencies. The first of these is *reliability*. Organizations have unusual capacities to produce collective products of a given quality repeatedly. In a world of uncertainty, potential members, investors, and clients may value reliability of performance more than efficiency. That is, rational actors may be willing to pay a high price for the certainty that a given product or service of a certain minimum quality will be available when it is needed. Reliability depends on the *variance* of performance (including its timeliness) rather than its average level.

Organizations have higher levels of reliability than ad hoc collectives in two senses: one cross-sectional and the other temporal. Cross-sectional reliability means that an outcome chosen at random from a population of organizations will have a lower variance than one chosen at random from a population of other kinds of producers. Temporal reliability means variability over time in the quality (including timing of delivery) of an outcome is lower for those produced by organizations than for those produced by ad hoc groups. Overall we argue that the distinctive competence of organizations is the capacity to generate collective actions with relatively small variance in quality.

Organizations have a second property that gives them an advantage in the modern world: *accountability*. The spread of general norms of rationality in the modern world (Weber 1968) and a variety of internal and external contingencies demand that organizations be able to account rationally for their actions. Testing for accountability is especially intense during organization building, the process of initial resource mobilization. Potential members want assurance that their investments of time and commitment will not be wasted. When membership involves an employment relation, potential members often want guarantees that careers within the organization are managed in some rational way. Potential investors (or supporters) also assess accountability. In fact, the profession of public accountancy arose in the United States in response to the desires of British investors in American railroads for assurances that their investments were being managed in appropriate ways (Chandler 1977). Demands for accounting rationality in this narrow sense are both widespread and intense in modern societies. For example, the federal government will not allocate research grants and contracts to organizations that have not passed a federal audit, meaning that they have given evidence of possessing the appropriate rules and procedures for accounting for the use of federal funds.

Accountability testing is also severe during periods in which resources contract. Members and clients who would otherwise be willing to overlook waste typically change their views when budgets and services are being cut.

In our judgment pressures for accountability are especially intense when (1) organizations produce symbolic or information-loaded products (e.g., education, branded products versus bulk goods)—see DiMaggio and Powell (1983), (2) when substantial risk exists (e.g., medical care), (3) when long-term relations between the organization and its employees or clients is typical, and (4) when the organization's purposes are highly political (Weber 1968). Our arguments presumably apply with special force to organizations in these categories. Still, we think that pressures toward accountability are generally strong and getting stronger. The trend toward litigating disputes and pressures for formal equality in modern politics intensify demands for accountability. All organizations seem to be subject to at least moderate levels of accountability testing.

We argue that the modern world favors collective actors that can demonstrate or at least reasonably claim a capacity for reliable performance and can account rationally for their actions. These forces favor organizations over other kinds of collectives and they favor certain kinds of organizations over others, since not all organizations have these properties in equal measure. Selection within organizational populations tends to eliminate organizations with low reliability and accountability. The selection processes work in several ways. Partly they reflect testing by key actors and environments in the organization-building stage. Potential members, investors, and other interested parties apply tests of reliability and accountability to proposed new ventures. Such testing continues after founding. Unreliability and failures of accountability at any stage in a subsequent lifetime threatens an organization's ability to maintain commitment of members and clients and its ability to acquire additional resources.

> *Assumption 1.* Selection in populations of organizations in modern societies favors forms with high reliability of performance and high levels of accountability.

When does an organization have the capacity to produce collective outcomes of a certain minimum quality repeatedly? The most important prerequisite is so commonplace that we take it for granted. Reliable performance requires that an organization continually reproduce its structure; it must have very nearly the same structure today that it had yesterday. Among other things, this means that structures of roles, authority, and communication must be reproducible from day to day.

> *Assumption 2.* Reliability and accountability require that organizational structures be highly reproducible.

In general organizations attain reproducibility of structure through processes of institutionalization and by creating highly standardized routines. The first solution, institutionalization, is a two-edged sword. It greatly lowers the cost of collective action by giving an organization a taken-for-granted character such that members do not continually question organizational purposes, authority relations, etc. Reproduction of structure occurs without apparent effort in highly institutionalized structures. The other edge of the sword is inertia. The very factors that make a system reproducible make it resistant to change.

As a brake on structural change institutionalization applies both to the organization as a whole and to its subunits. But what about the diversity among sets of differentiated activities within the organization? Some kinds of organizations perform diverse sets of activities, sometimes in parallel and sometimes sequentially. Military organizations provide a striking example; they maintain "peacetime" and "wartime" structures.[1] Similarly, labor unions gear up for organizing drives or for waves of strikes and then return to more placid bread and butter collective bargaining. Manufacturing firms sometimes concentrate on redesigning products and at other times concentrate on marketing an extant set of products. Each phase of organizational activity involves mobilizing different kinds of structures of communication and coordination. In a real sense these kinds of organizations can be said to use different structures in different phases.

Does this mean that these organizations have somehow escaped inertial tendencies? We think not, at least from the perspective of attempts at building theories of organizational change. These organizations have multiple *routines;* they shift from one routine (or set of routines) to another in a fairly mechanical fashion. We think that organizations have high inertia both in the sets of routines employed and in the set of rules used to switch between routines.

According to Nelson and Winter (1982, p. 96) routines are the "source of continuity in the behavioral patterns of organizations." They are patterns of activity that can be invoked repeatedly by members and subunits. One way of conceiving of routines is as organizational memory—an organization's repertoire of routines is the set of collective actions that it can do from memory. But Nelson and Winter emphasize that organizations *remember by doing*. Like knowledge of elementary algebra or high school Latin, collective knowledge, which is the basis of organizational routinues, decays rapidly with disuse. This fact implies that organizations face the classic specialism-generalism dilemma in deciding how

[1]Janowitz (1960) discusses the various conflicting demands of organizing military activities in peacetime and war. Etzioni (1975) discusses the shifts in control problems that arise in armies and labor unions as a result of such changes.

many routines to maintain at any fixed level of resources. But generalists (those with many routines) are not subject to less inertia in the manner in which they adapt to environmental change in the sense that they still use a limited number of routines. As Nelson and Winter (1982, p. 134) put it:

> it is quite inappropriate to conceive of firm behavior in terms of deliberate choice from a broad menu of alternatives that some outside observer considers to be "available" to the organization. The menu is not broad, it is narrow and idiosyncratic. . . . Efforts to understand the functioning of industries and larger systems should come to grips with the fact that highly flexible adaptation to change is not likely to characterize the behavior of individual firms.

We think that it is a reasonable first approximation to think of organizations as possessing relatively fixed repertoires of highly reproducible routines. Then the argument of this paper can be applied either to the organization as a whole, where the issue is the diversity of the repertoire, or to the individual routine. Thus, we argue that the properties that give some organizations reproducibility also make them highly resistant to structural change, whether designed or not. Resistance to structural change is a likely by-product of the ability to reproduce a structure with high fidelity:

> *Assumption 3.* High levels of reproducibility of structure generate strong inertial pressures.

The three assumptions form the core of our first argument. Taken together they imply:

> *Theorem 1.* Selection within populations of organizations in modern societies favors organizations whose structures have high inertia.

This theorem reverses the role claimed earlier (Hannan & Freeman 1977) for structural inertia in organizational ecology and evolution. It states that structural inertia can be a *consequence* of selection rather than a precondition. All that is required is that some organizations in an initial population have high levels of reproducibility (hence high levels of inertia) and that selection pressures be reasonably strong. Under such conditions, selection pressures in modern societies favor organizations whose structures are resistant to change, which makes selection arguments all the more applicable.

In addition to varying by aspects of structure, the strength of inertial forces may also vary with life cycle phase, size, and complexity. The remainder of the paper considers these issues.

Life Cycle Variations in Inertia

Newly created organizations apparently have lower levels of reproducibility than older ones. As Stinchcombe (1965) pointed out, new organizations typically have to rely on the cooperation of strangers. Development of trust and smoothly working relationships takes time. It also takes time to work out routines. Initially there is much learning by doing and comparing alternatives. Most observers of industrial organization assume the existence of a learning curve that has a steep slope initially but quickly bottoms out. Such arguments also apply to nonmarket organizations such as labor unions or universities. It also takes time for an organization to acquire institutional reality to its members and to become valued in its own right.

> *Assumption 4.* Reproducibility of structure increases monotonically with age.

> *Theorem 2.* Structural inertia increases monotonically with age. (From Assumptions 2 and 1)

> *Theorem 3.* Organizational death rates decrease with age. (From Assumption 4 and Theorem 1)

Theorem 3, often called the "liability of newness" hypothesis (Stinchcombe 1965), has been well documented empirically. (For recent evidence, see Freeman, Carroll, & Hannan 1983.) Death rates appear to decline approximately exponentially as organizations age. This finding suggests that reproducibility rises roughly exponentially with age over the early years in an organization's life.

Processes of external legitimation also take time. Although an organization must have some minimal level of public legitimacy in order to mobilize sufficient resources to begin operations, new organizations (and especially new organizational forms) have rather weak claims on public and official support. Nothing legitimates both individual organizations and forms more than longevity. Old organizations tend to develop dense webs of exchange, to affiliate with centers of power, and to acquire an aura of inevitability. Thus, processes of institutionalization in the environment and exchange relationships with key sectors of the environment may account for the relationships stated in Theorems 2 and 3. The argument to this point cannot distinguish between the internal and external sources of the relationships.

Size and Inertia

We argued above that dampened response to environmental threats and opportunities is the price paid for reliable and accountable collective action. If this argument is correct, organizations respond more slowly than individuals on average to environmental changes. However, some organizations are little more than extensions of the wills of dominant coalitions or individuals; they have no lives of their own. Such organizations may change strategy and structure in response to environmental changes almost as quickly as the individuals who control them. Change in populations of such organizations may operate as much by transformation as selection.

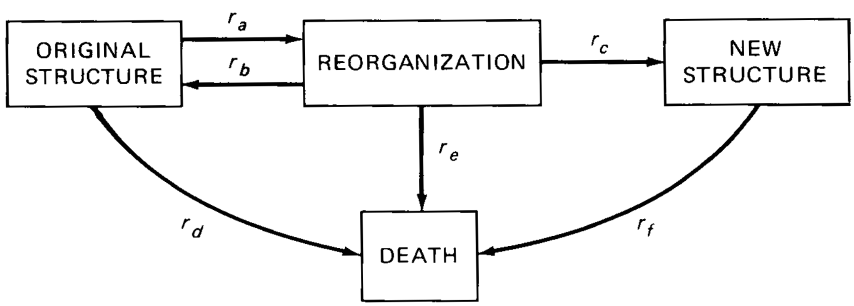
But a large organization can be a simple tool of a dominant leader only when the leader does not delegate authority and power down long chains of command. Failure to delegate usually causes problems in large organizations. Indeed, the failure of moderate-sized organizations is often explained as resulting from the unwillingness of a founder-leader to delegate responsibility as the organization grew.

One way to conceptualize the issues involved is to assume that there is a critical size, which may vary by form of organization (and also, perhaps, by age), at which failure to delegate power sharply limits viability. In such a threshold model, organizations may be quite responsive below the threshold level of size. Above the threshold, organizations tend to have higher inertia. Or the relationship between size and inertia may be roughly continuous. Downs (1967, p. 60) argues that for the case of public bureaus ". . . the increasing size of the bureau leads to a gradual ossification of its action . . . the spread and flexibility of its operation steadily diminish." Whether there is a threshold, as we have suggested, or a continuous relationship, as Downs suggested, it seems clear to us that size does affect inertia.

> *Assumption 5.* The level of structural inertia increases with size for each class of organizations.

Assumption 5 seems to suggest that selection arguments are more appropriate for large organizations than for small ones, contrary to widespread opinion (Aldrich 1979; Perrow 1979; Scott 1981; Astley & Van de Ven 1983). However, the situation is more complex than this. The likelihood that an organization adjusts structure to changing environmental

Figure 1 State space for the process of fundamental change in organizational structure. (The r's are instantaneous transition rates.)



circumstances depends on two factors: the rate of undertaking structural change and the probability of succeeding in implementing change, given an attempt. Assumption 5 suggests that the first quantity, the rate of attempting change, is higher for small organizations. But what about the second quantity?

It is helpful in answering this question to complicate the model slightly. Fundamental reorganization may sometimes occur gradually and imperceptibly. But sometimes sharp breaks with the past can be discerned, and one can identify the approximate time of onset of the reorganization. In such cases it may be helpful to introduce a new state into the model: the state of attempting fundamental reorganization. Figure 1 depicts the possible transitions in the expanded state space. The parameters associated with each transition, the *r*'s, are instantaneous transition rates. In terms of this representation, Assumption 5 states that the rate of moving to the state of reorganization increases with size. But it says nothing about the other rates.

The processes of dismantling one structure and building another make organizational action unstable. Consequently, the variance of quality increases and timeliness of collective action declines during reorganization.

> *Assumption 6.* The process of attempting reorganization lowers reliability of performance.

Assumptions 1 and 6 together imply

> *Theorem 4.* Attempts at reorganization increase death rates.

Organizations undergoing structural transformation are highly vulnerable to environmental shocks. Large size presumably enhances the capacity to withstand such shocks. Small organizations have small margins for error because they cannot easily reduce the scope of their operations much in response to temporary setbacks. Indeed, the claim that death rates decrease with size is nothing more than a restatement of the idea advanced earlier (Hannan & Freeman 1977) that longer time spans must be used to study replacement in populations of large organizations.

Assumption 7. Organizational death rates decrease with size.

We assume that size has qualitatively similar effects on all three death rates in Figure 1: r_d, r_e, and r_f. Thus, small organizations are assumed to be more likely than large ones to enter the state of reorganization but are also more likely to exit this state by death.

Finally, there is the issue of success at implementing change (the rate of moving from "reorganization" to "new structure"). An organization undertaking reorganization can successfully make the transition to the new state or it can drift back to its original structure, assuming that it does not die. The model in Figure 1 contains two rates that pertain to these processes: r_c, the rate of moving to the new structure, and r_b, the rate of returning to the old one. The effects of size on these rates is unclear. On the one hand, the greater inertia of large organizations might lower the rate of success at reorganization. On the other hand, success at reorganization might depend on the magnitude of resources applied to the task. Since large organizations typically have more resources than small ones, this line of reasoning suggests that the rate of achieving structural change increases with size.

The relationship between size and the rate of structural change is indeterminate in our theory for two reasons. The first is ignorance about the effects of size on rates of completing structural reorganization, conditional on having attempted it. The second source of indeterminacy is the implication that small organizations are more likely to attempt structural change but are also more likely to die in the attempt. Although our analysis does not offer an answer to the main question about size and inertia, it does not support the widespread view that ecological arguments are particularly appropriate for the study of change in populations of *small* organizations.

The model in Figure 1 may be substantively interesting in its own right, assuming that approximate information on dates of leaving states of reorganization can be obtained. It provides a framework for addressing a variety of questions about inertia and change. It has the advantage of transforming what have been mainly rhetorical questions about the applicability of the ecological perspective into specific research questions.

Consider again the question of life cycle variations discussed in the previous section. Recall that we assume that reproducibility increases with age (Assumption 4) because routines become worked out, role relations stabilize, etc. What effect, if any, does structural reorganization have on these processes? We think that reorganization is sometimes tantamount to creating a new organization (with a given level of resources). When reorganization is that fundamental, work groups are reshuffled bringing strangers into contact, routines are revised, lines of communication are reshaped, etc. In this situation reorganization robs an organization's history of survival value. That is, reorganization reduces the reliability of performance to that of a new organization. The stability of the previous structure does not contribute to reducing variability with new sets of procedures, role relations, etc.

If internal processes are solely responsible for the tendency of organizational death rates to decline with age (Theorem 3), the death rate for an organization that has just entered the state "new structure" should be no lower than the death rate of a completely new organization with that structure (and levels of resources). In this sense, reorganization sets the "liability of newness" clock back to zero.

> *Assumption 8.* Structural reorganization produces a liability of newness, raising the death rate to the level of newly created organizations.

The argument in the preceding paragraphs can be viewed as one way to formalize some long-standing notions about organizational crises. Child and Kieser (1981, p. 48) put the issue as follows: "To some extent, a crisis successfully overcome may represent a new birth, in the sense that changes initiated are sufficiently radical for a new identity to emerge." We suggest that such questions be viewed in terms of shifts in age-dependencies in organizational death rates.

External processes may also account for the tendency of death rates to decline with age. For example, we mentioned the tendency for organizations to acquire legitimacy simply by virtue of longevity as well as the fact that it takes time for organizations to develop enduring exchange relations with key actors in the environment. Some sorts of changes in strategy and structure strain external relations, especially when the changes imply a shift in ostensible goals. But simple structural reorganization, without any apparent change of goals, does not rob an organization's history of its value for public legitimacy and does not necessarily upset exchange relations with the environment. Old organizations can presumably count on their existing exchange partners for support during and following such structural change.

If the liability of newness reflects internal processes, the death rate will jump with structural changes. But if the decline in the death rate with age reflects mainly the operation of external processes of legitimacy and exchange, the death rate will not jump when structural changes do not imply a change in basic goals. That is, arguments about internal and external processes lead to different predictions about the effect of structural reorganization on the death rate. Therefore, the study of such effects may shed light on the relative importance of internal and external processes in accounting for age-variation in the death rate in selected organizational populations.

Finally, there is no reason to suspect that the death rate declines with duration in the state "reorganization." Quite the contrary—as the length of time over which reorganization is attempted increases, the costs (especially the opportunity costs) of reorganization increase. As the fraction of organizational resources devoted to reorganization increases, the capacity of the organization to produce collective products declines along with its capacity to defend itself from internal and external challenges. Hence protracted periods of reorganization disrupt organizational continuity and increase the risk of death.

> *Assumption 9.* The death rate of organizations attempting structural change rises with the duration of the reorganization.

A model consistent with this assumption is the classic Gompertz model:

$$r_e(t \mid t_r) = \theta e^{\kappa(t - t_r)},$$

where t_r is the time of entering the state of reorganization and $\kappa > 0$. This sort of model can perhaps elucidate another claim in the organizations literature. March (1981, p. 567), referring to the work of Hermann (1963) and Mayhew (1979), states that

> organizations facing bad times will follow riskier and riskier strategies, thus simultaneously increasing their chances of survival and reducing their life expectancy. Choices that seek to reverse a decline, for example, may not maximize expected value. As a consequence, for those that do not survive, efforts to survive will have speeded up the process of failure.

It is hard to imagine how an action can both increase a survival probability and increase the death rate in conventional models for the death rate (since life expectancy is a monotonically decreasing function of the death rate). However, the framework introduced above is consistent with this sort of pattern.

Consider the case in which the death rate of organizations in some environment rises precipitously at a certain moment t_1 (due perhaps to some discontinuous change in the environment). Death rates of organizations that retain their structures will gradually decline to an asymptote that is considerably higher than the asymptotic rate in the old environment.

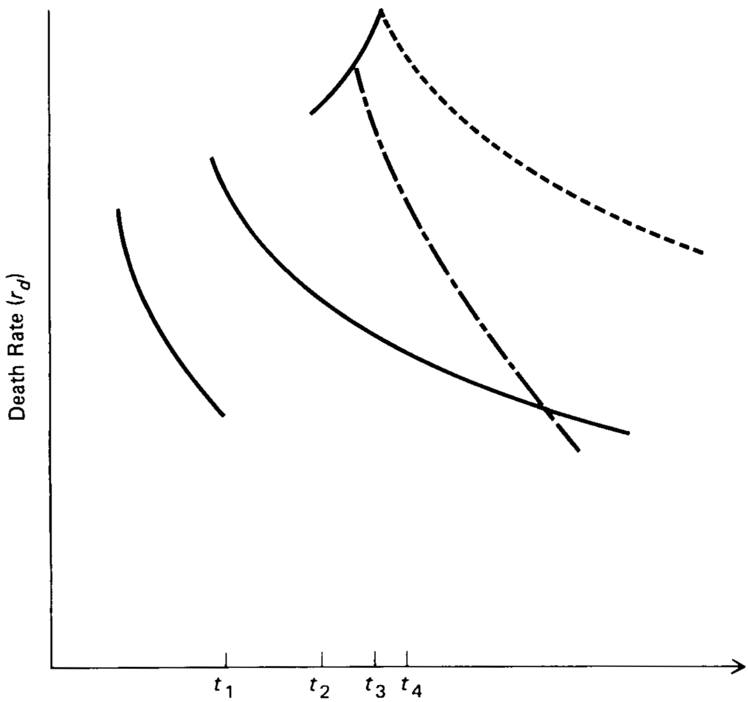
Suppose that some organizations in the population attempt structural change at t_1. Now consider two kinds of trajectories of death rates by age. The dashed trajectory in Figure 2 depicts the death rate of an organization that successfully implements the new structure at t_3. The dotted trajectory pertains to an organization that reverts to the old structure at t_4. In a collection of histories like those in Figure 2, one would see that strategic action to promote survival exposes organizations to great risks (thereby "reducing its life expectancy"). But because the death rate declines rapidly with duration in the new structure, a successful transformation eventually leads to a lower death rate (seeming to "increase chances of survival")—even lower than the death rates of organizations that retain the original structure. However, it is not clear that structural change necessarily increases unconditional life expectancy. This depends on the various rates. Still, introducing the competing risks of death and reorganization allows one to deal systematically with this complicated problem.

Environmental Change, Size, and Inertia

Assumption 5 states that large organizations are less likely than small ones to initiate radical structural change. Does this mean that larger organizations have greater inertia, as Downs (1967) and others have claimed? If inertia is equated with low absolute rates of initiating structural change, it does. When inertia is viewed in comparative terms, as we argue it should be, the relationship of size to inertia is more complicated than the literature has indicated.

According to Assumption 7, the death rate declines with size. This statement is equivalent to the proposition that time-scales of selection processes stretch with size, as we noted earlier. One way to visualize such a relationship is to consider environmental variations as composed of a spectrum of frequencies of varying lengths—hourly, daily, weekly, annually, etc. Small organizations are more sensitive to high-frequency variations than large organizations. For example, short-term variations in

Figure 2 Hypothetical death rate functions for a population of organizations exposed to a shift in selection pressures at t_1.



The solid decreasing curves represent the death rates of organizations that retain their strategies and structures. The rising solid curve represents the death rate function of organizations that undergo attempts at reorganization at time t_2. The dashed curve represents the new (better adapted) strategy and structure at t_3. The dotted curve represents the death rate function of organizations that revert to their old strategies at t_4.

the availability of credit may be catastrophic to small businesses but be only a minor nuisance to giant firms. To the extent that large organizations can buffer themselves against the effects of high-frequency variations, their viability depends mainly on lower-frequency variations. The latter become the crucial adaptive problem for large organizations. In other words, the temporal dimensions of selection environments vary by size.

We proposed above that inertia be defined in terms of speed of adjustment relative to the temporal pattern of key environmental changes. Although small organizations are less ponderous than large ones (and can therefore adjust structures more rapidly), the environmental variations to which they are sensitive tend to change with much higher frequency. Therefore, whether the adjustment speeds of small organizations exceed those of large ones compared to the volatility of relevant environments is an open question. One can easily imagine cases in which the reverse is true, in which elephantine organizations face environments that change so slowly that they have relatively less inertia than the smallest organizations.

Complexity and Inertia

The complexity of organizational arrangements may also affect the strength of inertial forces. Although the term "complexity" is used frequently in the literature to refer to the numbers of subunits or to the relative sizes of subunits, we use the term to refer to patterns of links among subunits. Following Simon (1962), we identify a simple structure with a *hierarchical* set of links, which means that subunits can be clustered within units in the fashion of Chinese boxes (what mathematicians call a lattice).

In hierarchical systems flows (of information, commands, resources) are localized: An adjustment within one unit affects only units within the same branch of the hierarchy. Simon (1962) argued that hierarchical patterns appear frequently in nature ("nature loves hierarchy") because the probability that a complex assembly is completed in an environment subject to periodic random shocks is higher when stable subassemblies exist, as in a hierarchy. More complex structures do not have many stable subassemblies and thus are vulnerable to shocks during the whole developmental sequence.

Recent work on population ecology supports Simon's argument. For example, May (1974), Šiljak (1975), and Ladde and Šiljak (1976) show analytically and with simulation experiments that ecological networks are destabilized when links (of predation, competition, or symbiosis) are introduced. Both the number of links and the complexity of the pattern affect stability.

We think that similar arguments apply to structural change within organizations. When links among subunits of an organization are hierarchical, one unit can change its structure without requiring any adjustment by other units outside its branch. However, when the pattern of links is nonhierarchical, change in one subunit requires adjustment by many more subunits. Such adjustment processes can have cycles; change in one unit can set off reactions in other units, which in turn require adjustment by the unit that initiated the change. Long chains of adjustments may reduce the speed with which organizations can reorganize in response to environmental threats and opportunities.

Complex systems have slow response times not because they are any slower than simpler systems in detecting environmental threats and opportunities but because the process of adjustment takes longer. In terms of the framework developed in earlier sections, this argument implies:

> *Assumption 10.* Complexity increases the expected duration of reorganization.

That is, once a complex organization has begun structural change, it will tend to be exposed to a longer period of reorganization than a simpler organization attempting similar changes. Assumptions 9 and 10 imply:

> *Theorem 5.* Complexity increases the risk of death due to reorganization.

A complete analysis requires consideration of the effects of complexity on rates of initiating change and of its effects on success in implementing change (as we discussed above in the analysis of the effects of size). We are not yet ready to make any claims about effects of complexity on these rates. Still, the result in Theorem 5 suggests that population ecological analysis might be more appropriate for explaining change in populations of complex organizations than in populations of simple ones because complexity increases inertia by at least one mechanism. This result, like that on size, disagrees with the conventional wisdom.

Conclusions

We have attempted to clarify when it is reasonable to assume that organizational structures have inertia in the face of environmental turbulence. We have argued that selection pressures in modern societies favor organizations that can reliably produce collective action and can account rationally for their activities. A prerequisite for reliable and accountable performance is the capacity to reproduce a structure with high fidelity. The price paid for high-fidelity reproduction is structural inertia. Thus, if selection favors reliable, accountable organizations, it also favors organizations with high levels of inertia. In this sense, inertia can be considered to be a byproduct of selection. Our argument on this point may be considered an instance of the more general evolutionary argument that selection tends to favor *stable* systems (see Simon 1962).

Of course, the claim that selection favors organizations with high inertia is not a warrant for assuming that most organizations have high inertia. Selection pressures often may not be strong enough to screen exhaustively for the "most fit" organizations. Moreover, most organizational populations are replenished more or less continuously by an inflow of new members. Younger organizations tend to have less inertia than older ones, and new organizations are more likely to adopt structures that differ greatly from those that would dominate any steady-state of the process subject to selection and closed to new entries.

Organizational selection operates on many dimensions besides reproducibility of structure. If selection pressures on specific features of structure are sufficiently strong, organizations with the characteristics appropriate to the environment are favored even if they have relatively low levels of reproducibility.

By the same token, environments in which change is turbulent and uncertain may not constitute a systematic regime of selection. The traits that are favored may shift frequently enough that no clear trend emerges. Such settings may favor organizational forms that can take quick advantage of new opportunities and the appearance of new habitats. The capacity to respond quickly to new opportunities presumably competes with the capacity to perform reliably and accountably. (Brittain & Freeman 1980; Freeman 1982). Such dynamics may dilute the importance of reliability and accountability in organizational selection. We will address these issues in subsequent papers.

For all of these reasons, it is not sufficient to assume that selection processes favor organizations with high inertia and to proceed as though observed populations contain only such organizations. These considerations lead naturally to consideration of systematic variation within populations in the strength of inertial pressures. Existing theory provides some insights into these matters. One line of reasoning, which we pursued, suggests that inertial pressures increase with age—that organizations tend to ossify as they grow older. We suggest that the more fundamental process is that reproducibility increases with age. It follows from our general perspective that the death rate declines with age.

The effects of size on inertia are problematic in our revised theory. We think that analysis of the effects of size on inertia must consider several kinds of transition rates. One is simply the rate (in an absolute time scale) of attempting fundamental structural change. Another transition concerns success in implementing change. There is also the effect of attempting change on the death rate. We argue that small organizations are not only more likely than large ones to attempt change, but are also more likely to die in the process. Without further information on the magnitudes of the rates, it is not clear whether small or large organizations have higher overall rates of successfully implementing change. Our analysis suggests that it is premature to conclude that ecological theory may be applied more readily to small than large organizations. Clearly this matter deserves more theoretical and empirical attention.

Comment

HOWARD ALDRICH

I want to emphasize a point that Hannan and Freeman have made about the paradigm shift their work represents. In 1974, at the International Sociological Association meetings in Toronto, Hannan and Freeman presented a paper called "The Population Ecology of Organizations." Things haven't been the same since that time! They pointed out the very high death rate of organizations compared with what the literature in the 1960s would have led us to expect. For a representative cross section of organizations, the death rate is about one in ten per year, and for new organizations it is over one in two per year. Such high rates are a substantial challenge to the traditional way of conceptualizing organizations, which treats them as lasting forever.

Organizational change can be conceptualized as three interrelated processes: variation, selection, and retention (Aldrich 1979). First, to produce change in any living system, variation is required, whatever its source—purposive change, creativity, luck, chance, errors, mistakes, or sabotage. Second, selection processes choose among the variations that have been generated. Third, directed change is impossible unless there is also a mechanism for retaining the selected variations. Hannan and Freeman describe the latter as organizational inertia.

I would like to make several points. Hannan and Freeman's general theme concerns the relative balance between external selection (in this case, pressures toward accountability or reliability) and internal variation. My first point focuses on the relationship between original organizational structures and what Hannan and Freeman called "the reorganization state." They argued that we may distinguish between a state called "reorganization" and one called "initial structure," and that it is the movement from initial structures to states of reorganization that poses a hazard to organizations. Their argument poses conceptual and empirical questions. A sizable school of organizational theorizing does not separate these states, arguing instead that organizations are constantly changing on a day-to-day basis. "Variation" is thus something that takes an organization to a different state every day.

Against that, Hannan and Freeman argue that once you have gotten over the initial hump of the liability of newness, such day-to-day variations don't affect organizational survival chances. Therefore, Hannan and

Freeman believe it is possible to distinguish the two states conceptually. The empirical issue, however, is *how* one distinguishes these two states. The paper says that the reorganization state is announced by the people in the organization; it is a consequence of scanning the environment and then making a substantial change. Without that clue, making the two-state distinction is problematic. How will one know that reorganization has actually occurred? At this point in the development of the model, measuring the distinction is a fairly substantial stumbling block.

My second suggestion for further elaboration concerns the sources of pressures for reproducibility. Hannan and Freeman have emphasized external pressures for accountability and reliability. I would place more emphasis on *internal pressures.* I have noted elsewhere that aging organizations encounter substantial social pressures toward internal compatibility between divisions, between members, over selection of employees, the socialization of employees, promotion policies, and so on (Aldrich & Fish 1981; Aldrich & Auster, forthcoming). Quite a few *internal* processes, therefore, limit variations. So the issue must be raised of the extent to which a structure is free to vary independently of external constraints. Under what conditions could internal compatibility, as opposed to external relevance, be a dominant force in structural retention? We also should consider organizations' relative autonomy or freedom from environments, and the extent to which organizations have the ability to ward off possible pressures toward change.

My third point also focuses on external constraints. Evaluating agents outside organizations have fairly explicit criteria for reproducibility and accountability. They look for financial statements, for routine operating procedures, and for well-understood mechanisms and structures. I think this implies that entering the state of "reorganization" may not be so problematic if organizations mimic or imitate other organizations, following fads and fashions in the population they compete in, such as management by objectives, profit centers, and matrix organizations (DiMaggio & Powell 1983).

My fourth point is more an appreciation than a criticism. This model helps us understand the enthusiasm of chief executive officers of big organizations for expansion by acquisition rather than by internal growth. Acquisition enables organizations to bypass the state of reorganization or get through it in one fell swoop, instead of internally diversifying and going through the risk of changing internal structures to meet changing conditions. Some of the strategies used by large corporations to protect themselves against the risks Hannan and Freeman have identified include buffering, loose coupling, internal venturing, the creation of autonomous divisions, and the subsidizing or supporting of spinoffs. Hannan and Freeman's model helps us understand these actions as strategies for avoiding the risk of the reorganization state.

I want to raise two more issues. One has to do with, paradoxically, where the "environment" is in this model. For a model based on population ecology, the environment enters in a surprisingly muted form. First, it enters as a constant. Hannan and Freeman argue that selection in modern society favors structural inertia. Thus, there are constant pressures toward developing reproducible structures with reliability and accountability. Second, environment enters as a variable. In the next to last section of the paper, Hannan and Freeman introduce the environmental dimension of "variability"—high frequency versus low frequency—arguing that there is an interaction between variability and size. Large organizations are affected by low frequency variability (and are able to ride through high frequency variability) and small organizations are affected by high frequency variability. My suggestion for the future development of this model is to add the other dimensions that have previously informed their work: grain, and uncertainty. For the moment, at least, the environment is in their model mainly as a constant and only marginally as a variable.

I have one final point. Last night, we invoked the names of Marx, Weber, and Durkheim. One Durkheimian implication of Hannan and Freeman's work struck me this morning. What are the implications of an increasing concentration of economic resources for societal adaptability? What happens in twenty years from now when we have the Fortune 400, or 350? Here Mancur Olson's point (in his conference paper) is highly relevant. He noted that in many societies organizational and industrial variation has been heavily constrained by external coalitions of organizations which have shared standards of reliability and accountability. This leads into a related question, namely, where do new organizational forms come from? The implication of Hannan and Freeman's paper is that we can't look to older, larger organizations for new organizational forms. Is it possible to design public policies to facilitate the formation of new forms of organizations, to counter what looks like a potentially maladaptive trend?

General Discussion

Peter Blau: The logic of what you say is very intriguing because it provides a new perspective. But there seem to be a couple of problems. One theorem says that inertia increases survival chance. But is inertia itself not an adaptation? How do you distinguish the theory of adaptation and the theory of selection?

John Freeman: It is useful to make a distinction between whether the *form* of organization that you are studying survives and whether *individual* organizations in those populations which manifest the form survive. Inertia is a way of saying that individual organizations are not adaptable. However it may be that the very properties that make those *individual* organizations inert are adaptive from the point of view of *forms* surviving over a long period of time. For example, you can find Mexican restaurants being founded and failing monthly in a geographical area but Mexican restaurants as a kind of organization persist, so you can say that this form of organization is adaptable.

James Coleman: I have two questions suggested to me by particular cases. First, consider General Motors. The assembly plants in General Motors take two organizational forms. The traditional organizational form is one with a fairly clear hierarchy. The other form is new, patterned after the Japanese, and called quality-of-work-life programs. Now, the death rate for the first form is higher than for the second and the birth rate is lower, so that the form with quality-of-work-life is expanding in the population of assembly plants in General Motors. The assembly plant is part of an organization. Can one look at parts of organizations from a population ecology point of view? And if the answer to that is yes, then the question is, Just where does rational choice end and population ecology begin? In this case the environment for this part of an organization is a single decision-maker.

The second question has to do with city manager forms of government. They seem to be very unstable compared with a mayor-council form of government and seemed to disappear fairly rapidly after they came into being. When I looked at this some years ago, I concluded that the reason was that they did not have a good internal mechanism for the expression of grievances and dissent by the population. Dissent had to be expressed through overturning the system. The assumptions in your paper don't have very much to do with the kind of proposition that I was just stating, that is, the internal structure of the organization affecting its survival rate.

John Freeman: Let me talk first about the General Motors assembly plants issue. I think you could study the competition between those two forms of assembly plants the way we propose. It might very well be that the answer you would come up with would be uninteresting, namely, the CEO of General Motors says, "I want this kind and not that kind and it's just because that's how I want to do it." Then what you would find from our kind of research wouldn't tell you very much. However, I would not presume that there is a single decision-maker within General Motors who decides on things like that. General Motors is a somewhat amorphous polity in which the politics of these decisions are extremely complicated in exactly the way we described in the paper. You might not lose very much by ignoring the fact that there is this behemoth called General Motors which fumbles its way through the world and dominates its environment as much as it adapts to it.

Lastly, it's clear that one way of dealing with an uncertain world is to form federations of organizations or superorganizations and that the consequence of that is to bump the selection process up a level of analysis. Perhaps what is manifested in selection among corporations here is going to take a societal form in Japan, and maybe the whole Japanese structure will blow up one of these days for exactly the reasons you pointed out with the mayor-council form of government. There is no general answer to your question. It's a question of the research situation that you are studying and where your theoretical bets are about whether you are going to learn something.

Anthony Oberschall: My question has to do with noneconomic organizations, such as churches, sects, denominations, and cults. If you think your theory can be applied to these organizations what would you mean by reliability and accountability? If you look at it from the product end, you see that a lot of people want old-time religion and that they are not so choosy about organizational forms. They are willing to go to camp revival meetings, to little churches; they are willing to get it from television preachers, and so on.

Michael Hannan: We hope that we are addressing this question by studying labor unions, which are collective activities of exactly the kind that operate in a nonmarket context. Nelson and Winter, who are doing alternative kinds of evolutionary models of organizations, keep challenging us to explicate exactly what the selection environment is for organizations like labor unions. I have never seen that as a problem. It seems to me that labor unions are competing with each other for the time and loyalty of members. It's quite clear if you read the history of labor unions that competition is often brutal, sometimes more brutal than the competition between unions and management that attempts to crush unions. They are also competing for legitimacy, and in certain points in the history of this movement that's the crucial struggle. There are clearly environments in which events occur that favor various forms of organization, and there are alternative principles of union organization that have thrived at different periods. The most notable is the craft versus industrial form of organizing, or the many successful utopian unions in the 1800s. These different forms were all competing with each other. We choose the union example because it seemed so obvious that there is this competitive struggle going on. It seems to me that there is also intense competition for money among church organizations, and these TV preachers are in very much a market kind of context.

John Freeman: One brief footnote. There was a paper in the *American Sociological Review* last spring by Miller McPherson that looked at voluntary associations in an ecological way, using, in fact, some of the models we do. For him the key resource for the voluntary associations he studied was not money. It was people's time and membership. So I would like to emphasize that the environment does not have to have anything to do with money.

Harrison White: We are all grateful to you because you are opening our intellectual horizons. I wonder if it wouldn't help to open them up even a little further.

What I want to suggest is a third line. There is another kind of evolution that is hard to handle in your present framework. In this evolution units aren't dying, instead they coagulate into bigger units. Let me give an example from business. Spence and Potter, two economists, have done a classic study of corn wet milling where the actors are getting together and completely redefining their identities. It is not just reorganization, it's everybody saying "I'll cut off my hands and legs and we'll reshuffle them, because if you take all my arms and I take all your legs, we are in better shape." Wouldn't it help to have an explicit framework for that as well?

Mancur Olson: Some organization theory is a little bit like a murder mystery in which the victim is killed for no reason at all. That is to say, one doesn't get any sense of the reasons or individual motives that account for the existence of a particular organization and the characteristics it has. Suppose I look at a physician's office and find that the physicians are organized as a corporation. Presumably behind this organizational form are such causes or reasons as the tax code, which gives the physicians an incentive for becoming a corporation. At another time the environment or the laws might change, and then a different organizational structure would emerge.

Now, when we get to large and well-established organizations, it is sometimes useful to think of the organization itself as having a motive and a desire to survive. Presumably at a corporation like General Motors, which Jim Coleman brought up, there are so many people with specialized human capital that is suited to General Motors, and so much physical capital that matches the skills and organization of the people, that we couldn't imagine General Motors disappearing without lots of people being hurt a lot. So there is, if you will, a social structure at General Motors, and that social structure's "motives" can be understood if we have, in turn, explained or derived it from looking at the purposes of people involved in it. But in general, why shouldn't we bring in motives more in organizational theory and thereby get a better sense of why organizations crop up, change, and die?

Michael Hannan: It seems to us helpful to wonder whether there is a tight coupling between motives of participants and the collective action of organizations. We are struck with the frequency with which unanticipated consequences of organizational action dominate the intended consequences. In trying to link organizations to large-scale social change, it seems too complex at the moment to try to link it to rational choice. This is a strategic choice, and we see no theoretical inconsistency.

Robert Eccles: I, like others, find this an interesting way of looking at organizations—a different way than I am used to. But I have a problem which I also had in reading Aldrich's book. What do you mean by "form" and what do you mean by "reorganization"? Sometimes form is a type of organization—say Mexican restaurants—but it is also a particular way of organizing a Mexican restaurant. In most Mexican restaurants there is a waiter or waitress and there is a cook. You could have waiters and waitresses take the order, go back and cook it, and bring it out to you. Is that a different kind of form or is that the same? The question of reorganization is related to this. I can think of many different types of reorganizations. You could go from being a low-cost producer to specialties. You could go from a functional structure to some kind of multidivisional structure or you could bring in a new CEO, who would fire the whole top management team, as sometimes happens. Could you clarify for me what you mean by "form" and what constitutes a reorganization?

John Freeman: What you need at bedrock is a set of rules which define when an organization is in the population of interest and when it's out. That is all an organizational form has to be. Ron Burt and I have talked about defining forms and related niches in a networking way that is quite reminiscent of what Ed Laumann presented yesterday. But you can also do it in a more naturalistic way, as we did in the restaurant study where we asked ourselves what kinds of categories people use when they want to go to a restaurant.

Michael Hannan: Ultimately the definition of forms will be successful only if the definition has isolated populations that have different "fitness functions," that is, functions that relate probability of success or growth rates to environmental variations. If we could observe those directly, that would be the basis for choosing forms.

The question regarding reorganization could be answered in those terms as well. Does the new structure have a different fitness function associated with it? In each of the empirical research projects, we make guesses about these functions. In a sense, we are in the same situation as early biological evolutionary theorists when they based their work mostly on distinctions made by amateur naturalists carefully studying swamps and marshes and lakes, and so forth. There are many observers around this world, but we often don't exploit them enough. For example, there is a large class of business analysts that does nothing other than study the organizational forms in the steel industry or whatever and try to make judgments about them. If you can find situations where most industry analysts would agree there are three different ways to organize an industry "X," that's probably a reasonable basis on which to begin. I would trust that much more than what would come out of a cluster analysis of the interlocks between firms.

Howard Aldrich: This is one issue where there is some heterogeneity in the theory group. Hannan and Freeman are taking an inductive approach to the definition of forms. At the other extreme, Bill McKelvey takes a very deductive approach. He has written a book called *Organizational Systematics,* published last year. I joined him in a variation on that statement, arguing that forms should in part be arrived at through historical studies. Thus, there is also a middle ground between inductive and deductive. But McKelvey has worked out a fairly elaborate program to get at forms deductively. Whether that will succeed depends in part upon how many people are convinced to try the methods advocated in his work. So there is a friendly amount of competition in the field on precisely this question of form.

Interpretive Sociology

Public Problems as Phenomena: The Shape of a Humanistic Social Science

JOSEPH R. GUSFIELD

There is a form of mordant humor illustrated by the story of two Frenchman who, following World War I, are trying to explain that disastrous set of events. "We wouldn't have been in the war if it weren't for the bicycle riders and the Jews," said the first. "Why the bicycle riders?" asked the second. The first one replied, "Why the Jews?" It is a vein of irony in which an explanation is proffered seemingly assuming order and consistency in the world. The punchline is the reverse: a world of caprice, whim, and random unpredictability.

In this paper I want to consider this question of the orderliness of the social world as a central problem in sociological theory; one which divides sociologists and lies at the ground, if not the root, of many issues of perspective and method. I want to do so primarily with reference to recent research at several levels of generality: in relation to my recent work on drinking and driving, to mine and others' work on alcohol issues, and to the general study of that field labeled "social problems." Following this descent into the caves of empirical work, I want to return to the clouds of abstract talk to discuss the implications for questions of method. In the final section of the paper I will draw these matters to a close with a description of the sociological product that follows from my analysis.

The Subject-Object Relationship

In the version of science that has been so influential in social science, a real and knowable world of fact is postulated; a world that exists independent of actors and observers. Observers are capable of discovering such facts. In one form or another, the stimulus and the response can be delineated by the observer. Controls can be placed on observations and experiment or quasi-experiment performed. Concepts can be utilized in a manner that enables generalization and theory to explain the world of fact. In an opposite perspective within social science a more idealistic conception of the subject's relation to the object is assumed. From this viewpoint, subjects play a much more active role in selecting, shaping, and defining the objects to which they respond. What actions mean to the actor is not given to the observer. They involve interpretations, first by the actor and second by the observer, the social scientist.

The problem of meaning then becomes a central issue in social science. If the experience of the actor cannot be understood by the observer without knowing what the object means to the actor, then he or she must develop special ways of interpreting actions. It is not the noumena—the thing in itself—but the phenomena—the experienced object—that is the basic datum of the sociologist.

The Social Construction of Reality

The realization of multiple realities sets the stage of inquiry into how it is that social order is at all possible (Schutz 1970, chap. 12; Garfinkel 1967, chaps. 1, 2; Berger & Luckmann 1966). An emphasis on the part human beings play in constructing or creating the reality of their worlds underlies this form of understanding. For example, a conventional explanation of human action in terms of norms and roles implies a passive actor, guided by fixed rules. A perspective drawn from social constructionism emphasizes the activity performed by the actor and actors in defining the situation as one in which one of a number of possible roles are playable. It may, and often does, involve a negotiation between interacting persons to determine what the situation is and what role will be acceptable (Cicourel 1973, chap. 1).

The same swath of human activity can take on diverse meanings, can be framed in diverse ways. Goffman's analysis of framing devices showed the modes in which a given "reality" could possess different meanings: as make-believe, as ceremony, as drama, as contest, as technical redoing (Goffman 1974, chaps. 1–3). The model in life-drawing classes usually puts on clothes during a "break" to control the definition of his or her nudity as "technical." So, too, physicians characteristically cover portions of the anatomy "irrelevant" to the medical activity (Emerson 1970). In these ways we define our situations and proscribe our roles. Language plays a significant role in developing frames. Kenneth Burke has been seminal in calling attention to the dramatistic keys by which we develop "ways of placement": scene, act, agent, agency, purpose. Emphasis on one or another of these keys changes the situation, as dramatists demonstrate (Burke 1945, chap. 1).

Human beings not only act but are capable of reflecting upon their action, the actions of others, and the possibilities of alternative actions. Action is not the mechanistic outflow of pre-existing patterns. We are constantly assessing situations and monitoring ourselves and others. There is an interaction among imagined pasts, presents, and futures.

Reflection on events makes it possible to explain or account for our acts as well as those of others and in turn to move toward new events. Our experience of the world is both immediate and reflexive—imagined in retrospect. It makes it possible to construct the situation in reconstructing it; to provide order and explanation through the ideational activity that is the central character of the human. Action and experience are interpreted: given definition and meaning in the interaction of subject and object.

Interpretive Sociology

This approach to social theory has been around for a long time in sociological circles. The seed only came to flower profusely in recent years. It is seen in diverse colors in the focus of attention on how meaning and order arise, on the processes and products of interpreting "raw" sense-data into the patterns of orderly existence and experience. Mehan's study of teacher-student interaction in elementary school classrooms is illustrative of the search for the "constitutive rules" that make that interaction possible and that are almost always taken for granted by teachers (Mehan 1978). Such prior frames as turn-taking, answering, and recognizing the question form are assumed in the ordered interaction, although the videotapes revealed that not all the children uniformly had internalized such rules.

In recent years the flowers have become bushes and trees and many studies have emerged in one way or another "phenomenological" in taking as their subject matter not the objective world, apart from the interpretations of the actors and even the observers, but the object as an experienced, interpreted, created event. Thus, Aries investigated what childhood has been perceived as in different centuries (Aries 1962). Mehan's and Emerson's work, above, is indicative of efforts to investigate how it is that order is created, as Murray Davis's recent study of sexuality is, in part, a study of how it is that events are interpreted as sexual or nonsexual (Davis 1983). Foucault's studies of the forms of thought generating the perception and character of madness and of clinical medicine form still another instance of a focus on the modes of interpretation, or in his term *epistémes* (Foucault 1965, 1973).

The central feature of these and many other such studies is that what is conventionally seen as an objective fact of reality is discovered, or reinterpreted, as a product of the interaction between object and subject. Childhood, sexuality, class interaction, madness, clinical medicine are not "givens." They make sense within the filters of particular cultures and their categories for understanding and creating those objects and events. It is this recognition that is the central feature in my studies of drinking-driving, in recent work in alcohol research and in the reconstitution of social problems that I will describe.

Accounting for Public Problems: The Phenomenon of Drinking-Driving[1]

Why drinking-driving? It occurred to me that a study of how and why drinking-driving was seen as a major cause of auto accidents and deaths would be appropriate chiefly because there appears to be such unanimity among public officials about the irresponsible, perhaps criminal, character of the act, its relation to auto deaths, and its preeminence as a source of danger in American life. Scholars and intellectuals, conservatives and radicals, shared common beliefs on the issue. Such closed ranks make a sociologist attentive. How is such unanimity possible in a society otherwise so filled with divergent and conflicting voices?

[1] My approach is congruent with that now being undertaken in the sociology of science by a number of sociologists. It is a "strong program" in the sociology of knowledge, paying attention to how the content of science is creatively constructed. (The term is that of Bloor 1976.) This application of the sociology of knowledge is exemplified in several recent studies of both "hard" and "soft" sciences (Latour & Woolgar 1979; Barnes & Edge 1982).

Dimensions of Study

The literature on drinking-driving is, as you might imagine, voluminous. Several years ago an exhaustive review of drinking-driving studies found over 300 separate publications (Cameron 1977). With very few exceptions, the studies of drinking-driving assumed the existence of a possible problem area and sought the body of fact about drinking and driving that would analyze and/or predict its occurrence. My interest was not in continuing this genre of research and its questions. Rather I was interested in studying the consciousness of drinking-driving as a public problem (Gusfield 1981). To place the phenomena of the social problem into the realm of the problematic is to assume that the reality might be different, that drinking and driving are not "naturally" elements of importance in understanding automobile safety. A rigorous insistence on the multiple possibilities for interpreting reality places all knowledge and all social arrangements in doubt and makes it necessary to examine the processes by which consciousness, cognition, and order are possible. With such an examination, the character of the "social problem" undergoes a change, as I hope to show.

The Structure of Public Problems

We recognize, with even a slight exposure to historical material, that not only are particular social problems specific to historical periods, but that the delineation of social problems is itself specific to history, especially to the modern age. In his classic paper, "Social Problems That Are No More," Ian Weinberg (1974) points out that many ills that flesh was heir to in premodern periods were dealt with, or not dealt with, at the level of the family. They existed as private problems. The modern period, Weinberg pointed out, contains bureaucratic and professional "referral structures" which make the social and public character of such ills possible. These referral structures facilitate selecting drinking as a focus for understanding, explaining, and predicting auto accidents. Indeed, the conventional study of drinking-driving assumes that it is not foolish to treat auto "accidents" as if they are not, in fact, accidental; that they possess characteristics of pattern, explainability, and prediction. How then is the explanation created of auto accidents as, in a significant part, a result of drinking-driving? The data on auto accidents collected and reported attend largely to the characteristics of the motorist and the road. The documents of government agencies, the journalistic accounts of auto safety, the speeches of legislators and political officials, and the scientific studies of drinking-driving are also dominated by the central figure of the motorist, the driver. In Kenneth Burke's pentad of terms for explanation (scene, act, agent, agency, purpose), the emphasis is on the agent—the key in a psychological framing of events (Burke 1945, chap. 1.) A sociological frame would place attention on contexts: in Burkean terms, on scene.

The research studies are congruent with the focus on the motorist. There are many studies of *drinking drivers*—their demographic attributes, their occupational and class characteristics, their criminal and problem-drinking histories, their accident records. There are very few of *drinking-driving*: the social interactions before and during driving, the goals of the journey, the institutional responsibilities for accidents, as well as the physical scene of such driving.

The categories assumed in the formulation of drinking-driving research are in one sense intuitive but in another sense are cultural: shared assumptions among the world of professionals and among their audiences.

THE SCIENTIFIC ORDER: FICTION, RHETORIC, FACT. How is the factual certainty of the causal role of alcohol in motor accidents created? An examination of the research documents is necessary as one way to see how the scientific knowledge of alcohol and auto accidents is constructed. One of the central fictions in the drinking-driving field is what I call the "isometric fiction"—the treatment of one datum as the equivalent or index of another.

Prior to the invention of mechanical means to test for the chemical composition of the blood, the sober or inebriated state of the motorist arrested for illegal drinking-driving was evidenced solely by the arresting officer's testimony, a most unreliable source. With the development of the breathalyzer in the late 1920s, a Blood Alcohol Level became prima facie evidence of the condition. (It is set in most American states at .10:100 milligrams of alcohol in 100 milliliters of blood.) Its use in legal work "translates" or transforms physiochemical measure into a psychological measure—a judgment of the motor reflexes and state of risk-taking of the motorist, independent of the specific situation, experience with alcohol, state of health, character of road, and other features of the situation and the person that bear on fitness. The B.A.L. becomes a fiction used for practical purposes. It becomes identical to the psychological state which constitutes the concern of the law and the research. For example, here is a quotation from another research report: "About one-eighth of the drivers had been drinking to an extent great enough to impair their driving performance—0.05 percent b.a.l. or higher . . . "

(Wolfe 1975, p. 46). The qualifications of doubt in specific situations has been dropped and the measure has emerged as identical to what it indexes. It is presumed that every time a driver who died in an auto accident had a B.A.L. above the legal limit he or she was at fault in the accident. The use of the term "drunken" to describe drivers in turn substitutes an imagined, common-sense condition of drunkenness for the more complex state which is involved in the measurement procedures.

The organization of data collection and definition, the language of reporting, and the assumptions used in reaching definitive conclusions enable the observers, at the scientific level and at the lay public level, to make sense of the world of auto safety as crucially a matter of drinking-driving. The order and pattern in explaining auto deaths assumes the cloak of certainty woven by the looms of interpretation, rhetoric, and art. The cognitive order thus created is the foundation of the moral order that establishes official policy in this area.

THE CRIME OF DRINKING-DRIVING. During the riots following the assassination of Martin Luther King there were several public suggestions calling for police to shoot at looters. With sarcasm, the former Attorney General Ramsey Clark called instead for shooting "drunken drivers," whose act he described as "a far deadlier and less controllable crime" (*Los Angeles Times,* 1975). Law and public statements uphold a moral image of the drinking driver fixed as a type—the "killer drunk." The "killer drunk" becomes the paradigmatic case of drinking-driving in the same sense that Mircea Eliade maintains that myths provide the paradigmatic acts which ensure a reality as fixed and certain by the constant repetition of the mythic story (Eliade 1963).

The focus of both law and public condemnation on the drinking-driver is, in turn, also a selective process of inclusion and exclusion. Other possible "realities" in the spectrum of conceptions of auto safety are excluded. In the marketplace of ideas less individualistic, more contextual or mechanical elements are minimized. The little attention to such alternatives as safety belts, auto design, road condition, or legal liabilities narrows the scope of consciousness of auto safety.

THE NEGOTIATION OF REALITY. At the level of public pronouncement and law, drinking-driving is condemned as crime, not in the same category as a traffic offense. Yet direct observations of how police enforce drinking-driving law, how defense and prosecuting attorneys decide charges, and how judges determine sentences and suspend sentences leads me to characterize actions at more local levels of law and behavior as a negotiatory process. At the level of day-to-day behavior drinking-driving is treated less as a criminal offense than as a "folk-crime," another traffic infraction. More recent observational studies of how drinking-driving does, and does not, emerge as a topic of attention in public bars has further led me to conclude that drinking and driving is the normal, expected behavior among male bar patrons (Gusfield, Kotarba & Rasmussen 1984.)

THE RITUAL ELEMENT IN LAW. In trying to make sense of this discrepancy between public definitions of the drinking-driving problem and the routine, semiprivate definitions, I have interpreted the public definition, discussion, and legislation concerning drinking and driving as a form of ritual or ceremonial action. Although drinking-driving policies often use the language of a rational choice between alternative counter-measures, they are more readily interpretable as a ritual drama whose relationship to stated goals is diffuse, problematic, and irrelevant. The means chosen in the public arenas of law and politics are poor utilities to achieve deterrent effects. But such means-ends relations are inappropriate to evaluate actions which are symbolic rather than literal. The technical advice and programs couched in the framework of rational choice are thus often in contrast to the dramatic and rhetorical actions of politicians and government officials. To interpret policies in the realm of drinking and driving as means to ends, in the conventional sense of a rational choice, is to fail to understand how policies are also, or otherwise, events in and of themselves and unrelated to extrinsic ends. In this procedure the sociologist makes it possible to think otherwise about auto safety, to widen the scope of consciousness, to examine assumptions more critically, and to recognize the moral and political aspects of what are otherwise seen as acceptance of technical fact.

The reconceptualization of drinking-driving that I have described above is a case of the new experiencing of a social problem. What I want to stress in this section of the paper is that such definitions, perceptions, and explanations of phenomena as matters of fact are not solely academic, research issues. The existence, definition, and explanation of a social problem is a matter fought over, sponsored by, and organized among groups and organizations.

The titles of two recent works point up the direction of this emphasis on the active role of persons in creating the facts to which they respond: "Social Problems as Social Movements" (Mauss 1975) and "Constructing Social Problems" (Spector & Kitsuse 1977). This direction has been a significant theme in new perspectives in alcohol studies and in the more general approach to a theory of social problems.

The Alcoholism Movement

The dominance of the deviant drinker, the alcoholic, as the focus and even definition of the public problem of alcohol and the conception of that deviance as a disease called "alcoholism" has a history to it. It represents a change in the public discussion of alcohol and did not occur until the 1930s in the United States. It was not delivered out of the womb of a natural condition so overwhelmingly "there" as to be unquestioned. It is not a result of technical discoveries or new facts.

For most of the nineteenth century and until the repeal of the Eighteenth Amendment in 1933, the use of alcoholic beverages was a political issue in the United States. While there was concern for the "chronic inebriate," the object of temperance and prohibitionist movements was drinking in its various forms and manifestations. Even perceptions of alcohol as a causal agent did not occur until the nineteenth century. Colonial America did not explain accidents as "caused" by a substance, alcohol, but focused instead on the drinker as the crucial agent (Levine 1978). Although a conception of "chronic inebriation" was developed in the late eighteenth century and its status as a "sickness" discussed during the nineteenth, personal problems with alcohol were dominantly seen as matters of choice, will, or mental disturbance (Conrad & Schneider 1980, chap. 4).

Following Repeal in 1933 the public conception of alcohol problems underwent profound changes. The Protestant churches lost their power and their inclination to strive for control of drinking in America. New groups emerged during the next twenty years who worked toward overcoming the stigma attached to what came to be known as "chronic alcoholism" and toward obtaining medical treatment of that condition (Room 1978; Gusfield 1982). Alcoholics Anonymous and its byproduct, the National Council on Alcoholism, the emergence of professionals in alcoholism treatment and their associations, and the development of an academic interest in alcohol studies, especially at the Yale Center, are especially salient, both for their political activities and for their impact on public opinion. It is this collective activity that is referred to as "the alcoholism movement." A new definition of the alcohol problem emerged from these activities.

The resulting transformation in the meaning of alcohol as a public problem has two dimensions. First, as the alcohol issue was politically defused, alcohol problems became those of the deviant drinker, the person whose drinking is held responsible for his troubles. Thus, the cure of the alcoholic became the goal of the movement and alcoholism the definition of alcohol problems.

The second consequence of the alcoholism movement is summed up in the disease concept. Here language is itself significant. The term "alcoholism," while coined in the nineteenth century, was not widely used in the United States until after 1920. The emergence of the disease concept was not a result of new discoveries in physiology or medicine. It did, however, offer both a transformation in the conception of the causal and the political responsibility for deviant drinking. The medicalization of alcohol problems shifts attention to treatment of alcoholics and redefines a moral problem as one of medicine, of sickness, and thus not open to stigma. Even to name the problem defines its character. The "chronic inebriate problem" locates it in the person and is then open to stigma. The "problem of alcoholism" locates it in a condition external to human will and then not open to stigma. Consider the difference between "madness" and "mental illness."

In the past decade there have been strong signs of changes in the definition of alcohol problems. Critiques of the medicalization of alcohol problems, developing policies of preventive social controls, and a wider spectrum of definition of alcohol issues are now emerging (Beauchamp 1981; Gusfield 1982).

The Social Construction of Public Problems

This description of new approaches in alcohol studies is one instance of a general approach to the study of social problems from a phenomenological or social constructionist perspective. In the seminal statement of this perspective, Spector and Kitsuse (1977) write:

> Rather than investigate how institutional arrangements produce certain social conditions, we examine how individuals and groups become engaged in collective activities that recognize putative conditions as problems, an attempt to establish institutional arrangements. . . . In contention throughout the social problems producing process . . . are the definitions of reality that groups and organizations assert. . . . [p. 72]

Cicourel's study of juvenile delinquency, Douglas's study of suicide, and Lindesmith's study of opiate addiction, as well as Becker's study of marihuana induction, are well-known examples, as are the other studies already mentioned above (Cicourel 1968; Douglas 1967; Lindesmith 1947; Becker 1953). Such studies do not assume a fixed object as the subject of study but instead examine how meaning has emerged. Such self-awareness brings to public discussion an ironic, critical perception which enables the participant in the process to see the possibility of alternative realities, of other ways of conceiving the situation. It is in this sense that such a sociology has its humanistic character and function.

Causal Analysis and Understanding

Since Dilthey, and especially since George Herbert Mead, the admonition to see the world from the standpoint of the subject has appeared to invoke a mystical process.

But "understanding" is itself no simple matter. The ethnographer does not overcome the difficulties of intersubjectivity. The problem of knowing the categories of thought and the intentions of respondents by total immersion in the field is a baptism without salvation (Cicourel 1964, 1968, 1973).

Clifford Geertz has described both the problem and one solution to it in some practical reflections on ethnographic field work (Geertz 1979). In discussing the anthropologist's problem in seeing the native's point of view, Geertz distinguishes between "experience-near" and "experience-distant" concepts. The former are those the informant would himself use while the latter are those employed in professional pursuits. Confining oneself to "experience-near" concepts, writes Geertz, will leave the anthropologist "awash in immediacies as well as entangled in vernacular." Using "experience-distant" concepts will leave him "stranded in abstractions and smothered in jargon" (p. 227). The ethnographer, Geertz continues, cannot perceive what the natives perceive, no more than they can see the world as he sees it. The ethnographer can, however, do something else: "What he perceives—and that uncertainly enough—is what they perceive with or by means of or through or whatever word one may choose. In the country of the blind, who are not as unobservant as they appear, the one-eyed is not king but spectator" (p. 278).

The difficulty, as Geertz recognizes, is the now familiar one of the hermeneutic circle. In order to grasp the meaning of parts, you must understand the whole, the context. But in order to understand the whole, the context, you must grasp the meaning of the parts. The observer must go back and forth, from part to whole and back again (Hoy 1978, pref. and chap. 1). What finally may result is an interpretation which appears coherent and fits the known material. Such a product is not the laws, propositions, or fixed probabilities we are familiar with as scientific conclusions.

This begins to lead me toward the humanistic attributes I promised in my beginning.

Toward a Humanistic Science

The thrust of the interpretive or creative sociologies is distressing to those who have conceived of the social sciences in the paradigm of the Enlightenment. A model of natural science, whether historically correct or not, has informed social scientists in the self-understanding of their role in modern life. A humanistic science is at variance with that model. Its relation to the social and political world and its mission as a scholarly discipline lies closer to that of the humanistic disciplines of history, literature, and philosophy, but yet has its own status as a way of knowing.

The Natural Science Model

Two aspects of the natural science model as social scientists have utilized it are relevant to this analysis. The first is what I call the engineering theme. By this I refer to the anticipation that knowledge will lead to control. Whether you think of the Marxian theory of historical evolution of economies and political institutions or of the mundane role of drinking-driving studies in creating a "state of the art" on which to base legal policy, the mission of the social sciences has been substantially similar. It has sought to provide the basis for intelligent human policy through discovering the true state of things, generalizations which have the "out-thereness" quality that settles doubts, debate, and intellectual conflict. The search for laws and propositions predicated on a logic of cause and effect implies controllability and the ability to create an engineering of policy.

The second aspect of the natural science model is in the search for the right answer. The existence of more than one solution is anomalous in scientific procedure. That it is possible at any moment is a situation to be eradicated as soon as possible. The cumulative character of science is a major mark of the natural science model as it has operated in sociology. The general failure of cumulation and the lack of consensus about interpretations is part of the conflict, diversity, and disarray that sociology presents to the intellectual world today.

The Humanistic Model

What appears as disarray to some can also be viewed, I submit, as both the essential character of social, as distinct from natural, science and a source of richness. The very creativity which produces the multiple realities that are troublesome for positivist method is also the means through which the observer discovers the multiple meanings laden in raw data. In his/her own constructions of the first-order constructs of the subjects, the sociologist brings to the study of behavior the diversity of possible significances and standpoints from which behavior takes on relevance and meaning. The orientation toward social problems, including the study of alcohol, is illustrative of this. It is in the activities and definitions of groups and organizations that the phenomena of social problems are constructed and not from unmediated effects of objective conditions on passive publics.

What the sociologist brings to data, as occurs in the field of social problems, is in turn a new conception of the material, or the problem. If my recent research achieves anything in the public, as well as the academic, arena, it does so in forcing a new vision of the drinking-driving problem in which that definition of auto safety is seen as one among a number of potential interpretations and its current formulation one among several ways of conceiving the phenomena. The response "I never thought of it that way before" is what the humanistic model in social science is about.

We are not reduced to equating scholarly detachment and patient observation with political bombast and the common sense of the common man. In much of sociology the act of the observer is more akin to the judge in a legal setting than the experimenter in the laboratory. We judge judges for their exposure to alternative viewpoints, for their awareness of the limits which their own culture and self-interests place on judgments, for their capacity to understand and hold a multiplicity of meanings. Similarly, we judge sociologists not alone, often not at all, by the correctness of their conclusions but by the scholarly qualities of exposure to other possible views, by the logic of their arguments, by their ability to reflect upon their own interpretations. In that self-reflection lies the possibility of a credible yet creative social science.

THE MODEL OF THE TEXT. Another way of stating the humanistic model is suggested by Paul Ricoeur's "model of the text" (1978, chap. 10; 1979, chap. 2). Ricoeur maintains that human action can be treated as a text. As such the reader, the observer, is confronted not only with the multiple meanings subjects give to actions but with the multiple meanings inhering in the action. Texts contain possibilities and point toward new avenues: ". . . like a text, human action is an *open work*, the meaning of which is 'in suspense.' It is because it 'opens up' new references and receives fresh relevance from them, that human deeds are also wanting fresh interpretations which decide their meaning" (1979, p. 86).

In sociology a great deal of what we do is of this nature. The discovery of an intellectual problem, the invention of a new concept, the emergence of a new metaphor are profound ways in which sociologists affect the temper of their times. Functional theory, I submit, has made its major contribution less through discovering "truth" than through opening up the many potential significances and relevances of action under scrutiny. With a different vocabulary and a different perspective about sociology, Merton's classic essay on manifest and latent functions is an instance of a humanistic social science of powerful impact. The social scientist has become, in many ways, the source of public definitions of the situation.

The impact of the social sciences lies as much in the language of its work as in the research which is being reported upon. The metaphors in use become the furniture of our imagination and lead us into thinking about society in one way or another. Consider the metaphors in the concept of "social stratification." As I have tried to show for drinking-driving, a literary art is both essential to and unavoidable in the practice of transmitting science. How that art is conducted is significant for the product we produce.

A HUMANISTIC SCIENCE. My analysis puts the sociologist in the same seat with the literary and art critic, the philosopher and the historian. Yet there are several vital differences which preserve the scientific character of the enterprise and are crucial to its methods. It is not the open, often skeptical, and dispassionate nature of science. That is no more a property of the natural scientist than of the humanist. It is the common meeting place of good scholarship.

Sociology retains, in all its versions, the attitude and the activity of the empirical scientist. It is in the effort to obtain data, to experience the events and actions about which we write and study that sociology is not the same as the humanities. Our data are metaphorically a text. They are by no means only verbal data.

The quest for the natural event seen in its situated occurrence implies an empirical orientation which tries, however limited that attempt may be, to reach out toward the "raw event." In most accounts of the social construction of interaction and institutions there is an implied natural event which is undergoing definition, construction, organization. Empirical study is one hopeful way to try to get beyond the phenomena, beyond the filter of culture, interests, and language to the irreducible noumena.

Whether or not such "obdurate facts" are discovered, the social scientist is as much bound by his/her data as the literary critic is by the text. There are constraints inhering in the data that limit the plausibility and the credibility of interpretations. The more paradigms, the more perspectives, the more the very conception of data reveals its partial character and its debt to the cultural, ideological, or other forms which limit the possibilities of study.

What is implied in my paper is a dialogic view of the social sciences. It is in the disarray, the diversity, the constant clash of studies and conclusion that much of the value of the social sciences can be found. What they contribute and create is less a stock of correct answers, of universal concepts or cumulating laws, than an awareness of possibilities. In enabling us to better reflect upon our assumptions, in expanding our imagination, in uncovering the relevance of seemingly irrelevant conditions and actions, that is no small accomplishment.

Comment

JOHN KITSUSE

As Gusfield's paper documents, he commands a detailed body of information about the social and cultural organization of the factual and moral status of alcohol and its use in American society. I consider his work on the drinking-driver problem an excellent example of the social construction of social problems.

Having noted this theoretical solidarity, I would like to comment on some of the issues Gusfield touches upon in his characterization of what he refers to as "interpretive sociology." It may be helpful to frame these comments with a quotation from Richard Zaner, a philosopher of the phenomenological persuasion. He says, "a philosophical theory that is incapable of accounting for its own possibility is a radically mistaken one. Similarly, any empirical scientific theory whose philosophical foundations are not thus self-accountable is in serious difficulty." Reading Gusfield's paper as a discussion of how interpretive sociology might give an account of the presuppositions of its theorizing, I would like to pick up on some of the issues he poses.

As a theoretical view based on the phenomenology of Schutz, interpretive sociology generates a principle of method and an acute self-conscious attitude in the social scientist in his or her observations and theorizing about the social world. This injunction to be self-conscious is general, based on the recognition of the social scientist as human and thus an instrument of perceiving, observing, feeling, thinking, and acting in the world like the actors who are the subjects of his research. More specifically, disciplined self-consciousness is conceived as the stance of the disinterested scientist who, according to Cicourel, is required to replace his personal biographical situation with a scientific situation.

This disciplined self-consciousness might be viewed as an analogue to the "control of variables in the natural science model of research" to which Gusfield refers. From a methodological point of view, the interpretive social scientist's control enables him or her to move further into the world of the actors to explore the particularity of their understandings of that world. In contrast, the natural science control of variables is the warrant for holding these particularities constant for purposes of various analytical manipulations of the data.

As I read Gusfield's discussion of the differences between the actors' interpretive process and that of the social scientist, he conceives of it to be a matter of the relative degree of being self-critical. He says that

> the interpretive practices that characterize the common world of human beings also apply to the sociologist. Selectivity, taken-for-granted assumptions, and human interests are attributes which the scientist shares with the natural attitude of the common man. This does not mean that the scientific is not any different from the natural attitude. The sociologist is more self-critical, more willing to hold matters in abeyance, to search for alternative explanations. Nevertheless his/her interpretations are also one among the multiple, possible interpretations. To make sense out of the actions or data observed, the sociologist must utilize some scheme, some interpretive practices which create the order that constitutes his/her account of the observations.

Now it seems to me that the difference between the scientific and natural attitudes is not that the sociologist is more self-critical than the everyday actor, but *what* the sociologist is self-critical about. His or her task is to develop a theory about the social realities observed, including the common-sense theories of the actors who are the objects of study. In attending to this task, the sociologist's self-critical stance is to alert him or her to the requirement that the second-order constructs, that is, the theorizing, remains faithful to and congruent with the first-order constructs recorded in the data. Schutz states that "the thought objects constructed by the social scientist in order to grasp this social reality have to be formed upon the thought objects constructed by the common sense thinking of men living their daily life within their social world." Such an injunction offers the occasion for what Joe Gusfield has called in his paper the "obdurate facts argument" that critics direct against the theoretical perspective.

I think of this argument as the "Do you mean to tell me?" argument, as in "Do you mean to tell me that those people in Jonestown drank those portions of Kool Aid at the direction of their leader?" Or "Do you mean to tell me that the bus drivers in San Francisco refused to handle transfer slips because they are afraid of contracting AIDS?" On the evidence, what can we say but "Yes, that is what I mean to tell you." If these are the data, the first-order constructs, are they to be conceived as bizarre delusions? If so, what are the methodological rules that enable the social scientist to distinguish them from what he or she might consider more reasonable interpretations of social reality?

The method of the participant observer that distinguishes interpretive sociology and that can produce evidence of such thought objects and activities may call the social scientist's credibility as "objective observer" into question. He or she may be seen as manifesting the classic symptoms of going native, a condition that undermines the grounds of his or her scientific situation. "Poor old Sam, he's disappeared into the everyday world of multiple social realities." The condition makes him capable of describing that world no better than the common man and disqualifies him to engage in the task of rendering a second-order interpretation of his experience.

The question, then, concerning the constraints imposed on the thought and action by the "obdurate facts" should first of all be directed to the social world of the actor, not to the professional world of the social scientist. Without doubt the facts "have the force of constraining thought and action for common men and social scientists alike," but an interpretive sociology must provide the theoretical possibility that even what Schutz calls the paramount realities of the physical world may hold different meanings for them. While the interpretations of the social scientist may be constrained and controlled by the scientific canons of logic and proof, the interpretations of the common man about which the social scientist seeks to give an account may be much less restricted, much more variable.

General Discussion

Joseph Gusfield: Phenomenology has often been criticized for doing away with any kind of definitive statement that one can make about what is happening and consequently reducing the role of the observer to reporting on what he finds from his relationship to subjects. I am very reluctant to take away from sociologists the kind of activity that tries to give some sense of what are the consequences of actions. What I have been admonishing here (and I do find a link between the interpretive sociology, the phenomenologists, and the symbolic interactionists) is to get as close as we can to the raw experience which is involved; to recognize what it is that we are doing when we are making those interpretations, and what kinds of assumptions we must make about the first-order constructs of the people we study. Otherwise we get ourselves into that kind of solipsism that phenomenologists have often rightly been accused of.

Allen Grimshaw: I would prefer a structural explanation for some of the drinking-driving issues in terms of Jim Coleman's fourfold macro-micro junctures that he is going to talk about tomorrow. Take, for example, what is called "pilot error" in air transportation. What goes on in the assessment of culpability in air traffic is that there are, among others, five principal corporate actors: the pilots' union, the carriers, the insurers, the aircraft manufacturers, and the airport air traffic control system. Two of those corporate actors, the pilots' unions and the carriers themselves, have stakes in finding that it is not pilot error and those stakes are very substantial. In Coleman's typology, you look at drivers atomistically and it might be useful to see how those relations can be defined in terms of a set of corporate actors.

Joseph Gusfield: I think what you are suggesting is quite right, that the definition or the version of reality that gets promulgated and gets accepted has much to do with the nature of the social actors who are involved. I would not deny that, but I am interested in how that interpretation gets done. And, on that basis, suggest different ways of looking at problems connected with alcohol.

David Heise: I am interested in your use of the word "evil," like "the evil drunk driver." I did hear you say "moral" as opposed to something else, but I don't see how you are coming at the definition of evil, the moral issue in a different way.

Joseph Gusfield: What I am coming at is the conception of drinking-driving as a serious criminal action. You find it in the organizations developed in the last two or three years—MADD (Mothers Against Drunk Driving), the call for severer punishments, in short the effort to portray the individual drinking driver as equivalent to that particular model of the killer drunk. And in that sense you are dealing with what I call a moral issue.

David Heise: Let me just persist a little bit. Would you say that taking blood content to be an "as if" for drunkenness is the same kind of "moralizing" as taking the "killer drunk" as an interpretive model?

Joseph Gusfield: Just two years ago the state of California and a number of states in the last two or three years in the United States have passed what are called "per se laws." Under the per se law, you cannot rebut the evidence that someone has been under the influence of alcohol if their blood alcohol level reaches a certain point, .10 in most states in the United States. There is a wide range in susceptibility to influence based upon alcohol. It depends upon how much experience you have with it, how much you have had to eat, your weight, etc. The "as if" quality can be seen as an index that the law has decided upon because it is the only convenient way they can handle the matter. But what I and others have found in research on drinking is a tendency to assume always the most alcohol hostile orientation; that anything that might provide a justification for either moral or legal condemnation will be accepted.

Harrison White: I think you are drawing a false dichotomy. You keep pushing natural science away as if it was that different. I think that they are living with very similar kinds of uncertainties. At the same time, you may not drive the phenomenological point far enough. David Phillips has found that after a TV news or media illustration of a suicide or an execution or one of these other things, there is a noticeable rise in, among other things, accidents on the highway. It would seem to me this raises two specters of third-order phenomenological problems, of your interpreting how they interpreted the reports.

Joseph Gusfield: Phillips has demonstrated that increases in stories about death of various kinds lead to death, sometimes of the same kind and sometimes other kinds. But there is no way he can explain what is happening. A wide variety of problems is involved when you ask: What constitutes an automobile death? For example, if somebody dies 90 days after an automobile crash would it be considered an automobile death or not? Some jurisdictions will so consider it; some will not. If they find some relationship, is it a suicide or is it an accident? By and large police have not been terribly concerned with that and if they find an automobile involved at all, they take it as an accident unless they can find a note. So it is a question of how, in turn, the police and the coroner interpret or determine the situation. I have been more interested in the character of it as a public problem than in the question of the accidents themselves.

Arthur Mann: Is this moral judgment unique to the United States? Take a place like Sweden where they drink very heavily, or the Soviet Union or Ireland. Do you happen to know if in those countries the authorities place a moral judgment on drunkenness?

Joseph Gusfield: I think there is nothing particularly astounding in the moral judgment about drinking that has existed in the United States. I was more interested in the relation between the technical order and the moral order. Sweden has had a strong temperance movement for a long time, still has, and so has Finland. Both approximate the sense of condemnation found in the United States. But there are other countries where that is not case.

Mancur Olson: I enjoyed the presentation and even found it quite charming. When it was over I wondered why I had enjoyed it so much. On reflection I think there is a key to why I liked it.

Quite often you spoke of myths, such as the myth that blood alcohol gives a good measure of a complex phenomenon like drunkenness. I was expecting that the myths would be contrasted with truths, or at least with less mythical ideas. But that expectation was not met. You went on to play down the idea of changing policy, even though you had suggested that existing policy was based on myths. How we conceived of or reacted to or felt about the problem was more important for you. I had a little trouble relating that to your earlier complaints about myths, because if we were doing something based on myths we must have been making some mistakes. Then you went on to talk in a rather denigrating way about the facts that are out there, and you were running down a central characteristic of all science, that is, the idea that science and scholarship should seek truth from facts and that we should try to verify theories and ideas and hypotheses.

At that point it finally occurred to me why I liked your talk so much: It was like a superb sermon. Shouldn't you then call it not a humanistic approach to social science but a theological approach to social science?

Joseph Gusfield: By "myth," I certainly don't mean that something is necessarily false. I mean by that a story, a picture that serves as a way of crystallizing a set of ideas about something. The condemnation of the drinking driver, the severe punishment, is really to be understood as a way of dramatizing and expressing a set of feelings about the phenomenon of drinking-driving. The effort to suggest policies which are rational and prudent ways of dealing with that problem is not so relevant to that aspect of it. It is just as if we talk about capital punishment from the standpoint of whether it is an effective deterrent, without recognizing that it may indeed be a way of expressing emotions, of revenge, of retribution. And, I guess, trying to focus on this expressive aspect is partly theological.

Theda Skocpol: I heard you say that the image of the killer drunk is dramatized in public statements when people make general remarks about how terrible drunk driving is, and it is used to construct laws which result in the arresting of a fair number of people. But you also said that when they got into the courts, people just set that frame aside and the lawyer and the accused and the judge deal with the situation as if they really knew what was going on. What I found interesting is the distinction between these two constructions. All legal proceedings have an element of moral drama; Durkheim certainly taught us that. And yet, when courts get down to deciding whether or not Professor "X" who was drunk on the way home from the department Christmas party should really be sent to jail, they set aside the moral drama and deal with facts in almost a scientific manner.

Joseph Gusfield: I think you put it very well. What I called the negotiated reality at the level of courts and the public drama that goes on do not quite relate to each other. One level is facing up to the practical problem of immediate and particular people and the other is dealing with a dramatic and public action.

Organization Theory

Firm and Market Interfaces of Profit Center Control

ROBERT G. ECCLES AND
HARRISON C. WHITE

The Puzzle of Decentralization

Large American manufacturing firms have widely adopted some form of decentralization or divisionalization in past decades (Chandler 1962; Vancil 1979; Haspeslagh 1983). Why? Our argument will be a cross-sectional one; it explains, without depending on imitation or history, why decentralization makes sense here and now for current executives of large manufacturing firms. Our argument also is a structuralist one: For both the innovators and later adopters, divisionalization must be interpreted with special reference to the context of what the other firms in a sector of the economy were doing.

Vancil (1979) conducted one of the few extensive and systematic surveys of decentralization practices, and his work has been an important guide and landmark in the development of our argument. Vancil's focus was the autonomy, real and perceived, of the lower manager under decentralization. We concentrate on the control perspective of the senior central executives—whether as individual Chief Executive Officer (CEO) or Chief Operating Officer (COO), etc., or as some form of group executive.

Our approach rests on, and part of its novelty consists in, our view of the subordinate managers as being controlled through their locked-in participation in social mechanisms, which we call interfaces, with other like managers.

Vancil is construing the relation between top executives and managers of divisions as one of domination, and he further focuses upon how this structure of domination is seen from below. We argue that this exactly misses the point of the decentralization innovation. The control relation between executives and lower managers is converted by decentralization into a different form, from domination to exploitation. "Exploitation" means that the lower manager is constrained by the objective features of social context to carry out a pattern of action attuned to the interests pushed by the controlling group, here top management, without the latter needing to exert direct domination. In support of this argument we cite our reading of the evidence that it is fairness, and not autonomy, where most difficulties of upper management now center.

Our underlying interest is social control with modern, Weberian rational organization. Udy (1970), Wallerstein (1980), and Braudel (1982), among others, have pioneered this topic. We argue that modern forms of control are flexibly decoupled from constraints widespread in that society and are distinctively geared to large size. In part, large size may flow from technological and efficiency reasons (e.g., Woodward 1965; Chandler 1977), but in our view modern rational control is a leverage process which works on and presupposes a disjunction between one social level and another, so that the larger the organization, the better. More energies are channeled by modern rational controls precisely because they are designed to be indirect controls, controls which presuppose and use the enormous energies of self-reproducing social mechanisms such as interfaces. Modern rational control in industry, as elsewhere, is essentially a piggyback operation in which domination, though present, does not account for the distinctive gains in scope and effectiveness over traditional, socially imbedded organizations. We attempt in this paper to contribute to this general line of argument by focusing on a particular, and we believe crucial, level in an industrial form of modern organizations.

Firms and Markets

Following present American business convention, let us call the relatively autonomous divisions of a firm which are at issue *profit centers.* Our central idea is that any concrete part of our big business economy is organized around profit centers as the atoms, with two different and simultaneous forms of combination into molecules. Profit centers from diverse firms build together a market interface; profit centers from diverse markets together form each big manufacturing firm. This is not some passive and symmetric matter of architecture. On the contrary, this system articulates chief executive strategies, and it manipulates and presupposes natural crescive networks of support and alliance among managers inside and across firms. Chief executives have in effect delegated detailed operational control of each profit center to the joint discipline of the interface constituted by comparable profit centers from other large firms. A profit center manager first and foremost must attend to maintaining a viable niche in his market interface.

Each chief executive surveys a portfolio of divisions, profit centers from different markets. The crucial authority to commit new resources for shaping market niches remains with the chief executive, including the endlessly considered options of going into and getting out of markets. Entry and exit by producers of markets is much rationalized, since it is only a separable part of a big firm and not the whole entity that is to be created or disposed of. Each chief executive is continually balancing off one with another of his profit center's futures. What he expects of a profit center depends both on the overall context of the environment of that market and on the particular niche his profit center has won.

Other chief executives are carrying on the same analyses; so the array will never be still as actions by one big firm trigger changes in context and competition for others. One firm's acquisition is another firm's divestiture. No two firms are likely to have the same array of market memberships, so that it will be hard to find clear segmentation of the economy into sectors. In contrast to the perspective of economics, here the essential pressure of control is seen as coming not from the "other side" of the market, but from one's own side of producers. At the same time, this control has an inescapable flavor of collaboration. For example, these big firms often can find goods needed by one profit center in the output of another of their centers, thus bypassing the market interface: But if you always find sources inside, where are you going to turn for quotes of bids to keep your system calibrated?

It has been customary to contrast firms and markets as *alternative* forms of social organization for allocating resources (Arrow 1974; Williamson 1975). In firms the allocation and control of resources is through a hierarchy of authority. In markets it is through the price mechanism. However, those who have examined actual formal organizations and transactions between them have shown that hierarchies can have properties of markets (Chandler 1962; Vancil 1979) and that markets can have properties of hierarchies (Corey 1976, 1978; Eccles 1981; Stinchcombe 1982). The former is illustrated by the organization of a firm into a number of profit centers which are modeled after the markets of economic theory. The latter is illustrated by relationships which continue over time, are not determined solely by the price mechanism, and even have elements of hierarchical authority.

We argue that the emergence of the multi-profit-center firm has established a new unit of analysis, the profit center, which is at the foundation of a theory of both markets and hierarchies (see Salter & Weinhold 1979; Porter 1980). Since markets are composed of profit centers and so are firms, the position of a profit center in its market will be affected by its position vis-à-vis the other profit centers in the firm; conversely, the position of a profit center in the firm will be affected by its position in its market.

Portfolio Planning

We propose that the self-sustaining interface property of markets applies in some way to the firm as a set of profit centers, with its senior executives as the buying side. We argue that portfolio planning is a management technique for designing this interface. This interface is a mechanism of control by the CEO, and one based on identifying the position of each profit center in its own market. Thus, a "hierarchy interface" is built. The CEO can be conceptualized as a customer of profits or cash flow in return for certain investments in profit centers. The profits center managers are producers of these profit flows. Thus, the hierarchy interface presupposes all of the separate market interfaces which these profit centers participate in.

After this "hierarchy interface" is negotiated, positions of individual profit centers in their markets may shift, thereby disturbing the existing market interface. This sets forces in motion to restore a self-sustaining market interface, which may be different from the original one. But the resulting position of the profit center in this new market interface may disturb the hierarchy interface, as its position vis-à-vis the rest of the profit centers in the firm is now changed. It is clear how the CEO's problem of control and the problem of creating self-sustaining interfaces which yield markets (White 1981b) are intimately related. Figure 1 diagrams this phenomenon.

The phenomenon can be illustrated by examining profit center $PC_{D,5}$ in the middle of the matrix shown in Figure 1. It is one of nine profit centers in Firm D and is part of hierarchy interface D. It is also one of seven profit centers competing in Market 5 and is part of market interface 5. Changes in either interface can change characteristics of this profit center, which in turn affects the other interface.

The multi-profit-center structure has become especially common in large organizations. Wrigley (1970) estimated that 86 percent of the 500

Figure 1 Firm and Market Interfaces of a Profit Center

Firms

Markets	A	B	C	D	E	F	G
1	$PC_{A,1}$			$PC_{D,1}$	$PC_{E,1}$		$PC_{G,1}$
2		$PC_{B,2}$		$PC_{D,2}$	$PC_{E,2}$	$PC_{F,2}$	
3		$PC_{B,3}$		$PC_{D,3}$		$PC_{F,3}$	$PC_{G,3}$
4	$PC_{A,4}$	$PC_{B,4}$	$PC_{C,4}$	$PC_{D,4}$	$PC_{E,4}$		
5	$PC_{A,5}$	$PC_{B,5}$	$PC_{C,5}$	$PC_{D,5}$	$PC_{E,5}$	$PC_{F,5}$	$PC_{G,5}$
6			$PC_{C,6}$	$PC_{D,6}$	$PC_{E,6}$		
7				$PC_{D,7}$	$PC_{E,7}$	$PC_{F,7}$	$PC_{G,7}$
8	$PC_{A,8}$	$PC_{B,8}$		$PC_{D,8}$	$PC_{E,8}$		
9				$PC_{D,9}$	$PC_{E,9}$	$PC_{F,9}$	$PC_{G,9}$

largest industrial companies (the so-called *Fortune* 500) had such a structure in 1967. Rumelt (1974) estimated that the use of this form by the 500 largest industrial companies increased from 24 percent in 1949 to 51 percent in 1959 to 79 percent in 1969. (These figures are for product division, geographic division, and holding company structures. The first is the vast majority of cases in all years.) In a 1976 survey of the *Fortune* 1000 largest industrial companies, Reece and Cool (1978) found that 95.8 percent of the 620 respondents had a multi-profit-center structure.

The conventional explanation given for organizing the firm into profit centers is that when a company diversifies into many products and markets, the increased complexity overwhelms the decision-making capacity of the CEO. The solution is to have lower-level managers make the general management decisions which involve the tradeoffs and integration of functional resources (R&D, engineering, sales, manufacturing, etc.) as if they were the CEO of a small firm.[1] Thus, revenue growth,

[1] The organization of the firm into profit centers may go several levels deep. In some large firms there are groups which are composed of a number of divisions, which in turn are composed of a number of product lines. Groups are typically composed of divisions which are related in some way, such as technology, products, or markets. The degree of interrelatedness of products in a division is even higher. See Vancil (1979).

profitability, and return on investment become primary measures of performance.

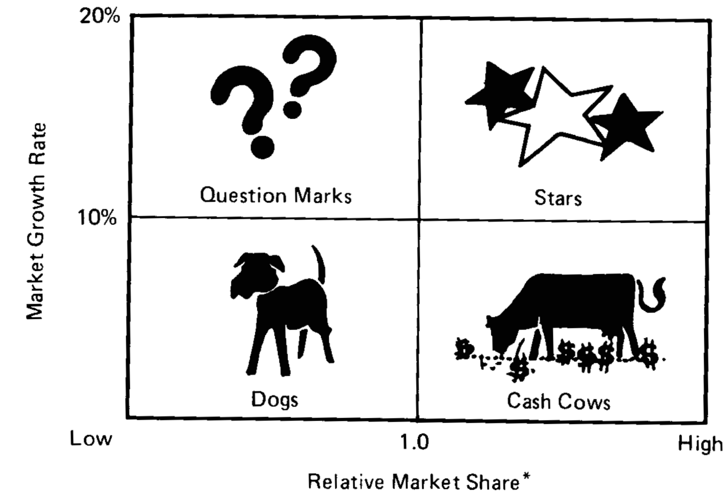
While it is possible to decentralize responsibility for operating management—i.e., the day-by-day decisions that must be made to deliver the product to the market—it is not possible to decentralize resource allocation decisions. Both operating expenses and development expenses are incurred by the profit centers. The CEO and the rest of the top management group must decide how to allocate scarce resources among the many profit centers. This problem is exacerbated by the very advantage of the multi-profit-center organization. Since it relieves the CEO of detailed involvement in the different businesses, he does not have the same type of in-depth knowledge, however imperfect, he would have had under a functional structure in an undiversified firm. This problem is exacerbated by the tendency for profit centers to proliferate to lower and lower levels. It is not uncommon in large diversified companies to find tens or even hundreds of profit centers. When this is the case, how is the CEO to make decisions on how to allocate limited resources across the various profit centers?

One possibility, which is both simple and has a certain appearance of fairness, is to divide the total pool of available resources across profit centers as a percentage of sales, with all profit centers receiving the same percentage. The major problem with this is that greater market opportunities may exist for smaller profit centers than for larger, which may justify a disproportionately large investment in them. But how can these determinations be made given the existence of many different profit centers? Providing the CEO with detailed information on each profit center is not a solution. He can never understand the profit center's market in the same depth as its manager. But the fact remains that somehow resources must be allocated.

A number of portfolio planning approaches exist. Some of the most popular ones are reviewed by Grant and King (1979) and Abell and Hammond (1979). Bruce Henderson, founder of the Boston Consulting Group (BCG), initiated the growth-share matrix which illustrates how these techniques use market interfaces to create hierarchy interfaces for obtaining CEO control (Henderson 1980). Hamermesh (n.d.) discussed the administrative challenges of this technique. Figure 2 shows the growth-share matrix. On the vertical axis is the growth rate of the market, typically divided into high (greater than 10 percent) and low (less than 10 percent).[2] On the horizontal axis is the relative market share of

[2]Henderson developed his approach in the years around 1970, and concrete details already have an anachronistic flavor.

Figure 2 The Boston Consulting Group's Growth-Share Matrix



Note: Adapted from Abell and Hammond (1979).
* Market share relative to largest competitor.

the company's profit center in terms of the largest competitor in the market. Ratios of greater than one are considered to be large market shares, and less than one are considered to be small.

From this matrix four quadrants are identified by Henderson. "Cash cows" are profitable products with little growth potential; the profits generated from these businesses are used to fund others. Investment can be in "stars," product/markets where the company has a large share but rapid growth requires high investment. Investment can also be in "question marks," product/markets with high potential but which require large investments because of both market/growth and the low market share of the company. The BCG approach emphasizes the importance of market share based on the concept of the experience curve, or the tendency for costs to fall with increases in cumulative volume and thus market share. The remaining quadrant, "dogs," contains product/markets where the company has a weak position and where there is little growth potential. These should be divested.

Although the growth-share matrix is conceptually simple, it illustrates how the markets of the various products are used to create a hierarchy interface, with the dispersion required for the latter being provided by the fourfold (or more) discrimination of profit centers. Markets are compared in terms of growth rates, and then profit centers are compared in terms of both their market shares and the growth rates of their markets. The ordering of profit centers that results, combined with the assumption (that has some empirical support) that large market share is the key to profitability, can assist the CEO in making resource allocation decisions. The CEO knows that as such ordering emerges it induces new energies of maneuver among his subordinates.

Of course, profit center managers soon learn the nature of this game. They spin informal networks of information and disinformation among themselves and via other managers. Since it is they who are in the best position to provide data on market growth rates and market share, they have the opportunity to define markets in such a way that their business does not appear to be a dog. Thus, while an aggregate view of a business may result in a "dog" classification, a more refined view which recognizes several different markets may uncover segments which are more attractive and where the company has a strong position. Abell (1980) expressed this in terms of the importance of creativity in defining a business, i.e., a market:

> The importance of market share as a determinant of profitability has been widely publicized and recognized in management and academic circles. In addition, it has been incorporated into several formal planning approaches as a critical variable with which management can influence profits. To view it as *the* primary strategic lever is an easy and seductive step.
>
> The contrasting view can be stated as follows: strategy formulation requires a more creative act. Market share is the result of such an act, not strategy itself. Business strategy pivots on *defining* the business in a way that leads to competitive superiority in the customer's eyes. [pp. 3–4]

There are two consequences of this approach to defining a business. First, profit center managers are forced to understand the external environment in some depth. Their search for competitive advantage contributes to the formation of the markets described above. Second, and a consequence of the first, there is a proliferation in the number of profit centers in terms of separately identified market segments, although not necessarily in terms of separate organizational units. This makes the technique of portfolio planning increasingly necessary in order to establish resource allocation priorities. As a result, the CEO retains control while at the same time knowledge of external markets lower in the organization is increased. And this knowledge contributes to the formation of self-sustaining interfaces which define and support these markets.

Portfolio planning contributes to the formation of a hierarchy interface by specifying the appropriate degree of dispersions among profit centers. Clearly the CEO does not want all of the profit centers to be located in the same quadrant of the growth-share matrix, nor can they be in most cases. The CEO is concerned with the growth rate of the profit center relative to the growth rate of the market since if the former is below the latter the profit center will lose market share. Thus, there are limitations on how many profit centers can be in the star or question mark regions since they require relatively high investments to keep pace with market share. In order to have the money to invest in profit centers in these regions, the company needs enough cash cows generating sufficient profits to fund these opportunities. And obviously the company wants to limit the number of dogs, particularly if they represent cash drains. Thus, the company wants a spread of profit centers across at least three of the quadrants. But it is inevitable, that, over time, as markets evolve and change, some of the company's businesses will fall into the dog region, or else some part of an existing profit center will be identified as being in this region as a result of the portfolio planning exercise.

When the CEO decides to increase the company's market share in a particular business, such as a star or question mark, or to decrease it, such as in a dog (perhaps as part of a plan to get out of the business) this may change its volume, cost structure, and quality. Thus, maintaining the hierarchy interface affects the market interfaces. But those changes affect other companies which are also concerned with *their* hierarchy interfaces. These latter, in adapting, cause further changes in markets, in which still other firms have profit centers competing, and these changes may require adjustments in the hierarchy interface of the original company, etc.

By our argument, portfolio planning is a natural consequence of the widespread implementation of multi-profit-center structures within firms. An economy composed of many firms making many products which compete in many markets is qualitatively different from the economy of classical microeconomic theory or from theories of monopoly and oligopoly which focus on single markets.[3] Today's economy is built on transactions between profit centers, which are embedded in firms, not upon transactions between single-product firms.

[3]William Baumol brought this to economists' attention, and he has led them in their usual pursuit of rationalizing away new phenomena by conjectural elaborations of their familiar existing rhetoric of optimality; for a survey, see Bailey and Friedlander (1982).

Portfolio Planning Survey

Support for this argument is provided by a survey on the use of portfolio planning in large diversified firms conducted by Philippe Haspeslagh (1983) and sponsored by the *Harvard Business Review*. He surveyed the *Fortune* list of the 1,000 largest industrial firms and found "that portfolio planning approaches are widespread among large diversified industrial companies and being increasingly introduced" (p. 305). These techniques were developed in the 1960s and became increasingly popular by the 1970s, by which time the vast majority of large diversified companies were organized into a number of profit centers. Haspeslagh estimated "that, as of 1979, 36% of the *Fortune* 1000 and 45% of the *Fortune* 500 industrial companies had introduced the approach to some extent. Each year during the last five years, another 25 to 30 organizations have joined the ranks" (p. 305).

Use of this technique is less common in single-industry companies or companies where it is difficult to treat businesses independently of other businesses because of common technologies or manufacturing processes (e.g., aerospace, tobacco, mining, and furniture). It is also used infrequently in conglomerates which allocate resources in this fashion without the need of a formal tool. "Conglomerates, on the other hand, speak portfolio planning prose like Monsieur Jourdain in Molière's *Le Bourgeois Gentilhomme*—sans savoir" (p. 313).

Portfolio planning is most common in diversified companies whose businesses are related "because of the difficulty they have in assessing the strategic performance of each of their businesses and allocating resources selectively" (p. 311). Haspeslagh estimated that 75 percent of these companies were or planned to use portfolio planning at least at the corporate level. Companies in this category included those in the chemicals, electronics and appliances, paper and fiber products, food, and petroleum refining industries.

Haspeslagh compared companies which did *process* portfolio planning (a well-developed form of portfolio planning in which top management explicitly negotiates strategic missions with strategic business unit managers) with those which did not in terms of issues of most importance during the annual planning review. For those which did no portfolio planning, next year's profit objectives were of greatest importance, followed by long-range profit objectives. For those which did process portfolio planning, long-range profit objectives was first (of nine objectives), long-range resource allocation was second (sixth for no portfolio planning companies), long-range sales objectives was third (fifth for the others), and next year's profit objectives was sixth. *Thus, portfolio planning is not being used as a mechanism for obtaining greater involvement in the profit centers on a day-by-day basis but rather to create distinctions among profit centers in order to develop a self-sustaining interface.*

The emphasis on the long term is necessary for building a self-sustaining interface intended to be stable for some period of time; portfolio planning is being used to make current resource allocation decisions to achieve long-term objectives in a kind of informal linear programming approach. In fact, Haspeslagh criticized many companies for focusing primarily on capital investment (costs charged over a number of years) and not items which are expensed such as R&D, marketing, and applications engineering (costs charged in the year they are incurred).

No less important is being able to create a spread or array of distinctiveness among the profit centers, giving each its distinctive strategic mission. Since the mission must be negotiated with the CEO, and vis-à-vis the strategic missions of the other profit centers, the CEO retains greater control; he would have less control if all profit centers were expected to achieve some level of profitability and then, as long as they did so, were left alone. Any lack of dispersion would reduce the opportunities for CEO control.

Transfer Pricing

In most real firms, leaving aside extreme conglomerates, profit centers cannot be seen as separate beads held together on a string of relative position in portfolios.[4] Units of whatever kind within the larger firm are not as cleanly independent of each other, even in operations, as the simplified financial formulations of portfolios would suggest. Transfer pricing—how vexing issues of sourcing and pricing flows among units within the firms are resolved—is a natural arena for viewing the interdependencies upon which are laid down the multidivisional structures which we argue serve major purposes of control.

In theory, a profit center manager has all of the necessary functional resources to produce the products for the markets he competes in. In practice, profit centers often obtain resources from other profit centers (as well as from corporate staff and line functions). One profit center may

[4]See Eccles (1983) for a more detailed discussion.

produce a good which it sells externally that is also an intermediate good for a product sold by another profit center. Since profit center evaluation requires identifying all costs incurred in producing the good, a *transfer price* must be established on this profit center exchange. Since this transfer price affects the measured performance of the profit center managers it is often the source of conflict.

Transfer pricing is a subject which has received extensive treatment in the economic and accounting literature, although few conclusions of practical significance have emerged. (See Hirschleifer 1956 for the first formal economic treatment of the problem, Anthony 1965 for a critique of the economic approach, and Anthony & Dearden 1980 for a managerial approach.) Not one of these treatments explains the data on transfer pricing practices in one of the few empirical studies of this problem provided by Vancil (1979) in the course of his monograph on decentralization.

Transfer pricing, like portfolio planning, is an inevitable result of the multi-profit-center organization and existed from its inception. Chandler (1962) discussed the emergence of this problem in the early 1920s at DuPont and General Motors when the multidivisional structure was created. References to this problem can be found earlier than 1930 (National Association of Cost Accountants 1925). It is one aspect of the problem of creating a self-sustaining hierarchy interface, and we think it should be seen as a mechanism of CEO control.

We have argued that actual firms and markets should be analyzed in terms of the extent to which each contains the properties of both the markets and the hierarchies of economic and sociological theory. In particular, markets connote decisions freely made by actors based on price and volume. A buyer chooses from whom he will purchase a needed good depending upon the terms offered, compared with those of others (White 1981a, 1981b). In its limiting form the same price is determined by the market across all volumes and if a seller is just a bit higher he will have no customers. In practice, sellers provide a range of prices reflecting differences in volume, quality, service, dependability, etc., and the buyer makes his choice based on price and other factors (Corey 1976, 1978). In any case, price and volume terms determine source.

The introduction of cost, revenue, and profit measures into a hierarchy is the introduction of market concepts. Thus, the multi-profit-center firm is a hierarchy with substantial market characteristics since the profit centers are modeled after the firms of economic theory. Transfer pricing practices can be understood in terms of the degrees to which elements of hierarchy and market are contained within the firm. These degrees reflect differences in the CEO/profit-center interface which are a consequence of the control needs of the CEO. These degrees also correspond to differences in the relationships between the profit centers.

We want to distinguish three different ranges of operation that can be found in the hierarchical control interface among the profit centers of a firm subject to the CEO. At one extreme would be a conglomerate or holding company where different units have little interconnection; a less extreme example would be GE. "Competitive" is suitable as a label for this range, where the limitations are most severe on the CEO's ability to understand in any detail the products, markets, and technologies of the profit centers. Consequently, he must exercise control primarily in terms of outcomes from the profit center and thus use strictly financial measures as the common basis for evaluating performance and making resource allocation decisions. These are the firms Haspeslagh identified as having used a de facto portfolio planning approach before it became a formal technique.

In this *first* range for hierarchical interface, the "competitive" range, emphasis on outcomes rather than on specifying the actions to be taken reflects a strong presence of market characteristics compared with hierarchy characteristics. The transfer pricing policy for inter-profit-center relationships consistent with this form of control is to let the profit center managers determine between themselves whether or not to engage in a transaction. They will negotiate a favorable price, a "competitive" price, or the transaction will not take place. As in the ideal markets of economic theory, price will heavily influence the exchange decision.

We call the *second* range of hierarchical interface "integrated." It is found in firms having much less diversity among profit centers, which are typically organized into vertical chains through integration within the firm of multiple steps of the production process. At each stage some of the output is sold externally and some is transferred internally for further processing. When there is a high degree of vertical integration among closely related businesses the inter-profit-center transactions must be mandated by the CEO to preserve the required coordination necessitated by interdependence. This is a strong use of hierarchy relative to market, since price is irrelevant to both the buying and the selling profit centers. But since the transaction is between profit centers it is important to measure costs. By using a full-cost transfer price which does not include a profit margin, the total costs required to produce the product of the buying profit center are determined as if this profit center were producing the intermediate good for itself. The selling profit center is treated as a profit center only for the product which it sells externally, where it has the autonomy to make transaction decisions.

The most sophisticated use of transfer pricing for control is found in the *third* range of interface, which we call "collaborative." This range occurs in firms which are high in both market and hierarchy characteristics. The term "collaborative" reflects the simultaneous presence of both competitive and cooperative aspects. This third range abuts both the others: vertical integration is high, which results in mandated internal transactions. Diversification is high due to an emphasis on defining, in as much detail as possible, the separate strategic business units, even though they may share common technologies or manufacturing processes. In firms of this third range, diversification is a result of the creative act of defining businesses noted by Abell (1980), as compared with the more inherent diversification of firms producing products unrelated in terms of technologies, customers, and competitors. The emphasis on diversification requires market price transfers in order to reflect as closely as possible the profitability of the profit center as if it were a stand-alone business.

The extreme difficulty of agreeing upon market price when a transaction is mandated (how can it be a market price if it was not determined in a market transaction?) creates conflict which becomes a mechanism for CEO control in this third range of operation for the interface within the firm. The profit center managers have an incentive to seek the CEO's involvement in adjusting the market-based transfer prices. They provide him with useful information when doing so, such as prices being offered by external suppliers and paid by external customers.

Vancil's survey (1979) of 291 manufacturing firms, discussed earlier, lets us discriminate among these three ranges of interfaces using data on transfer pricing. The best data are on transfer pricing *method,* and we put aside the degree of autonomy in selecting transactions. The firms used in his study were reclassified into the three ranges of competitive, integrated, and collaborative operation of their hierarchical control interfaces.[5] There are two basic considerations in discriminating ranges. First is the extent of transfers between profit centers. Second is the transfer pricing method.

Table 1 shows the extent of transfers between profit centers for the different ranges. Collaborative organizations have the greatest amount.

[5]The reclassification involved in assigning firms to each of these types of interfaces was somewhat complex since some of Vancil's categories (1978) included five types of interfaces. Measures of diversification and vertical integration were used to make assignments. For details see Eccles and Carley (1983).

Table 1 Extent of Transfers in Three Ranges of Interface for Hierarchical Control

(Chi-Square = 65.162 Sig. = .000 DF = 8)

	Competitive	*Integrated*	*Collaborative*	*Total*
No Transfers	18.0%	43.5%	5.5%	18.6%
	(17)	(30)	(7)	(54)
1–3%	33.0	8.7	14.1	18.9
	(31)	(6)	(18)	(55)
4–7	18.1	18.8	23.4	20.6
	(17)	(13)	(30)	(60)
8–15	18.1	13.0	25.0	19.9
	(17)	(9)	(32)	(58)
≥ 16	12.8	15.9	32.0	22.0
	(12)	(11)	(41)	(64)
Total	100.0	100.0	100.0	100.0
	(94)	(69)	(128)	(291)

For competitive interfaces the lack of interdependence among diversified businesses is reflected in smaller amounts of profit center transfers. Interfaces in the "integrated" range also show lower inter-profit-center transfers, but for a different reason: the interdependencies are contained within single profit centers. Vertical integration on top of diversification, the distinguishing characteristic of organizations in the collaborative range of interface, necessarily results in extensive inter-profit-center transfers. As shown in Table 1, 57 percent of the 128 firms in this range have 8 percent or more internal transfers; 32 percent of these firms have 16 percent or more.

Table 2 reports transfer pricing method. One problem with these data is that "negotiation" was used as a method in the response set on the questionnaire. More properly, it describes a process for using a method, something Vancil himself later recognized. It is not unreasonable to assume, however, that respondents interpreted negotiation as referring to the kind of valuation that occurs in a competitive market exchange—the negotiation of a price. This can be contrasted with a more hierarchically determined market price, something more like "list price." Variable cost approximates that marginal cost competitive price of microeconomic theory and would be expected in the competitive range for control interface. Hierarchical exchanges require that fixed costs be included.

In Table 2 collaborative and competitive interfaces are distinguished from integrated interfaces in transfer pricing method. The latter make relatively extensive use of full cost in those cases where there are transfers between profit centers. This method is much less popular in the other two ranges, where the more market-like valuations of cost plus profit, market price, and negotiation are used. As would be expected, negotiation is most popular in the competitive range, where the profit centers determine the price of the transfer through a negotiation that does not involve the CEO. In contrast, market price, a market-based valuation which permits some hierarchical intervention, is more popular in the collaborative type. While these results are at a high level of aggregation, and there are a number of methodological issues which cannot be discussed here, they are suggestive. (See Eccles and Carley 1983 for a more extensive discussion.)

Table 2 Transfer Pricing Method by Range for Hierarchical Interface

(Chi-Square = 76.385 Sig. = .000 DF = 12)

	Competitive	*Integrated*	*Collaborative*	*Total*
No answer	14.9%	40.6%	1.6%	15.1%
	(14)	(28)	(2)	(44)
Full cost	12.8	29.0	22.7	21.0
	(12)	(20)	(29)	(61)
Cost plus profit	14.9	8.7	15.6	13.7
	(14)	(6)	(20)	(40)
Market price	21.3	13.0	35.2	25.4
	(20)	(9)	(45)	(74)
Negotiation	28.7	7.2	16.4	18.2
	(27)	(5)	(21)	(53)
Variable cost	4.3	0.0	5.5	3.8
	(4)	(0)	(7)	(11)
Combination	3.2	1.4	3.1	2.7
	(3)	(1)	(4)	(8)
Total	100.0	100.0	100.0	100.0
	(94)	(69)	(128)	(291)

Conclusion

Markets and hierarchies are not mutually exclusive social forms for the allocation and control of resources. Rather, they should be seen as concepts, derived from social forms of the past, which have been used to create now-existing social forms which intermix the pure types. The multi-profit-center firm is a prominent example. Decentralization of a hierarchy into a number of profit centers creates the necessity for a self-sustaining hierarchical interface internally, analogous to the external market interface confronting each profit center separately. Portfolio planning is one way of creating this hierarchical interface. Since the hierarchy still exists[6] control must be maintained over the interrelationships on the profit center side of the interface. Thus, within the firm there is a need for transfer prices, another example of the use of a market characteristic.

We conclude with some speculations. The control aspect of portfolio planning is further seen in the apparently curious fact that "in practice (except for considerations of capital investment and, of course, the strategic planning system itself) companies do not alter formal administrative systems in accordance with the strategic missions of SBUs (strategic business units)" (Haspeslagh 1983, p. 316). Haspeslagh felt that, in time, companies which did portfolio planning properly in fact would make this alteration. However, the very lack of consistency or congruency between portfolio planning and such administrative systems as performance measurement, evaluation, and reward may be part of the source of CEO control. Without this fit, profit center managers are less able to define a position in the portfolio with objectives that they know they can achieve and thus set up a game they know they can win. Our position is supported by Hamermesh's in-depth clinical analysis (n.d.) of how portfolio planning was managed in several companies.

More concretely, if a manager is designated as in charge of a "cash cow" and rewards are defined which are appropriate for cash cows, such as cash flow to corporate headquarters, he or she has an incentive to continue to define the business as a cash cow and not to look hard at the markets for ways to redefine parts of it that are something else, such as a "question mark." He or she can become complacent and pay more attention to managing the system than working hard at defining the profit center's role in the market interface. Some inconsistency between mis-

[6]As opposed to the complete decomposition of a hierarchy into profit centers which really are independent firms; this appears to be a new phenomenon through leveraged buyouts and other forms of decomposing firms.

sion and rewards keeps the profit center managers on their toes looking for ways to bring about a fit, rather than gaming the system of rewards; at the same time the inconsistency is giving the CEO control.

One of the limitations often asserted about portfolio planning is that managers do not like to be designated as being in charge of certain kinds of strategic business units. This argument typically takes the form that it is hard to provide motivation for managers in charge of dogs or even cash cows. But again, we would argue that this labeling is an important part of creating distinctions and dispersion. Complete acceptance of portfolio planning throughout the organization would have the same result as when any system or structure no longer requires the CEO to resolve conflicts: a loss of control.

A second key fact may be that once a profit center gets in trouble it tends to lose autonomy. This is a practice that tends to be taken for granted, to be treated as self-evident. But when it is missing budget or otherwise runs into unanticipated trouble, a subordinate unit could make the most creative use of maximum leeway. Under portfolio planning, as rhetoric plus reality of executive control, the control is tight. We argue that careful comparative studies could show that operational autonomy in this behavioral sense is *more* pronounced in old-fashioned centralized functional firms. Indeed, we think that it may have been just the increasing autonomy of functionally socialized departments that was a main trigger of the heralded shift to "decentralization."

Haspeslagh asked, "Does portfolio planning constitute a step forward in the management of diversity, or is it simply a passing phenomenon?" (p. 327). He noted that sometimes new systems are used to address current problems and when the problems are solved the need for the new system goes away. Our response to this question is that portfolio planning may be both. It is a step forward in managing diversity, perhaps an inevitable one, given the invention of profit centers. But on a broader scale and over a long period of time it may prove to be a passing phenomenon as the economy of profit centers provides the conditions for the emergence of new social-economic forms, just as the firms and markets of the nineteenth century were the foundation for the emergence of today's more complex interpenetrations of hierarchies and markets.

Comment

JOHN PADGETT

Ever since the pioneering empirical work of Alfred Chandler, the investigation of multidivisional firms has been transforming our understanding of what organizations are and of how they operate. Williamson began the theoretical task by undermining the classical Weberian dichotomy between markets and hierarchies. Through his transaction cost analysis of contracting systems, Williamson underlined the fact that current business management is a mixture of market and hierarchy principles. Hybrid organizational forms such as diversification, profit centers, performance evaluation, transfer pricing, subcontracting, joint ventures, and conglomerates all point to the pervasiveness both of profit-oriented exchange relations within economic units and of overtly coordinated authority relations across economic units. Traditionally, massive organizational boundaries were thought theoretically to segregate hierarchy principles within from market principles without. With Chandler and Williamson, and with economic facts, these organizational boundaries have come tumbling down—so much so that the previously easy semantic identification of "economic production unit" with "the firm" is no longer unproblematic.

As the last section on transfer pricing makes clear, the present Eccles and White paper is directly within this Chandler-Williamson tradition. At the same time, however, Eccles and White are trying to take the next step. It is not only that markets and hierarchies creep into each other's domain, their argument goes; but also that, in the present context of multidivisional firms at least, each is a constituent of the other. The anatomical meaning of this statement is contained in the matrix of Figure 1. Firms are sets of profit centers. Markets are orthogonal other sets of profit centers. Each profit center, therefore, is a member of a market and a firm simultaneously. For Eccles and White, this matrix is intended to be more than an accounting device. The set of profit centers which can congeal into a firm depends upon those profit centers' roles within their individual markets. Conversely, the set of profit centers which can congeal into a market depends upon those profit centers' roles within their individual firms. The main implication of interest to those who think of markets as primeval is that market structure and firm structure *mutually* constrain and influence each other. This is standard sociological reasoning about multiple reference groups applied to an economic phenomenon, but it will be surprising to Eccles and White's straw men economists, who find it easier to contemplate market structures influencing firm structures than the reverse.

If this perspective is to prove useful, empirical predictions must eventually be forthcoming. In previous work on markets as role sets (summarized briefly in this paper), White derived a whole battery of predictions about stability conditions for market existence, about "terms-of-trade" pricing schedules, and about concentration ratios of production units within markets. Nothing comparable is delivered here, but it helps interpretation of what is at stake to realize that with this paper White is beginning to transpose his earlier reasoning about markets into the realm of firms. Over and over again, it is the formal homology between markets and firms, rather than their difference, which is emphasized: Markets and firms both are role sets ("interfaces") of profit centers, engaged competitively in transactions with (and hence defined with respect to) a single third party. For markets, the third party is the (aggregated) consumer, with purchase transactions mediated through terms-of-trade pricing schedules. For firms, the third party is the chief executive officer, with investment transactions mediated through portfolio planning (or more generally through budgets). This homology gives rise to a new view of firm structure: Instead of the traditional emphasis on vertical arrays which place CEOs and profit center managers in a single group unified by the social boundary of "insiders" versus "outsiders," Eccles and White place the emphasis on horizontal arrays of profit center managers, within and across firms, unified by the social cleavage between them and the higher strata CEOs. The unitary actor rhetoric of "organization in its environment" is thereby scrubbed in favor of a rhetoric of elite control.

Even with the project in its current state of development, however, there are a number of predictions about firm and market structure latent in the paper's discussion of portfolio planning. Given the crucial (and perhaps unrealistic) assumption that the Boston Consulting Group technique is typical of portfolio planning methods, these predictions appear to be the following:

1. Multidivisional firms will be "balanced" in composition between cash cow profit centers (largest profit centers in markets with growth rates less than 10 percent) and question mark/star profit centers (profit centers in markets with greater than 10 percent growth). There is no prediction about the relative firm composition of question marks and stars, but in the long run multidivisional firms will not contain any dogs (less than largest profit centers in markets with less than 10 percent growth).

2. Since all CEOs are striving for the same internally balanced firm composition, market composition will differ radically depending upon whether aggregate market growth is high or low. If growth is low, only the largest producing unit will be a member of a multidivisional firm. All the rest, presumably, will be old-fashioned single-product firms. If market growth is high, on the other hand, most producing units will be swallowed up into multidivisional firms.

3. Aggregate growth rates across markets will be more volatile in economies dominated by multidivisional firms than in economies composed largely of single-product firms. This is a consequence of (2): as market growth rate rises, multidivisional firms faddishly rush to invest in question marks and stars. But as market growth rates decline, multidivisional firms divest the dogs.

It remains to be seen whether these and future predictions prove accurate. But at least this new conceptualization offers the promise of new payoffs.

On the more critical side, I have comments both at the level of the paper itself and at the level of the market project of which this paper is a part. The sections on portfolio planning and transfer pricing are poorly integrated and partially contradictory (see, for example, the first sentence of the transfer pricing section). There appears to be an empirical reality that underlies this poor fit. The portfolio planning discussion applies most naturally to those multidivisional firms which the transfer pricing section labels "competitive." The components are not profit centers per se, but rather "strategic business units," among which vertical integration linkages are minimal. The transfer pricing discussion applies to a lower order of profit center aggregation, at which significant material pass-throughs exist. The result is that the transfer pricing section has little of the market embedding theme of the portfolio planning section and the portfolio planning section has little of the organizational structure complications of the transfer pricing section.

A more important consequence of this disjuncture is that the meaning of "control" is not entirely clear. In the portfolio planning section, the image is of indirect "ecological" control, in which a role frame is established within which profit center managers monitor each other. In the transfer pricing section, the image is of more direct FDR-style interventions, in response to induced conflicts which percolate to the top. This ambiguity may point to the possible fact that profit centers serve two control functions simultaneously. Profit centers as strategic business units are managerial devices to position the production units within their markets. Profit centers as departments are managerial devices to position production units within the firm. These two role frames may not be consistent nor may their structures be homologous, in part because they do not operate at the same level of aggregation. The contradictions and tensions produced by this overlay raise fascinating future research questions to investigate.

On a more general level, my query is whether this line of research (and many of the other lines represented in this volume) has truly resolved the boundary or partition problem. The "interface" mechanism presupposes the existence of some closed set, the stability properties of which can be investigated. Or to put the point another way: Proto-groups, be they markets or firms, are axiomatically assumed, the actualization of which is the collective action question. This axiom presupposes that all of the actors in the "potential" role frame have somehow come to have the same set of others in their (perhaps cross-cutting) reference groups. This theoretical procedure is of course quite common, but I argue that it finesses the problem of deriving those cleavages which define social groups in the first place, and it forces the researcher to import in an ad hoc fashion auxiliary attributional facts to make operational the "potential group."

General Discussion

Michael Hechter: I want to ask you to clarify. I confess I don't understand the dependent variable in your theory. I don't understand what it is you are trying to explain. Is there a simple kind of *Y* variable that fits into an equation that this theory attempts to explain?

Robert Eccles: There are three different things we want to explain: the proliferation of divisionalized form, the existence of portfolio planning, and the transfer pricing practices that exist. In the classic way of thinking about it, firms really don't need to divisionalize, and yet many do just that. The existence of portfolio planning is the second dependent variable. That hasn't always existed. It was introduced in the early 1960s by a company called the Boston Consulting Group, and it has become as popular in large firms as this multidivisional form; and we would argue that first came the multidivisional form for control, which then created its own control problems, primarily on resource allocation. Portfolio planning came in to solve those problems. Transfer pricing practices are the third dependent variable.

Peter Blau: You seem to be contradicting what Harrison White started off with. He said you want to explain why the chief executive officer does it, and he said it is control. I would have thought that the amount of control he achieves through these three mechanisms you have outlined is a dependent variable.

Harrison White: That's what Eccles is saying. Portfolio planning is a technique that these chief executive officers have adopted, evidently to get control. But there is a further step. In addition we want to erase a false question. The false question is: When is it a firm and when is it a market? Firms and markets exist only in interpenetration with one another. It's a false conceptualization brought about by atomistic thinking of very good economists who have an imagery of firms and markets.

That may give you a flavor of the kinds of things we are trying to get at.

Unidentified Speaker: The firm/market matrix you set up doesn't seem to be stable. This morning we were told that the firms in that array are in constant flux and I think that one could argue that the markets are also in constant flux. I wonder if your solutions take that into account?

Harrison White: Yes. We claim that the identities of the firms and the markets are fairly stable. These are huge firms and they don't disappear all that quickly, and the markets don't either. What you are talking about is this constant change of who is in which market. Just to stay in the market you've got to stay alert, you've got to keep investing, and many tradeoffs have to be made. As a result, profit centers are quite often simply disappearing. Having moved out of that market will change the array, the portfolio. On the other hand, I don't want to mislead you. We are trying to capture a fairly definite kind of market and make predictions, and therefore it's certainly not extremely flexible and may not be flexible enough to do the whole job.

Gary Becker: The economics literature on the principal-agent problem also deals with control, and decentralization and profit centers are consistent with the implications of that literature. Yet, if limited information is crucial to principal-agent conflict, how can the paper assert that information acquisition is a secondary aspect of the motivation for decentralization? A second question is: How does your analysis of markets relate to an equilibrium analysis of the sorting of differentiated firms and consumers, as in the hedonic literature?

Harrison White: The essential aspect of agency is the need for control, and as I see it information is secondary when you look at control. I argue for this by going through a comparative historical array of where agency came from and where you use it. Control is a configuration of organization, a really large organization, in which it is embedded.

Gary Becker: If the chief executive officer has sufficient information about these units, wouldn't that solve the control problem?

Harrison White: You've got it wrong. If he's got control, he will automatically get the information. Good chief executive officers are lazy. They set it up so that people are forced to generate the information. The profit center managers are caught in a position where they are necessarily generating information. If you go at it from an information point of view, you're not taking an executive point of view but a staff point of view.

Gary Becker: Isn't that saying the same thing? You're trying to decentralize the profit units in order to elicit the correct information for the CEO. That's the principal/agent approach.

Harrison White: Let's put it this way. The principal/agent approach is ingenious and it's correct. But it's missing the point. It would never occur to them to look at the embedding in a complex organizational system. To the extent that you want to get at explaining configurations of very large organizations and how they interrelate it's a very long way around to go the principal/agent relation and try to build it up from there.

With regard to the second question you were asking, let me just say that every piece of technical apparatus we used has come right out of good, neoclassical economics, of Spence, of Lancaster, of Kirk. There's certainly no disagreement about that part, it's just what you choose to look at as the strategic thing.

John Freeman: It seems to me that these two brands of theory work in different circumstances. Where Harrison White's theory is most likely to work is in highly uncertain and highly diverse markets because precisely under those circumstances the volume of information that the agents would have to gather for the CEO would be most likely to overwhelm the system. Where you have very similar markets that are very stable, you build a bureaucracy with routinized accounting. So maybe the difference between what White is advocating and what Becker is suggesting has to do with the degree of uncertainty. Uncertainty is exactly where economic theory tends to have the most trouble.

Harrison White: As you well know, having had practice in business, there's no such thing as getting real effective control through routine, bureaucratic accounting mechanisms alone. That's one of the reasons why, when you go inside the firm, you will find it penetrated throughout its length with market aspects and interface aspects.

Paul Hirsch: I think your approach is remarkably radical for the discipline. For one thing, to ask what drives executives? If I go to the library and look up industrial sociology, nothing is said about executives. They present the view from the assembly line. What's exciting about your perspective is you're asking the question of how the markets and environments of firms vary, which is something the discipline really didn't address very much. We usually look at how work is organized irrespective of what market it's in.

Theory of Social Change

Explaining the Origins of Welfare States: A Comparison of Britain and the United States, 1880s–1920s

THEDA SKOCPOL AND
ANN SHOLA ORLOFF

Social scientists have been "worrying about social change," as Charles Tilly puts it (1983, p. 1), since the emergence of the various disciplines in the nineteenth century. The enduring concern to understand the roots and consequences of the industrial revolution has certainly not been misplaced. Yet sociologists, even more than other social scientists, have labored under serious misconceptions. All too often, according to Tilly (1983, p. 6), sociologists have assumed that " 'social change' is a coherent general phenomenon, explicable *en bloc.*" And they have postulated that the "main processes of large-scale social change take distinct societies through a succession of standard stages, each more advanced than the previous stage."

Note: We are grateful to James Coleman, Lewis Coser, David Knoke, Barbara Laslett, Robert Liebman, and Charles Tilly for comments that we took carefully into account in making revisions of our paper for publication. Another version of this argument, with additional evidence and discussion of alternative hypotheses, appears in Orloff and Skocpol (1984).

In fact, he counters, " 'social change' is not a general process, but a catchall name for very different processes varying greatly in their connection to each other" (Tilly 1983, p. 7). No irreversibly progressive logics can be attributed to such processes as they have actually occurred. During the last several hundred years of Western and world history, the concentration of capital in economic production and the consolidation of national states have been perhaps the most fundamental processes of large-scale, long-term social transformation. Yet even with respect to these two, "stage theories of social change assume an internal coherence and a standardization of experiences that disappear at the first observation of real social life" (Tilly 1983, p. 7). In order to better understand the varying impacts of capitalist accumulation and statemaking, sociologists must examine and compare concrete groups, places, and sequences of events. Through such comparative and historical investigations, it becomes evident that capitalist accumulation and statemaking have proceeded differently in distinct times and places and have intersected with one another in unstandardized—and sometimes surprisingly unprogressive—ways.

Tilly's caveats about the study of social change are especially applicable to one major kind of twentieth-century transformation, "the development of the modern welfare state," which has heretofore been conceptualized and explained in remarkably evolutionist and progressive terms. Since World War II, scholarship on public pensions and social insurance has been dominated by political sociologists whose sympathies have been predominantly liberal or social-democratic. Explicitly or implicitly, most have understood the modern welfare state as an inexorable and progressive phenomenon—perhaps launched and "completed" sooner by some nations than by others, yet appearing in recognizable stages in all national societies as a necessary and irreversible concomitant of industrialization, the expansion of the working class, and the institutionalization of mass political participation. From the perspective of the 1980s, with its resurgence of conservative thought within and beyond academia, such views are increasingly difficult to sustain in arguments about the present and future of public social expenditures. Yet we will demonstrate that they have also—all along—been equally misleading about the cross-nationally and historically varying *origins* of those public social-expenditure programs that, taken together, constitute the modern welfare state.

In developing our own point of view on the origins of modern welfare efforts, we shall also make much of another important insight offered in passing by Tilly. In order to carve out and defend their own precarious niche among the various professional social-science disciplines, he points out (1983, pp. 19–20), sociologists developed a vested interest in grounding dynamics of change in "society" conceived as "a thing apart" from "the state." Nowhere has this been more true than in the literature on the origins and development of modern public welfare programs, where industrialization and urbanization, perhaps modified by the struggles of social classes, have been the chief explanatory variables invoked.

As we are about to demonstrate, however, the rhythms and forms of public welfare provision are only loosely tied to industrialization or capitalist development. The roles of classes and social groups active in welfare politics need to be understood not simply in terms of their socioeconomically conditioned interests and capacities, but more directly in terms of the situations of such classes and groups within historically changing state structures with autonomously active public officials. If sociologists are to theorize effectively about the modern welfare state, they must not only set aside untenable assumptions about social change as a progressive and standardized phenomenon. They must also cease leaving states and their effects outside of the analytic purview of the discipline.

Explaining the Emergence of Modern Welfare States

Comparative research on the development of modern welfare systems has proliferated in the last twenty years, but very little of it has focused directly on the timing and forms of the origins of modern social insurance, public assistance, and pension programs. Instead, researchers have primarily sought to generalize about national variations in the post–World War II period, focusing in particular upon the growth of public social-welfare spending and the expansion of the population coverage of programs already in place. Broadly speaking, two families of arguments have set the terms of debate as such cross-national research has proceeded. These are "logic of industrialism" arguments and "working-class strength" arguments. Scholars in each camp have offered ideas about when, how, and in what sequence modern social benefits programs originated historically.

Logic of industrialism arguments about the development of modern welfare states emerged from cross-sectional, aggregate-quantitative studies of large numbers of nations (see for examples: Cutright 1965; Jackman 1975; Kerr et al. 1964; Wilensky 1975, chap. 2). Although specific arguments vary, such studies argue that thresholds or processes of national economic development, demographic dependency ratios, and the sheer longevity of programs are far stronger variables for explaining cross-national variations in social spending and program coverage than regime types or political or ideological variations within regime types. Despite their ahistorical research designs, extrapolations about causes of welfare state origins and the likely sequences of modern social programs have been put forward in logic of industrialism studies. The argument goes as follows: Either industrialization itself or else the decline of the proportion of the work force in agriculture creates new needs for public spending to protect old and young dependents, the unemployed, the disabled, and the sick. Urbanization and the aging of a population accelerate the economic trends. At the same time, industrialization creates the new wealth to make collective social provision possible, and new organizational means through which social benefits can be demanded and delivered. Breakthroughs toward the modern welfare state should therefore come as nations develop economically; as development proceeds, subsequent policies will build upon and "fill out" early beginnings, adding new programs and completing the population-coverage and raising the benefits of existing programs (cf. Cutright 1965).

A variety of "political economy" arguments about welfare states crystallized in the 1970s, partly in critical response to the logic of industrialism school. Scholars developing these arguments have also done cross-national quantitative studies of welfare spending (and sometimes also welfare financing) in the post–World War II era, but they have concentrated on explaining variations among twelve to eighteen rich capitalist democracies. At this intermediate range of comparison, levels of economic development and demographic variables no longer account for so much of the explained variation. Some studies have attributed relatively higher levels of welfare spending to combinations of national "openness" to the international economy plus the presence of strong institutions for "corporatist" bargaining among interest groups and the state (Cameron 1978; Wilensky 1976, 1981; Katzenstein 1982, 1983). Yet the largest cluster of political economy studies (e.g., Korpi 1978; Korpi & Shalev 1980; Castles 1978, 1982; Stephens 1979; Martin 1973) have put forward what we shall label "working-class strength" arguments about welfare state development.

In these studies, a definite historical dynamic is implied or asserted, using Swedish Social Democracy as the ideal type case of highest and most redistributive social spending. The welfare state is said to be shaped by class-based political struggles over the use of the state to benefit the working class economically and organizationally. A fully developed welfare state is a product of

> a highly centralized trade union movement with a class-wide membership base, operating in close coordination with a unified reformist-socialist party which, primarily on the basis of massive working class support, is able to achieve hegemonic status in the party system. To the extent that these criteria are met, it is hypothesized that the welfare state will emerge earlier, grow faster, and be structured in ways which systematically favor the interests of labor over those of capital. [Shalev 1983a, p. 11]

According to the logic of this argument, especially in its boldest form as put forward by John Stephens (1979, p. 89), "the welfare state is a product of the growing strength of labour in civil society," and the way to explain its historical origins as well as its subsequent growth is to examine variations in the emergence of centralized trade unions, and variations in the strategic ability of trade unions to create—or reorient—electoral political parties into tools for the acquisition and use of state power to effect welfare measures.

Whatever their relative merits for explaining welfare patterns in the post–World War II period, neither logic of industrialism nor working-class strength arguments have emerged unscathed in the handful of genuinely historical studies that have addressed the emergence of modern welfare states in two or more nations.

Collier and Messick (1975) address logic of industrialism hypotheses that welfare states have emerged as a byproduct of economic development or socioeconomic modernization. Working with data on fifty-nine nations that "had formal political autonomy with regard to domestic policy at the time of first adoption" (1975, p. 1299), these researchers found no support for level of economic modernization (measured by work force in agriculture, work force in industry, or real income per capita) as a necessary and sufficient condition for the initiation of social security programs, and only very slack support for the idea of economic modernization as a necessary threshold, since the "least modernized nations" in their sample initiated social security with less than 5 percent of the work force in industry, over 80 percent in agriculture, and less than $51 per capita income (in 1961 dollars). According to Collier and Messick, processes of international diffusion of social security programs from more to less developed nations account for overall world patterns in which later adopting nations—especially those adopting since 1922—initiate social security at much lower levels of economic development than earlier adopters. They attribute the ease of such diffusion to the "larger role of the state in later-developing countries" (1975, p. 1313). Regardless of social needs or demands, Collier and Messick suggest

(1975, p. 1310), political authorities in later-developing countries may realize that social security policies represent easy ways to tax citizens and useful tools for coopting labor movements.[1]

By disaggregating their data, Collier and Messick note a reverse pattern of international diffusion among the earliest adopters of social security: Before 1908, later-adopting countries tended to be more socioeconomically advanced than earlier adopters. Flora and Alber (1981) have gone on to explore these patterns among European nations using more nuanced measures of welfare legislation and "modernization" than those applied by Collier and Messick. Their study examines the timing of seventy-four "core" social insurance laws in twelve European nations. Overall, they confirm Collier and Messick's descriptive findings, showing (Flora & Alber 1981, pp. 61–63) that successive European nations "established their systems at a slightly higher level of socioeconomic development" (measured in terms of industrialization and urbanization) and "generally at a much higher level of political mobilization" (measured in terms of the electoral participation of the working class). Moreover, developmental thresholds of socioeconomic change or electoral mobilization are found to have "weak explanatory power" (1981, p. 65) in accounting for the timing of welfare breakthroughs. Only by exploring the combined effects of socioeconomic change and political mobilization are Flora and Alber (1981, p. 69) able to argue that seven out of their thirteen nations introduced social insurance at "similar" levels of overall "sociopolitical development."

Beyond this finding (which depends on averaging highly disparate, offsetting levels of socioeconomic development versus political mobilization), Flora and Alber find they must abandon any developmentalist scheme and group countries in terms of types of regimes and political systems. "In parliamentary democracies," they note (1981, p. 72), "the extension of the suffrage clearly increased the propensity to introduce insurance systems." But regardless of suffrage levels, such constitutional-dualistic monarchies as those of Germany and Austria "tended to introduce social insurance schemes earlier (in chronological and developmental time) than the parliamentary democracies. . . ." Several possible reasons for this important pattern are mentioned, including the "greater need to solidify the loyalty of the working class" in the face of "a growing and hostile labor movement that threatened the legitimacy of the non-parliamentary political regimes," and the fact that "the constitutional-

[1] For two case studies that bear out some of Collier and Messick's speculations, see Malloy (1979) on Brazil and Spalding (1980) on Mexico.

dualistic monarchies had already developed stronger state bureaucracies capable of administering [social insurance] systems and preserving a paternalistic heritage" (Flora & Alber 1981, p. 70.)

If empirical comparative tests of logic of industrialism ideas about the emergence of modern welfare programs have pointed toward the need for new hypotheses focusing on the pioneering initiatives of constitutional-dualist monarchies and on the socioeconomically "precocious" uses of state power in late-industrializing nations, complementary pointers have emerged from the less definitive tests to which working-class strength arguments have so far been put. Existing research suggests that the issues about the role of the working class are far too complex to be encompassed in any model that posits a straightforward causal sequence from working-class demands and capacities, however measured, to the emergence of modern welfare programs.

Pryor (1968, p. 173) found that the existence of social insurance programs in 1913 across seventeen industrialized nations could be predicted solely by measuring the extent of trade unionization, but this simple relationship does not hold up for 1905 or earlier (cf. Stephens 1979, p. 115, with the patterns of welfare development reported in Flora & Alber 1981). Moreover, Pryor remarked (1968, p. 475) that his finding did not necessarily mean that social insurance programs were brought about by unionization, for both could be the result of the mobilization of working-class political power. Stephens (1979), the foremost advocate of the working-class strength approach to welfare state development, focuses his attention on whether an early alliance emerges between centralized industrial unions and a social-democratic, labor-based political party. Yet his explanatory model is too closely tied to Swedish developments from the 1930s onward to do justice to variations in welfare state development from the 1880s through the 1920s. In the chapter of his book (1979, chap. 5) where he presents comparative historical accounts, Stephens does not discuss Germany or other nations that were nonparliamentary before World War I. More telling, in his account of British history (1979, pp. 140–49), Stephens virtually ignores the crucial role of the Liberal party in launching the welfare state between 1908 and 1911, and he fails to acknowledge (let alone explain) why Britain was well ahead of Sweden in public social spending during the 1920s, even though British unions were less centralized and inclusive of the labor force than Swedish unions, and despite the fact that the British Labour party attracted less working-class support than the Swedish Social Democratic party.

Careful historical studies by Heclo (1974) and Shalev (1983b) explicitly address the role of unions and labor-based parties in welfare state development. Heclo compares in detail the concrete policy processes by which public provision for the elderly and the unemployed developed in Britain and Sweden from the nineteenth century to the present. He finds no consistent role for unions or left parties—or for any classes, socioeconomic groups, or political parties—in bringing about or elaborating social policies. While the demands, needs, or expectations of social groups and political parties may (along with industrialization itself) help to dramatize the need for new policies—in ways that Heclo does not really try to explain—actual policy innovations come, he claims, from reform-oriented "middlemen" operating at the interface of states, parties, and social groups, as well as from civil servants who are familiar with administrative problems and possibilities. Moreover, Heclo stresses that policymakers communicate ideas, experiences, and (positive or negative) models across national boundaries, and that they above all orient themselves to the legacies of prior state actions:

> [This] is not to suggest that all development is the replay of past policy; but what is normally considered the dependent variable (policy output) is also an independent variable (in an ongoing process in which everything becomes an intervening variable). . . . Policy invariably builds on policy, either in moving forward what has been inherited, or amending it, or repudiating it. [Heclo 1974, p. 315][2]

In another highly relevant study, this one more macroscopic and structural than Heclo's, Michael Shalev (1983b) juxtaposes the development of the Israeli welfare state and the expectations of the working-class strength model. If ever there was a case that should fit the model perfectly, Israel is it, for the Israeli state itself, along with all of its social policies, was fashioned by a trade union confederation, the Histadrut, in collaboration with a social-democratic party, the Israeli Labor party. Yet Shalev shows that social protections in Israel have always had patterns of coverage, financing, and delivery at odds with the universalistic and generous provisions for the entire working class predicted by working-class strength arguments that take Swedish Social Democracy as the benchmark case. The Israeli labor movement's distinctive historical ac-

[2] This point holds for the attitudes of class-based groups toward possible social policy departures, as well as for the attitudes of reformers, state officials, and political leaders. It suggests why it might well be analytically fruitless to posit transhistorically and cross-nationally fixed class (or other economic) interests with respect to social welfare. Classes and economic interest groups, like others, "take their interests where they find them" (Heclo 1974, p. 299) in given historical and political contexts.

tivities as a manager of Jewish immigration to Palestine, as a channel for overseas financial contributions, and as a founder and builder of the Israeli national state must be understood, he argues, to explain the emergence and patterns of Israeli social-welfare programs.

Of course the Israeli case could simply by treated as "deviant" with respect to an otherwise adequate working-class strength model. But Shalev wants, instead, to use "the failure of the . . . model in the Israeli context . . . as a challenge to . . . develop an alternative theoretical perspective sufficiently broad that it can incorporate the Israeli case" (1983b, p. 12). He continues in a vein that resonates well with the other historical findings we have just surveyed:

> The essential weakness of the . . . [working-class strength] model . . . is the passive role it assigns to the state itself, whose actions are assumed to be reducible to the interests of constituent groups in civil society with the greatest political leverage. . . . Several factors are consequently neglected. (1) The interest of the parties which control states in using the resources of the state to perpetuate and enhance their political capacities—as opposed to the purer motive of responding to the interests of constituents, which parties in power can do in varying degrees and in varying ways. (2) The interest of the politicians and bureaucrats who man the state in managing the citizens and resources within their orbit to defend the state as an organization and control mechanism. (3) The very different sets of problems, and resources for confronting them, . . . characteristic of states . . . because of their varying points and forms of articulation with the world system of states.

In sum, cross-nationally and theoretically oriented historical studies addressing working-class strength arguments complement the more general, quantitative findings of Collier and Messick (1975) and Flora and Alber (1981). To understand the timing and forms of emerging modern welfare states, these various studies suggest, we must shift attention from socioeconomic development as such, and from the demands and organizational capacities of the industrial working class alone, toward the overall governmental, political, and historical contexts within which working classes and their organizations interact with other social and political forces.

A Comparison of Britain and the United States

If the literature on the emergence of modern welfare states points toward new emphases on states and political contexts, this nascent analytic agenda is still very ill defined. And the possibility remains that it should include only special sorts of nations—monarchies, late industrializers, and embattled Israel—where state authorities have necessarily had "abnormally" large roles to play in social-welfare politics. Perhaps liberal capitalist democracies are still best encompassed by adjustments of the existing hypotheses about industrialization or working-class strength as the underlying determinants of the emergence of welfare states.

A comparison of social politics in Britain and the United States between the 1880s and the 1920s offers an excellent opportunity to test existing hypotheses and develop new ones for a pair of liberal-democratic cases that ought to fit society-centered theories, if any do. A straightforward case can be made for the validity of comparing Britain and the United States in this time period. During the decades following Bismarckian Germany's pioneering institution of social insurance programs in the 1880s, reformers and politically active leaders in all of the major industrial-capitalist nations of the West investigated and debated how similar or alternative modern welfare programs might be devised to meet the needs of their own countries. In both the United States and Britain, innovators of similar class and occupational backgrounds participated in this transnational reform ferment (Heclo 1974, pp. 310–11; Skocpol & Ikenberry 1983; Lubove 1968; Mann 1956, p. 675; Hennock 1981). In due course, similar legislative proposals for workers' compensation, old-age pensions, and health and unemployment insurance emerged in both countries. The early twentieth-century United States, in short, was part of the same community of policy discourse as Britain.

But similar policy proposals met with sharply different fates in these two nations. As Table 1 summarizes, Britain launched a full range of modern social programs before World War I. Industrial accident insurance was promulgated in 1897 and 1906; old-age pensions were instituted in 1908; and health and unemployment insurance were started in the National Insurance Act of 1911. Subsequently, from the conclusion of World War I through 1925, the vast majority of the British industrial working class was brought into the world's first unemployment insurance system and a broad-gauged old-age and survivors' insurance system was instituted. Britain actually became during the 1920s (according to one recent measure in Flora & Alber 1981, p. 55) the world's leader in overall "welfare state development."

Meanwhile, in the United States during the Progressive Era of roughly 1906 to 1920, proposals for health insurance, unemployment insurance, and old-age pensions failed to pass. Indeed, modern welfare legislation progressed in only two areas: by 1919, laws requiring employers to have industrial accident insurance passed in thirty-eight states and, between

Table 1 Major Legislation on Workers' Compensation, Old-Age Pensions, and Social Insurance in Britain and the United States Before 1929

Britain	*United States*
1897 Workmen's Compensation Act	
1906 Workmen's Compensation Act (extended to cover industrial diseases and more occupations)	
1908 Old-Age Pensions Act	
1911 National Insurance Act (health and unemployment insurance)	1911–19 Mothers' Pension Laws in 39 states (7 more states by 1931) 1911–19 Workers' Compensation Laws in 38 states (2 more states by 1929)
1920 Unemployment Insurance Act (extended coverage to all manual workers)	
1921 Unemployment Insurance Act (non-contributory benefits)	
1921 Unemployed Workers' Dependents Act (added benefits for spouses and children)	1922–28 Old-Age Pension Laws in 6 states (all county-optional)
1925 Widows', Orphans' and Old-Age Contributory Pensions Act	
1927 Unemployment Insurance Act (established rights to open-ended benefits)	
Source: Ogus 1982, pp. 151–54.	*Sources:* Brandeis 1935, pp. 575–77; Leff 1973, p. 401.

1911 and 1919, thirty-nine states also passed mothers' pension laws primarily designed to help respectable working-class widows to provide for their children (Brandeis 1935, pp. 575–77; Leff 1973, p. 401). By a few years into the 1920s, moreover, it was apparent even to discouraged progressive reformers that the United States was rejecting public social benefit programs in favor of the ideal of "welfare capitalism"—the notion that the needs of the American people could be met by economic growth and private corporate benefits, supplemented by local or voluntary group efforts in emergencies (Brandes 1976; Brody 1980). Political gains for modern social-welfare legislation during the 1920s were restricted to the establishment in only six states of laws enabling—but not requiring—localities to create and fund minimal pensions for the very needy elderly (Lubove 1968, p. 136).

We seek to pinpoint why Great Britain was among the world's welfare pioneers, while the early twentieth-century United States failed to adopt similar kinds of modern welfare legislation. A comparison of Britain and the United States allows us to control for the effects of liberal values and ideologies. The "strength of liberalism" in America has often been used by researchers as an excuse for not applying cross-national explanatory models to the "exceptional" U.S. case before the 1930s (cf. Birch 1955, p. 27; Collier & Messick 1975, p. 1313; Grønbjerg, Street, & Suttles 1978, pp. 4–5; Kaim-Caudle 1973, pp. 184–85; Rimlinger 1971, pp. 62–71). But Britain, like the United States, was permeated by laissez-faire liberal ideals in the nineteenth century. In both nations, these ideals survived among popular and elite voluntary groups into the early twentieth century, while among some educated elites "new liberal" values justifying the pursuit of collectivist welfare goals made parallel headway in the United States and Britain alike (Orloff & Skocpol 1984, pp. 733–35). Indeed, it was in Britain, not the United States, that the most spectacular instance of voluntary group resistance to the coming of the modern welfare state occurred. The British friendly societies, predominantly working-class benefit groups that enrolled well over half of the male population at the turn of the century, delayed passage of public old-age pensions for over two decades (Treble 1970; Gilbert 1966, pp. 165–221). Thus, there is no reason to treat the United States as unusually liberal in comparison to Britain. And there is every reason to bring these two liberal nations in to the same comparative analytic framework to account for their diverging patterns of actual welfare legislation despite similar ideological and intellectual environments in the nineteenth and early twentieth centuries.

By comparing British breakthroughs to failed proposals for comparable reforms in the United States, this research aims to make a strategic contribution to the previously surveyed explanatory debates about the emergence of modern welfare states. We will use this study of two liberal-democratic and capitalist-industrial nations to make plausible the thesis that explanations for the emergence of modern welfare states in *all* nations—not just monarchies, late-industrializers, and other "deviant" cases—can be revised to place less emphasis on social needs and demands and more emphasis on the conditions of what might be called "political supply" by state and party managers.

State Structures and Social Politics in Two Industrializing Liberal Democracies

Most political sociologists, whether pluralists, Marxists, structure-functionalists, or adherents of other perspectives, consider government activities to be expressions of—or responses to—social demands. Organized groups, including political parties, are conceptualized as vehicles for the expression of such demands, which are seen as culturally or socioeconomically rooted. After groups or parties weigh in at the political arena, some perhaps more effectively than others, governments generate policy outputs to meet the social demands. The overall process therefore looks something like this:

Socially Rooted Demands ⟶ What Groups or Parties Propose ⟶ What Governments Do

Here we are going to work with a more historical, structural, and state-centered view of how the political process works, a view drawing on ideas from Shefter (1977), Heclo (1974), and Skocpol (1985). From Heclo (1974) we take especially the insight that policy-making is an inherently historical process in which all actors consciously build upon or react against previous governmental efforts dealing with the same sorts of problems. This means that the goals of politically active groups and individuals can never simply be "read off" their current structural positions. Instead, the investigator must take into account meaningful reactions to previous policies. Such reactions color the very interests and goals that social groups or politicians define for themselves in politics.

From Skocpol (1985), we take an emphasis on the dual importance of states—defined as administrative, coercive, judicial, and representative institutions—in the political process. States may be sites of autonomous official action, not reducible to any social group pressures. Both appointed and elected officials have organizational and career interests of their own, and they devise and work for policies that will further those interests, or at least not compromise them. Equally important, the organizational structures of states indirectly influence the politics of all groups in society. Definitions of what is feasible in politics depend in part on the capacities various groups attribute to existing or emerging state structures. And the policy outcomes for which groups strive may depend on perceptions of how well various policies would be administered. In addition, historically changing state structures also affect the very political organizations through which policies can be collectively formulated and supported. On this latter point in particular we draw heavily on a seminal article about political parties by Martin Shefter (1977).

Shefter investigates why it is that some parties—and, indeed, often entire systems of competing political parties—operate by offering followers "patronage" jobs and other kinds of divisible payoffs out of public resources, while other parties and party systems offer ideological appeals and collectively oriented programs to groups, classes, or "the nation" as a whole. The usual answers have to do with the inherent cultural proclivities or socioeconomic preferences of given nationalities or social classes: e.g., Irish people and peasants want patronage, so that is what their governments and parties offer. But Shefter provides evidence against such answers and instead highlights the effects on parties' modes of operation of historical sequences of state bureaucratization and electoral democratization.

In some European absolute monarchies, state bureaucratization preceded the emergence of electoral democracy (and even the emergence of parliamentary political parties in some instances). When electoral parties finally emerged in such countries, they could not get access to the "spoils of office" and therefore had to make programmatic appeals, based on ideological world views and (if their prospects of forming governments with some authority were good) promises about how state power might be used for policies advocated by organized groups in their targeted constituencies. On the other hand, in countries where electoral politics preceded bureaucratization, parties could use government jobs as patronage. Later, there might be struggles over how to overcome "political corruption" in order to create a civil service free from patronage; and if reform succeeded, political parties might then change their operating styles toward more programmatic appeals. However, if patronage was established (or survived into) a fully democratized polity, it was, according to Shefter, extraordinarily difficult to uproot thereafter. For mass electorates, and especially the electoral politicians devising appeals to them, would have a stake in using government as a source of patronage.

The kind of thinking that underlies Shefter's argument is worth capturing in a diagram that can stand in opposition to the more typical mode of reasoning in political sociology portrayed above:

State Formation →	How Parties →	Social Demands
(sequence of bureaucratization and democratization)	and Administrations Operate	Selected for Political Response

Our own model of politics synthesizes this mode of reasoning with the traditional line of reasoning in political sociology outlined above. The combined model can be diagrammed as follows:

State Formation →	How Parties and Administrations Operate →	Policy Possibilities Selected for Political
	↓	
Socially Rooted Demands →	What Groups or Parties Propose →	Response

Applying this model and building upon Shefter's arguments about "patronage and its opponents" can help us to explain why Britain and the United States in the early twentieth century were so different in their propensity to adopt new programs of public social spending. When the international ferment over modern social insurance and pensions spread to Britain and the United States around the turn of the century, it caught these nations and their social and political elites at very different conjunctures of political transformation. Essentially, these were two liberal polities that had moved from patronage-dominated politics toward public bureaucratization at different phases of industrialization and democratization. Britain already had a civil service, programmatically competing political parties, and legacies of centralized welfare administration to react against and build upon. Modern social-spending programs complemented the organizational dynamics of government and parties in early twentieth-century Britain. The United States, however, lacked an established civil bureaucracy and was embroiled in the efforts of Progressive reformers to create regulatory agencies and policies free from the "political corruption" of nineteenth-century patronage democracy. At this juncture in American history, modern social-spending programs were neither governmentally feasible nor politically acceptable.

From Oligarchic Patronage to a Modern Welfare State in Britain

Britain's polity in the nineteenth century started out as a liberal oligarchy, ruled by and for landlords (cf. Namier 1961; Thomson 1950, p. 21; Webb 1970, pp. 53–57). During the course of the century, this polity underwent several intertwined transformations, which lay the basis for the Liberal welfare breakthroughs of 1908–11: the expansion of national administrative activities, especially in the realm of social-welfare policy; the reform of the civil service; the step-by-step democratization of the parliamentary electorate; and transformations in the modes of organization and electoral operation of the major political parties.

For quite some time, British historians have recognized the "Victorian origins of the welfare state" in an administrative sense (Roberts 1960). Laissez-faire may have been the charter myth of nineteenth-century British government, but despite the reality of free-market and free-trade economic policies, in the realm of domestic social-welfare policy the implementation of the New Poor Law called for administrative supervision and social planning on a national scale. The 1834 New Poor Law was radical not only in its substantive precepts embodying the ideals of market capitalism, but also because it established a central authority, the Poor Law Commission, to supervise local poor relief institutions, and substituted elected boards of guardians governing groups of parishes for the local magistrates who had formerly monopolized supervision of the poor (Roberts 1960, pp. 43–45; Heclo 1974, pp. 55–58). The very structure of welfare administration in the unitary British polity meant that tensions generated by the workings of the local workhouses, asylums, and sick wards, and by the implementation of other policies to deal with the poor, inevitably and recurrently generated pressures for *national* debates, investigations, and policy changes. This was all the more true once welfare administration was brought under direct parliamentary and ministerial controls in 1847, especially as the gradual democratization of the electorate through the reforms of 1832, 1867, and 1884 brought the middle class and the upper ranks of the working class more fully into routine local and parliamentary politics.

Meanwhile, important changes also occurred in the workings of the British civil service as a whole. In the eighteenth and early nineteenth centuries, oligarchic patronage predominated and "the public services were the outdoor relief department of the aristocracy" (Smellie 1950, p. 69). But industrialization and urbanization, along with the geopolitical exigencies of maintaining British imperial domains and coping with growing international economic competition, generated pressures for the British government to become more efficient and technically competent than patronage would allow (Smellie 1950, pp. 69–70). Reform advocates were initially frustrated by those with a vested interest in the existing system (Finer 1937, pp. 45–49; Cohen 1941, chap. 7). Yet, finally, proposals for civil service reform succeeded politically in the 1870s. Prior changes in universities made them plausible as agencies for training and credentialing civil servants (Cohen 1941, pp. 81–83). And once it became clear that working-class political influence might grow as the electorate expanded, the landed and business groups and the existing governing elites of Britain came together in order to maintain the elite civil service on a new basis (Shefter 1977, pp. 434–37; Greaves 1947, pp. 21–32).

In turn, civil service reform, along with step-by-step electoral democratization, had important implications for the organization and operations of the political parties. With the credentialization of the civil service, the parties had to stop relying on elite patronage and develop new methods of raising funds and rewarding activists and new ways of winning votes in an expanding electorate. In the 1870s and 1880s, both the Liberal and Conservative parties created constituency organizations and at the same time began to formulate programs to appeal through activists to blocs of voters and financial subscribers (Hanham 1959; McGill 1962; Douglas 1971, pp. 1–17; Shefter 1977, pp. 438–41).

We are now in a position to see how the historical legacies of the New Poor Law, along with the overall structures of ministerially directed parliamentary democracy in place in Britain by the early twentieth century, facilitated the Liberal welfare breakthroughs of 1908 to 1911. From the 1890s onward, there was widespread elite and popular disgruntlement with the way members of the respectable working class who became impoverished due to old age, ill health, or unemployment were handled by poor law institutions. National politicians, Conservative and Liberal alike, became interested in reforming or replacing the New Poor Law to deal better with the problems of the "worthy poor" (Harris 1972; Collins 1965). Some of their concerns were generated from administrative dilemmas within established programs as well as by the threat to the whole edifice of local government finance posed by the rising costs of the poor law. Other concerns arose from the obvious political fact that the votes of working-class people—and the support of their organizations, the unions and friendly societies—had to be contested by parties engaged in increasingly programmatic competition.

The Liberal welfare reforms crystallized by two routes in the context we have outlined. In the face of the national campaign waged by the National Committee of Organized Labour on Old-Age Pensions, the Liberals devised their noncontributory and need-based old-age pensions as a tool of programmatic competition with the Conservatives and as a way to cement their party's alliance with the Labour Representation Committee. Then proposals for unemployment and health insurance came through initiatives from Liberal Cabinet leaders allied with civil administrators at the Board of Trade and the Treasury. For unemployment and health insurance alike, intragovernmental elites took the initiative in persuading both working-class and business interests to go along. Once this persuading was done and the Cabinet was set on its course, moreover, the discipline of the Liberal party in Parliament ensured passage of the National Insurance Act, and there were no independent courts to which disgruntled parties could appeal.

Stepping back to put these policy departures in broader context, we must underline that the administration of social spending as such was not fundamentally problematic for British elites in this period. The "corruption" of patronage politics was behind them, and disputes were now focused on levels and forms of spending. All established party leaders were concerned about the socioeconomic efficiency of Britain in a competitive world and—even more—about how to attract or retain trade union and working-class electoral support in a democratizing polity with programmatically competing parties. In this context and conjuncture, pensions and social insurance looked like good ways to circumvent for the respectable working class the cruelties, inefficiencies, and costs of the New Poor Law of the nineteenth century. Such policies also looked like appropriate programs to appeal to—and, in the case of the social insurance measures, newly *tax*—the working class, involving them more fully in the life of a united nation, yet under the hegemony of enlightened professional middle-class leadership.

The Struggle Against Patronage Democracy and the Limits of Social Spending in the United States

America's polity in the nineteenth century has been aptly described (Skowronek 1982, chap. 2) as a polity of "courts and parties" operating in a multiply tiered federal framework. The courts adjudicated rights of contract and private property. Meanwhile, highly competitive political parties provided a modicum of integration across the various levels and branches of government. Crucially, U.S. electoral politics was fully democratized for white males in the Jacksonian Era. Thus, political parties were able to rotate the "spoils of office" to reward their cadres as the parties swept into and out of office in the constant rounds of close-fought elections characteristic of nineteenth-century American democracy (Shefter 1978; Keller 1977, chaps. 7, 8, 14).

The entire system worked best at all levels when governmental outputs took the form not of programs devised to appeal to organized collectivities, but of politically discretionary ***distributional*** policies, such as financial subsidies or grants of land, tariff advantages, special regulations or regulatory exceptions, construction contracts and public works jobs (McCormick 1979). For through such policies it was possible for the patronage-oriented parties to extend their preferred style of recruiting and rewarding cadres into effective strategies for knitting together elements of popular and business constituencies that were ethnically, geographically, and socioeconomically diverse.[3] Ideal sets of distributional policies combined measures that raised revenues—or created jobs—with those that allocated them.

The post–Civil War pension system was an excellent example of a policy generated by the distributional proclivities of nineteenth-century patronage democracy. It allowed the Republicans, especially, to confer on individuals in many localities pensions financed out of the "surplus" revenues from the constantly readjusted tariffs they sponsored to benefit various industries and sections of industries (McMurry 1922, p. 27). Between the mid-1870s and the 1890s, the original laws mandating disability pensions for northern Civil War veterans were repeatedly liberalized, and mounting numbers of applicants signed up. After 1890, any veteran with 90 days service or more could receive a federal disability or old-age pension if he became unable to perform manual labor or reached the age of 62. From the 1880s to the 1910s, between one fifth and one third of federal expenditures went for "Civil War pensions," and congressmen, especially Republicans, continually vied to hand out more. As many as one out of two elderly native-born white men in the North ended up as beneficiaries of this largesse generated by America's patronage democracy.

In significant contrast to Britain, there was no national poor law in the United States, either in theory or in administrative fact. Instead, mixtures of Elizabethan and New Poor Law practices were institutionalized in diverse forms in different states and, above all, in thousands of individual local communities, where the prime responsibility lay both financially (as in Britain) and legally (Tishler 1971, pp. 4–6, 100). Reactions against poor law practices would not so readily converge into a series of national debates, as they did in Britain.

[3] Indeed, these diversities—especially different ethnic identities in the American working class—were actually reinforced, and in a sense produced and reproduced, by the workings of patronage democracy, with its efforts to "balance tickets" among various ethnic representatives (see Katznelson 1981, pp. 58–72; Shefter forthcoming).

The more consequential reactions to federal patronage democracy were, however, directed against what increasingly came to be seen as its thoroughgoing "inefficiency" and "corruption." As the United States became a truly national economy and society in the decades after the Civil War, problems faced by public policy-makers challenged the distributional style of patronage democracy, and vociferous demands emerged for civil service reform. The initial proponents were "Mugwumps," mostly upper and upper-middle class reformers located in the Northeast, especially Massachusetts. Very much like the successful British civil service reformers of the 1870s, the Mugwumps wanted public administration to be taken out of patronage politics, so that expertise and predictability could prevail. At first, however, the Mugwumps' programs of civil service reform and ideas about new administrative functions—functions that would take over tasks of regulation from the courts as well as tasks of conflict resolution from the parties—made headway only in minimal "patchwork" fashion (Skowronek 1982, pt. 2). Party politicians enjoyed secure roots in the mass white-male electorate and were clever at coopting reform ideas for their own purposes, taking only a few marginal official positions at a time out of the patronage system (Skowronek 1982, pp. 68–84). In sharp contrast to the situation in Britain, there was no impending threat of further electoral democratization to prod social and political elites into civil service reform.

Not until the Progressive Era of the early twentieth century did administrative reform—along with new kinds of public policies—really make significant headway in the United States, and then more at municipal and state levels than at the national level (Schiesl 1977). Social demands for new kinds of collective policies "in the public interest" broadened out from the very elite ranks of Mugwumpry to include the growing ranks of the educated, professionalizing middle class and (in many places) farmers and organized workers as well (Wiebe 1967, chap. 5; Buenker 1978, chap. 6). New governmental functions of business regulation, consumer protection, provision of efficient urban services, and improvement of conditions for industrial workers were all very much on the agenda of public debate in the Progressive Era. Yet there were always the nagging questions: Would party-based politicians take up new challenges that obviously did not fit well with their preferred patronage-oriented, localistic, and distributional styles of operation? Indeed, throughout the Progressive Era, the common denominator of all reform remained the struggle against patronage and political corruption, and other more substantive programs for economic and social-welfare purposes were likely to gain broad support only if they also appeared to further structural transformations away from the ills of patronage democracy (McCormick 1981).

The legacies of nineteenth-century patronage democracy and the conjuncture of its crisis in the Progressive Era created a much less favorable context for advocates of old-age pensions and social insurance in the early twentieth-century United States than the one enjoyed by their counterparts in Britain. Most basically, there was the sheer weakness of public administration, due to the original absence of state bureaucracy in America, the limited achievements of civil service reform in the nineteenth century, and the dispersion of authority in U.S. federalism. In contrast to the situation in Britain, there were in the early twentieth-century United States no influential high-level public officials strategically positioned to formulate new social benefits policies with existing administrative resources, press them on political executives, and work out firm compromises with organized interest groups. Typically, reforms in the Progressive Era were not autonomous initiatives from either civil servants or politicians. They were usually pressed upon state legislatures by broad coalitions of reform and interest groups (Buenker 1978).

This made sense not only because of the weakness of public administration, but also because this was a time when party organizations as such were weakened, even though the Republicans and Democrats remained jointly dominant. Moreover, U.S. political parties operated differently from British parties. They were democratic, patronage-oriented parties that found their established ties to electoral constituents and business interests under attack by reformers, not programmatic parties looking for new policies to attract organized labor. The challenge for elected U.S. politicians was to find ad hoc ways to propitiate reform-minded pressure groups, while still retaining the loyalty of (often ethnically organized) working-class supporters.

As Democratic and Republican politicians alike often discovered, the socioeconomic reforms that best fit this formula were primarily *regulatory* laws rather than public social spending measures. A few articulate reformers in the Progressive Era campaigned for health insurance, unemployment insurance, and even (though least often) for old-age pensions. And some politically active groups, including unions, supported these ideas. But, in general, it was very hard to get support from broad coalitions of groups—especially given the reluctance of upper- and middle-class people to champion social spending measures in the United States. In a period when the struggle against "political corruption" was still very much on the agenda, these people doubted that social spending measures could be implemented honestly and feared that they might well reinforce the hold over the electorate of patronage politicians.

Such fears were strongest and most openly expressed on the issue of old-age pensions, which would have been noncontributory governmental handouts very much like the post–Civil War pensions. Back in 1889, leading Mugwump and President of Harvard Charles Eliot had denounced the Civil War pensions as "a crime . . . against Republican institutions" because they "foisted . . . perjured pauper[s] . . . upon the public treasury" (quoted in McMurry 1922, pp. 34–35). Echoes of this revulsion reverberated into the Progressive Era and made even leading reformers wary of pensions. For example, Charles Henderson, who wanted to believe that the logic of the Civil War pensions pointed toward more universal modern social protections, was forced to acknowledge that the "extravagance and abuses of the military pension system have probably awakened prejudice against workingmen's pensions" (1909, p. 227).

The clearest acknowledgment of the difficulties that the Civil War pensions and the legacies of American patronage democracy presented to social insurance advocates came from Henry Rogers Seager, a professor at Columbia University and prominent member of the American Association for Labor Legislation. In lectures and a book that appeared in 1910, Seager revealed why it was that Progressive reformers, even advocates of social insurance like himself, were unwilling to mobilize union leaders as British reformers had done to press for what would probably have been the most popular social spending measure of all, noncontributory old-age pensions:

> Our experience with national military pensions has not predisposed us to favor national pensions of any kind. Giving full weight to the fact that the number of aged persons is strictly limited, there is still the danger that, if once embarked on the policy of granting annuities of the public treasury to private citizens, pressure would be brought to bear on Congress to lower the age limit and increase the annuity, and . . . this might lead to unwise extensions of the policy in both directions. [Seager 1910, p. 145].

This brings us to our explanation for why two modern social-welfare measures—workers' compensation and mothers' pensions—actually *did* pass, each in almost forty states, during the Progressive Era. The ultimate pressure that pushed through both of these new social-welfare programs came from the wide publicity magazines and investigatory commissions gave the issues and from the lobbying of state legislatures carried on by broad coalitions of reform and interest groups (Leff 1973, pp. 400–413; Berkowitz & McQuaid 1980, pp. 35–36; Tishler 1971, pp. 126–27). But the pressure could never have built up in the first place had not elites and the political public generally been broadly receptive to these ideas—receptive through forms, such as official commissions, that were simultaneously used to denounce or delay the other major modern welfare reforms.[4] Why was this?

One feature of both types of laws that recommended them to reformers, elites, and Progressive Era political publics was that they involved little or no public spending, certainly not of the order that would be required for old-age pensions or partially state-funded, British-style social insurance. Workers' compensation laws merely required businesses to insure their employees against injuries; in all but a few states this could be done through private carriers or even through self-insurance (Brandeis 1935, pp. 581–87). Mothers' pension laws largely restricted themselves to helping morally "worthy" widows; only 25 percent of the states provided any of the financing; and all the laws were "local-option," leaving it to individual countries or towns to decide whether to help mothers, exactly which ones, and at what levels of aid (Lubove, 1968, p. 99).

More important, neither workers' compensation nor mothers' pensions constituted a wholly new departure for public action in America; certainly they did not mandate (as pensions and social insurance would have done) new fiscal functions for barely established civil administrators or for potentially "corrupt" party politicians. Rather, these laws reworked activities already being handled in the American polity by the courts. Workers' compensation laws removed disputes over compensation for injured workers from the common law and the courts and typically placed settlements under the supervision of regulatory agencies that were set up to monitor payments by businesses to injured workers or the dependents of deceased workers (Berkowitz & McQuaid 1980, pp. 37–40). No brand new public jurisdiction was established; the venue was simply changed.

Mothers' pensions also involved the courts—this time, interestingly enough, both as proponents of change and as loci for trustworthy administration of social spending. As things were, judges—including especially those of the new juvenile courts established in the Progressive Era—had

[4]For example, in Massachusetts, where the Commission on Old-Age Pensions (1907–10), chaired by Magnus Alexander, reported strongly against state action (Linford 1949: chap. 1), a 1910–11 Commission on Compensation for Industrial Accidents (on which Alexander also served) proposed and drafted the state's Workmen's Compensation Act of 1911 (Asher 1969, pp. 466–69). And a 1912–13 Commission on the Support of Dependent Minor Children of Widowed Mothers recommended and drafted the state's widows' pension legislation, enacted in 1913 (see *Report of the Commission on the Support of Dependent Minor Children of Widowed Mothers* 1913).

to decide about the removal of the children of poor mothers when they could not adequately provide for their offspring. Especially where respectable widows were concerned, removal from the home was coming to be a heinous decision, for the idea was growing that children needed maternal care. Juvenile court judges in Illinois and Missouri helped to initiate the nationwide movement for mothers' pensions (Leff 1973, pp. 400, 405; Lubove 1968, pp. 99–100). Some reformers worried, as did Charles Henderson, that even the small sums involved in mothers' pensions might become "another kind of spoils for low politicians" (quoted in Tishler 1971, p. 153; see also Leff 1973, p. 404). But most of the new laws obviated this worry by putting juvenile court judges in charge of deciding who should receive pensions and supervising the performance of the recipients thereafter (Leff 1973, p. 401; Lubove 1968, p. 99). Judges, not politicians or bureaucrats, took charge of administering most of these tiny new flows of social spending instituted in the Progressive Era.

After the victorious multiple-state campaigns for workers' compensation and mothers' pensions during the U.S. Progressive Era, very little additional legislative movement toward a welfare state would occur again, until the coming of the Great Depression and the New Deal transformed the political and administrative universe in the states and the nation (Skocpol & Ikenberry 1983). Only then would the stage at last be set for the "belated" coming of a kind of modern welfare state to America—or, if one looks at it in a less evolutionist way, for a return to a federal commitment to social spending on a scale commensurate with the post–Civil War pensions originally extended by America's nineteenth-century patronage democracy.

Conclusion: From Social Demand to Political Supply

Between the 1880s and the 1920s, Britain and the United States were two liberal democracies whose political publics participated in the transnational debates about social welfare that spanned Europe, the Americas, and Australasia. Unprecedented problems and capacities thrown up by industrialization encouraged people in both nations to consider new kinds of public welfare policies, and similar legislative proposals for pensions and social insurance emerged in Britain and the United States alike. Yet the outcomes were very different: Britain launched her modern welfare state before World War I—becoming, indeed, the world's first full-scope welfare state through her pioneering initiatives in unemployment insurance. But the United States rejected the possibility of instituting (either nationally or in the states) modern welfare programs that would entail large-scale public expenditures.

Political sociologists have long tended to assume that governmental actions flow out of socioeconomic development and that they respond to economically or culturally rooted social demands—especially, in a democracy, to the expressed preferences of popular majorities. According to this pervasive perspective, if a welfare state does not arrive or "develop," it must be because there is no socioeconomic need or because, despite the need, "the people" or "the working class" are unwilling or unable to demand the appropriate policies.

Finding such arguments empirically inadequate to account for British and American social politics between the 1880s and the 1920s, we have developed alternative arguments from a more state-centered perspective. The structures and modes of operation of states and political parties changed historically in Britain and the United States according to the different sequences and forms of democratization and bureaucratic reform in the two nations from the nineteenth into the early twentieth century. In turn, such macroscopic political transformations were crucial because—to state the point in general terms—the structures and modes of operation of states and parties affect what might aptly be called the conditions of "political supply" of social-welfare policies.

This happens in two different but equally important ways. First, even in a liberal-democratic polity, much of what happens depends on what political leaders offer—either directly by administrative or executive initiative, or in a more roundabout way through the efforts of electoral politicians to define, and then compete to meet, social demands. Second, all politically active groups formulate their goals in part by taking into account how officials and politicians are likely to perform in relation to given kinds of policies. This means that state structures and party organizations matter not only because they affect the activities and capacities of politicians and officials, but also because they affect the expectations and demands of all the various social groups—including especially the most socioeconomically powerful and culturally influential groups—active in political struggles and discussions.

To the extent that this perspective has proved convincing in our detailed comparative analysis of British and American social politics in a watershed period, it suggests that the emergence of modern welfare programs in capitalist liberal-democracies might be brought into the same overall analytic framework as welfare developments in polities where state authorities or statebuilding exigencies have played much more apparent roles than in Britain and the United States. Much comparative-historical and theoretical work remains to be done before this possibility can be achieved, and we do not want to imply that things happen in exactly the same way in authoritarian versus liberal-democratic polities. Yet this examination of two cases that ought, more than any others, to be explicable in terms of our received socioeconomic theories of "welfare state development" certainly underlines the need—and opportunity—for analytic reorientation.

Comment

ARTHUR MANN

It is flattering to be the only historian asked to the Thomas and Znaniecki Conference on Contemporary Social Theory. I also welcome the invitation to respond to the Skocpol-Orloff comparative study of the foundations of the British and American welfare states in the years surrounding World War I. Their paper bears on a synthesis I am trying to write of the American Progressive movement. The term is an umbrella term that was coined just after 1910 for a multitude of reform movements that had been agitating the country since the end of the Spanish-American War. Continuing through World War I, they placed so distinctive a stamp on their times that historians later named the first two decades of this century the Progressive Era in American history.

The Progressive movement in the United States stood essentially for three objectives: political reform, the regulation of the economy, and social welfare. But why, with respect to outcomes, did Americans in Woodrow Wilson's generation do more in regard to the first two than the third? The question takes on an extra dimension when one observes that in Lloyd George's Great Britain of the same time the emphasis was the other way around.

The paper by Theda Skocpol and Ann Orloff takes such a transnational view in asking questions about the Progressive movement. Although their ultimate objective of contributing to a general theory of social change is not mine, their immediate quest relates directly to my own inquiry. They seek an explanation for why, in the opening two decades of this century, British New Liberals were more successful than American Progressives in setting up a welfare state. I find their account very helpful and therefore wish to be useful to them in turn. It is in that spirit of reciprocity that I address the approach, sources, and significance of their work.

The logic of their approach (or methodology as sociologists say) strikes me as sound. They first show that socioeconomic processes do not account for Britain's head start over the United States. Why not? The two processes most often cited in grand theories on the rise of a welfare state—"the logic of industrialism" and "working-class strength"—were more or less the same in both countries. Nor can one contend that America lagged behind the former motherland because of a greater attachment to classical liberal values: limited government, individualism, and voluntarism. Skocpol and Orloff hold those factors constant, too, for the two sides of the Atlantic.

All that being the case, one must look for a solution to the problem outside the industrial revolution, class relations, and traditional values. Skocpol and Orloff find it in political structures. Far more so than the United States, Britain had a unitary government; each of its major political parties was ideologically uniform; the Poor Law needed overhauling; and a civil service had evolved by 1900 to administer a welfare state. Not only was the United States different in those respects, but its "patronage democracy" had become so corrupt and inefficient that even American proponents of social insurance hesitated to entrust welfare to public officials.

So far, so good. Skocpol and Orloff have set the stage for cabinet ministers and civil servants to reform Britain from above. That such types were the prime movers for the welfare state over there is something that historians have known for some time. But in locating the impetus for change in political structures, Skocpol and Orloff have to put a finer edge to their quarrel with theoreticians who give top priority to socioeconomic processes.

Before you can have legislation for factories and factory workers, you've got to have factories; and these last did, in fact, come into being with the industrial revolution. Skocpol and Orloff of course know this. Their task is to sharpen their distinction between necessary and sufficient causes.

My next point, about sources, is inseparable from substance. What is one to read in order to get inside the minds of the Britons who, from the top down, laid the foundations for a welfare state? Although Skocpol and Orloff address the question, they pass over it too quickly. More than that, they occasionally slip into the reification trap. It is well and good to say that the welfare state was a stage in nation-building, but stages don't design or build. Designers and builders do.

Theda Skocpol and Ann Orloff therefore ought to pay closer attention than they have to the actors in the history they recount. As the manuscript now stands, it is by no means clear how or why politicians as diverse in social background as Lloyd George and Winston Churchill came to agree with the socialists Sidney and Beatrice Webb on using the power of the state to establish a national minimum. Nor have Skocpol and Orloff thought to ask why Conservatives in Commons and the House of Lords ratified that departure in public policy. New Liberals, Fabians, Tories—what were their motives?

Possible answers abound in the theoretical literature. They range from the Marxist putdown of sham concessions to the proletariat to Morris Janowitz's appreciation of a genuine concern for the dignity and self-esteem of fellow citizens. Which of the many possibilities, if any, fit the actors at issue? If I ask Theda Skocpol and Ann Orloff to be specific, it is because I think specificity is in order. Students of the past, whether they call themselves historical sociologists or historians, have the responsibility of recreating the dead as they probably were, and then of making them talk to the living in ways that the living will understand. Properly done, that humanistic rendering of the past invests generalizations with a concrete persuasiveness.

The Skocpol-Orloff manuscript contains few references to biography, yet that genre usually gives the reader a greater sense of contemporaneity with bygone ages than do theoretical, general, or monographic works. The same is true, in many instances, of what historical actors themselves wrote or spoke, often in heat of battle. Theda Skocpol and Ann Orloff have studied some of those primary sources; their work would profit from their considering more of them.

My remarks on types of evidence apply to the American side of the Atlantic, not just to the British. From my own reading of the sources, Skocpol and Orloff have yet to prove that American Progressives were not more attached than British New Liberals to individual effort and voluntary action. I am impressed by the number of American reformers who, coming from small-town and rural backgrounds, distrusted not only bureaucrats but bureaucracy. Nor were they, on the whole, as class-conscious as their British counterparts. This last difference is of particular importance because of the Skocpol-Orloff contention that the major thrust for reform came from above.

To get at what I am getting at, consider Beatrice Webb and Jane Addams. Born within two years of each other, they emerged as the foremost female advocates of social justice in their respective countries, and today each is still known in the English-reading world for an outstanding autobiography written while still active. Their reminiscences demand attention, in the context of our conference, for the light they shed on unlike impulses in British and American reform.

Not the least dissimilarity has to do with class. "As life unfolded itself I became aware," Webb recalled, "that I belonged to a class of persons who habitually gave orders, but who seldom, if ever, executed the orders of other people" (Webb 1926, p. 42). There is no comparable statement in Addams's book. On the contrary, she remembered her own girlhood as that of "a western American who had been born in a rural community where the early pioneer life had made social distinctions impossible" (Addams 1910, p. 38).

Although both women changed over the years, neither outgrew her respective origins. The upper-crusted English Fabian viewed public affairs from the top down, whereas Hull-House's Illinois founder upheld the American tradition of a common citizenship, on which her father, a self-made man and Lincoln disciple, had reared her. Decades later, while in England to study its social movements, Addams rejected their "class-conscious" approach to the problems of the poor as contrary to her country's spirit (Addams 1910, pp. 23–42).

All this notwithstanding, the Skocpol-Orloff paper makes an incremental contribution to learning that is also exciting. Rarely does one have the pleasure of reading so deft a piece of historical detective work. By means of the comparative method, Theda Skocpol and Ann Orloff have identified and assembled the various parts of the puzzle. Unless I am mistaken, the important ones are in place. It remains for the authors to have another go at the actors who left an indelible mark on the first two decades of our century.

Howsoever we might wish otherwise when we look enviously at the natural and physical sciences, it will be a long time, if in fact the time ever arrives, before history or sociology becomes cumulative. But there is another, more positive way to describe our respective fields and the relationship between them. Both are amplifying disciplines about things human; the members of each should therefore resound against the other. I take it that Jim Coleman meant just that when he invited me to take part in your deliberations.

Comment

DAVID L. FEATHERMAN

In my necessarily brief discussion of the fine-grained, penetrating historical analysis of the emergence of welfare policy in Britain and the United States by Skocpol and Orloff, I wish to draw attention to several apparently parallel developments in metatheories which motivate inquiries about social change, on the one hand, and about individual change, on the other. Both developments in theoretical perspectives reflect an approach that can be called "contextualistic."

Skocpol and Orloff refer to Tilly's recent summary of theoretical thinking about social change in setting out their own point of departure from these, perhaps limiting, assumptions. They might also have cited extended works by Eisenstadt, Nisbet, and others who have commented on the burden of "evolutionary" paradigms and of "historical progressiveism" in the history of social theory. For example, Nisbet's *Social Change and History* (1969) offers a critique of the typological approach in sociology. In this genre of analysis, society becomes an intellectual abstraction from empirical, naturalistic instances. This abstraction, society, manifests itself in one of several qualitative forms according to some structural logic, usually hierarchically sequential in nature, which is at base a classification into which empirical instances of societies can be placed with more or less "goodness of fit." The sequence of hierarchy reflects a teleological or teleonomic mechanism, by incorporating the realization of either some exogenous set of ultimate constraints that "draw" the society forward progressively or some endogenous, genotypical factors for structural-functional transformation that "drive" societal forms through the hierarchy. Taken as a whole, this paradigm views social change as universal, end-state oriented, irreversible, and manifested in hierarchical sequences. In many respects it takes a closed-system approach in which the main "driver" of change is endogenous; the driver is the process of adaptation to perhaps random but, nevertheless, exogenous "shocks" to the social system. As such, the social system tends to be cast as a passive reactor in its environment even as it remains subject to metamorphosis from within.

In the case of individual human development, challenge to the intellectual hegemony of a so-called biological-growth paradigm of ontogeny has been vigorous over the past decade and a half (e.g., Baltes, Reese, & Lipsitt, 1980). In a variety of social-behavioral sciences a life-span perspective has emerged as an alternative model; it is a "contextualistic" approach to developmental change (Featherman, 1983).

Skocpol and Orloff's comparative analysis of national political systems also reveals a similarly complex, interactive contextualism, as applied to organizational rather than person units of analysis. It suggests that human systems evolve in a sometimes reversible sequence of states or phases (e.g., during the Progressive period in the United States the demise of de facto old-age, survivor's, and disability pensions); it suggests that the particular pattern of emergence of the welfare state is not only "age-graded and history-graded" but reveals many idiosyncratic "life-history" influences (hence the rich fabric of unique historical influences that distinguish Britain and the United States); and suggests that state bureaucracy and leadership are themselves actively implicated in the initiatives of and resistances to the development and diffusion of welfare provisions within the state.

One last general observation about the study of social change as paralleled by contemporary problematics in the study of individual change. One might ask, having read a richly textured contextualistic analysis such as in the Skocpol-Orloff paper, "Is the study of societal development still feasible?" To my thinking, it is neither necessary nor desirable to give up the concept of development as applied either to individuals or to societies. To be sure, not all change is developmental; but if one were able to analyze the necessary conditions under which entities changed in a developmental pattern—as distinct from nondevelopmental change over some interval—that would add to the power of the theory (Featherman & Lerner, 1985).

But how to do that within the relativistic, contextualistic paradigm? We have given up the handy a priori templates for sequential change qua development—for example, the logic of intrinsically induced, ineluctable societal transformations as in the thesis of industrialism, or Werner's orthogenetic principle of ontogeny (1957). We cannot simply compare societies against some ahistorical, acontextual hierarchical classification to determine their developmental status, or whether over some interval they have "developed." In my own recent writing (Featherman 1985; in press) I have both suggested a reconceptualization of the concept of development and recommended the adoption of a method for its study within a contextualistic framework. The concept of duration dependence and the methods of dynamic modeling (e.g. Tuma & Hannan 1984) provide useful diagnostic tools in the discrimination of developmental process within the general domain of change. These approaches also seem promising in assisting the analyst to interrelate embedded processes of change that may be developmental (Featherman & Petersen 1985). That is, it permits one to represent the hypothesis that developmental change in some entity arises because it is entrained by its context which induces a particular dynamic in resonance to the context's dynamism.

I would offer this quantitative, comparative approach as a complement to the historical, comparative one in the Skocpol-Orloff paper. The additional value of the former is its insistence on representing extra-entity relationships (e.g., selective processes) as well as intra-entity relationships (e.g., adaptive processes) as the *raison d'être* of developmental change—social or individual. Further, it offers a formal definition of developmental change without a commitment to progressivism while providing a definitive differentiation of development from change in general.

General Discussion

Joseph Ben-David: I would like to follow up Arthur Mann's suggestions. When I listened to this paper one thing which came into my mind was that Bloomsbury was within walking distance of Whitehall. But Washington—at that time—was in the boondocks and there was no Bloomsbury in the United States. I think this was a basic difference in the social structure of the two societies.

Wasn't it also important that this was a period of very rapid immigration in the United States? Many of these immigrants came from eastern Europe and southern Europe—that is, they came from places where the government was authoritarian, tyrannical, and cruel. Therefore, those classes which would have been most interested in these reforms had probably least trust in any government, especially if it was central. They developed functional substitutes for central welfare legislation, in the form of ethnic organizations. I also wonder whether the trade unions themselves in the United States might not have contributed to the lagging of welfare legislation by preferring to develop their own pension schemes.

Theda Skocpol: First of all, in this period, it's not a question of providing for the lower portions of the working class. European pension or social insurance programs were not aimed at the very poor or even the not so well off within the working classes. Second, we have found no evidence for the United States that immigrant benefit societies created a situation in which immigrants were not willing to take help from government. Indeed, they were quite willing to take help at the local level. That is one of the legacies of patronage democracy that created problems in the eyes of the elites. Finally, the point about intellectual ties to government circles is one that interests us a lot. We didn't just look at the United States at the national level. In many of the key metropolitan areas where reform was being debated for state-level policy innovations, there were very close relationships between the university elites and public officials. Certainly that was true here in Illinois, and it was true in New York and Massachusetts as well.

Ann Orloff: I want to underscore the point that it was the "respectable" working class who were being taken care of by the programs in Europe.

By contrast, in Massachusetts the Civil War pensions went disproportionately to "native whites" and it is precisely when that cohort died off that agitation for old-age pensions began in the United States.

Joseph Gusfield: The crucial part in America of course is the failure to have brought in some substitute for the Civil War pensions once they reached their end. And you seem to attribute this to the Progressive Era and the attack on patronage, which was the concept of a welfare system at that particular point. But why did those progressives within the movement who would have sponsored a more bureaucratized system of old-age pensions not win?

Theda Skocpol: We were able to find progressives, including the most reform-minded, who balked at old-age pensions precisely on the grounds that they represented free-flowing spending that might feed political corruption. These reformers would have had to be the source of energy because they were the supporters of other social measures. If they balked at old-age pensions, that strengthens our argument about the failure of the middle-class reformers to coalesce with labor in the United States. This failure contrasted with the cross-class coalition that pushed successfully for old-age pensions in Britain.

Michael Hechter: Most of the comments have been asking you whether you controlled for a whole host of independent factors that might cause this variation in policy outcome and I frankly am a little skeptical of your ability to deal with these questions using your research strategy of comparing three units. We heard from Ed Laumann yesterday something about the complexity of the policy formation process even in one country let alone comparing it with two, and you're talking here not only about a single policy but a panoply of social policies. I'm just pessimistic about how far you can go with your research.

Theda Skocpol: I am certainly not going to argue with you that a hundred cases is better than two in some abstract sense. This study does make a very sustained effort, through a whole variety of comparative strategies, to approximate the kinds of controls that are needed to eliminate the contending counterarguments. The fact that we make comparisons across policies to examine certain kinds of arguments can turn the multiplicity of the policies into an advantage. I think all one can do with these kinds of things is to do the best you can with the comparative historical material, to examine in detail the precise causal mechanisms that alternative theories postulate. I'm not sure that highly quantitative studies that build extremely simple-minded variables into the analysis can be as rich a source of theoretical innovation as this kind of study.

Harrison White: As I heard you, the fact that the Americans were using regulatory measures at all came out as a by-product or a leftover. The more I think about this the more fascinating I find the question. Why regulation? I think Professor Mann would be happy to tell you it was very important and one of the few cases of a long-lasting cultural influence. Those of us who are studying the American economy see the leftovers of that regulation still coming in. Where did it come from? I don't think you have told us, partly because you are so focused on welfare. I'm putting in a plea to write another paper and get at regulation. It's extremely interesting.

Ann Orloff: It seems regulatory legislation served a function for politicians similar to the welfare state legislation in Britain. They were able to appeal to middle-class and working-class groups with that kind of legislation.

Theda Skocpol: If one were to take up the problem of explaining why regulatory legislation was so prolific in this period and why American regulation takes the distinctive form it does, that would be a challenge at least equal to the formation of the American welfare state.

Sociolinguistics

Language Structure and Social Structure

WILLIAM LABOV

The past three decades have witnessed a great deal of scholarly activity under the label of "sociolinguistics."[1] Yet the barrier between sociology and linguistics remains as firm as ever. In their studies of speech communities, linguists have as often as not tried to create their own sociology, with curious results; and a vanishingly small number of sociologists

[1] The present form of this paper is a slightly revised version of the one I distributed at the conference. Although there are many ways that the results of the conference might change what I have said here, those changes are best presented in the form of future research. I am particularly grateful for the comments of Allen Grimshaw, whose discussion is based on a less complete version of the paper distributed in advance. I have not attempted to make all of the changes that his remarks might call for, since that would produce a circularity that would be puzzling to the reader of this volume. I have taken advantage of the good offices of my sociological colleague Teresa Labov, however, and made a number of improvements in response to her final reading of the manuscript.

have made use of the tools of linguistic analysis.[2] On the sociological side, this is not too severe a limitation. A great deal of important work has been done in the sociology of language where the data take the form "X speaks language Y," without any further analysis of Y (e.g., Deutsch 1966; Fishman 1966; Lieberson 1965). Impressive advances have been made in studies of conversation on the basis of our ordinary understandings of meanings of words and sentences (e.g., Sacks, Schegloff, & Jefferson 1974). On the linguistic side, the presence of the barrier may have had more serious consequences.

Formal linguistic studies have usually been based on introspective judgments from the theorist, or from an informant who responds to direct questions about language. This approach advances the study of formal structure quite rapidly, and it is an indispensable first approach to a little-known language. But such introspections can hardly provide the bedrock of an objective science, and formal theorists have suffered from a surprising lack of awareness of the social factors that bear upon their judgments. The end result is a model of language that reflects the temperament of the theorists: an optimal mechanism for representing logical propositions, free from any historical or social influence.[3]

The linguistic studies that are the basis of the present report represent a different approach, which includes efforts to overcome the barrier between sociology and linguistics. (1) The studies are descriptions of speech communities rather than individuals; (2) they are based on recorded interviews that approach as closely as possible the speech of everyday life; and (3) they involve multivariate analysis of the variation that is inevitably found in such records. For this report, I will draw on

[2] This is a strong statement to make after some twenty years of interdisciplinary effort, but it is not easily qualified. Some reasons for the disciplinary barrier are suggested in "Crossing the Gulf Between Linguistics and Sociology" (Labov 1978). The lack of transfer applies to the techniques that linguists use in their main line of business: phonology, morphology, syntax, comparative linguistics. In the area of pragmatics, discourse analysis, and speech act theory, which crosses sentence boundaries, linguists have no privileged position, and the most important contributions have come from others: e.g., Grimshaw's work on verbs of manipulation (1981), Goffman's papers on replies and responses and "response cries" (1981), and the work of Sacks and Schegloff mentioned below.

[3] The tendency of formal linguists to abstract from social and historical events has been formulated most recently by Chomsky in his Pisa lectures as a guiding principle: "that the theory of core grammar, at least, is based on fundamental principles that are natural and simple, and that our task is to discover them, clearing away the debris that faces us when we explore the varied phenomena of language and reducing the apparent complexity to a system that goes well beyond empirical generalization and that satisfies intellectual or even esthetic standards" (1981, p. 14).

twenty-five studies of this type carried out since 1960 in urban and rural communities in the United States and Canada, Latin America, western Europe, the Near East, India, and China.[4] To make these materials accessible and useful to sociologists, I will reduce the linguistic details to a minimum. This is a practical move, since there is a high degree of independence between the social and linguistic factors that influence linguistic variables.[5]

In presenting these results, I will make use of several sociological frameworks for viewing complex systems. I will review the major findings of studies of the speech community: first in a structural-functional perspective that was automatically assumed in earlier work, and then within a framework that takes conflict as a fundamental feature of society. By reducing the results to their most general outlines, I hope to make them available for the general discussion of the utility of these competing perspectives.

The Stability of Linguistic Stratification

A limited number of social variables have been found to control sociolinguistic variation: socioeconomic class, sex, age, ethnicity, race or caste, and urban-rural status. The most regular and stable patterns have been shown by the first two, which will be considered in this section.

Class and Style Stratification

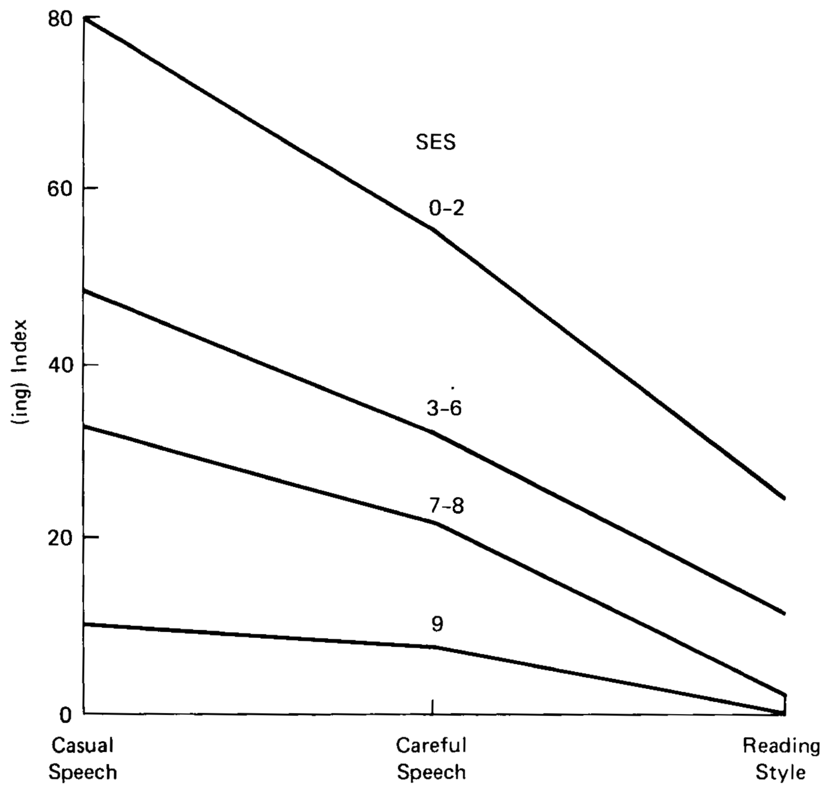
Figure 1 is a display of the social and stylistic stratification of a sociolinguistic variable in New York City. It is based on the recorded speech of eighty-one residents of the Lower East Side, from a secondary survey of a random sample constructed by sociologists who were concerned with employment opportunities for youth.[6]

[4] A review of recent studies of urban speech communities is given in the Postscript to the 1982 printing of Labov 1966.

[5] This appears most clearly in the multivariate analysis of data sets that include both internal and external variables. When alterations are made in an internal, linguistic independent variable, compensating changes are found in other linguistic variables, but none in the external social factors such as sex or age, and vice versa (Sankoff & Labov 1979).

[6] This was Mobilization for Youth, a large-scale assault on the problem of juvenile delinquency, based on the program developed by Cloward and Ohlin (1960) for maximizing opportunity structures. The social class index was constructed by John A. Michael.

Figure 1 Class and style stratification of (ing) in working, living, etc., for white New York City adults. Socioeconomic status scale: (low) 0–2, 3–6, 7–8, 9 (high). Vertical axis is percentage using the nonstandard form in. *Horizontal axis displays styles of speech.*



The variable is the pronunciation of words ending in unstressed /ing/ and will be indicated as (ing). The well-known variants are *ing* and *in* (often spelled in'). The vertical axis shows the percentage of the nonstandard form *in*. The horizontal axis displays three divisions of the interview materials: casual speech, which comes closest to the vernacular style of everyday life; careful speech, characteristic of question-and-answer routines; and reading style.[7]

[7] Further levels of formality were used in these interviews, with an even stronger focus on language: the reading of word lists and judgments on minimal pairs.

For each style, Figure 1 shows the mean values for subdivisions of the status hierarchy, following a set of objective indicators of socioeconomic status established in the primary survey: an equally weighted index of occupation, education, and income.

The pattern of Figure 1 has been reproduced in other studies of (ing) in the United States, Canada, England, and Australia.[8] Figure 2 shows (ing) for Norwich, a medium-sized city in the east of England studied by Peter Trudgill (1974). Similar patterns have been charted for other stable sociolinguistic markers in English; the Spanish of Puerto Rico, Buenos Aires, and Bahia Blanca; the Portuguese of Rio de Janeiro, Sao Paulo, and Belo Horizonte; the French of Montreal, Paris, and Tours; the Persian of Teheran; and the Arabic of Amman (see Postscript to 1982 printing of Labov 1966).

The regularity of these patterns can be assessed in several different ways. They show that the speech community is heterogeneous: At each contextual style, the population is differentiated by the linguistic variable. Yet the community is also homogeneous: each social group follows the same pattern of style shifting.

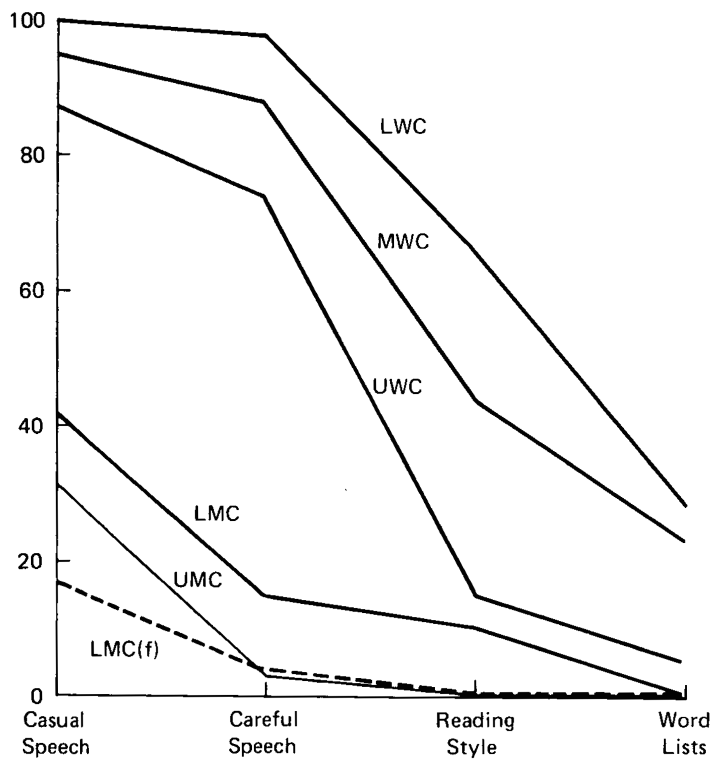
The overall pattern may be described as "orderly differentiation" of the speech community. It reflects a uniform recognition of the prestige norm: in the case of (ing), the *ing* variant. This uniformity of evaluation is reflected in controlled experiments in the field: In self-report tests, subjects show a regular shift toward the prestige norm;[9] they also show a high level of agreement in "matched guise tests," where their unconscious social attitudes toward the linguistic features are conveyed as personality judgments.[10] Uniform treatment of the variable also rests on a

[8] Fischer (1958) studied (ing) in a New England town; Labov (1966) in New York City; Trudgill (1974) in Norwich; Cofer (1972) in Philadelphia; Mock (1979) in the Ozark Mountains; Cheshire (1982) in Reading, England; and Shopen and Wald (1982) in Australia.

[9] A "self-report test" is a field experiment that registers the effect of a norm of the dominating society of a speaker's perception of his/her own speech. A tape recording of four or five variants of a variable is played and the listeners select one as closest to the way they usually speak. See Labov (1966, chap. 12) and Trudgill (1972) for a more extended development of this technique for registering covert prestige in British English.

[10] Subjective reaction tests are based on the "matched guise" technique developed by Lambert (1967). Subjects hear a series of short passages spoken by different speakers and are asked to make personality judgments or rate the speaker on a variety of socially oriented scales (job suitability, friendship, toughness). The same speakers recur in the series in different guises, and the subject's ratings are effectively independent. In linguistic adaptations of this method, efforts have been made to get increasing control of the stimulus, so that it is

Figure 2 Class and style stratification of (ing) in Norwich (adapted from Trudgill 1971).



uniform structural base. For each speech community, the definition of the variable involves specific linguistic details. The words that are included in the (ing) variable will vary from one community to the other, although the sociolinguistic pattern of the *ing* and *in* variants remains the same. A speech community may then be jointly defined by a uniform set of linguistic structures and a uniform set of evaluative norms.

---

possible to say precisely what listeners are reacting to (Labov 1966; Labov et al. 1968; Labov et al. 1972). In the most recent developments, natural speech is altered in individual parameters, and resynthesized. This technique has been used to identify the specific features that lead to the identification of a speaker as black or white (Graff & Labov 1983).

Note that Figure 2 shares all of these properties with Figure 1. Yet the quantitative data are also sensitive enough to reflect a sharper separation of middle-class and working-class speakers than we find in the United States, reflecting the greater separation of the two groups on both a symbolic and economic plane.

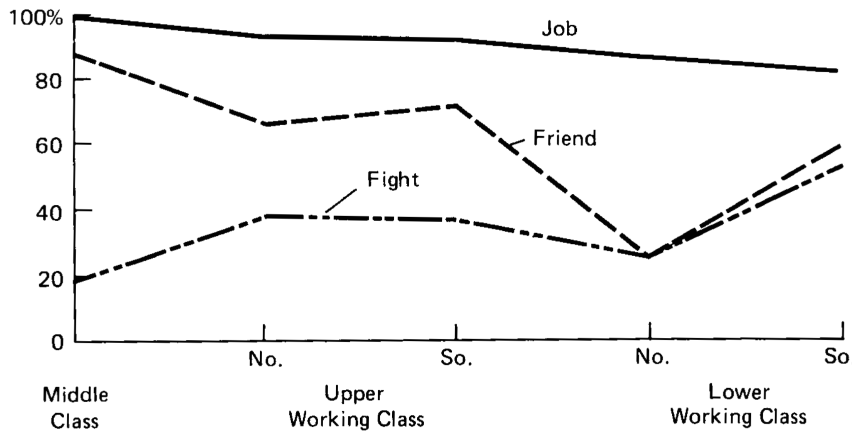
Not every linguistic variable becomes a sociolinguistic marker: many alternations show no social variation at all. Those that do develop stable social stratification become aligned with the socioeconomic hierarchy in a regular way and develop equally regular stylistic stratification. But if the entire community recognizes the same overt norms for the variable, what prevents the gradual triumph of the prestige norm and the elimination of the stigmatized norm?

The long-term stability of sociolinguistic variables is one of their most remarkable features. In some cases, their history extends across social transformations and revolutions: not over decades and centuries, but over millennia. The variable (ing) appears to be a continuation of the alternation of two suffixes of Old English: the noun-forming suffix *-ing* and the participle-forming suffix *-ind*, which became by regular sound change *in*. Although the *ing* form has been the written standard since the fourteenth century, the *in* variant has served a succession of social classes as an alternate associated with spoken as against written style, and as a symbol of their opposition to the dominant social class.

Such long-term stability implies the existence of an opposing set of covert norms, supporting the overtly stigmatized form. Although the dominant prestige norm is usually the only one recognized in formal experimental conditions, covert norms have been detected in self-report tests and in matched guise subjective reaction tests. Figure 3 shows the responses of black adults in South Harlem to passages spoken by a middle-class and a working-class black speaker. The vertical axis shows the percentage of subjects who rated the middle-class speaker higher than the working-class speaker on a seven-point scale. The horizontal axis shows the mean responses of middle-class subjects, upper-working-class subjects raised in the North and in the South, and lower-working-class subjects raised in the North and in the South. The top curve shows responses to the question, "What is the highest job this person could hold, speaking as he does?" The community as a whole rates the middle-class speaker higher, but there is a steady decline in consistency from middle-class to lower-working-class subjects from the South.

The bottom curve shows responses to the question, "If this person was in a street fight, how likely would he be to come out on top?" It is the inverse of the first curve, yielding a clear indication of an opposing covert norm responsible for the stability of sociolinguistic stratification. The

Figure 3 Percent rating middle-class speaker higher than working-class speaker on 3 scales for 5 social groups (from Labov et al. 1968; p. 242).



intermediate curve shows responses to a third question: "If you knew this person for a long time, how likely would he be to become a good friend of yours?" For the highest three social groups on the left, the responses to this question are close to the "job" question. For the two lowest groups on the right, there is a switch-over, and a close association with responses to the "fight" question. Thus, we have evidence for the opposing norms that are responsible for the stable equilibrium, and also an indication of a split in the community in alignment toward one or the other of these two norms.

Sex Stratification

The first quantitative study of (ing) was Fischer's study (1958) of children in a New England town. Girls used significantly more *ing* than boys. This finding is repeated in the many larger studies of (ing) that followed—in New York, Norwich, Ottawa, Philadelphia, and Sydney. The pattern can become so extreme that in some Philadelphia neighborhoods, there is no overlap in frequencies between the group of women speakers studied and the men. For stable sociolinguistic variables, it is generally true that women use more prestige variants than men and fewer stigmatized variants. This finding holds for all English speech communities investigated, for Latin America and Western Europe, and it appears to be true in China as well. It is characteristic of women in remote rural areas as well as in large cities.[11] The effect appears in all social classes, although it is maximal in the second highest social class and sometimes minimal in the lowest social class. In Figure 2, the mean values for women of the lower middle class in Norwich are shown by a dotted line: they are considerably lower than the lower middle class men, and overlap the mean values for the upper middle class for both sexes.

The pattern is not universal. In recent sociolinguistic studies in India, Iran, and Jordan, men have been found to use the prestige variant more than women.[12] We seem to be dealing here with a southwestern and southern Asian *Sprachbund,* a linguistic area that has developed a distinct pattern of sexual differentiation, possibly under Muslim influence.

A Structural-Functional Interpretation

When stable sociolinguistic patterns were first discovered, they were seen as structures that function in society to identify and differentiate speakers in a rational manner. The stratification of social classes in regard to a given variable allows the listener to make some estimate of the education, occupation, or residence of a speaker after a few minutes of conversation. The patterns of style shifting permit an estimate of the stance taken by the speaker in regard to the listener, which may be translated into judgments about friendliness, willingness to accommodate, or seriousness of intent. It was suggested that without such "orderly differentiation," speech communities are handicapped (Weinreich, Labov, & Herzog 1968).

We can unify the dimensions of style and class stratification by viewing them both as ways of establishing social distance between the speaker and the listener (Bailey 1973). Sociolinguistic variables are then seen as devices that allow one to adjust social distance in the speech situation: without them, social distance is more fixed, with consequent limitations on social interaction.

The uniformity of sociolinguistic behavior across ethnic boundaries in an urban community testifies to the current functioning of the system. Second-generation ethnic groups are not differentiated from third-

[11] See, for example, the study of the mountain village of Ucieda in northern Spain by Holmquist (1982).

[12] This appears in Modaressi's study of Teheran (1978) and Abdel-Jawad's investigation of Amman (1981). The reversal of the position of men and women appears in all social classes, including college-educated women who play a major role in public life.

generation groups, except for linguistic traits that are disappearing from public life and are preserved only through transmission within the family (Allen 1973).

On a broader scale, a system of sociolinguistic variables may be seen as a functional correlate of social mobility or an instrument for mobility. When all speakers have access to the underlying structures and norms, it is possible for them to adjust their use of the variables as they shift their places in the social hierarchy.

When speakers do not share a common set of norms and structures, sociolinguistic variables can operate as barriers to social mobility. The southern British sociolinguistic variable (h) stratifies the population without a common structural base: many working-class speakers do not know which words contain *h* in the standard dictionary. We then have hypercorrection in formal speech—*Hi say, That's hawfully good of you*—with consequent limitations in social mobility.[13]

Racial Differences Within the Speech Community

The discussion so far has presented a view of a speech community with a uniform set of norms and structures governing stable sociolinguistic variables, with orderly differentiation of the use of those variables. All of this applies across class, ethnic, age, and sex differences among speakers within the white community. There are dramatic differences when we come to racial differences in language.

Convergence and Divergence: The Black English Vernacular

In the discussion to follow, the term "Black English" will be used to refer to the entire range of forms of English used by black people in the

[13] In stable societies with a long history of stable sociolinguistic stratification, we find comparatively little hypercorrection. In the United States and England, we rarely find speakers pronouncing *th* in place of *d* in their formal speech ("I put it *thown there*") although the frequency of phonetic [d] for *th* may be quite high in informal speech. Abdel-Jawad (1981) found no hypercorrection in Amman; although speakers often used the phonemes /k/, /g/ or glottal stop in place of the classical /q/, they always knew which words originally contained /q/. These speech communities thus operate with a uniform structural base. But hypercorrection is quite common throughout the English speech community whenever speakers acquire superposed dialects.

United States. The term "Black English Vernacular," or BEV, will be used for the particular form of Black English that is used by the majority of black citizens of the United States, as children growing up, and as adults in their intimate conversation with their peers. "Vernacular" is a technical term in sociolinguistics, used to indicate the first linguistic system acquired by native speakers, over which they have the most automatic control, which can be used with the minimum attention given to audio-monitoring. Every speaker has a vernacular; the vernacular of many middle-class black speakers is much closer to the standard English of the classroom than that of working-class blacks in the inner city. But the vernacular of the inner cities is remarkably uniform throughout the United States, and many middle-class black speakers have a strong passive knowledge of this system.

The grammar of BEV has been studied in some detail since the early 1960s, and it is perhaps better known than any other nonstandard vernacular. There were many sharp controversies over how different it was from the grammar of other dialects, what its origins were, and whether it formed an impediment to reading and writing standard English. A consensus on these points has gradually emerged, supported by both quantitative and qualitative evidence. This consensus can be summed as three propositions:

1. The grammar of BEV shows the effects of an Afro-Caribbean origin, with many typological similarities to the Creole languages of the Caribbean. Most of these features are in the verbal system, particularly the subsystem of tense and aspect.[14]

2. Many grammatical and cultural features are shared with dialects of the southern United States, particularly in the area of syntax. Here it appears that influence has operated in both directions.

3. The grammar of BEV has gradually converged with the grammar of other dialects. At one time, BEV treated adjectives like verbs, and the grammar had no place for *is* in an expression like *He tall.* But at present, mature BEV speakers produce sentences like these as optional extensions of contraction: *He is tall* → *He's tall* → *He tall.*

So far, the situation seems promising for the integration of BEV and other dialects into a single system that will operate like the stable sociolinguistic variables with a common structural base. Many believe

---

[14] "Tense" refers to a set of grammatical categories that locate an action in a point in time—like past, present, and future. "Aspect" is a more subtle type of category, which describes the contour of the action in time: repeated, extended, completed, beginning, etc. The English progressive and present perfect are aspectual categories of standard English.

that there are advantages to a "bidialectal" situation, where speakers use their vernacular for intimate conversation with their peers, but shift easily along a continuum to communicate with more distant groups in formal settings. But there are indications that BEV is diverging from other dialects in many respects.

Current research on the tense and aspect system of BEV shows that it is typologically allied to West African and Caribbean languages that stress aspect at the expense of tense; but not one of the BEV aspect markers shows the same semantics as the corresponding Caribbean form. The complex meanings that have developed in BEV have not been reported for any Creole language.

The auxiliary *been* is used in the Caribbean to indicate the past tense. In BEV, it has been developed to mean that the action referred to occurred in the past, a relatively long time ago, and that it is still true at present:

> A: That's a new coat you've got.
> B: I've *been* own(ed) that for two weeks.

Such semantic features could have developed in the United States a century or two ago; or they may be present but undescribed in the Caribbean. There is no absolute proof that we are dealing with linguistic divergence. Harder evidence of divergence in progress has appeared in our current studies of "third-singular *s*" in North Philadelphia.[15]

All previous studies of BEV show the absence of any fundamental rule of subject-verb agreement that places an *s* after the verb in the present tense if the subject is a third-person-singular noun phrase: *I go, You go,* but *He goes.* Experimental evidence shows that young black children in the second grade can interpret this *s* as a marker of the present, but they do far worse than chance in trying to use it as a marker of singular versus plural: On a forced choice, they will interpret it as a sign of the plural (Torrey 1983). In North Philadelphia, a new development has appeared in the grammar of a group of young adult speakers who have minimal contact with white society. The suffix *s* on the verb has become for them a marker of the past narrative, used for all persons (*I goes, we waits for him, he comes over*), but hardly ever in present contexts (Myhill &

[15]The research reported here was supported by NSF as part of a study of "the influence of urban minorities on linguistic change" from 1981 to 1983 at the Linguistics Laboratory of the University of Pennsylvania. I am indebted to my colleagues Wendell A. Harris, Sharon Ash, David Graff, and John Myhill for the results outlined here.

Harris 1983). No speaker over 30 shows this pattern; it is not present in the speech of adolescent youth in Harlem recorded in the 1960s. This is a clear confirmation of continuing divergence from other dialects; it is also the first report we have had of a regional difference in the grammar of BEV.

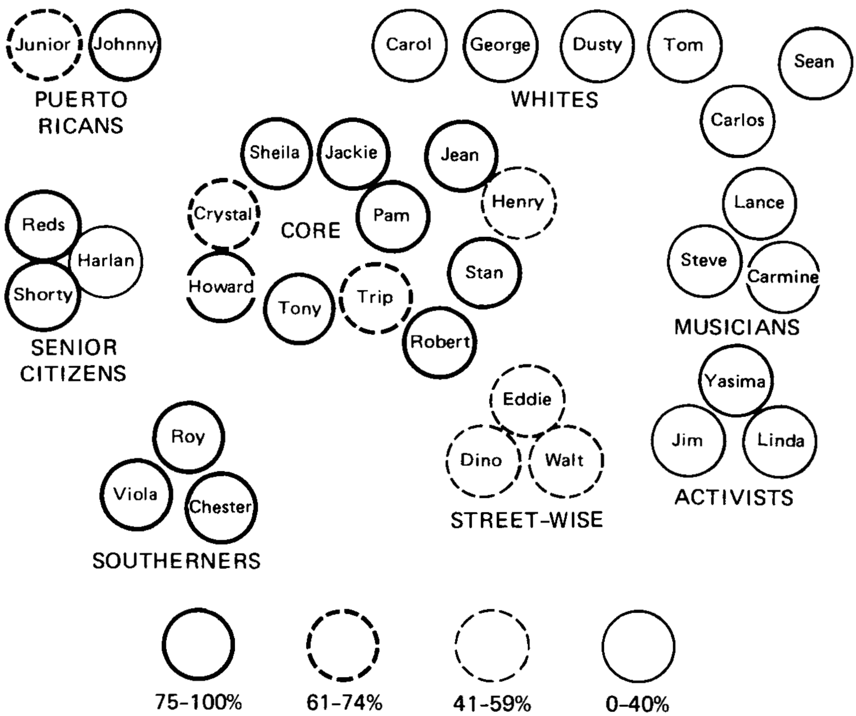
It is not hard to see that the social and economic conditions for divergence are present in North Philadelphia: an increasing rate of residential segregation in a city that is now 38 percent black, and an increasing rate of unemployment (over 50 percent for this age group). This does not mean that black people are not making social progress in Philadelphia. There is a steady increase in the numbers of upwardly mobile blacks in the areas to the west of North Philadelphia, where mixed residential areas provide a transition belt between lower-class blacks and upper-class whites. In addition, there are many other routes to contact between whites and blacks that lead to the penetration of the grammar of BEV by the rules of other dialects.

Figure 4 shows how third singular /s/ is treated by groups who have followed alternate routes of social mobility in the black community: routes that are not associated with higher education or movement into the middle class. The thickness of the circle for each individual indicates the frequency of absence of the standard grammatical marker. In the center are seven speakers from the core vernacular group: young adults with minimal white contacts.[16] Minimal use of the standard form is shown by this central vernacular group, by black senior citizens from the inner city, and by blacks raised in the South. But on the right-hand side of the diagram there appears a penetration of standard forms into the speech of people who have otherwise retained their black identity in all obvious respects: musicians engaged in jazz, rhythm and blues, and rock; militant blacks involved in political protest movements; and older street-wise blacks who have experience with confidence games, forgery, and burglary.

It appears that any mechanism that leads to greater personal contact with the dominant culture will lead to modifications of the vernacular grammar in that direction, irrespective of ideological bias or conscious effort. It is all the more striking that the majority of inner city blacks are diverging from and not converging with the dominant linguistic pattern.

[16] In his 1979 study of BEV in Pacoima in southern California, Baugh showed that adults who worked with vernacular speakers and lived with vernacular speakers showed the most consistent forms of language. Until that time, the deepest, or most "basilectal" forms of BEV had been recorded from adolescents.

Figure 4 Social types in the North Philadelphia black community in relation to absence of third singular verb marking.



The numbers of black people in the segregated areas continue to increase, income differentials between black and white continue to increase, and educational achievement in the all black schools remains very low. The linguistic situation correlates with the formation of what has been called "a permanent underclass."

The Problem of Accessibility

The fact that black and white sections of the speech community differ in their linguistic habits does not mean that they belong to two separate systems. A structural-functional approach to the situation described in the last section might easily show that the linguistic divergence is a predictable adjustment to polarities in the social structure. Sharper opposition between black and white, even greater antagonism, does not break structural relations. It is a commonplace since Simmel that conflict unites the antagonists in a web of social relations.

However, there are two signals of structural incompatibility that could lead us to posit distinct systems in conflict rather than a single structure that contains the two in structural opposition. One is continued divergence, indicating a fundamental disequilibrium. The other is the inability of the antagonists to recognize the conflict. I have pointed to the first condition in the black-white conflict of the inner cities. It might seem that the second condition does not hold, since both blacks and whites know that they speak differently, can imitate each other to some extent, and discuss the matter freely.[17] In recent years, there has been a great deal of public discussion of "Black English," and the term itself now has a certain social standing.

Yet there is strong evidence that the fundamental black-white linguistic opposition lies well below the level of social awareness and is inaccessible to the ordinary mechanisms of social negotiation and resolution.

Inaccessibility to Introspection

In general, introspective judgments about the meaning of verbal auxiliaries are weak, and there is little agreement about their meanings in the standard language. But the obscurity of the BEV particles is more profound, even for linguistic analysis. Our technical knowledge of the verbal system of BEV has emerged by a slow process of interaction between black and white linguists. Whites do not perceive many black forms until overt discussion focuses attention on them. Blacks do not know that the verbal forms they use are not used by whites as well as blacks, and can recognize the opposition only when whites bring it to their attention.[18]

In everyday social life, the role of grammar is even more submerged. The grammatical elements that differentiate black from white systems do

---

[17] The contact we are dealing with must be meaningful interaction, in the sense that the behavior of one person will affect the life chances of another. It must be remembered that all of the speakers involved here are exposed to standard English for at least five hours every day on the mass media. But those who show the influence of other dialects are those who engage their speakers in ways that might yield some profit for them.

[18] In our current research in North Philadelphia, we recorded a number of imitations of whites by blacks. Most of these imitations were wide of the mark, if we take an accurate rendering of the white Philadelphia vernacular as the goal. But they all contained some standardized features attributed to whites which were related to the crucial differences between the systems, as confirmed by our experimentally controlled subjective reaction tests. On the other hand, Puerto Ricans who have adapted most of the features of BEV show no perception of the black-white oppositions in language.

not appear in novels of black authors who otherwise represent vernacular speech quite well; nor in situation comedies that in other respects show a wide range of black and white styles; nor in interviews that in other respects come close to the speech of everyday life.

The Transparency of Grammar

It was pointed out above that high frequency of contact with the dominant society alters the grammatical system of many blacks so that they can no longer be said to be speaking BEV, from a linguistic point of view, but some other intermediate form of Black English. Yet many black politicians and professionals whose vernacular is quite remote from BEV can deliver public talks in a style that community members label as "street talk," "ghetto English," or "Black English." Linguistic examination of this kind of performance shows that the effect is achieved by the use of a small number of cues: primarily intonation and lexical choice, with very little change in grammar.

Our research over the past few years has focused on the black-white linguistic interface in Philadelphia, on the speech of whites who have acquired a thorough knowledge of black culture and acceptance in black society. At the top of Figure 4 are symbols for four whites in this category. Their use of third singular *s* typically shifts very little from the white norm. Of the third from the right, a black friend said, "If he turned his back, you wouldn't know if he was black or white." But this impression has nothing to do with the grammar that he uses. Third singular *s* is rarely absent in his speech. In general, the social construct "speaks dialect X" is quite remote from the linguistic description of the underlying system.

Camouflaged Grammatical Features

The emergence of "camouflaged" forms of BEV is a process of selection through the gradual elimination of overtly stigmatized elements, and the creation of new grammatical meanings that are less accessible to observation (Spears 1980). When a black speaker says, "He come tellin' me . . . ," both black and white listeners may hear it as a slight variant of the general colloquial form "He came saying to me . . . ." But after prolonged observation, it appears that blacks also say, "He come comin' here," and "He come goin' over there." The *come* used by blacks is a part of the verbal system which means,"He has the nerve to present himself doing . . . ." Like many of the items mentioned above, it is *camouflaged*. Blacks use the form, even in middle-class, polite society, without being heard as using the language of the street.

Some Consequences of Inaccessibility

Given the distance between the social and linguistic definition of "speaks dialect X," it might follow that the inaccessibility of the linguistic structure is of little importance. By definition, the social definition of the language or dialect is accessible to ordinary people. What difference does it make, then, if speakers of different dialects have different underlying structures that they aren't aware of?

There is one area of social life where this difference is critical. When a child enters school and faces the problem of learning to read and write the standard language, unknown and unrecognized differences in the structural base can interfere with the cognitive process of learning to read and deepen the problem of establishing sound social and emotional relations to the school system. This is perhaps even more true for unrecognized differences in the rules for the use of language.[19]

Conflict Without Contact

We have seen that the existence of conflict within the speech community cannot itself present a challenge to a structural-functional interpretation. Linguistic conflicts only rarely lead to genocidal termination of social relations, even in Belgium, Spain, or India. Languages may be terminated, but the sociolinguistic oppositions will be continued within a new language. Language differences may provoke violent antagonisms, but people know how to deal with the linguistic emblems of contempt, rudeness, refinement, and rejection. Members of a speech community often have a rich and detailed knowledge about forms of speech that they would never use: they are fully competent to know when they have been insulted. Curses, taboo words, and ritual insults are available to express conflict within a system that can recognize and absorb it.

Such social reactions are based on recognition. There is no structured way for members of a speech community to evaluate linguistic differences that are not recognized. The structural-functional interpretation of social conflict is difficult to pursue when there is no way of labeling, categorizing, and reacting to the positions of each side. When we look at the divergence of black and white vernaculars in the inner city, we see conflict without contact, and therefore conflict without structure.

[19]A great deal of evidence on this point was brought forward in the Ann Arbor Black English trial (Labov 1982). Yet it must be admitted that research still falls short of specifying how the social and cognitive factors interact to produce the observed result.

This is not true of all aspects of racial conflict. We have a relatively clear view of the competition for jobs, for domination of the market in music and sports, for control of the sidewalks and the city streets. On the other hand, uncharted linguistic and cultural differences between black and white form a gap in the social structure. In geologic terms, we are dealing with a fault, where structural mismatch is the result of subterranean forces. If the disequilibrium continues without correction, we may be dealing with a rift in the social bedrock.

One might think that such discontinuities are commonplace in multilingual communities. But long-standing multilingualism provides a structure exactly contrary to the one we have been discussing. Gumperz studied such a stable multilingual village in Maharashtra, where three languages—Hindi, Marathi, and Kannada—had been maintained for centuries, and people switched rapidly from one to the other. The vocabularies of the three languages were quite distinct. But the phonetics and the grammatical categories of all three had converged, so that people were using the same linguistic machinery to produce three different languages, simply by switching the words that they used (Gumperz & Wilson 1971). In the black-white dual speech communities of the northern cities, the reverse holds true: the surface vocabularies are largely the same, but there is structural divergence below the surface. Whites often think they understand what is being said in the black vernacular, because the words are recognizably English, but their understanding may be wide of the mark. Misunderstandings are increased by the growth of camouflaged features, where the same words are used for different grammatical purposes.

The view of the speech communities that emerges might be termed "disorderly heterogeneity." We are dealing with two distinct linguistic communities, who share neither a common set or norms nor a common structural base.

External Conflict: The Pressure from Outsiders

The linguistic divergence of the black and white communities in the northern cities has two aspects. On the one hand, there is a rapid evolution of the local sound pattern (Labov, Yaeger, & Steiner 1972; Labov 1980). In all the northern cities, we find that these sound changes are confined to the white community: Blacks do not participate in them. When upwardly mobile blacks reduce the distance between themselves and whites, it is by their participation in the general network standard, not the local speech pattern. On the other hand, the majority of blacks in the inner city who are segregated from the white community and the black middle class show increasing divergence in their linguistic forms: through preservation and further development of the Black English Vernacular grammar.

The sound changes that dominate the white speech communities can be seen as intensified claims to local rights and privileges (Labov 1963, 1980). One reason for this intensification is the reaction to pressure from another speech community, defined as alien and exterior, whose members exert a potential claim to those local rights and privileges. On Martha's Vineyard (Labov 1963), there were two such exterior communities: the Native Americans and the Portuguese. In New York (Labov 1966) and Philadelphia (Labov 1980), the same role is played by blacks and Hispanics.

The excluded community is not without resources. Excluded from local rights and privileges, the black community draws support from its own traditions, norms, and talents. While the white vernacular community forms a tight-knit local reference group, the black community allies itself to a national reference group, with the striking uniformity characteristic of the Black English Vernacular.

The urban speech communities of the United States recognize a second national reference norm: the speech of upper-middle-class professionals, especially those whose occupation deals with language. This subsection of the speech community generally strives to avoid local identity, since the rights and privileges it claims are not awarded locally. There is a general convergence on a Standard English grammar, based on the written form, and a "network English" pronunciation, which is governed by the principle that local features are to be avoided.

Overview

A thorough-going structural-functional approach to language could be applied only if linguistic systems did not undergo internal change and development. But the fact of language change on a large scale is not consistent with the view of language as a structure that achieves the major communicative needs of human beings through a stable resolution of opposing pressures. We must anticipate a high level of conflict within the linguistic structure as a consequence of the association of newer and older forms with particular social groups, symbolic investment in those alternative forms, the development of prestige norms and social stigma, and ideological support of those developments.

The integrity, or cohesion, of the sociolinguistic structure would be maximal if (1) sociolinguistic oppositions were symbolized only by surface forms—sounds and words—or (2) speakers had access to the most systematic forms of linguistic oppositions. We have seen evidence here that neither of these conditions holds. It follows that speakers are not free to resolve or negotiate sociolinguistic oppositions as they would political or economic issues.

We can therefore expect that the speech community will show a high degree of structural uniformity, and orderly differentiation, whenever there is a high degree of social and economic integration and social oppositions are symbolized by superficial features: differences in sounds and differences in vocabulary. When these conditions do not hold, we can expect to find continuing divergence below the level of social awareness, gaps in the sociolinguistic structure, and misperceptions on the cognitive level.

Finally, it should be pointed out that neither of these situations can be automatically accepted as favorable or unfavorable to the interests of members of the communities involved. Although lack of cohesion and structural mismatch work to the disadvantage of the black community in the inner cities of the United States today, there are other situations where relative isolation from the main stream of social life has positive consequences. For language as well as other forms of social behavior, it seems necessary to take an objective view of the processes that lead to cultural diversity and cultural integration. Commitment to the advancement of any given social group does not necessarily imply commitment to one or the other.

Comment

ALLEN GRIMSHAW

William Labov has used a background which includes training in both "autonomous linguistics" and sociological methods and theory to contribute to our understanding of core issues in both disciplines. His past work has included both macro studies of phenomena of stratification (and mobility) and linguistic and social change; much of this research has employed analyses of results of extensive surveys of phonological production. He has also done micro studies of social interactional processes and rules; this research has attended to a variety of features of speech (phonological, syntactic, prosodic) in "comprehensive discourses analyses." Labov started his paper by observing that sociologists know little about linguistics and linguists less about sociology. He continues to prove himself an exception by the work reported here, however, in which he turns to the study of processes of social conflict, and specifically to analysis of "camouflaged" linguistic differentiation as a marker of social distancing.

There are two contrasting views on what is happening to American English (aside from the perennial view of prescriptivists that it is deteriorating!). The first is that it is undergoing homogenization as a consequence of widespread exposure to the mass media, particularly radio and television. The second, held largely by whites (academics, for example) who are beginning to interact with black Americans entering the main stream, is that black and "white" English are becoming increasingly differentiated; this latter view is based primarily on fairly salient (and often stereotyped) differences in pronunciation and lexicon. Labov agrees that there are differences in pronunciation and words in Black English Vernacular (BEV) that are taking on new meanings. The linguistic data he employs for the analyses reported here are syntactic rather than phonological; his claim is that there are differences between BEV and Standard American English (SAE) *which are not accessible to naive speakers and not obvious even to trained linguists.* These differences, as he shows in the exemplary case of absence of third singular verb marking, are most manifest (and increasing) in the speech production of blacks who have minimal contacts with whites.

How can such subtle differences be a source of socially negative outcomes? In the case of schooling, Labov observes, the courts have concluded that the differences result in unequal educational opportunity for BEV-speaking children. In the case of adolescent and adult speakers, others have demonstrated that very modest differences in speech production can be highly disruptive of communication, *particularly when those in interaction think they are speaking the "same" language.* The combination of minor differences in prosodic production, similarly minor differences in grammar, and different notions of social appropriateness in, for example, styles of conflict talk, can have major effects in generating frustration in talk involving speakers distinguished by these differences—without their being aware of the source and nature of their joint difficulties. A first implication of this work for sociological investigations of conflict is that we should perhaps attend more systematically to the possibility of analogously camouflaged (or not consciously recognized) differences in other aspects of social life. Such group-linked (as contrasted to individual) differences may exist for a wide range of attitudes, values, behaviors, and personal attributes (e.g., odors resulting from dietary differences—or differences in conceptions of personal space). The idea is not novel (vide Goffman); Labov offers a persuasive demonstration.

Labov notes that one of the reasons that the differences he discusses remained undiscovered for so long is that economic, political, and social discrimination against blacks and accompanying patterns of segregation sharply reduce the likelihood that members of the groups marked by hidden but critical linguistics differences will have more than restricted exposure to social interaction (and extended exchanges of talk) with whites; he has mentioned elsewhere that the scarcity of linguistically trained speakers of socially disvalued languages has made the research task more difficult. An irony in Labov's report is his finding that militant blacks involved in political protest are more like whites in their speech; the finding is suggestively similar to Robin Williams's that militant blacks were less hostile to whites than their less integrated fellows (Labov invokes Simmel on conflict as a generator of social relationships).

Labov has looked at subliminal linguistic differences between socially distant categories. A second possibility which sociologists might profitably explore is whether analogous camouflaged differences may not exist between groups in closer social adjacency, such as genders or categories differentiated by their relations to the means of economic production. It seems to me that these dimensions of Labov's current work are rich with exciting possibilities for sociological students of social conflict. It further occurs to me, although I have not worked out the implications, that there may be some interesting clues about the maintenance of these differences as markers of group boundaries in Olson's formulation (this volume) of the notion of "exclusivity" as a factor in the formation and maintenance of groups—particularly those with "caste-like" features.

In the brief time remaining to me I can do no more than list some things Labov has said which seem to me either to be particularly provocative and suggestive and/or to require further specification.

1. Labov remarks that "a limited number of social variables have been found to control sociolinguistic variation: socioeconomic class, sex, age, ethnicity, race or caste, and urban-rural status." I don't think he would himself deny the importance of *social context* of talk as a critical variable for such variation as code selection or choice of strategies of verbal manipulation; I have found the *sociological* variables of relations of power and of affect, and utility, extremely productive for study of micro-dimensions of sociolinguistic variation.

2. I am unclear as to how social class differences in choice of linguistic variants serve as symbols of "opposition to the dominant social class"; this seems to me to impute a class *awareness* and use of the differences which has not been demonstrated.

3. Labov's report that women are *not* universally more likely than men to use prestige variants suggests a number of interesting questions about the position of women in societies manifesting different gender patterns.

4. I don't know what it means to say that sociolinguistic patterns have been seen as "structures that function in society to identify and differentiate speakers in a rational manner"—*unless* Labov's observation is a methodological and not a theoretical one. I am similarly unclear on what it means to say that speech communities are "handicapped" if they do not have "orderly differentiation."

5. I believe that Labov is quite correct in remarking that sociolinguistic variables can be seen "as devices that allow one to adjust social distance in the speech situation." It is well to keep in mind, however, that overcompensation and errors in code adoption may lead interlocutors to conclude that the person with whom they are talking is either "socially pushy" or condescending.

Labov's original paper contained an extremely suggestive discussion of some of his work on linguistic (and social) change. It is regrettable that it had to be cut; I hope sociologists will seek out his work on the topic. There are additional specifics in what has been retained which could be mentioned. It will be clear that I have found this paper, as I usually do Labov's work, richly stimulating and filled with demonstration of the losses our two fields experience by failing to work more closely together.

General Discussion

Mancur Olson: Older blacks in the Philadelphia ghetto had speech that was less different from most American whites than did younger blacks in the same ghetto? I find it a most important and depressing finding that the older blacks, growing under a system of legal segregation, should in some respects diverge less from the rest of American society than teenagers growing up now in the big northern city ghetto.

William Labov: I am equally discouraged. All previous discussions of this issue have been in terms of the opposite tendency, "decreolization"—that at one time Black English was a Creole similar to the Caribbean Creoles and that it has gradually moved closer and closer to the other dialects. That holds true on many features of the language. But on two or three points the evidence is clear to me that we are getting divergence. Some are "camouflaged" features that sound like mainstream English but actually work in different ways and carry different meanings. There are also radically different developments in the tense and aspect systems. For example, in the black vernaculars now developing, there appear to be many sentence types that have no obligatory marker of tense to distinguish past or present. That is a radical change from mainstream English that demands a tense marker on every sentence.

Anthony Oberschall: I would like to make a comment on Grimshaw's statement that he doesn't find a cost-benefit analysis for the adoption of language very useful. I remember reading a remarkable book by Eugen Weber called *Peasants into Frenchmen.* He reports that as late as about 1860 something like 30 percent of France's population did not actually speak French. They spoke all kinds of other languages, like Breton, Catalan, Basque, etc. By about 1900 everybody in France spoke French.

Allen Grimshaw: What I said was that the development of language is not subject to a cost-benefit analysis.

William Labov: I agree with Grimshaw. You have described a situation where individuals change their linguistic allegiance. Susan Gal (1980) has done a detailed study of this in the relationship between Austrian German and Hungarian, and she has shown how the economic forces lead people to move in that direction. The symbolic capital that Bourdieu refers to can be quantified and dealt with in terms of economic advantage to the individual. But in terms of the driving forces that move a language consistently over thousands of years in a certain direction, this cannot be done. That remains a mystery for linguists and everybody else.

Ronald Burt: It seems there are two feasible ways in which relations might have an effect on linguistic behavior. Most of your metaphors are communication metaphors: two-step flow, socialization metaphors, the classic argument of clique formation, cohesive group. You get more contact, you bus people to various places, they talk to one another, they change. But when I look at your graphs of language variables by social position (age, race, socioeconomic status, and so on), I see evidence of a slightly different social process—the process of positional influence. The positional argument isn't "I do what I'm told or what I learned in communication," but rather "I do what's proper to my position."

William Labov: The question of position as opposed to contact is an important issue because there are societal norms for speech which are absorbed by people from the mass media. However, they do not have the direct effects on the linguistic system that contact with the next highest or next lowest status group does. In cases of status incongruence, they produce profound linguistic insecurity. In my study of the Lower East Side of New York City, a plumber who had a very high income and had moved into a middle-class project showed the highest degree of linguistic insecurity in speech. It's true that the transmission of linguistic traits demands first of all the importance of a high frequency of contact but it has to be contact that is structurally meaningful. Second, there must be some recognition that desirable traits are associated with that other group, which, as Grimshaw pointed out to me, may include unconscious recognition. Third, the group must somehow register the belief that the acquisition of those traits would make a difference in their life chances. Those beliefs are not anything you can test by direct inquiry. But that's the model we've been working with up to now.

Allen Grimshaw: Let me add that there is a close association of class aspirations and hopes for mobility with patterns called hypercorrection. The people who have aspirations for mobility, whatever they may be, are more likely than the people in higher prestige groups to speak in what they perceive to be the proper way.

Randall Collins: I want to point out something which I think is theoretically significant about what Labov and Grimshaw are saying. Insofar as we have anything like a theory of ethnicity, it's been a history of theories of assimilation. How does ethnicity come about and go away? Labov is showing us some structural features which may be producing that phenomenon we call ethnicity or some kind of status group structure of long enduring force. In a certain sense, ethnicity is constantly being changed or being created as well as disappearing.

William Labov: The dimension of ethnicity in American society is only a minor factor in all our linguistic studies. Almost all linguistic traits pass freely across the boundaries between Italian, Greeks, Irish, Ukrainians, and Poles in Philadelphia, in New York, or in Australia. But ethnicity as defined by racial barriers in the United States shows the opposite pattern; most linguistic changes stop short at racial boundaries. We always have a problem with European scholars who say, you don't mean by ethnicity what we mean by ethnicity, because ethnicity has been skewed by interaction with the color situation in the United States.

Social Psychology

Modeling Symbolic Interaction

DAVID R. HEISE

Symbolic interactionism (e.g., McCall & Simmons 1978; Stryker 1980) provides insights into how humans understand social events. Definitions of situations categorize settings and the people in them, narrowing a person's attention to a constrained range of phenomena, a restricted set of identities and objects that guide understanding and anticipation of social events. Events are created and interpreted to confirm the situational meanings provided by the definition of the situation. People create events to establish identities, to maintain identities, and to restore damaged identities. Expectations for another's behavior reflect his or her

Note: Returning to the University of Chicago twenty years after graduate studies gives me an opportunity to thank my professors, several of whom are attending this conference. I studied with Salvatore Maddi, Elihu Katz, James Davis, Mayer Zald, Peter Blau, Edward Shils, and they have my gratitude for intellectual guidance. Above all, I want to thank Fred Strodtbeck for the support he gave me and acknowledge how influenced I was by the loving presentations of social psychological theories which he gave in his role of teacher.

identity, and if an observed event does not confirm an identity, then the person may be reinterpreted, labeled, so that the event does confirm a new identification.

These notions provide a general framework for understanding the generation of sociocultural knowledge, but there are problems to solve. What aspects of social classifications are confirmed or disconfirmed by reality? What processes are involved in confirmations? How are discrete, qualitative social classifications abstracted and used to generate information beyond what is available in the classifications themselves?

Affect Control Theory

Affect control theory (Heise 1977, 1978, 1979) provides a model of social relations developed from symbolic interactionism and the premise that social events are constructed and reconstructed so as to confirm the meanings of social classifications. The theory proposes that affective associations are crucial dimensions of meaning, that affective dynamics underlie the confirmation-disconfirmation process, and that people translate back and forth between qualitative classifications of the perceived world and a fluid domain of affectivity where creativity functions.

Sentiments

Osgood (1962) reviewed a variety of evidence demonstrating that three basic dimensions are involved in affective response. The Evaluation, or attitudinal, dimension—measured by semantic differential ratings of good versus bad, nice versus awful—represents the pleasant-unpleasant aspect of affect. The Potency dimension—measured by contrasts of powerful-powerless, big-little—represents the sense of ascendancy versus vulnerability in affect. The Activity dimension—with rating scales like fast-slow, lively-quiet—represents affective arousal. Extensive cross-cultural research (Osgood, May, & Miron 1975) showed that these three dimensions of response—EPA—are universal, present in people around the world, regardless of culture or language.

Every social classification—setting, identity, trait, mood, status characteristic, behavior, etc.—carries a sentiment that can be assessed by asking what the classification connotes in terms of goodness, powerfulness, liveliness. Sentiment implies a fundamental feeling about something, not just a fleeting impression, so sentiments are estimated by averaging across repeated measurements. When focusing on a homogeneous cultural or subcultural group, ratings of a social classification can be averaged over the focal population in order to measure social sentiments.

Semantic differential scales for these three dimensions have been used to measure sentiments for thousands of social classifications. At present, large EPA (Evaluation, Potency, Activity) dictionaries are available for three populations—undergraduates in the U.S. South (Smith-Lovin & Heise 1983), Canadian undergraduates (MacKinnon 1983), and working-class Catholic youths in Belfast, Northern Ireland (Smith-Lovin, Heise, & Willigan 1983). The American dictionaries contain EPA profiles for 765 social identities, 440 attributes, 600 behaviors, and 345 settings. The Canadian and Irish dictionaries contain EPA profiles for about 600 social identities and 600 social behaviors. Table 1 shows some settings, identities, traits, and behaviors for which sentiments vary widely among U.S. undergraduates. Numerical values roughly correspond to assumed-interval codings of -4 to $+4$ with good, powerful, lively on the positive side (the actual metric is somewhat more refined, based on successive-intervals scaling—Heise 1978).

Table 1 suggests possible gender subcultures among the American undergraduates. A "disciplinarian" is viewed less favorably by males than by females; "imprison" is more lively for females than males; "feminine" is potent for females, not for males; "industrious" is more lively for males than females; "abandon" is more lively for females than males. (In Table 1, gender differences of 0.8 are significant at the .05 level in a two-tail test.)

Transient Feelings

How can social events confirm or disconfirm sentiments? Affect control theory proposes that fundamental sentiments are compared with transient feelings that result from events in order to assess whether the event was confirming or not. These outcome feelings—varying along the same EPA dimensions as sentiments—provide a basis for appraising confirmation-disconfirmation.

A confirming event is one that generates transient feelings that nearly match fundamental sentiments. "A son and a grandparent were together in a village, and the son thrilled the grandparent" is an example. Southern male undergraduates who were presented with this event description said the Son was good, very slightly powerful, and lively, and this is close to their fundamental sentiment about a Son—the way they rated Son outside of the context of any event, as shown in Table 2. The rating of the behavior, Thrill, in the context of the event was fairly similar to the rating of "to Thrill someone" presented in isolation. Respondents' feeling about the Village where the event happened also was similar to rating "a Village" out of context. Respondents rated the Grandparent somewhat dif-

Table 1 Sentiments for Some Social Classifications

	Male			Female		
	E	*P*	*A*	*E*	*P*	*A*
Settings						
Ballgame	1.97	1.46	1.65	1.72	1.53	2.00
Drive-in movie	0.83	−0.38	1.26	0.79	0.15	1.25
Mob	−1.46	2.23	1.88	−1.23	1.86	2.01
Gay Bar	−0.90	−0.66	1.38	−0.80	−0.28	0.97
Identities						
Athlete	1.33	1.82	2.13	1.00	1.78	1.48
Child	1.42	−1.48	2.31	1.94	−1.10	2.52
Disciplinarian	−0.62	1.02	−0.93	0.25	1.56	−0.23
Slut	−1.81	−1.18	1.81	−2.03	−0.53	1.18
Traits						
Industrious	1.98	1.74	1.89	1.54	2.09	0.96
Feminine	1.29	−0.29	0.70	1.81	0.81	0.28
Ruthless	−1.75	1.00	1.40	−2.26	1.40	1.25
Rude	−2.08	−1.09	1.17	−1.46	−0.49	1.02
Behaviors						
Thrill	1.68	1.51	1.75	1.37	1.25	1.45
Agree with	0.98	0.38	0.02	0.55	0.34	−0.58
Imprison	−1.46	1.64	−0.53	−1.12	1.50	0.70
Abandon	−2.55	−0.78	0.94	−2.60	0.54	0.32

ferently from the way they rated Grandparent out of context, but not differently enough to make the event disconfirming.

"A coward and a roughneck were together in a fight, and the coward soothed the roughneck" is a disconfirming event. This Coward impresses respondents as being neither good nor bad, slightly weak, and slightly lively—quite different than the out-of-context rating of a Coward, as shown in the bottom part of Table 2. The event also creates incongruous feelings about the Roughneck, who becomes excessively nice, too weak, and too quiet. The event somewhat disturbs feelings about Soothing.

Table 2 Affective Dynamics in Two Events

		Fundamental Sentiment	*Transient Feeling*	*Difference*	*Sum of Squared Differences*
		Confirming Event			
Son	E	1.67	1.17	−0.50	
	P	0.08	0.37	0.29	
	A	1.76	1.40	−0.36	0.46
Thrill	E	1.68	0.82	−0.86	
	P	1.51	1.17	−0.34	
	A	1.75	1.46	−0.29	0.94
Grandparent	E	2.24	1.32	−0.92	
	P	0.71	−0.01	−0.72	
	A	−1.77	−1.22	0.55	1.67
Village	E	1.19	0.81	−0.38	
	P	−0.35	−0.07	0.28	
	A	−0.67	−0.39	0.28	0.30
Event Total					3.37
		Disconfirming Event			
Coward	E	−1.45	0.10	1.55	
	P	−2.24	−0.64	1.60	
	A	−0.51	0.42	0.93	5.83
Soothe	E	1.73	0.96	−0.77	
	P	1.09	0.73	−0.36	
	A	−1.26	−0.38	0.88	1.50
Roughneck	E	−1.83	−0.61	1.22	
	P	1.25	0.57	−0.68	
	A	2.09	0.63	−1.46	4.08
Fight	E	−1.00	−0.77	0.23	
	P	0.66	1.17	0.51	
	A	1.82	1.62	−0.20	0.35
Event Total					11.76

Only the setting, a Fight, is undisturbed. These deflections of feeling create nonconfirmation.

We can quantitatively measure how much the disconfirming event deflects transient feelings (E′P′A′) away from sentiments (EPA) by computing a sum of squared differences. Table 2 shows that the total of 11.76 deflection units far exceeds the equivalent measure computed for the confirming event above: 3.37 units of deflection.

Affective Dynamics

What processes allow people to generate social knowledge which they have not learned explicitly? Affect control theory proposes that the creative aspect of social process is based on affective dynamics.

Events are perceived in terms of social classifications which evoke affective associations. An event may produce emotions that signal how the transient feelings compare with fundamental sentiments. Continuing waves of affective dynamics occur as the confirmation principle operates on both fundamental sentiments and transient feelings. Outcomes are affective determinations that can be characterized as EPA profiles, like the profile representing an appropriate behavior in the given circumstances.

A formal model to represent these processes has been developed, based on empirically derived impression-formation equations.

REACTIONS. A simple social event can be viewed as consisting of an actor, a behavior, and an object person, and the affective impact of an event can be predicted from current feelings about these elements. For example, for a person currently feeling that a Mother and Child are quite good and Hurting is quite bad, it is predictable that the Mother Hurting her Child would be viewed as bad.

Predictions are obtained with equations which accept quantitative measurements of pre-existing feelings and yield quantitative descriptions of outcome feelings. The equations are derived from data like those in Table 2—in-context and out-of-context ratings for the elements in a sample of events. Structural equations for predicting outcomes may be estimated by regressing in-context measurements on out-of-context measurements. Interaction effects turn out to be important, so products of out-of-context variables also are included as predictors. This procedure produces equations like the following.

$$A'_e = -.43 + .39A_e + .48B_e + .15B_eO_e$$

This says that the perceived goodness of the actor in an event (A') is a function of the initial attitude toward the actor (A), attitude toward the behavior (B), and a balance effect (BO) described by multiplying goodness of act with goodness of object (for example, bad onto bad is good). This particular equation accounts for 86 percent of the variance in outcome impressions of goodness-badness (Gollob 1968).

Equations have been obtained for predicting how a given event changes prior feelings of goodness, power, and liveliness into outcome feelings along the same dimensions for actors, behaviors, objects, and settings (Heise 1969, 1970, 1978; Heise & Smith-Lovin 1981; Smith-Lovin & Heise 1982, 1983).[1]

PROACTIONS. Proaction equations are derived from the reaction equations under the assumption that people construct events that will produce impressions which confirm fundamental sentiments about event elements. For example, a normal mother would not hurt her child because the act disconfirms the fundamental goodness of Mother and Child; instead she would choose an act like Assist which reinforces the meanings of Mother and Child and Assisting.

The analytic problem is to use the reaction equations as a basis for predicting new behavior that best confirms sentiments, given current feelings about the actor and object. The pre-event sentiments and feelings toward the actor and object are taken as givens. The unknown behavior should minimize deflections of post-event transient feelings away from fundamental sentiments. Algebraic expressions are set up to represent the differences between future outcome feelings and fundamental sentiments, employing the reaction equations as predictors of outcomes, so only pre-event quantities are involved. The expressions are differentiated to find their minima, resulting in proaction equations that define the EPA profiles for the required behavior in terms of pre-event feelings and sentiments about the given actor and object.

The minimization logic also deals with another event construction problem—redefining people who are involved in events that do not confirm their identities. For example, a mother is not expected to hurt

[1] Current equations based on 515 event descriptions are more complex than the one above, with more interaction terms and cross-dimensional effects (e.g., a lively behavior reduces actor goodness). Newly developed statistical procedures (Smith-Lovin & Heise 1983, app. D) were used to adjust variances and covariances of predictor variables in order to minimize estimation bias due to measurement errors, and a maximum likelihood algorithm was used to obtain parameter estimates.

her child, so a woman who does hurt her child might be labled with a negative identity. In this case the analytic problem is to define the kind of actor who would be confirmed by a given event. Applying the same procedures as before results in event reconstruction equations that define the EPA sentiment profile for the required actor in terms of sentiments about the given behavior and object. Examples later will show some of the effects implied by the equations.

AMALGAMATIONS. Translation between social classifications and affective representations might be more complicated than considered so far. A definition of the situation might assign a person an identity (e.g., Doctor) along with a status characteristic (the Old Doctor), a personality trait (the Extroverted Doctor), or a mood (the Hostile Doctor). Redefining a situation after an unexpected event might involve assigning a new identity, or the original identity might be modified by a status characteristic, trait, or mood to understand why the event occurred.

Structural equations can be obtained to describe how identities and modifiers amalgamate. In an exploratory study, Averett (1981) obtained EPA ratings for identities and modifiers out-of-context and for modifier-identity combinations. Then equations were derived for predicting the combination ratings from the out-of-context ratings. The following, for example, is the equation for the evaluation outcome.

$$A_e = -.30 + .61M_e - .13M_p - .17M_a + .50I_e - .01I_p - .05I_a$$

This says that evaluation of an amalgamation (A_e) is mainly a function of modifier evaluation (M_e) and identity evaluation (I_e), but modifier potency and activity also contribute. Averett showed that structural equations for amalgamations are similar whether the modifier is a status characteristic, trait, or emotion word (the above equation was obtained by combining all three kinds of data). She also demonstrated that amalgamations operate in reaction equations exactly as do singular identities.

The modifier-identity equations describe how modifier and identity sentiments combine into a melded sentiment. An inverse process is involved in selecting modifiers for partial reidentifications. Analytically, the problem is as follows. An identity profile is given by the original definition of the situation, a computed profile which would explain a given event is provided by proaction equations, and the modifier-identity equations must be solved for the modifier profile which would combine with the original identity profile to produce an amalgamate equivalent to the computed profile. This is the equation for the required modifier evaluation.

$$M_e = .55 + 1.72A_e + .32A_p + .36A_a - .87I_e - .18I_p - .16I_a$$

For example, an observer might want to modify the identity of a son (EPA: 1.7 0.1 1.8) who disobeys his father. The appropriate profile for someone who disobeys a father is computed from proaction equations with actor as the unknown, giving −1.1 0.8 1.0—this is the required profile for an amalgamation of Son and some modifier to account for the event. Substituting the given profiles into the last equation, we find that the adjective should have an evaluation of −2.6.

Affect control theory hypothesizes that emotions are sensations of congruency and noncongruency between one's identity and the self-impressions produced by recent events. When people talk about a discrepancy between a person's self-sentiment and transient feelings, they do so in terms of emotion labels. The inverted amalgamation equations serve a second function in describing emotion attribution processes. They allow us to compute the EPA profile for an emotion word ($M_eM_pM_a$) that could be combined with the identity profile ($I_eI_pI_a$) in order to describe a current transient feeling ($A_eA_pA_a$).

EXAMPLE. The following example illustrates how the various equations in affect control theory can be used to model social interaction.

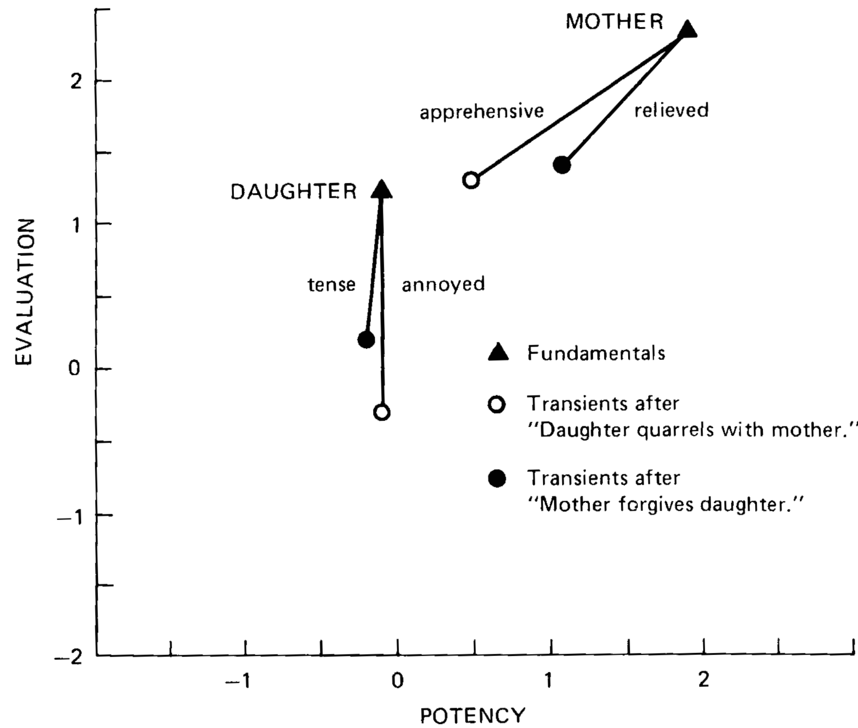
The example focuses on a mother and daughter and on how the mother might respond to an unexpected behavior from the daughter. The entire analysis was conducted using EPA dictionaries from southern U.S. female undergraduates; EPA profiles are given in parentheses. Affective dynamics for the event sequence are graphed in Figure 1. The graph shows evaluation and potency changes; activity dynamics are relatively unimportant in this interaction.

A woman, taking the role of Mother (fundamental: 2.3 1.9 0.0), is with her Daughter (1.2 −0.1 0.9).

To begin, we initialize transient feelings equal to fundamentals. A routine analysis would solve the proaction equations for behavior profiles and look these up in an EPA dictionary in order to define the kinds of acts that are sentiment-confirming in the situation. Such an analysis indicates that a mother normally might engage in nice, strong acts (2.0 1.7 0.1) toward a daughter—behaviors like Assist, Educate, Support, Aid, Cheer, Interest, Guide, Encourage. A daughter theoretically would engage in nice, lively acts (1.3 0.1 1.4) toward a mother—Rally, Amaze, Speak to, Idealize, Alert, Astonish, Appease, Exalt.

Suppose, however, that the daughter is in a bad mood, and she Quarrels with (−1.1 0.0 1.3) her mother. Applying the reaction equations, we find that the daughter's negative act reduces the mother's transient

Figure 1 Affective Dynamics for an Event Sequence



goodness and potency (to 1.3 0.5 −0.1). The inverted amalgamation equations are applied to find the profile of an emotion word that can combine with Mother and produce the transient profile. The result (0.5 − 1.0 0.0) corresponds to an emotion attribution of Apprehensive for the mother. From the mother's standpoint, the behavior also disturbs the daughter's transient (to −0.3 −0.1 1.2), and an emotion analysis suggests that the mother would suppose the daughter feels Annoyed.

The mother might try to understand the event better by reconceptualizing her daughter. Since the event is only moderately disconfirming (5.7 units of deflection), the mother probably would not seek a full identity shift but rather a modification of the daughter identity. First we solve the proaction equations to get the profile for the kind of actor who would quarrel with a mother (− 1.1 0.0 1.3). Then we apply the inverted amalgamation equations to get the profile (−.2 −0.1 0.5) for a modifier that could combine with Daughter, yielding the reidentification profile as an amalgamate. The modifier profile is looked up in a dictionary of modifiers, and we find that the mother might consider her daughter Inconsiderate. This redefinition of the daughter would reduce deflection units to 3.1.

Alternatively, the mother could try to recover through action. To predict what the mother might do, we apply the proaction equations to the original sentiments and to the mother's and daughter's transients after the first event, solving for a behavior profile. The result (2.1 2.3 −1.4) corresponds to acts like Calm, Soothe, Pray for, Forgive. If the mother were to forgive the daughter, she would improve her transient (1.4 1.1 0.0), and an emotion analysis indicates that the mother then would feel Relieved. The mother also would see the daughter's transient as being somewhat improved (0.2 −0.2 0.9), although the mother would suppose that the daughter still feels Tense.

OPERATIONALIZATION. The affect-control model operationalizes situational logic by requiring categorizations of relationships and participants when the situation is defined. Currently, relationships are categorized as Verbal or Physical (to deal with touch-behaviors) and also as Primary (admire, hate), Economic (pay, hire), Managing (command), Fixing (heal), and Training (instruct). Identities are categorized as Male or Female, and also as Casual (friend, dope), Ascribed (adolescent), Legal (judge, criminal), Trade (doctor, prostitute), and Sociosex (wife, adulterer). All words in the behavior and identity dictionaries have been assigned binary codes corresponding to these categories, and a word is retrieved for event construction only if it fits the configuration that was given as an input.

For example, "medicate" is coded 0100010—not verbal, physical, not primary, not economic, not managing, fixing, not training—and this behavior can be retrieved only in relationships that are characterized as physical and fixing. "Thief" is coded 1110110—male, or female, casual, not ascribed, legal, trade, not sociosex; thus, the identity is recalled in reidentification analyses only when the person initially was characterized as either male or female, and able to take on casual, legal, or trade roles in the situation.

The categorization systems are provisional. Future research will seek a more refined system for classifying relationships, following empirical leads of Triandis (1977) and Marwell and Hage (1970) and the theoretical lead of Parsons' pattern variables (1951). Taxonomic methods (Werner and Fenton 1970; Spradley 1979; Heise 1982) may be fruitful for producing a more refined system of componential analysis.

Evaluation of the Theory

Computer Simulations

Computer simulations to expedite computations and dictionary searches have been part of affect control theory since the theory was first formulated in the early 1970s. Simulation analyses in affect control theory receive natural language specifications of a social situation and produce natural language predictions concerning the behaviors and reconceptualizations that might occur. Three analyses in Table 3 illustrate computer simulation capabilities as of 1983. The Doctor-Patient analysis deals with an occupational role. Mugger-Victim analyses involve a deviant role. The second Mugger-Victim analysis indicates how a particular setting might influence behavior.

The Doctor-Patient analysis is set up with a male seeing himself as Doctor. (External inputs are italicized in the table.) The program retrieves the male EPA profile for Doctor, 1.8 2.1 −0.3.

The doctor sees his interaction partner as a Patient, male EPA: 0.1 −1.7 −0.8. The doctor's relationship to the patient is defined as verbal, physical, managing, fixing, so only behaviors of these types will be retrieved in defining his conduct. The second person is a female who sees herself as Patient and the man as Doctor. The program retrieves the female EPA for patient, 0.2 −1.6 −1.0, to represent her sentiment toward self, and the female EPA for Doctor, 1.7 1.8 −0.3, to represent her sentiment toward the doctor. The patient's conduct is limited by a definition of her relationship to the doctor as verbal, physical, primary, exchange.

The program computes an EPA profile defining behavior for the doctor that would best confirm his sentiments toward self and other, 1.4 1.1 −1.0. A dictionary search indicates that this profile fits Considering, Soothing, Cautioning, Calming, Listening to, Esteeming, Counseling, Medicating. The program also computes the profile for sentiment-confirming behavior of the patient toward the doctor, 0.4 −1.1 −0.3, and behaviors fitting this profile are Worshipping, Revering, Submitting to, Idolizing.

Doctor soothing the patient is treated as an event that occurs. This generates 1.2 units of deflection for the doctor, 1.0 units for the patient, the two being slightly different because of differences in male and female EPA profiles. The doctor's transient feelings after the event theoretically can be characterized by emotion labels fitting the profile 1.8 0.5 −0.6: Sympathetic, Peaceful, Satisfied, Relieved, Affectionate, Grateful, Warm, Kind. The doctor's transient feelings toward the patient are interpreted

Table 3 Simulations Based on Affect Control Theory

Example 1
Doctor sees self as:

doctor	1.8	2.1	−0.3

In general, Doctor can take roles which are:
male, trade

Doctor sees Patient as:

patient	0.1	−1.7	−0.8

Doctor sees the relationship as:
verbal, physical, managing, fixing

Patient sees self as:

patient	0.2	−1.6	−1.0

In general, Patient can take roles which are:
female, casual

Patient sees Doctor as:

doctor	1.7	1.8	−0.3

Patient sees the relationship as:
verbal, physical, primary, exchange

1: Doctor might [1.4 1.1 −1.0] consider Patient (or soothe, caution, calm, listen to, esteem, counsel, medicate her).
2: Patient might [0.4 −1.1 −0.3] worship Doctor (or revere, submit to, idolize him).

Assume Doctor does *soothe* Patient and the setting is unnamed. Then:

Doctor feels [Deflection units = 1.2] [1.8 0.5 −0.6] sympathetic, peaceful, satisfied, relieved, affectionate, grateful, warm, kind.
Doctor supposes Patient feels [1.2 0.2 −0.1] grateful, relieved, sympathetic, hopeful, satisfied, sentimental, affectionate, peaceful.
Patient supposes Doctor feels [2.0 0.7 −0.5] grateful, sympathetic, kind, compassionate, satisfied, serene, warm, peaceful.
Patient feels [1.0] [1.2 0.3 −0.1] relieved, grateful, sympathetic, sentimental, serene, pleased, satisfied, warm.

1: Doctor might [1.2 1.7 −1.0] counsel Patient (or medicate, console, explain, calm, consider, instruct, reassure her.)
2: Patient might [.4 −1.6 −0.2] [undefined] Doctor.

Example 2
Mugger sees self as:

mugger	−2.9	0.4	1.8

Table 3 (Continued)

In general, Mugger can take roles which are:
male, casual, legal, trade
Mugger sees Victim as:

victim	−0.3	−1.8	−0.4

Mugger sees the relationship as:
verbal, physical, primary, exchange, managing
Victim sees self as:

victim	0.1	−1.8	−0.6

In general, Victim can take roles which are:
female, casual, legal
Victim sees Mugger as:

mugger	−2.1	1.0	1.5

Victim sees the relationship as:
verbal, exchange

1: Mugger might [−2.0 0.0 2.4] haze Victim (or molest, bully, insult, taunt, curse, jeer, laugh at her).
2: Victim might [−0.2 −0.3 −0.7] doubt Mugger (or disbelieve, placate, sweet-talk, implore, beg, evaluate, patronize him).

Assume Mugger does *bully* Victim and the setting is unnamed. Then:

Mugger feels [0.6] [−0.6 1.0 0.8] possessive, belligerent, furious, lustful, angry, cocky, defiant, aggressive.
Mugger supposes Victim feels [0.3 −0.4 0.0] shocked, snowed, spacey, anxious, surprised, grateful, relieved, blue.
Victim supposes Mugger feels [−0.8 0.5 0.7] belligerent, possessive, annoyed, contemptuous, furious, vengeful, angry, coy.
Victim feels [0.3] [.03 −0.5 −0.0] tense, apprehensive, sorry, blue, snowed, apathetic, dissatisfied, regretful.

1: Mugger might [−2.0 −0.2 2.4] molest Victim (or taunt, laugh at, haze, jeer, curse, doublecross, aggravate her).
2: Victim might [−0.2 −0.3 −0.9] placate Mugger (or disbelieve, doubt, beg, evaluate, patronize him).

Example 3
The setting is: *mob*

Mugger EPA:	−1.5	2.2	1.9
Victim EPA:	−1.2	1.8	2.0

Mugger sees self as:

mugger	−2.9	0.4	1.8

Table 3 (Continued)

In general, Mugger can take roles which are:
male, casual, legal, trade
Mugger sees Victim as:

victim	−0.3	−1.8	−0.4

Mugger sees the relationship as:
verbal, physical, primary, exchange, managing
Victim sees self as:

victim	0.1	−1.8	−0.6

In general, Victim can take roles which are:
female, casual
Victim sees Mugger as:

mugger	−2.1	1.0	1.5

Victim sees the relationship as:
verbal, exchange

1: Mugger might [−1.8 0.0 0.1] neglect Victim (or spoil, shame, haunt, undermine, discredit, mislead, forsake her).
2: Victim might [−0.2 0.1 −0.3] disbelieve Mugger (or implore, doubt, evaluate, sweet-talk, coax, patronize, refuse him).

Note: Inputs are italicized; triplets of numbers are EPA profiles.

by supposing that he attributes consequent emotions to her which fit the profile 1.2 0.2 −0.1: Grateful, Relieved, Sympathetic, Hopeful, Satisfied, Sentimental, Affectionate, Peaceful. The analysis indicates that the patient arrives at different emotional interpretations than the doctor because of gender differences in sentiments, though the results are similar in this case because the sentiments are not very different.

The program then derives the sentiment-confirming events that might occur next, given that transients were changed by the doctor soothing the patient. The doctor now would have to act less nice, more powerful, more quiet, and verbal implementations of the EPA profile are more instrumental than they were on the first round. The patient would have to act weaker than she might have done on the first round, so much so that no entries in the behavior dictionary provide a reasonably close implementation for her affective determination.

The next analysis involves a male Mugger with a female Victim. The mugger-victim relationship is defined as verbal, physical, primary, exchange, managing. The victim to mugger relationship is defined as verbal and exchange. Computations indicate that the mugger's sentiment-confirming act would be bad and lively: Hazing, Molesting, Bullying, Insulting, Taunting, Cursing, Jeering. The victim could confirm her definition of the situation by quiet acts toward the mugger; words retrieved are Doubt, Disbelieve, Placate, Sweet-talk, Implore, Beg, Evaluate, Patronize. Assuming that the mugger bullies the victim, he might feel Possessive, Belligerent, Furious, Lustful, while the victim feels Tense, Apprehensive, Sorry, Blue. Having bullied, he might act slightly weaker, and the victim's behavior options on the second round are slightly quieter than they were on the first round.

The last example in Table 3 presents the mugger and victim again, but this time in the specified context of a Mob. The analysis suggests that the mob setting might substantially reduce the activity of the mugger's behavior, to the point where he might Neglect or Forsake the victim. The explanation is that the mugger has to maintain the meaning of Mugger, Victim, and Mob all together, with Mob contributing to impressions produced by any event he creates. His actions are muted as a net result.

The simulations indicate that affect control theory can account for normal role behavior and for creative responses to deviant acts—either sanctioning or labeling. Simulated emotions are reasonable enough to encourage the idea that emotions are subjective signals about how events have created deflections from fundamental sentiments. Simulations with specified settings are provocative enough to hypothesize that settings integrate directly into event processing, beyond limiting how a situation is defined.

Empirical Studies

Wiggins (1980) tested the accuracy of some predictions from affect control theory by presenting college student respondents with two sets of twelve scenarios like the following. The first example focuses on behavior intentions; the second example focuses on behavior expectations.

> You (a student) are in a room with a doctor and a nurse. The doctor compliments and encourages you and then leaves. How likely is it that you would do each of the following things to the nurse?

> You (a student) are in a room with a doctor and a nurse. The doctor compliments and encourages the nurse and then leaves. How likely is it that the nurse would do each of the following things to you?

Different scenarios were created by choosing evaluatively positive or negative identities for the two nonrespondent characters (e.g., Doctor and Nurse), and by making the priming actions (e.g., compliment and encourage) evaluatively positive, negative, or neutral. Behaviors presented for likelihood ratings were results from simulations for all scenarios.

Wiggins correlated the respondents' likelihood ratings with the total amount of deflection produced by each behavior option in each scenario. In the case of intentions, the correlation was $-.50$ for male respondents and $-.36$ for female respondents (negative correlations indicating that smaller deflections go with higher likelihoods, as predicted). For behavior expectations, the correlations were $-.71$ for males and $-.67$ for females. Thus, respondents' indications of what they expect in interaction sequences correspond substantially with the deflection principle in affect control theory.

Wiggins carried out an experiment to ascertain whether affect control theory predicts actual behavior. A nonintuitive prediction from the theory was the focus: If a person with a positive self-identification experiences a negative event, then he or she should act unusually pleasant toward a valued interaction partner, but unpleasant toward a disvalued interaction partner. Behaving positively toward a valued other offers a way of recovering one's spoiled identity; avoiding positive behavior toward a disvalued other keeps one's predicament from getting worse.

This was tested by setting up a pseudo-study of interpersonal communication with university students as subjects. An actress playing a secretary appreciated or humiliated subjects while they were in the presence of their communications partner—another actor who was identified as a positively valued other (university student) or a negatively valued other (delinquent). The secretary left the room after her manipulation action, and the communications partner then elicited behavior from the subject by asking "What time is it?" and "What do we do now?" Subjects' behaviors before and after the appreciation-humiliation manipulation were videotaped and judged later by other students as to pleasantness or unpleasantness. Variations in subjects' behaviors corresponded to the predicted pattern. In particular, humiliated subjects acted friendly and ingratiating toward another student but cool and unfriendly toward a delinquent.

Thus, available empirical tests support ideas in affect control theory. Of course, the theory applies to so many different kinds of social phenomena that it will take years of empirical testing to get a full picture of where the theory's strengths and weaknesses are.

Figure 2 Levels of Analysis in Affect Control Theory

Quantitative Model	Numerical values for interactants and setting on standard variables	→ Numerical computations for people's responses to events and for construction of new, confirmatory events	Numerical values defining feelings, behaviors, new identities
	↑ Translation by dictionaries		↓ Translation by dictionaries
Qualitative Model	Definitions of situation in terms of standard classifications	→ Imposition of cognitive constraints to maintain logical consistency between new productions and original definitions	Verbal specifications of emotions, events, reidentifications
	↑ Interpretation of analyst or informant		↓ Evaluation: subjectively by analyst or informant; or by observations, experiments
Phenomena of Interest	People's perceptions of social scenes		People's expectations, intentions, attributions, labelings

Summary

Levels of analysis in affect control theory are diagrammed in Figure 2. People's perceptions of social scenes are the focus of analyses. These are incorporated into analyses by identifying classifications that participants themselves might make in defining a situation. The qualitative classifications are translated into quantitative values of goodness, powerfulness, and liveliness. A formal model is applied to input profiles in order to derive new profiles representing the outcomes of affective dynamics. The outcome profiles, translated back to qualitative classifications of people or behaviors define event options. Options are further narrowed on a qualitative basis to maintain logical consistency between predicted events and participants' definitions of the situation. Final results are natural language descriptions of events, reidentifications, and emotions that participants should expect at the scene. These can be evaluated by an analyst, by a member of the group being studied, or through observational or experimental studies.

The theory's quantitative model was developed by focusing on simple reaction processes. Measurements on the universal EPA dimensions of the semantic differential were made for systematic samples of events, and equations were derived by applying structural equation modeling procedures to the data, with detailed attention to measurement errors. Proaction equations were derived from the reaction equations through the use of theoretical assumptions and mathematical analysis.

The theory's qualitative model depends on large dictionaries of settings, identities, attributes, and behaviors. The dictionaries allow lay classifications to be used as input and output in analyses, fostering "common sense" evaluations of theoretical productions. Items in the dictionaries are scaled on the three semantic differential dimensions to permit translations between qualitative and quantitative variates. Additional semantic codings of words in the dictionaries focus results from the quantitative model and allow predictions to be constrained to what is possible in a given situation.

Affect control theory proposes that people operate in terms of cultural classifications. A definition of the situation assembles the classifications that are relevant at a particular scene and supplies the sentiments that are to be maintained. Laws of affective dynamics, operating with syntactic principles, extract the accumulated experience that has been condensed into sentiments and allow people to use this information creatively in response to novel circumstances. Distributional rules censor affective products and produce sensible, relevant outcomes. People do not learn which specific events are permissible and which are not. Rather they generate expectations as needed from the meanings of cultural classifications.

Comment

FRED L. STRODTBECK

I once heard Robert Frost describe a theory that he believed he followed in writing poetry. He'd reported that he would think up a good line, "Two paths diverged in a snowy wood." Then, this line, standing alone, controlled the lines that followed, subject to certain artistic constraints. My companion at that talk concurred that Frost may have written his poems in that way but he doubted that *Paradise Lost* was so composed.

My guess is that many of you have similar reactions to Heise's microsocial model. For this reason, a part of my task is to be sure that Heise's *medium* does not interfere with your understanding of Heise's *message*. To start historically, W. I. Thomas (1928), in his address as president of the American Sociological Society, stated:

> In approaching problems of behavior it is possible to emphasize—to have in focus of attention for working purposes—either the attitude, the value, or the situation. The attitude is the tendency to act, representing the drive, the affective state, the wishes. The value represents the object or goal desired, and the situation represents the configuration of the factors conditioning the behavior reaction.

He also commented:

> It is possible to work from the standpoint of adaptation. Any one of these standpoints will involve all the others, since they together constitute a process. [pp. 154–55]

During the intervening fifty years, the idea of feedback between elements in the attitude-behavior equation has persisted and been strengthened. The cognitive dissonance phase of social psychology (1957 to 1969) enabled the field to become more explicit about the conditions under which successful enactment changes beliefs in what are called counterattitudinal situations. Heise's responsiveness to this and related developments can be illustrated in his restatement (1979) of balance theory in terms of *actors* and *events:*

> First, events produced by good actors are felt to be somewhat nicer than those provided by bad actors. Second, a good act produces a positive impression and a bad act produces a negative impression, regardless of other considerations. Third (the balance principle), goodness comes from good acts on good objects or from bad acts on bad objects; negative impressions result from good acts on bad objects or bad acts on good objects. [p. 18]

The key to understanding what Heise has accomplished is the insight that he has combined the thought styles of Heider and Thomas. It is error to focus on the fact that Heise's system enables him to produce quantitative predictions of what will result when a wide range of actors, acts, and situations are combined. The quantification he produces is wholely subordinate in importance to the fact that the quantification can be used for retrieval. I suggest and attempt to demonstrate that *retrieval* is the dynamic element in Heise's system.

Has Heise diminished free will, has he done away with the playfulness and planfulness of social action, or, more sociologically, has he broken with the voluntaristic social theory of Weber and Parsons? You must answer such questions for yourselves. There are, however, three points I'd like to make: (a) Rational or purposive action does not exclude the sentiment and emotion with which Heise works; (b) the complex world of semantic meaning that Heise addresses should not be depreciated by being thought of as a nonrational residue of economic action; and (c) if voluntarism is to be operationalized on the microsocial level, what better way to do this than by visualizing an actor selecting one of the eight alternatives provided by Heise's system?

Heise has accepted the guidance of social psychological theory (in addition to common sense) in the imposition of categories and cognitive constraints. He uses Osgood's semantic differential dimensions as a filing rationale, and then uses the filing dimensions to characterize his results. His challenge to other investigators is: Can you find a better way to characterize the affective connotation of words? And one of the reasons that the evaluation of Heise's system is difficult is that, at the moment at least, there are no visible competitors.

To clarify the way the filing and the constraints are related, let me illustrate: Three terms such as dog (1.1, 0.6, 1.1), heart (1.6, 0.3, 1.1), and courage (1.7, 0.4, 0.6) are, in Osgood's dictionaries, good, slightly potent, and lively. They are therefore located within a tiny sphere in semantic differential space. Heise has dropped excess terms and reorganized his dictionaries into four semantic differential spaces: 765 *identities* (kinds of people), 440 *attributes*, 600 *behaviors*, and 345 *settings*.

This means that when he searches at plus one, evaluation, plus a half, potency, and plus one, activity, in the attribute dictionary he's not going to encounter animals or bodily organs. He'll find courage and the seven other words most similar in connotation. There are also cross-cutting constraints involving sex of actor. Thus, to restate the earlier point, retrieval under constraint is the dynamic element in Heise's theory for, after retrieval, he postulates an act of selection, guided by the rule of eight, in order for the interaction to continue.

Without more firsthand experience, it's hard to assess the potentiality of the method in suggesting experiments and guiding the analysis of natural texts. Some have assumed that emotion is a further signal to help recognize good from bad intent on the part of powerful or powerless others who are either active or passive. From such a perspective, one may conclude that Heise's work is relevant to conflict theory—particularly conflict management. It is not inconceivable that extensions of his methodology might be used to blot out affect concomitants during the management of interpersonal distance, say in the successive acts of spouses. The feedback loops between original ratings and applications to natural behavior are robust and have the inherent flexibility to come alive in quite unexpected ways.

General Discussion

Ronald Burt: I have a problem with your procedure because of results on related items that a set of students in one of Columbia's summer research seminars found. The question of their research was the meaning of cognitive space. We exposed people to a variety of hypothetical situations, created as vignettes using, on a much narrower scale, the same sort of conditions you do. Respondents were also asked to interpret their especially close relations with real people, on the same dimensions used to interpret the vignettes. If you looked at the semantic or cognitive space for real people versus the semantic space of hypothetical relations, you found that movement through the space was different. I had the feeling that there was regional rationality in the meaning of relationships—there were places in the total space within which one logic worked but when you crossed boundaries, different logics were invoked. The interesting substantive questions are then: What are the logics within and across these spaces? What are the boundaries that create logics for interpreting relations within regions of the space? Without knowledge of the boundaries, one should find many interaction terms, so that one gets very complex models because one has not taken the segmentation in the cognitive space into account.

David Heise: Without going into more detail it sounds like it is indeed a question of different positioning of these persons in the space and, exactly as you say, the interpretations that would come out of this kind of modeling would be based on very elaborate interaction terms. Your point is that there is qualitative variation of rules in the space. And I think that's a very interesting idea to study empirically.

Siegwart Lindenberg: The computer has selected my reaction to your paper: hopeful, emotional, astonished, and puzzled. My question has to do with the stability of sentiments. In your model, "mother" is always good, "thief" is always bad, and so on. Why do you think these sentiments are so stable? I think these sentiments change under certain constraints. For instance, take parents with a child and change their social situation. The parents are becoming richer, so that they can afford to go out a lot, which increases their opportunity costs for "parenting." This will probably change the meaning of all sorts of roles, for both the parents and the child. In your model, the meaning is there and we adjust our behavior in such a way that it would reproduce the meaning that we thought was there. But when do people adjust meaning to fit behavior and when do they adjust behavior to fit a given meaning?

David Heise: At present there is no specific part of affect control theory that tells how these fundamental sentiments change. But there will be. The additional theory will allow us to say how events that have occurred in the past are incorporated to change meanings. Then "rich" might be combined with "parent." In the middle-class subculture in which these ratings are obtained, "rich" is not particularly good. It is extremely powerful and it is kind of quiet and the net impact of combining "rich" with "parent" would be to move the expected behaviors in those directions.

In the short term meaning or sentiment is the independent variable, the variable from which all of these analyses derive. But in the long term, events are ecologically or rationally created. An event can become an exemplar and define the proper sentiments toward particular roles. The meaning of the exemplar event becomes an independent variable creating responses to a great variety of circumstances, including sanctioning responses and interactions with multiple persons.

Harrison White: I want to make two practical suggestions. Retrieval under constraint is the core of what's going on. I would like to locate that in the space of Chomsky on the one hand and artificial intelligence on the other. My feeling is that the initial presupposition you started with is very Chomskian: There are so many possibilities that there has to be a code and I am going to recover that code. But the Chomsky code hasn't worked out that well. It has proved to be limiting in many ways. I think it would be good for you to move away from that position toward artificial intelligence, which to my amazement is starting to work now. They have found mainly two simple things. First, you have to differentiate more; three is not enough. You want a more refined set of nodes and maybe collapse later. Second, you want much more cycling. I'm not that overwhelmed by any one of your theses and I think you don't yet have enough independent parts, and you have not been cycling enough. If you look at good AI work, it is running a long time before anything comes out. You are still in a way doing on a computer what you could do by hand.

The second suggestion is more practical. I would love to see you do the following. You will find a dozen playwrights in any city. I would love to see you generate some play scripts with your model and then go to a playwright and have him make a script out of this, a play for six-year-old kids in grammar school. Use him as an expert, and then you look and see what are the few things he did to turn this rather bland piece into something that is, at least for sixth-graders, a script.

David Heise: The playwright idea is very intriguing. I hadn't thought of that. I'll have to move on something like that. On artificial intelligence I absolutely agree. That I had considered before. Unfortunately, I have computer problems even now, regarding memory and computation time. But I think you are absolutely right in the cycling idea.

William Labov: As you know, all linguists have trouble with the Osgood approach, which has to be integrated in some way with the semantic component. People don't look up words by their affective components: they have to locate them first by semantic and syntactic features, and then make finer choices by affective features. But there's also a social dimension to word use that you might want to take into account. I've done some work on the uses of the word "child"—not in the kinship sense, but in sentences like "you're only a child." It appears to have no cognitive meaning here. Dictionaries only define this sense of "child" in a circular manner: "behaving in a childish manner." "Child" is used to assign social rights and privileges: It covers a very different age range in "Children keep out" and "Not for children's use." I don't see how this social use of language would be integrated into your model of the social use of language. There's certainly a difference between social and emotional expression. Harvey Sacks's work on categorization devices is another example of the social determination of word selection. I'm curious as to how those social uses of language can be integrated with the purely emotional dimensions that you use.

David Heise: As a matter of fact the categorization device is used in the theory to explicate the process of defining a situation. We have to assign identities to people that make sense with one another. And that's essentially what a categorization device is. Presumably one could get very specific kinds of models to deal with that and I feel confident they will in fact come from ethnomethodology.

James Coleman: I want to return to the theoretical paradigm that underlies this. The assumption is that people automatically behave so as to obtain optimum confirmation within limits imposed by the situation of recent happenings. In view of that I want to pay tribute to a man named Lecky who wrote a little book called *Self Consistency.* Lecky's notion is similar to that expressed in affect control theory. One of the examples he used was a mother who gives a child castor oil in orange juice so that the child will come not to dislike castor oil, or at least drink it. Instead what happens is that the child comes to dislike orange juice. The event can generate either of two reactions. Similarly with events which may lead to change in self-conception. What are the kinds of conditions under which events will force changes in self-conception? Even if there is this great tendency to behave in such a way as to confirm one's conception of one's self, that conception changes. What are the conditions under which self-conception undergoes change?

David Heise: This theory is not a theory of self, interestingly enough. This is a theory in which people are taking on continuously different situated identities. Within a situation, an identity could last a long period of time. For example, one could be a professor for decades and take on that identity within a certain context. And the meaning of "professor" as applied to oneself could change. If students kept falling asleep in one's class year after year, you might finally end up saying, " I must be a boring professor." That kind of reinterpretation is very much a part of the model.

James Coleman: Let me add one point. If this notion of affect control is as powerful as you suggest, it might well be that some of the things that Labov was investigating, for instance the difference between formal and casual language in the black community, could be generated by one's desire to confirm one's self-conception.

David Heise: It's an intriguing idea that one might code those variations and put them in a dictionary, find out what their profiles are, and then use that dictionary to analyze someone who is talking casually. A casual speaker then would be put into the model, and that should select different actions.

Social Movements

A Theory of Social Movements, Social Classes, and Castes

MANCUR OLSON

The Need to Construct a Theory with Elemental Entities

The behavior of most individuals is, of course, influenced to some degree by the social groups to which they belong. The information and beliefs to which children are exposed obviously vary from one social group to another. Since it usually costs something to acquire new information, the theory of my discipline of economics leads one to predict that the behavior of individuals will sometimes be influenced by the beliefs and perceptions they inevitably absorb in the social groups in which they happen to have been raised. For the sociologist such a prediction must be utterly banal, for the influence of socialization upon behavior is obviously one of the central concerns of sociology.

Important as the influence of social groups on individual behavior may be, it is clearly not enough to study this influence and how it varies from group to group. The social groups themselves need to be explained. Any logically complete social theory must explain why social groups exist

Note: Material quoted from *The Rise and Decline of Nations* (1982) is used with the permission of the publisher.

and why they may differ from one society or circumstance to another. The existence of some elemental social groupings, such as that of a mother and her infant, might be explained on biological or genetic grounds, and some social scientists might regard these biological or genetic factors as exogenous and thus outside the range of their professional concerns. But clearly all social groupings and group behavior cannot be explained this way. Thus, any social theory that claims to be balanced and conceptually complete must offer some explanation of the existence and variety of social groups.

If we are studying a federation, it is natural to begin explaining its existence and characteristics by looking at the constituent groups of which the federation is composed. Thus, we can learn something about why the United States came to exist and to have the characteristics it has by looking at each of the original thirteen colonies or states and at their deliberations about whether to form a union and about what type of union it should be. Similarly, the emergence and character of those groups or organizations that are not federations, but rather assemblages or associations of individuals, begins most naturally with the study of the individuals that compose them and the reasons which induced them to enter into the association or pattern of interaction that is observed.

Social groups, in short, should be explained in terms of the interests, needs, or other traits of the individuals of which they are composed. Individuals are, of course, the "primitive entities" or fundamental building blocks of social life. A social science which claimed to explain social groups without deriving this explanation from an understanding of individuals and their interactions would be like a science of chemistry without a knowledge of atoms and molecules. Assertions such as that "Italians chose a Christian Democratic government because they were influenced by the Roman Catholic church" or that the "middle class voted conservative to protect its privileges" can have some truth-value, but they inherently lack the homogeneity and clarity that well-founded assertions about individuals can claim. This methodological point, familiar as it may be, nonetheless needs emphasis here, because it partly accounts for the distinctive character of the theory that follows. There is also the possibility of confusion, since there are ideologies and political creeds that emphasize "individualism," but these creeds have nothing to do with the methodological preference for theories that are built upon primitive or fundamental entities rather than upon aggregates that are heterogeneous or imprecisely delineated.

This paper will accordingly set forth a theory of certain social groups and forms of social action that is derived from reasonable and familiar features of individual behavior and that may be tested empirically not only by comparing its predictions about social patterns or aggregates in different circumstances or societies with social reality, but also by determining whether the individual behavior that the theory postulates is consistent with our observations of individual behavior. The theory as presently formulated does not, however, encompass most social groups or forms of social interactions. It covers only what sociologists commonly call "social movements," and certain rigidities and barriers that are characteristic of social classes in a relatively strict and narrow sense, and castes of the kind that are found principally though not exclusively in India.

The theory of these social institutions was arrived at serendipitously in the course of my studies of collective action, economic growth, and macroeconomic fluctuations. The implications of this theory for social movements, classes, and castes have accordingly not been set out separately before. Although this theory of certain social groups can ultimately best be understood by studying the more general theory of which it is a part, it is nonetheless true that the parts of the theory that are most pertinent to the special concerns of sociologists are in scattered passages and endnotes in a book that is mainly about other matters. Thus, it might be useful to segregate the "social" implications of the theory in question in a separate paper, so that sociologists might somewhat more quickly determine whether the complete theory was worth their attention. For the most part, then, this paper merely extracts the scattered passages and endnotes in *The Rise and Decline of Nations* (Olson 1982) that are most pertinent to sociologists and attempts to relate them to the sociological literature and perspective.

Because of the severe limitations on the length of each contribution to this volume, this essay must assume that readers already know that the benefits of collective action—such as that undertaken by lobbies, collusions, and cartels—are collective goods that go automatically to everyone in some group or category, whether they have borne any of the cost of the collective action or not. It follows that, at least in sufficiently large groups, these benefits provide no incentive for individuals to make any sacrifices to help any organization for collective action working in the interest of a group of which they are a part, even if they understand how much they and others like them gain from this collective action. Thus, collective action is extremely difficult and problematic, and when it occurs, at least in large groups, will require "selective incentives" or individual rewards or punishments that are applied to individuals according as they do or do not share in the costs of the collective action.

Selective incentives are never available to scattered groups, such as consumers, taxpayers, the unemployed, and the poor, but eventually, with enough ingenuity or good fortune, may be worked out for some groups. In short, this essay must assume that readers are familiar with *The Logic of Collective Action* (Olson 1965), or have at the least read the summary of that book in chapter 2 of *The Rise and Decline of Nations.* Ideally, the reader should also have some awareness of how the sociologist's theory of mass movements has changed since the 1950s and early 1960s because of an improved understanding of collective action.

Implications of the Logic in Combination with Other Theory

If *The Logic of Collective Action* is correct, and is combined with conventional propositions from microeconomic theory and a few other assumptions that will probably not be controversial, a number of further implications follow. The connections to these further implications are set out in *The Rise and Decline of Nations.* I will strive here to give some intuitive sense of each of these implications in a few sentences.

Implication I is that no society can attain symmetrical or complete organization of group interests, so that it is and will be impossible for leaders of all groups to bargain together to obtain a "core" or Pareto-efficient allocation of resources. This follows trivially from the fact that some large groups do not have access to selective incentives. Implication II is that stable societies with unchanged boundaries will accumulate more organizations and collusions for collective action over time. As time goes on more of those groups that can organize will have enjoyed the fortunate circumstances and able political entrepreneurship needed to organize, whereas the interest of organizational leaders in maintaining their position ensures that organizations with selective incentives will not disappear unless destroyed by upheaval or war. Implication III is that "small" groups have disproportionate organizational power in all societies, but since they are not as slow to organize as large groups this disproportion is greater in societies that have lately suffered instability. This follows directly from the logic referred to above.

If the associations or collusions for collective action are small in relation to the society of which they are a part, they will gain little from trying to make their societies more efficient, because their members get only a minute fraction of the gains from a more efficient society (they are in a position akin to that of an individual in a large group). Similarly, they will gain from obtaining a larger share of the social output for their members, even if the social loss from the redistribution is a substantial multiple of the amount distributed to them. There are compelling reasons set out elsewhere for believing that most of the efforts of distributional coalitions have excess burdens that are large in relation to the amount they win in distributional struggle. Therefore, Implication IV is that special-interest organizations and collusions are largely distributional coalitions which on balance reduce the efficiency and income of the societies in which they are located. Implication V is that associations that encompass a large part of the societies in which they are located are severely constrained in seeking redistributions toward their own clients because their own members will bear much of the social costs; if an association represents half of the income-earning capacity of the country, its members will on average suffer half of the loss in social efficiency that results from the redistribution, and it will accordingly not seek any redistributions to its members which cost the society more than twice as much as the amount distributed to its members.

Implication VI is that distributional coalitions will make decisions more slowly than the individuals and firms of which they are composed (because they must make decisions either through by-laws or by unanimous consent bargaining) and will accordingly tend to have crowded agendas and bargaining tables. Implication VII is that distributional coalitions slow down a society's capacity to adopt new technologies and to reallocate resources in response to changing conditions, thereby reducing the rate of economic growth. This is due partly to the crowded agendas and slow decision-making and partly to considerations that are not intuitively obvious or susceptible to summary description. Implication VIII is that distributional coalitions are exclusive. Since this implication is particularly important for the analysis of social classes and castes, it will be discussed separately in more detail. The last implication, IX, is that the accumulation of distributional coalitions increases the complexity of regulation and the role of government, partly for reasons that are not immediately obvious nor capable of brief description.

Taken together, these implications provide parsimonious explanations of a wide variety of phenomena. They largely explain the "economic miracles" after World War II in Germany, Japan, and Italy; the "British disease" of slow economic growth; the slow growth of the northeastern and older midwestern parts of the United States since the war and the more rapid growth of the South and the West; the more remarkable and anomalous examples of economic growth and stagnation over the world since medieval times; the simultaneous inflation and unemployment of the 1970s and 1980s; the increasingly divisive character of political life in some long-stable societies; and the unusual degree of inequality of income in many unstable developing countries. These explanations have attracted a wider interest, at least among economists, than the class and caste rigidities that this paper will attempt to illuminate; and since these explanations spring from the same concise theory that will be used here, they add some strength to the argument in this paper. Unfortunately, there is not space enough to go into these matters here, so I shall focus exclusively on the social barriers that are of special concern to sociologists. To do this, we shall have to look rather carefully at Implication VIII.

The Incentive to Exclude

This implication is that, after they reach a certain point, distributional coalitions have an incentive to be exclusive. In the case of collusive oligopolists or others who operate in the marketplace, the reason is simply that whatever quantity an entrant would sell must either drive down the price received by those already in the cartel or alternatively force existing members to restrict their sales further. When it is created, the cartel must normally enlist all the sellers in the market if it is to succeed, but once it has done this each member of the cartel has an interest in sharing the limited sales at the monopoly price with as few other sellers as possible. Thus, there is a compelling incentive to exclude any entrant. If the number of physicians increases, for example, the earnings of physicians must decline if other things are equal, and in country after country one finds that the professional organizations representing physicians work to limit entry into the profession.

In the case of those distributional coalitions that seek their objectives by political action, the reason for exclusion is that there will be more to distribute to each member of the coalition if it is a minimum winning coalition. A lobby, or even a military alliance seeking spoils, will have less to distribute to each member if it admits more members than are necessary for success. In a world of uncertainty, the size of the minimum winning coalition may not be known in advance, in which case the coalition must, over a range, trade off lower payoffs per member against greater probabilities of success. There will nonetheless always be some point beyond which it must be in the interest of the existing members to exclude new entrants.

A governing aristocracy or oligarchy provides an illustration of the exclusivity of special-interest groups that use political or military methods. Imagine a country or historical period in which some subset of the population, such as the nobility or the oligarchy, dominates the political system. This subset has an incentive to choose public policies that distribute more of the social output to its own members. Except in the case where the aristocracy or oligarchy would increase its security if new members (for example, powerful rivals) were added, it will be exclusive: Every unnecessary entrant into the favored subset reduces what is left for the rest. The relevance of this argument is evident from the exclusiveness of governing nobilities throughout history. Any number of devices and emblems have been used to mark off ruling aristocracies from the rest of the population with the utmost clarity. The exclusivity is perhaps most dramatically evident when a ruling group is secure enough to pass its powers on to its descendants. In these cases there is, of course, great resistance to admitting anyone other than the children of the nobility or ruling group into the ruling group.

If the sons and daughters of the ruling group marry outsiders, and both the sons and daughters and the spouses of these sons and daughters are in the ruling group in the next generation, the ruling nobility will tend to double in size in the next generation; there will then tend to be half as much for each member. One possible solution is to allow only descendants of one sex, say males, to be in the ruling class in the next generation. But those members of the ruling group who have only or mainly daughters have reason to oppose such rules; even apart from natural concern about their daughters, they would lose their share of the future receipts of their ruling group. How can all families in the ruling group bequeath their share of the group's entitlement to descendants without making the value of a share in the entitlement decline by half or more with each successive generation?

One way they can do this is through rules or social pressures that enforce endogamy: If the sons and daughters of the ruling group are induced to marry one another, the growth of the ruling group can be constrained in ways that preserve a legacy for all the families in it. Again, the evidence that nobilities and aristocracies have resisted marriages to commoners and lower ranks generally is abundant, and from many diverse societies. Similarly, royalty's abhorrence in earlier times of marrying commoners can be understood as a rule that helped to limit the losses that a multinational class such as European royalty would have suffered had their numbers expanded exponentially. Individuals in privileged groups will also sometimes have personal reasons for marrying someone in the same privileged group, but these individual advantages of intra-group marriage do not explain the legal and social rules that say a marriage outside the ruling group is to be condemned *by others* in the group as a violation of *principle*. Such rules are best explained, as they are here, at least partly in terms of the interest of the group.

Collective action will also be easier if a group is socially interactive, so that "social selective incentives" can be used to motivate the provision of collective goods. The fact that everyone in the relevant group gets a uniform amount and type of the collective good and must put up with the same group policies also argues that groups of similar incomes and values are more likely to agree. Thus, distributional coalitions, once big enough to succeed, are not only exclusive, but also seek to limit the incomes and values of their memberships, and will therefore be particularly resistant to entrants that differ significantly from those already in the group.

Social Class Barriers

In recent times there has been a near-consensus, both among the British themselves and foreign visitors, that Britain's slow growth was due partly to a class consciousness that allegedly reduces social mobility, fosters exclusive and traditionalist attitudes that discourage entrants and innovators, and maintains medieval prejudices against commercial pursuits. Since Britain had the fastest rate of growth in the world for nearly a century, we know that its slow growth now cannot be due to any *inherent* traits of the British character. There is, in fact, a great deal of evidence that at the time of the industrial revolution Britain did not have the reputation for class differences that it has now. It is a commonplace among economic historians of the industrial revolution that at that time Britain, in relation to comparable parts of the Continent, had unusual social mobility, relatively little class consciousness, and a concern in all social classes about commerce, production, and financial gain that was sometimes notorious to its neighbors:

> More than any other in Europe, probably, British society was open. Not only was income more evenly distributed than across the Channel, but the barriers to mobility were lower, the definitions of status looser. . . .
>
> This was a people fascinated by wealth and commerce, collectively and individually. . . . Business interests promoted a degree of intercourse between people of different stations and walks of life that had no parallel on the Continent.
>
> The flow of entrepreneurship within business was freer, the allocation of resources more responsive than in other economies. Where the traditional sacro-sanctity of occupational exclusiveness continued to prevail across the Channel . . . the British cobbler would not stick to his last nor the merchant to his trade. . . .
>
> Far more than in Britain, continental business enterprise was a class activity, recruiting practitioners from a group limited by custom and law. In France, commercial enterprise had tradi-

> tionally entailed derogation from noble status. [Landes 1969, p. 39–122]

It is not surprising that Napoleon once derided Britain as a "nation of shopkeepers" and that even Adam Smith found it expedient to use this phrase in his criticism of Britain's mercantilistic policies.

The ubiquitous observations suggesting that the Continent's class structures have by now become in some respects more flexible than Britain's would hint that we should look for processes that might have broken down class barriers more rapidly on the Continent than in Great Britain, or for processes that might have raised or erected more new class barriers in Britain than on the Continent, or for both.

One reason that only remnants of the Continent's medieval structures remain today is that they are entirely out of congruity with the technology and ideas now common in the developed world. But there is another, more pertinent reason: Revolution and occupation, Napoleonism and totalitarianism, have utterly demolished most feudal structures on the Continent and many of the cultural attitudes they sustained. The new families and firms that rose to wealth and power often were not successful in holding their gains; new instabilities curtailed the development of new organizations and collusions that could have protected them and their descendants against still newer entrants. To be sure, fragments of the Middle Ages and chunks of the great fortunes of the nineteenth century still remain on the Continent; but, like the castles crumbling in the countryside, they do not greatly hamper the work and opportunities of the average citizen.

The institutions of medieval Britain, and even the great family-oriented industrial and commercial enterprises of more recent centuries, are similarly out of accord with the twentieth century and have in part crumbled, too. But would they not have been pulverized far more finely if Britain had gone through anything like the French Revolution, if a dictator had destroyed its "public" schools, if it had suffered occupation by a foreign power or fallen prey to totalitarian regimes determined to destroy any organizations independent of the regime itself? The importance of the House of Lords, the established church, and the ancient colleges of Oxford and Cambridge has no doubt often been grossly exaggerated. But they are symbols of Britain's legacy from the preindustrial past or (more precisely) of the unique degree to which it has been preserved. There was extraordinary turmoil until a generation or two before the industrial revolution (and this probably played a role in opening British society to new talent and enterprise), but since then Britain has not suffered the institutional destruction, the forcible replacement of elites, or the decimation of social classes that its Continental counterparts have experienced. The same stability and immunity from invasion have also made it easier for the firms and families that advanced in the industrial revolution and the nineteenth century to organize or collude to protect their interests.

Here the argument in this paper is particularly likely to be misunderstood. This is partly because the word "class" is an extraordinarily loose, emotive, and misleadingly aggregative term that has unfortunately been reified over generations of ideological debate. There are, of course, no clearly delineated and widely separated groups such as the middle class or the working class, but rather a large number of groups of diverse situations and occupations, some of which differ greatly and some of which differ slightly if at all in income and status. Even if such a differentiated grouping as the British middle class could be precisely delineated, it would be a logical error to suppose that such a large group as the British middle class could voluntarily collude to exclude others or to achieve any common interest. The theory does suggest that the unique stability of British life since the early eighteenth century must have affected social structure, social mobility, and cultural attitudes, but *not* through class conspiracies or coordinated action by any large class or group. The process is far subtler and must be studied at a less aggregative level.

We can see this process from a new perspective if we note that concerted action usually requires selective incentives, that social pressure can often be an effective selective incentive, and that individuals of similar incomes and values are more likely to agree on what amount and type of collective good to purchase. Social incentives will not be very effective unless the group that values the collective good at issue interacts socially or is composed of subgroups that do. If the group does have its own social life, the desire for the companionship and esteem of colleagues and the fear of being slighted or even ostracized can at little cost provide a powerful incentive for concerted action. The organizational entrepreneurs who succeed in promoting special-interest groups, and the managers who maintain them, must therefore focus disproportionately on groups that already interact socially or that can be induced to do so. This means that these groups tend to have socially homogeneous memberships and that the organization will have an interest in using some of its resources to preserve this homogeneity. The fact that everyone in the pertinent group gets the same amount and type of a collective good also means, as we know from the theories of fiscal equivalence and optimal segregation, that there will be less conflict (and perhaps welfare gains as well) if those who are in the same jurisdiction or organization have similar incomes and values. The forces just mentioned—operating simultaneously in thousands of professions, crafts, clubs, and communities—would, by themselves, explain a degree of class consciousness. This in turn helps to generate cultural caution about the incursions of the entrepreneur and the fluctuating profits and status of businessmen and also helps to preserve and expand aristocratic and feudal prejudices against commerce and industry. There is massive if unsystematic evidence of the effects of the foregoing processes, such as that in Martin Wiener's *English Culture and the Decline of the Industrial Spirit, 1850–1980* (1981).

Unfortunately, the processes that have been described do not operate by themselves; they resonate with the fact that every distributional coalition must restrict entry (Implication VIII). Class barriers could not exist unless there were some groups capable of concerted action that had an interest in erecting them. We can see now that the special-interest organizations or collusions seeking advantage in either the market or the polity have precisely this interest.

In addition to controlling entry, the successful coalition must, we recall, have or generate a degree of consensus about its policies. The cartelistic coalition must also limit the output or labor of its own members; it must make all the members conform to some plan for restricting the amount sold, however much this limitation and conformity might limit innovation. As time goes on, custom and habit play a larger role. The special-interest organizations use their resources to argue that what they do is what in justice ought to be done. The more often pushy entrants and nonconforming innovators are repressed, the rarer they become, and what is not customary is "not done."

Nothing about this process should make it work differently at different levels of income or social status. As Josiah Tucker remarked in the eighteenth century, "all men would be monopolists if they could." This process may, however, proceed more rapidly in the professions, where public concern about unscrupulous or incompetent practitioners provides an ideal cover for policies that would in other contexts be described as monopoly or "greedy unionism." The process takes place among the workers as well as the lords; some of the first craft unions were in fact organized in pubs.

There is a temptation to conclude dramatically that this involutional process has turned a nation of shopkeepers into a land of clubs and pubs. But this facile conclusion is too simple. Countervailing factors are also at work and may have greater quantitative significance. The rapid rate of scientific and technological advance in recent times has encouraged continuing reallocations of resources and brought about considerable occupational, social, and geographical mobility even in relatively rheumatic societies.

In addition, there is another aspect of the process by which social status is transmitted to descendants that is relatively independent of the present theory. Prosperous and well-educated parents are usually able through education and upbringing to provide larger legacies of human as well as tangible capital to their children than deprived families can. Though apparently the children of high-ranking families are occasionally apparently enfeebled by undemanding and overindulgent environments, or even neglected by parents obsessed with careers or personal concerns, there is every reason to suppose that, on average, the more successful families pass on the larger legacies of human and physical capital to their children. This presumably accounts for some of the modest correlation that is observed between the incomes and social positions parents had and those that their children eventually attain. The adoption of free public education and reasonably impartial scholarship systems in Britain in more recent times has disproportionately increased the amount of human capital passed on to children from poor families, and thereby tended to increase social mobility. Thus, there are important aspects of social mobility that the theory offered here does not claim to explain and that can countervail those it does explain.

I must once again emphasize multiple causation, and point out that there is no presumption that the process described here has brought *increasing* class consciousness, traditionalism, or antagonism to entrepreneurship. The contrary forces may overwhelm the involution even when no upheavals or invasions destroy the institutions that sustain it. The only hypothesis on this point that can reasonably be derived from the theory is: that if two societies were in other respects equal, the one with the longer history of stability, security, and freedom of association would have more institutions that limit entry and innovation, that these institutions would encourage more social interaction and homogeneity among their members, and that what is said and done by these institutions would have at least some influence on what people in that society find customary and fitting.

The evidence that there is a greater sensitivity to certain class or group distinctions and barriers in Britain than in various comparable societies, though apparently sufficient to generate near-consensus among observers, is unfortunately mainly informal and anecdotal rather than quantitative. There are mountains of casual evidence, and the evidence and perceptions of British observers appear to be in close accord with those of foreign visitors to Britain. One interesting example of this is the distinctive response of British commentators on the earliest versions of the present argument, written before the implications of the theory for class and group barriers were apparent to me. Most British commentators, however generous they might be, were quick to point out that my argument did not take account of the special characteristics of one or another of the British social classes, such as the alleged bloody-mindedness of the British working class or the allegedly aloof and anti-commercial attitudes of the British upper classes, or of the British class system as a whole. At first I resisted these criticisms as only ad hoc arguments; I foolishly overlooked the fact that not a single commentator from anywhere made any similar comments about any other country, and even somehow neglected memories from my own time as an American undergraduate at Oxford, which memories strongly supported my British critics' contentions that the British class system was distinctive and harmful to British economic growth. Finally, I realized that if my argument was right, the British critics who pointed out that one also had to take account of the class system also had to be right, and then generalized my theory. Once the theory was generalized to cover social rigidities, it was almost inevitable that the additional application of it that is set out later in this paper would come to mind.

Massive and consistent as the casual evidence that there is a distinctive class system in Britain seems to me to be, it is nonetheless useful to seek quantitative and systematic evidence as well. A consensus among observers obviously has meaning, but the perceptions of casual observers do not have the precision and ready comparability that is desirable. On the other hand, quantitative evidence of an incomplete or inadequate kind should not necessarily be given more weight than a vast amount of casual evidence.

Unfortunately, the types of quantitative studies that are now available do not provide appropriate tests of the hypothesis about social exclusion that my theory generates. One reason is that the hypothesis offered here relates to social exclusion or discrimination, but does not claim to explain any correlation between the status of parents and children that is due to different sized legacies of human and other capital. As the application of the theory later in the paper to castes should make clear, my theory, if correct, explains any systematic tendency to exclude or to discriminate against capable individuals who would like to enter some industry, occupation, or group; it does not, however, explain any differences in capabilities that are due to differences in upbringing and educational opportunities, except to the extent that these differences in turn are explained by the impact of the distributional coalitions in the parents' generation on the distribution of income. The quantitative studies of social mobility that exist now relate the social prestige of the occupations fathers had to the social prestige of the occupation of their sons. To the extent that the social prestige of the occupations that sons practice is due to the amount and kind of human capital they acquired, it is generally not explained by my theory.

A second reason why the existing quantitative studies do not provide a good test of my argument about social involution is that *these studies consider social mobility from one generation to another, and the majority of the distributional coalitions in Britain and other Western countries are not strictly multigenerational. That is, most of them do not restrict membership in the coalition to their own offspring*. The Indian distributional coalitions considered in the next section of this paper do this; so do some European coalitions such as the nobility and certain labor unions, but so far they appear to be less common in the West than single-generation coalitions. To the extent that membership in distributional coalitions is not passed on from one generation to another, the exclusion and discrimination inherent in these coalitions will not be captured by studies of the degree of association between the prestige of the occupation of fathers and sons.

A third reason why the studies of the association between the occupational prestige of fathers and sons do not offer a sufficient test of the present theory is that they leave out so much: differences in accent, dress, or style across different social groups; the role of inherited fortunes and titled aristocracies; the degree of resentment faced by uninvited entrants to established occupations or industries; the extent to which people are conscious of or sensitive about their social or class position; attitudes toward business; and attitudes toward entrepreneurship (which probably leads to the most dramatic changes in socioeconomic position). One measure of the significance of some of these variables that the existing quantitative studies leave out is the degree to which class and social position are correlated with allegiance to political party. Here it is significant that the association between socioeconomic status and adherence to the labor and conservative parties in Great Britain has been very much stronger than the corresponding association between socioeconomic position and affiliation with the Democratic and Republican parties in the United States (see Vanneman 1980). It might be objected that this British-American difference is due to the different nature of the political parties in the two countries rather than to any differences in the social structure, but since the political parties are in turn partly a reflection of the socioeconomic situation, and have the policies they have partly because of their desire to attract support, this objection is not convincing.

Although the existing types of quantitative studies of social mobility are by no means sufficient to test the present theory, they are nonetheless extremely useful for a variety of purposes. They also seem to show faint traces of the involutional process my theory describes. Treiman and Terrill (1975) find the rate of social mobility marginally lower in Britain than in the United States. Similarly, the data in papers by Erikson, Goldethorpe, and Portocarero (1979, n.d.) suggest that Sweden (whose more encompassing coalitions have a smaller incentive to exclude than their narrower counterparts in Britain) has somewhat more social mobility than England.

From Classes to Castes

My thoughts about how this theory relates to the caste system occurred when, quite by chance, I was reading Jawarharlal Nehru's *The Discovery of India* (1946). I would not have been surprised if Nehru, a practicing politician, would have celebrated the glories of India's ancient civilization and placed almost all of the blame for the country's problems on the British. Nehru naturally did point with pride to many of the great achievements of India and Indians, but what was most notable was that his praise was confined almost exclusively to certain periods of Indian history. Nehru offered a very different and negative view of Indian civilization at the time of the coming of the Moslems and of the conquest of India by the Europeans. India in this epoch was "drying up and losing her creative genius and vitality"; it was the "afternoon of a civilization . . . there was decline all along the line—intellectual, philosophical, political, in technique and methods of warfare, in knowledge of and contacts with the outside world, in shrinking economy." How, Nehru asked, could India's political freedom have been lost "unless some kind of decay has preceded it? A small country might easily be overwhelmed by superior power, but a huge, well-developed and highly civilized country like India cannot succumb to external attack unless there is internal decay, or the invader possesses a higher technique of warfare. That internal decay is clearly evident in India."

Nehru attributed the decay to "the static nature of Indian society which refused to change in a changing world, for every civilization which resists change declines." He reasoned that "probably this was the inevitable result of the growing rigidity and exclusiveness of the Indian social system as represented chiefly by the caste system." The caste system, he wrote, was a "petrification of classes" that "brought degradation" and is "still a burden and a curse." Nehru did not claim any originality for this diagnosis, and it has also been put forth by many others.

It is not, however, sufficient to explain the decline of India by the era of the Moslem and European invasions in terms of the caste system. We have not reached home until we have explained *why* India acquired the caste system when it did and comprehended the caste system in a theory that is testable against the experience of many countries.

One of the most common hypotheses is that the castes emerged out of guilds or similar organizations; most castes bear the names of occupations and there is evidence of guilds in earlier Indian history. Another common hypothesis is that visible racial differences among the indigenous peoples of India and between these peoples and the Aryan-speaking invaders were the source of the caste system; there are visible differences among some caste groups to this day, and the English word "caste" stems from the Portuguese *casta*, meaning "race." Yet another familiar explanation ascribes the castes to common lineages; a crucial feature of the caste system is endogamy, or the prohibition against marriage outside the basic unit of caste grouping, the *jati*, and many tribes have been incorporated into the caste system. It might seem that the theory here would focus exclusively on the hypothesis that the castes grew out of guilds, but the other two hypotheses are, if the theory offered here is right, also important, and we shall return to them shortly.

Traditionally caste groups were not only mainly occupational, but also exhibited all of the features of cartels and other distributional coalitions. They controlled entry into occupations and lines of business, kept craft mysteries or secrets, set prices monopolistically, used boycotts and strikes, and often bargained on a group basis rather than individually. The traditional caste system also had several other features that would be expected of distributional coalitions. One of these is that often groups rather than individuals change status. A caste group that enjoys prosperity will often rise gradually to a higher status, and may also decide collectively to adopt more restrictive ritualistic rules, thereby rising even in terms of the religious concepts of purity and pollution. Another is that Hinduism emphasizes the concept of *Dharma*, or the duties appropriate to the caste or group. Moral behavior, in other words, is defined not in a universalistic way, but in terms of obedience to the rules of one's caste or station; it is like "professional ethics" that rule out competition in a profession. Even the murderous thugs or other criminal castes were behaving consistently with their *Dharma* when they carried on their caste's activities.

None of the preceding is, however, an explanation of the prohibition against marriage out of the group that is such a basic feature of the caste system, nor does it explain any correlation of caste with racial or ethnic differences. For that we must turn to Implication VIII. That is, we must remember that distributional coalitions are characteristically exclusive and seek to limit the diversity of their memberships. We must ask how that implication would apply over a multigenerational time span.

Consider the situation of an older member of a profitable guild. As one of the co-owners of an advantageous coalition, the older member would have an interest in how he or his descendants might share in the future returns. He could pass the right to practice the monopoly occupation on to his sons and give or sell the right to practice the occupation to his sons-in-law. But even with a steady-state population the *number in the craft will double if both sons and sons-in-law are allowed to enter, and normally a doubling of the craft's membership would eliminate the gains of the cartelistic output restriction that gave the guild its value in the first place.* The same fundamental logic applies also to a caste that has disproportionate rights to village harvest, for its members also have an incentive to be in a minimum winning coalition. The only way a distributional coalition can retain its value over several generations is by restricting the children of members to marriages with one another or by disinheriting a large number of the children. I hypothesize that the Indian castes used mainly the first method. This brings us back to the discussion of Implication VIII, where the nobility and royalty of Europe were used to illustrate the exclusiveness of marriages (or bequests, as in primogeniture) that is essential to any successful multigenerational distributive coalition.

The Role of Racial and Ethnic Differences and of Shibboleths

Just as the origin of the caste system is often ascribed to guilds, so, as we recall, it is often also related to the racial diversity of India at the time of the Aryan migrations and also to lineage groups. These hypotheses are, as it turns out, also very much in keeping with the logic of Implication VIII. If a racially distinct distributive coalition gets started by alien conquerors, it will be able to preserve itself over many generations only by arbitrary rules of bequest such as primogeniture or through endogamy. If it is largely endogamous the differences in appearance will be preserved.

Any distributive coalition that uses endogamy to preserve its benefits over a multigenerational period must be large enough to avoid the problems of serious inbreeding. As the endogamous group gets larger, however, the difficulty of enforcing endogamy rises. How is this or that son to be restrained from marrying some especially appealing girl outside the group, or how are his parents to be prevented from making an especially advantageous marriage contract for him with relatively wealthy or powerful people outside the group? How can the astute outsider be kept from marrying into the coalition? If exogamy is not prevented, at least some of the families must lose their share of the coalition's future gains. If there are visible differences, it will be easier to determine who is in the group and who is not and to enforce the endogamy rule. Differences in speech, culture, and life style are also shibboleths that make it harder for the outsider to blend in. *The promotion of prejudices about race, ethnicity, culture, and intergroup differences in life style will also make the coalition work better.* The inculcation of these prejudices will increase the probability that the members will follow the rule of endogamy, and strengthen selective incentives by interacting socially only with their own group, of their own accord.

Although multigenerational distributional coalitions foster inefficiency, inequality, and group prejudice, it is nonetheless important to realize that some individuals and groups outside the society with these coalitions may improve their positions by joining that society, even if they enter at the bottom. Tribes without settled agriculture, for example, might in some circumstances have found that they would be better off by joining Indian society than staying out of it, even though they got the lowest status and were victims of distributional coalitions to boot. There have been many observations of such assimilation of tribal groups into India's caste system and they must help account for its great diversity.

In keeping with the scientific principle that the theory that explains the most with the least is the most likely to be true, any alternative explanation of the Indian caste system should also be capable of explaining some developments outside India. Similarly, the explanation offered here of the Indian caste system will be stronger if it not only explains diverse developments outside India, but also developments outside India that are similar to the Indian developments. To some extent the previous analysis of class rigidities was in that category, but there is still the problem that the traditional Indian caste system, with its endogamous marriage rules and its segregation of the different groups, lineages, or races, does not have many counterparts elsewhere. Perhaps surprisingly, one of the closest analogies is with South Africa.

The extraordinary system of *apartheid*, or racial segregation and discrimination, in South Africa is a relatively recent development. The severer forms of racial segregation and discrimination do not go back to the early days of the Boers in South Africa. On the contrary, it appears that there was more than a negligible amount of interbreeding between the Boers or other Europeans and the Africans. There is, after all, a large population of "coloured" or mixed-race people in South Africa today; the South African government treats them as a separate category and segregates them from Africans and Asians as well as Europeans.

A distinguished South African economist, W. H. Hutt (1964), has written a startling history of the evolution of progressively tighter systems of racial segregation and discrimination in South Africa. Hutt's account focuses closely on the mining industry in South Africa early in the present century. The mine owners and management needed labor and naturally preferred to get it at low wages rather than high wages. Since Africans had few other opportunities outside the traditional sector of African society, they were often available at low wages. The mine owners also drew upon the huge pool of low-wage labor in Asia, and for a time used indentured Chinese labor. European workers were also employed in the mines, mainly as foremen and skilled and semi-skilled laborers. It was clear that the far cheaper African laborers could at very little cost soon be taught the semi-skilled jobs, and the employers naturally coveted the savings in labor cost that this would bring.

The competition of cheaper African and Asian labor naturally did not appeal to the higher paid workers of European stock or their recently formed unions. There were strikes. In part because of these strikes there were a variety of changes in labor policy in South Africa. The Mines and Works Act of 1911, also called the Colour Bar Act, was passed. Although on a superficial reading relatively innocuous, as administered it constrained employers in their use of African labor in semi-skilled and skilled jobs. The regulations promulgated under the act prevented Africans in the Transvaal and the Orange Free State from entering a wide variety of mining occupations. They even specified *ratios* between foremen (whites) and mining laborers (Africans).

Disagreement about the ratios emerged. After World War I, the mine employers asked for a ratio of 10.5 Africans per white, whereas the labor union demanded 3.5 to one. A general strike in the Rand followed in 1922. This strike and the agitation that followed became a common cause of conservative Afrikaaners, and communist and socialist leaders, with all of them supporting the efforts to deny opportunities to the poorer Africans who were competing with white labor. The South African Labour party, modeled more or less after its British counterpart, prospered in the wake of the strike, and joined with the mainly Afrikaaner white supremicist Nationalist party in a coalition government. The Nationalist-Labour "Pact" government soon introduced the second Colour Bar Act, the Mines and Works Act of 1926. Hutt calls this "probably . . . the most drastic piece of colour bar legislation which the world has ever experienced." It was accompanied by a so-called "civilized labour policy," which limited opportunities for Africans still further. One of the devices used to keep African laborers out of jobs where they would compete with whites was the requirement of "the rate for the job." If the wage for a given job is fixed at a level attractive to Europeans, the employer has no incentive to seek African workers who would work for less. Apprenticeship rules under the "civilized labour policy" also had the effect of excluding Africans. Thus, the system, though it forced employers to adopt less profitable and more discriminatory policies than they preferred, brought substantial gains to organized white (and sometimes "coloured" and Asian) workers.

These and similar policies naturally drastically limited opportunities for African workers. Yet there was a continuing demand of Africans from farther north to enter, notwithstanding the policies against them. They came in at the bottom, and were victimized by the rules, but this was nonetheless better than the alternatives some of them had in the traditional sector. The analogy with the tribes that have been assimilated into the bottom of the Indian caste system is quite striking.

Let us now ask what necessary conditions must be met if the South African system, and the cartelistic gains it provides for many, are to be preserved over the long run. The system could not possibly survive over successive generations unless the demarcation between the races was preserved. If less favored groups could enter the more favored groups, as they would have massive incentives to do, wage differentials could not possibly be maintained. A continuation of the processes that generated the "coloured" population would make the system untenable in the long run, and even in what (by the standards of Indian history) would be the medium run.

That is not only an implication of the present theory, but the conclusion of the South African government as well. Just as the restrictions on the use of African labor in skilled and semi-skilled jobs increased over time, so did the rules categorizing all of the population into rigid racial categories and forbidding sexual relations, in marriage or otherwise, between them. (As this is being written, various pressures and fears appear to have generated official proposals marginally to change these rules.)

Undoubtedly any number of other causal factors have been at work in South Africa, and any account as brief and monocausal as this must be in many respects misleading. The purpose, however, is not to give a complete account, but to induce reflection about the sources of racial and other forms of discrimination. As others have argued before, the individual as a consumer, employer, or worker finds it costly to discriminate. The consumer who discriminates against stores owned by groups he finds offensive has to pay higher prices or suffer a lesser selection by shopping elsewhere. The employer who discriminates against workers of a despised group has higher labor costs because of it. The worker who does not accept the best job irrespective of the group affiliation of the employer is essentially taking a cut in pay. A similar logic applies to individual social interactions of other kinds. The fact that ***individuals*** find discrimination costly means that if individuals are free to undertake whatever transactions they prefer, there will be a constraint on the extent of discrimination.

Distributive coalitions of individuals, on the other hand, can sometimes gain enormously from discrimination. Any group difference that facilitates exclusion, by Implication VIII, will be advantageous. For periods of only a generation or two in length, the group differences can usually be considered as given, but over the centuries and certainly the millennia they cannot. In the long run, then, multigenerational distributional coalitions must tend toward endogamy. This is equally true of the South African whites, the Indian castes, and the European nobility.

Conclusion

This paper began with the argument that the existence of social groups and their characteristics need to be explained. An adequate theory of such groups must be deduced from testable assumptions about individual behavior. This does not deny that humans are fundamentally social animals or in any way minimize the importance of socialization and group pressures; it merely recognizes that social groups are themselves the result of human interaction.

This paper has not attempted to explain social groups in general, but only social movements, the exclusive character of certain social groups and their tendency to be composed of people of the same "class," the emergence of multigeneration guilds or castes, and the role that racial, ethnic, and shibboleth-generating processes play in this process. But the theory on which this paper is based does explain these phenomena parsimoniously and derives them logically from familiar and testable features of individual behavior. Admittedly, this paper suffers from the fact that most of the evidence about classes, castes, and social exclusion is drawn from only three societies. Nonetheless, it does provide a partial statement of a distinctive theory of some important sociological phenomena. I hope that this partial statement will be sufficient to persuade some sociologists that my theory is sufficiently pertinent to their professional concerns that they will want to study the theory as a whole, and not only the special applications of it that have been briefly described here.

Comment

MICHAEL HECHTER

Mancur Olson's paper is largely a plea for sociologists to take the theory he originally proposed in *The Logic of Collective Action* seriously. He tries to pique our interest by holding that the theory has important implications for the analysis of both economic development and of castes. I believe that his case for economic development, unfortunately an unpopular subject in American sociology and an even less popular one in American economics, is more convincing than his case for castes. In this discussion I plan to briefly comment on the applications he offers us and then to suggest that his theory has broad implications for sociology. Indeed, I would even like to propose that it touches questions at the very core of our discipline.

Let's briefly consider his examples. Olson's first argument is that distributional coalitions are inimical to economic growth. I find the logic here to be compelling, but I do wonder just how important these coalitions are as determinants of economic stagnation. The sketchy evidence presented in his recent book, *The Rise and Decline of Nations,* doesn't really permit much of an answer because Olson never compares their salience to that of other possible determinants.

On the other hand, his arguments that the Indian caste system arose to safeguard the public good of the members of wealthy castes, or that primogeniture and *apartheid* arose for the same reasons, leave me a bit uneasy because these explanations appear to rest on a most non-Olsonian form of group logic. All the examples refer to complex and society-wide institutions that must themselves be regarded as the outcomes of *prior* collective actions. He needs to tell us how free-riding was curtailed in the establishment of such institutions.

Many societies have been conquered by an alien race—the Anglo-Saxons by the Normans, the Han Chinese by the Mongols, for example—but there are relatively few examples of castelike stratification systems in world history. How does the theory explain this variation? Whereas it may account for the motivation of an alien conquering group to establish a ruling caste, clearly its success in so doing has to be a function of other constraints. What are these constraints?

Now, despite this problem, I want to go on to suggest that Olson's theory is more central to the core concerns of sociologists than either he or they have tended to appreciate. This is because the theory is about the conditions under which individuals participate in the production of public goods. The theory's key prediction is that because collective action requires costly selective incentives, its occurrence is far rarer and more problematic than had previously been assumed.

But there is another kind of public good that lies at the very heart of sociological theorizing and this is the phenomenon of social order itself. That social order is largely a public good was recognized intuitively, if not explicitly, by Durkheim and Parsons, among others, whose normative theories of social order were proposed as alternatives to the rational choice theories advocated by Hobbes and his followers. In lieu of selective incentives they spoke of sanctions; in lieu of free-riders they complained about deviants. Olson's logic of collective action formalizes Hobbes's behavioral assumptions but rejects his *deus ex machina* explanation of the rise of the state. In this respect the analysis is entirely consistent with Talcott Parsons's argument in *The Structure of Social Action*.

However, Parsons and rational choice theorists like Olson part company in their assessment of the sufficiency of selective incentives or sanctions as causes of the provision of a public good like social order. Olson argues that selective incentives are necessary and sufficient for the attainment of a public good in large groups. On the other hand, Parsons feels that more public goods—that is, more social order—are provided than can be explained solely on the basis of sanctions.

This question, I submit, is one of the core issues separating sociological from rational choice analyses. Whereas rational choice theorists attend only to environmentally determined constraints on individual action, sociologists turn again and again to such admittedly fuzzy notions as charisma, legitimacy, and internalized norms in their explanations of human behavior. They do so because they are often more aware of the costs that are imposed by measures to control free-riding than are many rational choice theorists.

Careful rational choice analysis reveals that selective incentives represent only one of three types of costs that must be incurred if free-riding (or deviance) is to be curtailed. Selective incentives are necessary but *insufficient* causes of an individual's decision to participate in the production of public goods. In addition there are two other costly requirements that many rational choice theorists tend to elide. First, groups must be able to detect whether an individual free-rides or does not do so. Call these attendant costs the costs of monitoring. Second, groups must ensure that each member receives the particular selective incentive which is appropriate to his or her past behavior—that is, rewards for participation in a collective action, punishment for free-riding. This activity entails costs of properly allocating the selective incentives. Now, when monitoring and allocation costs are added to the costs of providing the selective incentives in the first place, it's evident that free-riding in large groups is likely to be even a more pervasive problem than many rational choice theorists appreciate.

Can this problem be resolved within current rational choice theory? Perhaps. Groups can adopt certain kinds of structural arrangements—among them profit-sharing, or group rewards more generally—that give members incentives to monitor each other. At the same time, they can provide public sanctions that make a spectacle of free-riders who get apprehended and thereby convince others to walk the straight and narrow path.

But there is also a sociological solution. Perhaps under certain conditions individuals can be induced to maximize some collective rather than individual utility schedule. The issue crops up in various guises, in the attempt to develop "socialist man" in the USSR and the fact that all states seem to allocate so much money to the promulgation of ideology. The problem of the formation of preferences, or what sociologists sometimes call values, is most difficult, requiring excursions into modern psychology. Even so, there may be some point to devoting sustained thought to the question of preference formation. I suspect that the psychological assumptions underlying many rational choice theories will on inspection turn out to be rather antiquated and insufficient, particularly in that they place too great an emphasis on operant as against social and cognitive psychological mechanisms. For the moment, however, these issues are largely speculative.

General Discussion

James Coleman: I want to reinforce a point that Michael Hechter made, because it seems to me so important, given you are the author of this piece. It has to do with the logic of collective action. Consider the sentence in your paper: "The only way a distributional coalition can retain its value over several generations is by restricting the children of members of marriages with one another or by disinheriting a large number of the children." Or, "endogamy, which is necessary to the guild's continuation." You are treating collectivities as actors. In other words, you have engaged in exactly what your *Logic of Collective Action* argued against.

Mancur Olson: You're absolutely right that this is important but note it's not an inconsistency; it's an incompleteness.

James Coleman: Is not incompleteness all you've argued against in *The Logic of Collective Action?*

Mancur Olson: Fair point. The incompleteness is serious and should be particularly serious if one believes my earlier work. But if there *already is* a formal organization with selective incentives, the organization will normally act, at least to some degree, in the interests of its members. No one is surprised if a corporation like General Motors seeks profits that enrich its stockholders. Similarly, there is a long tradition among students in international relations, for example, of predicting what countries will do by looking at the "national interest." There are all sorts of dangers and problems with that, but theories of mutual deterrence, alliance burden sharing, and so on can still be useful. What I'm suggesting, then, is that it isn't astonishing that guilds, once they are organized, would in fact act in ways consistent with the long-run interests of their members and the survival of the guild.

Joseph Ben-David: You have an excellent theoretical argument about group action and the difficulty of group action. Unfortunately, most sociologists haven't learned it yet and I think reading your paper will help very many to get cured from one of the endemic diseases of our discipline. But I have difficulties with your theory of economic decline. You argue that the longer a country enjoys stability, the more it declines economically. By this logic, the country with the greatest economic troubles at this time would be Switzerland, for it is the country with the longest period of stability.

Mancur Olson: The case of Switzerland is discussed in *The Rise and Decline of Nations.* It's growing more slowly than Britain, though that piece of evidence is in turn weakened a lot by the fact that per capita income is much higher in Switzerland than in Britain.

There is an essay on this by Franz Lehner, a Swiss political scientist, in a book, *The Political Economy of Growth,* published by Yale University Press and edited by Dennis Mueller. Lehner argues that the influence of distributional coalitions in Switzerland is much attenuated by the restrictive constitution, which means that changes in economic policy favorable to any special-interest group are very hard to obtain given the constitutional obstacles. We do need more research before we can know how my argument fares in the case of Switzerland.

Chairman: We have not much time left and I suggest that Olson responds to the following questions at the end.

Mary Douglas: I would like to join your coalition. I see what's happening here and you're getting a lot of resistance. You are attacking a monopoly of the sociologists with a big takeover bid. Unless somebody did something of the kind you've done, which is to try to relate rational choice to social organization, the two would run in parallel paths and a great deal of thinking would be wasted. One of the complaints was that you were ignoring social structure but it seems that you have an account of marriage organizations, an account of property distribution, an account of morality. What I foresee for you are much different difficulties. You will never get proof. You've got a grand theory here, and what happens to grand theory is that you're bidding to become the dominant paradigm. But then the theory will be impossible to falsify.

Gary Becker: I found the paper interesting and stimulating, but I disagree with some fundamental assumptions. This disagreement indicates that economists are not always in a coalition against sociologists. You apparently do not consider how different groups interact and possibly offset each other in their competition for political influence. If a policy raised aggregate efficiency but harmed some groups (and, of course, benefited others), would those harmed have enough political organization and influence to block this policy? This question cannot be answered without a model of the competition or cooperation among groups. Since you do not provide such a model, you cannot determine whether pressure groups are an obstacle or an aid to economic growth and efficiency. In particular, if groups benefiting from policies that raised efficiency *tend* to have more political influence than groups harmed [as argued in G. S. Becker, "A Theory of Competition Among Pressure Groups for Political Influence," *Quarterly Journal of Economics,* Summer 1983], pressure groups might, *on balance,* raise output and efficiency by encouraging policies with positive effects on efficiency.

Joseph Gusfield: I am overwhelmed with admiration at your raw courage. For a period of Indian history, when almost every generalization is up for grabs, it takes great courage to make the kinds of generalizations you have made. The caste system for one thing has known a great deal of mobility. For example, several studies have shown that castes have not been able to maintain a great deal of exclusivity in factory jobs that were very much sought after. In addition, your theory doesn't explain the peculiarity of India in having a caste system.

Howard Aldrich: I want to follow up on Gusfield's point. Genetically, members of different castes in India look like part of a common population, which suggests that endogamy has not been a very effective way of subcastes preserving their boundaries.

Robert Eccles: Two questions: First of all, could you generalize at all about how long it takes for these coalitions or groups of collectivities to form? Second, does this choking off of economic growth occur while they're forming or do they reach a certain critical mass and at that point everything coalesces and then economic growth is choked off after that? What is the timing involved here?

Mancur Olson: I'm grateful for these questions. On Eccles' point about how long it takes: Organizing collective action takes time, but goes on faster in societies with modern transportation, communication, and technology. This is an important matter for further research and I would regard the theory in important ways as incomplete until we have done more work on this difficult question.

In many ways the most critical of the comments was from Joe Gusfield. Besides India, there are other societies with castes. But India has so many more castes and the question is why it has so many more. I do not have a full answer to that question. I offer my theory as a logical structure which generates hypotheses which facilitate detailed research. But I would call attention to the racial diversity of India, due partly to the Aryan invasion (after there was already an advanced society with settled agriculture) and also to the mountains and deserts of India, which divide that subcontinent into distinct regions across which there was relatively little intermarriage. This has meant that a variety of different peoples and races have been together in India over a very long time. Those communities with the knowledge and the land needed for settled agriculture presumably let people from the hill tribes, who were usually of different race or ethnicity, in to do the dirty work. Though admission to a community with settled agriculture presumably left the entrants better off than they were, those with power in the established community could nonetheless exploit the entrants to an enormous degree. This situation has been made much worse by the long tradition of indirect rule, wherein British, Moslem, and other rulers, instead of governing the villages, let them govern themselves, while exacting a collective tax on the whole village. This allowed the indigenous power structure of caste structure to survive. But I don't pretend that this is the whole story.

Gary Becker made a point that gets at the logic of the argument. One could put his argument this way: if we think of all of the groups in society as competing for privilege and special-interest legislation, wouldn't those groups that push proposals consistent with the social interest have an advantage? Now if *every* group could organize, his argument would stand and my argument would fall. However, I show in *The Rise and Decline of Nations* that there are very strong reasons for excluding universal or symmetrical organization; most important, some groups have access to selective incentives and other groups do not. To take one example, the workers in a particular mine can organize sometimes because there may be only one entrance to the mine and you can picket that entrance easily. But to picket all consumers of a given product all over the United States would require a picket line of prohibitive cost, thus consumers are not organized in any society of any size. So my immediate answer (and I wish there was time for a fuller one) to Gary Becker's point is that we do not have a society in which all coalitions are feasible. Accordingly we do not have an organizational structure that can give us a "core" or Pareto-efficient allocation.

Closing Address

Micro Foundations and Macrosocial Theory

JAMES S. COLEMAN

One of the few benefits that comes with organizing a conference is the freedom to break rules. Like a policeman going through a red light, it's not exactly right, but there's no one to tell you it's wrong.

A rule of the conference was that the papers "not be about social theory but be original contributions to the substance of social theory." The rule was obeyed everywhere except in the panels, where it was not intended to apply in the first place. It is this rule that I shall break. Since this paper is the final paper of the conference and intended to provide a kind of overview, it is perhaps appropriate that the paper focus on the character of social theory rather than on society. This makes possible an examination of each of the papers of the conference from the perspective I will sketch out. The latter part of the paper is devoted to such an examination.

That aspect of social theory which I want to address is its multilevel character. At the lowest level the system of organization with which social theory deals is the individual person; at the highest it is total societies. Furthermore, although much of the theorizing that is of interest to sociologists is at the higher or more encompassing levels of social organization, the observations we make are most often at the lowest level—that is, the behavior of individuals.

We are often engaged, then, in our social theorizing, in making transitions between levels, either with data or in the structure of the theory itself. In this paper I want to describe typical ways in which theory deals with these micro-macro problems and then to point out difficulties that arise with some of those ways, ending with some prescriptive statements about how the social sciences ought to make the transition.

Units Characterized by Variables

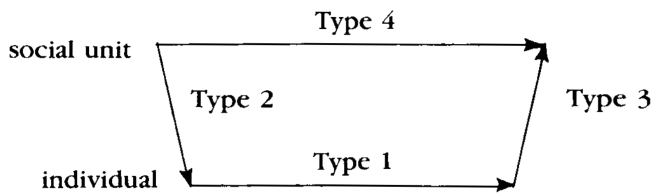
Most social theories can usefully be described as consisting of relations among variables. In a given relation, considered in isolation, there is an independent and a dependent variable, although the "independent variable" in a given relation may be endogenous, or internally determined, when the whole system of relations is considered. The question that may be asked—which is relevant to the movement between levels—is just what kind of unit the variable in each of these relations characterizes: an individual? a family? a city? a business firm? a school? a society?

To begin with a simplification, we can think of two types of units: an individual and a social unit, with all the cases listed above except individuals being social units. If we do that, there are four types of relations:

Independent Variable	*Dependent Variable*	*Type of Relation*
individual	individual	Type 1
social unit	individual	Type 2
individual	social unit	Type 3
social unit	social unit	Type 4

(I might have listed "micro" and "macro" in place of individual and social unit, respectively, but that may be subject to misinterpretation, since those terms are used with various meanings.)

The character of these relations can be usefully expressed by means of a diagram in which the upper and lower levels represent a social unit and an individual, respectively. The relations described above are shown by arrows leading from the independent to the dependent variable, and the arrows are labeled according to the type of relation.



Some theories may be composed solely of Type 1 relations or solely of Type 4 relations, although it is not possible to have theories consisting of interconnected relations that are only Type 2 or Type 3. A number of theories, I will suggest, consist solely of Type 1 and Type 2 relations, while a number of other theories consist of Type 1, Type 2, and Type 3 relations, and some consist solely of Type 4 relations.

Theories Based on Type 1 and Type 2 Relations

I will argue later in the paper that the central theoretical problem in sociology is the transition from the level of the individual to a macro level—the problem that economists call, in some of its forms, "aggregation," although the term is a misleading one. To get an idea of the character of this problem, I will first say something about work which does not involve it at all, theories involving only Type 1 and Type 2 relations.

Most quantitative empirical sociology takes individual behavior as the phenomenon to be explained. Most research on social stratification, for example, takes as the dependent phenomenon of interest an individual's occupational classification, or the prestige of that occupation, or the income of the job. "Status attainment research" is concerned with the effects of various characteristics of the individual, such as education, previous occupation, race, or sex, on current occupational prestige. It is also concerned with effects of various characteristics of that individual's family (such as parents' education, father's [or less often, mother's] occupation, number of siblings) on prestige. The first set of examples produces a Type 1 relation, while the second set produces a Type 2 relation, with the family as the social unit characterized by the independent variables. In rare cases social stratification research includes relations in which the independent variable is a larger social unit, such as the effect of average wage rates in the city of residence on the individual's income. These also are Type 2 relations.

It is clear also that some classic work in social theory consists of nothing more than Type 1 and Type 2 relations. Emile Durkheim's classic work, *Suicide,* for example, can be described as postulating and partially testing a set of paired relations taking the following form:

- a characteristic of society (or other social unit, such as the church) affects
- a psychological state of the individual (egoism, anomie, altruism) which affects
- a behavior of the individual (suicide or probability of suicide)

An example of such a pair of relations as proposed by Durkheim is: Protestant religious doctrine leads (through its individualistic orientation) to a psychological state that Durkheim called "egoisme," by which he meant, roughly, social isolation. This in turn leads to a greater likelihood of suicide. Here the pair of relations is first the effect of church doctrine on the psychic state of members, a Type 2 relation; and then the effect of the psychic state on behavior, a Type 1 relation.

This work of Durkheim is characteristic of what Sorokin called the "sociologistic school" of social theorists, which took some attribute of social organization as the determinant of variations in individual behavior. It is sociological not because of the dependent variable, but because of the independent variables. It is hardly to be despised, yet it is hardly complete or comprehensive as a social theory. It is a way of explaining individuals' behavior, but it does not provide an explanation for the behavior of a social system.

Difficulties in Certain Theories Involving Only Type 1 and Type 2 Relations

Some of the theories which take individual behavior as the dependent phenomenon of interest confront a special difficulty: the necessity of taking into account social organization in the very structure of a theory. I will illustrate the difficulty with research in social stratification.

Ordinarily, research in social stratification treats a change of job as if it were an individual decision: The determinants are background characteristics of the individual, aspects of life history which affect occupational mobility. The destination occupation is regarded as unlimited in number of open jobs; the taking of a new job of a particular type is analyzed in exactly the same way as the change of an attitude. Yet jobs are scarce commodities, and a new job is obtained only in competition with others.

In short, it is a matching process, carried out in a market structure. As in any such matching process, the final action depends not merely on the job-seeker's interest in this job, but also on his interest in other available jobs; and not merely on the organization's interest in this job-seeker, but also on its interest in other available job-seekers. In addition, these actions depend on the interest of *other* organizations in this and other job-seekers, and in other job-seekers' interests in this and other organizations. That is, the action depends intrinsically and directly on the distribution of other job-seekers and of other jobs, and of the distribution of interest at each point in these distributions.

If an appropriate theory of this matching process is specified, then no aggregation of individual behavior is necessary. The theory includes Type 3 relations, resulting directly from the matching process. The distribution of jobs held (or changes in the distribution) is directly given by the theory. Demographers and economists have made attempts to develop a model of a matching process for a structurally similar problem, the marriage market, in which a marriage depends upon parties on both sides of the market (see Sanderson 1980; Schoen 1983; Becker 1976; Gale & Shapley 1962; Roth 1983).

Some time ago, Harrison White's analysis of vacancy chains in the occupational structure drew attention to the importance of the occupational structure by taking the vacancy rather than the person as the unit of analysis. Yet this ignored that part of the structure—the persons seeking jobs—on which usual stratification research focuses exclusively.

What is most notable about the focus on individuals only (or on positions only) is the limits it imposes on the usefulness of stratification research results for theory. For example, a crucial element necessary for testing a functional theory of stratification is missing by our inability to analyze occupational change as a matching market. A functional theory of stratification includes actions by both those who award status to occupations and those who respond to it by competing for the occupations. With a conception of the occupational market as a matching market in which various resources (such as status) have value, this systemic or functional conception of stratification would be a step closer to being testable.

Theories That Include Type 3 Relations

In some sociological analyses, the outcomes of interest are intrinsically and obviously a social product and could not be studied as a mere aggregate of individual actions. These are the cases in which the dependent variable clearly characterizes a social unit, although individual-level variables may play a part in the theory.

An example is Max Weber's study, *The Protestant Ethic and the Spirit of Capitalism.* At one degree of detail, Weber is simply expressing a macrosocial proposition that I have labeled Type 4 in the earlier discussion: The religious ethic which characterized those societies that became Protestant in the Reformation (and particularly those that were Calvinistic) contained values that facilitated the growth of capitalist economic organization.

It is a virtue of Weber's analysis that he does not remain at the macro level, but moves down to the individual level. There, the single proposition breaks into three: one having an independent variable characterizing the society, with the dependent variable characterizing the individual (Type 2); a second with both independent and dependent variables characterizing the individual (Type 1); and a third with the independent variable characterizing the individual, and the dependent variable characterizing the society (Type 3). The propositions may be put, somewhat crudely, as:

(Type 2) a. Protestant religious doctrine generates certain values in its adherents.

(Type 1) b. Individuals with certain values (referred to in 1 above) adopt certain kinds of orientations to economic behavior.

(Type 3) c. Certain orientations to economic behavior (referred to in 2 above) on the part of individuals helps bring about capitalist economic organization in a society. (The central orientation to economic behavior is characterized by Weber as antitraditionalism.)

In this set of propositions, the third is of most interest, for it is the third which moves back up from the individual level to the societal level. Thus, in contrast to the research in status attainment, or Durkheim's work on suicide, where the phenomena of ultimate interest could be described as individual behavior, here the phenomenon of ultimate interest is clearly macrosocial, characterizing the society as a whole.

It is, however, in the third proposition that Weber's theory is weakest, for it is here that some combination of individual actions is necessary to generate a macrosocial outcome. The orientations to action of a worker in a capitalist enterprise are not the same as those of an entrepreneur, yet both are necessary to the enterprise. The orientations necessary to begin such enterprise are not the same as those necessary to continue it. In short, capitalist economic organization is a system of action, and to show how that system comes into being, or even how it functions once in being, a set of value orientations on the part of a population is not sufficient.

I want to suggest that the central problem of sociological theory is this problem, the problem of the Type 3 relation. It is poorly treated by

Weber in his *Protestant Ethic* and is ignored altogether as a problem in most social stratification research—as well as in a wide range of other social research.[1] It is the problem of moving from a micro level to a macro level, a problem described by economists as the problem of aggregation. It might be better described as a problem of "combination": how actions at the micro level combine to bring about an outcome at the macro level.[2]

In quantitative research that has theoretical aims, the same failure to pay attention to the Type 3 relation is evident. In this respect an evolution, which is only partially complete, has taken place in survey research. In its early days, with a focus on voting decisions, audience research, and consumer behavior, survey research (as exemplified by the work of Samuel Stouffer and Paul Lazarsfeld) was concerned almost wholly with Type 1 relations—within the individual—and secondarily with Type 2 relations. It was social psychology, practiced through the use of surveys of individual attitudes and behavior.

The second stage of this evolution is one in which the sample for the survey was a sample of well-defined social unit—for example, the American population aged 21 to 65. By this single change, the work became potentially relevant to macrosocial outcomes—for the sample now characterized a social unit about which statements might be made. A milestone marking this evolutionary stage is Blau and Duncan's *The American Occupational Structure* (1967). The uncompleted character of the evolution, however, is evidenced by the fact that the relations studied by Blau and Duncan—and by others working in this tradition—are wholly Type 1 and Type 2 relations. The Type 3 relation, from individual behavior to a macrosocial outcome, was implicitly assumed to be taken care of by simple aggregation up to the level of the social unit. Work in this

[1] Some would defend Weber on grounds that he was not attempting to account for the rise of capitalism, but only the "spirit" of capitalism. This, I believe, is a rather weak defense. For if Weber was attempting to account for the "spirit" of capitalism, is that to be regarded as a property of the society, that is, a shared norm, or as a constellation of beliefs on the part of the individual Protestant? If the former, Weber has failed to show the processes through which the individuals' beliefs give rise to the social norm (as well as to demonstrate the relevance of such a norm to the actual practice of capitalism). If the latter, one must ask just what is Weber's accomplishment, since under this interpretation it consists only of showing that a set of beliefs in the religious realm overlap with a set of beliefs in the economic realm.

[2] The Type 2 relation may be regarded as a second intellectual challenge to sociological theory, for it involves the effect of one unit (a social unit) on another (an individual). The process through which such an effect occurs should be explicated.

tradition ends by being curiously bifurcated: The *explanatory relations* developed through regression analysis of individual occupational status are all Type 1 and Type 2, giving a social psychology of status attainment; the *descriptive statistics,* showing occupational distributions, characterize the social unit—in this case, American society.[3]

Another stage in this uncompleted evolution is exemplified by *Inequality,* by Jencks and others (1972). Unlike Blau and Duncan's *American Occupational Structure, Inequality* did have a macrosocial focus. The focus of interest was on inequality of income in American society, a property of the social system, not of individuals. Yet almost all the analyses in the book focus on Type 1 and Type 2 relations. The dependent variables in turn are individual cognitive achievement of school children; level of educational attainment, measured by number of years of school completed; occupational status; and income. In the latter two analyses, educational achievement and attainment are independent variables. Jencks and his colleagues found little effect of an individual's education on his income; but this negative finding obscured the logical fault of the analysis: This result, or its opposite, does not imply anything about the effect of changes in educational inputs (macrosocial variables) on income inequality (another macrosocial variable). The latter may, for example, be related to the distribution of occupations, and an increase in the mean level of education would mean merely an increase in the educational levels of persons holding particular occupations. Neither the correlation between education and income nor the inequality of income need change.

Types of Relations in the Papers of This Conference

It is useful to locate the papers of this conference according to the typology of relations in social theory that I have used here. I will attempt that for each.

[3] It is ironic that the kinds of individual behavior studied in early survey research (voting, consumer behavior) that was not based on samples of a well-defined population are more appropriately aggregated to characterize a social unit than is job choice. While an individual's job change has ramifications for other jobs and other workers, by the vacancy created and the job that is filled, a voting decision or consumer choice is an individual act. Yet even here, extensive evidence indicates that the influences between this act and the acts of others engaged in similar choices are sufficiently great that a simple aggregation fails to show the Type 3 relation through which these actions combine to produce a social outcome. (See Katz & Lazarsfeld 1955; Rogers 1962; Coleman, Katz & Menzel 1966.)

BEN-DAVID: Ben-David and Zloczower's study of the nineteenth-century growth of experimental physiology was concerned with accounting for a macrosocial phenomenon, but did not simply attempt to establish a Type 4 relation between macrosocial variables. Indeed, to have done so would have required societal-level comparisons (e.g., to show that, in those societies in which X was absent, Y did not occur, . . .), which they did not depend on.

Then what was the methodological strategy? The answer is, I believe, similar to that in most good historiography, to go from the macro level *down* to the micro level, through Type 2 relations, and then back *up* again, through Type 3 relations. In Ben-David's case, this meant primarily the following path: Beginning with characteristics of the university system, primarily decentralization-induced competitiveness, they traced the effect of this on the motivation, behavior, and material success of young men with talent—a Type 2 relation, from macro to micro. From this they traced how the individual and joint actions of these new recruits, clustered in time and space, generated the growth of a new field—a Type 3 relation, from micro back to macro. Ben-David and Zloczower used what might be termed an economic market model to explain a macrosocial phenomenon, the functioning of the academic system as a system of action, and this led them from macro to micro and back up again.

LAUMANN AND KNOKE: The units in this paper appear not to be persons at all or corporate actors, but positions, and a structure of relations between positions. This means that the typology of Type 1–Type 4 relations that I have used in this analysis of social theory appears less applicable here.

Yet some aspects of the paper suggest just where network analysis is located with respect to other social theory. First, the empirical data collected by network theorists is on individuals who are members of the social unit or in some way associated with it. (In the analyses reported in the paper the "individuals" are themselves corporate actors, but from a methodological point of view, they are individuals relative to the social unit, the society as a whole.) The second point is that a portion of the data gathered about these individuals involves relations between them, the "structure" or "network" of network theory. In the first section of the paper, however, Laumann and Knoke state three postulates. Postulate II—as expressed in the statement that "similarities in status, attitudes, beliefs, and behavior facilitate the formation of intimate (or consensual) relationships among incumbents of social positions"—appears to be at the individual level; yet it is not a relation of Type 1. It constitutes a nascent relation of Type 3, with status, attitudes, beliefs, and behavior of individuals as independent variables, and a *relation* (or absence of it) as the dependent variable. This is not a large step toward the move up from individual to social unit, from micro to macro; yet it constitutes a start. A second proposition about the establishment of relations is contained in the discussion of "issue linkages": actors with interests in the same issue will be likely to establish relations.

With all this, it is clear that Laumann and Knoke are concerned with Type 3 relations: how to account for the actions of social units (communities, government policy-making) on the basis of characteristics of individuals (and often also of the relations between them). Whether this agenda of network theory can be accomplished by the methods it uses without conceptual leaps from pair relations up toward the macro level is another question; but this appears to be the goal.

OBERSCHALL: Oberschall's paper takes a macrosocial problem for explanation: the egalitarian and largely self-enforcing norms, rules, and laws in California gold mining camps. The explanation is based on an individual-level behavior principle, that of rational action. An implication of this principle which Oberschall uses to account for the normative structure is that those rules and norms will be established which minimize the costs of social transactions, including in this instance enforcement.

This becomes in effect a Type 3 relation in which an individual-level behavior principle on the part of members of a mining camp (together with ecological conditions imposed by nature) gives rise to a macrosocial outcome, that is, a particular normative system. What is missing is the equivalent of Walras's *tâtonnements* in a system of market exchange—the mechanism through which such a minimization of transaction costs is arrived at. This leaves unanswered such questions as, What about differential initial power or resources of different miners? Shouldn't this, by the same behavior principle, lead to differential reduction of transaction costs borne by those with greater and lesser initial power? And how extensive a change would be necessary in the ecological conditions in order that the normative structure change?

Despite the absence of a mechanism or process to show how the intentions of minimizing transaction costs are realized (and to support the argument that these intentions are realized), Oberschall's paper shows the outlines of how a Type 3 micro-to-macro transition may be made in social theory. The outcome to be explained is system-level behavior, and the explanation involves aims of individuals in a system of interaction.

BECKER AND TOMES: The Becker-Tomes paper is concerned with a relation in which the independent variable is the family's wealth, the dependent variable is that same family's change in wealth, and the two are connected through an investment decision. Thus, it is clearly a Type 1 relation.

Yet Becker and Tomes are interested in systemic questions: What prevents the *distribution* of income or wealth from becoming more and more unequal over time? And given that it does not, but instead the inequality is maintained at roughly the same level through a regression toward the mean, how does the investment decision partially offset this regression to the mean to affect a family's rise and fall over time?

Clearly these are Type 3 questions, despite the origins in a Type 1 relation. The dependent variable is the distribution of income, a characteristic of the society. One fact brings the paper's aims within the domain of problems ordinarily treated by sociologists: The "investment decision" that is at issue here is a family's intergenerational one. What amount of capital, human or material, or what amount of liability, will it bequeath to its offspring? This question is relevant to a number of questions which occupy the attention of sociologists, ranging from questions of intergenerational mobility to questions about family functioning, socialization, and "cycles of poverty." The Becker-Tomes paper is especially interesting for the contrast it presents to much work in sociology. I stated earlier in the discussion of social stratification research that the evolution of survey research from investigation of social-psychological Type 1 and Type 2 relations to macrosocial dependent variables is incomplete. The Becker-Tomes work indicates how, merely with a change in the investigator's focus of attention, an additional step in that evolution can be taken. It is possible in principle to move the occupational *distribution* forward from one generation to the next, to examine the steady-state distribution of occupational prestige to which current intergenerational mobility and fertility lead, and to study the projected rise and fall of American families within the distribution of occupational prestige. Sociologists who use survey methods appear sufficiently preoccupied with explaining individual behavior that this next step is seldom taken; and the Becker-Tomes paper does not carry matters sufficiently far to show how it might be done in ways that would connect with their theory.

HANNAN AND FREEMAN: The paper by Hannan and Freeman presents an excellent example of a theory containing only Type 4 relations. In their ecological theory of organizational change, the dependent variable of all relations (which are expressed as assumptions 1 through 10) is organizational survival. With only one exception, there is no resort to a discussion of processes internal to the organization, and this has nothing to do with individual persons; it has to do with the complexity of the links between subunits of the organization (assumption 10).

It is useful to contrast the Hannan-Freeman ecological theory of organizational change, which they characterize as the natural selection, or Darwinian theory of change, with the theories they discuss that involve adaptation or change of a given organization. The latter theories are based on rational action by occupants of particular positions in the organization; consequently, they necessarily involve both Type 2 and Type 3 relations—going down from the level of the organization as unit to the level of persons or positions within it and back up to the level of the organization as a unit. Hannan and Freeman's theoretical mechanism explicitly disavows individual action internal to the organization which could modify it; there is merely selective survival of different types of organizations. This selective survival process does, of course, affect a higher level of organization as well: the distribution of organizations with various characteristics in, say, a particular institutional area. Thus, although the propositions stated by Hannan and Freeman are all at the same level, that of the single organization, the nature of the process involved has implications for propositions at a higher social level. Only one such proposition is stated by Hannan and Freeman, as a deduction from assumptions 1–3: Selection within populations of organizations in modern societies favors organizations whose structures have high inertia.

GUSFIELD: The theoretical perspective of interpretive sociology described and pursued by Gusfield is one concerned in part with Type 2 relations, in which something outside the individual, ordinarily characterizing a social unit, affects the subjective perception of the individual. But the work is concerned largely with how this objective reality *does not* determine the subjective meaning. Thus, the work lies principally with a particular kind of Type 1 relation, in which certain unspecified characteristics of the individual (experiences, beliefs) affect the perception of reality or the meaning of particular events and circumstances that enter his awareness.

The work is, however, sometimes about norms or shared interpretations, such as the social definition which turns a phenomenon into a "social problem." Here is required a Type 3 relation, to move from an individual definition of the situation to a shared definition of the situation. It is here, however, that the work in this genre is weakest; it calls attention to the fact that socially shared meanings at one time and place are quite different from the socially shared meanings of the same objective reality at a different time and place. But it stops there, without asking, for example, what it is that leads drinking-and-driving to be generally regarded as a social problem at some times and places more than at others. The socially shared meaning can in fact be regarded as a kind of collective decision in which different meanings compete for dominance, and the question which remains unexamined is, What is the process through which a particular outcome is arrived at?

Here the work of the interpretive sociologist and that of the sociolinguist are especially close, for sociolinguistics is concerned with shared symbols, shared meanings, and common utterances. In particular, Labov's discussion of the conditions under which a subgroup will adopt the linguistic traits of another subgroup is relevant to the question of how one subgroup's definition of the situation comes to be widely accepted.

More generally, the multiplicity of meaning, and how one of several possible meanings of an event is socially selected, is one of the important and poorly understood problems of sociology. It is a classic case of a Type 3 relation: The shooting of a Korean airliner by a Soviet pilot has one meaning in the United States, another in the Soviet Union. Or the invasion of Grenada by American marines has at the outset several possible meanings, each with its own evaluation; over the period of a few weeks, these narrow down to a single meaning.

ECCLES AND WHITE: This paper takes as problematic the functioning of the multidivisional firm that is engaged in several markets. It does not remain at the level of the firm in its explanation, for it discusses the role of profit centers within the firm and the creation of internal pseudomarkets with transfer pricing among divisions. If the firm were a true market, the outcomes at the firm level would be produced by the interactions among divisions with differing resources and differing demands. But in Eccles and White's exposition, the firm-level outcomes are due to decisions of the chief executive officer, first in setting up such an apparently decentralized structure, then in establishing the basis on which transfer prices are arrived at, and finally in treating the set of profit centers like an investment portfolio.

In this respect, the Eccles-White paper carries out a particularly simple kind of Type 3 relation, one in which the macro-level (firm) outcome follows directly from the actions of a single individual—in this case, the chief executive officer. In this assumption, Eccles and White are in the tradition of the management decision-making branch of organization theory. This single-authority form of making the micro-to-macro transition is one characteristic of early and simple forms of societal governance. The Eccles and White paper contains some suggestion that the evolution of business firm governance will be, as has been the case in civil society, toward increasingly complex forms of self-governing systems.

SKOCPOL AND ORLOFF: Skocpol and Orloff's paper has a single dependent variable, clearly macrosocial: the governmental provision of welfare measures in a society. To see the nature of the theory, it is useful to contrast it with one of the other explanations they discuss. The "logic of industrialism" has, as they present it, two major sets of propositions of Type 2: First, industrialization (a property of the society) creates dependent persons separated from traditional dependency structures (such as the extended family and the community). Second, industrialism makes individuals wealthier and thus in principle able to finance welfare measures. The theory's missing link is in the Type 3 relation, through which the needs on the part of some persons and the wealth held by others somehow come together to provide welfare policies. The tests of the theory have, perhaps because of the missing Type 3 relation, remained at the macrosocial level in both independent and dependent variables.

Skocpol and Orloff's own theory has five variables at a macrosocial level. "State formation" affects "how parties and governmental administrations operate." This, together with "socially rooted demands," affects "what groups or parties propose," and these two latter variables affect the policies that are implemented. All relations in the theory are Type 4. The empirical problem is a difficult one, because Skocpol and Orloff use only a comparison of two countries, England and the United States. While they can show consistency of their empirical material from the two countries with the theory, the fact that the relations are all at the macrosocial level, with only two cases for comparison, means that the material cannot be conclusive, but only suggestive. Indeed, such societal-level macrosocial theories can hardly be conclusive, so long as all propositions remain that level (Type 4 relations), because of the small number of cases. Stronger tests can be made if it moves down from this level and contains relations of Types 1, 2, and 3. This is, I believe, the import of much of Mann's critique, when he asks the authors to examine the orientations of individuals in comparable positions in the two countries.

LABOV: Labov's paper on linguistic structure and social structure is especially interesting to examine from the perspective I am taking. Language is an explicitly social product, since it is the vehicle of communication between persons. Yet each expression is an utterance by a person. Thus, there is intrinsic movement between levels in the process by which language changes. This does not imply, of course, that a theory of language variations needs to dip down to the individual level. Indeed, remaining at a single level appears at first to characterize what Labov terms the structural-functional theory of linguistic variations.

Structural-functional theory reverses cause and consequence in a relation: Regarding the entity under consideration as a wholly homeostatic system allows structural-functional theory to view consequences (or desired consequences) as causes. Thus, in language, sociolinguistic variables (consequences of action) bring about a (desired) identification and differentiation of speakers. They allow one to adjust social distance in the speech situation.

As in most functional theory, its application to sociolinguistics implies some benefits for individuals, but nothing is shown about how these benefits lead individuals to bring it about. Thus, structural-functional theory in linguistics is an excellent illustration of an incomplete theory: It contains a relation of Type 2, in which a social product has a (desired) effect for individuals, but leaves completely implicit the Type 3 relations, through which the individuals act to initiate or maintain this desirable social product.

The other linguistic theory that Labov presents follows more straightforward lines of cause and consequence. It expresses (for example) the conditions under which a subgroup will adopt the linguistic traits of another subgroup. Here the relations all characterize the same unit: a subgroup of a speech community (in its relation to another subgroup). The propositions of the theory do not go down to a level below this or up to one above this. The observations, however, must necessarily be made at an individual level. This means that some care is necessary in aggregating to the level at which the propositions are stated. As Labov shows, age, sex, race, locality, and social class variations mean that aggregating across these boundaries will result in obscuring linguistic variations.

HEISE: Heise's paper on affect control theory processes is wholly concerned with individual behavior as a dependent variable. The general character of the theory is, as expressed by Heise, that "people behave so as to confirm notions that they have about themselves and others." Nearly all the propositions of the theory presented by Heise are of Type 1, involving other characteristics of the individual (e.g., attitude toward an actor), although some are of Type 2, involving environmental events. In addition, the words used by persons, which are an important part of affect control theory, are social products, with shared meaning. Thus, the predictions about behavior using language are implicitly Type 2 relations with the shared meanings as partial determinants of behavior. However, as with much social-psychological theory, the complexities of moving from micro to macro do not appear.

OLSON: Olson's paper takes throughout an orientation to social theory similar to the one I have taken in this paper, when I say that the central problem for social science is Type 3 relations, in which a macrosocial outcome is generated from individual characteristics.

First, drawing upon the argument of *The Logic of Collective Action,* Olson cautions against a common fallacy in the micro-to-macro transition: the assumption that individuals with like interests will join together to pursue a common interest. The conceptual leap from like interests to common action is the simple and inappropriate aggregation of an implicit Type 3 relation that I have discussed earlier.

But after cautioning against this, Olson does express a Type 3 relation through simple aggregation. Like interests of individuals (resulting from a similar position in society) *will* lead to formation of groups pursuing a common interest and excluding outsiders if the groups are small (Implication III). Olson calls such groups "exclusive distributional coalitions." Castes in India are examples of "exclusive distributional coalitions" which have arrived at rules (such as caste endogamy) which maintain the exclusivity (i.e., the size) over generations. But because of the failure to specify the conditions under which such groups will form, he has not "explained" how the caste system developed in India and not elsewhere. The "need" of an occupational group to restrict its size is an explanation that begs the question of how the micro-to-macro transition is made. The South Africa case is somewhat more instructive, for it gives some indication of the processes through which the caste system developed there.

Thus, taken as a whole, Olson's paper has a general orientation that is wholly compatible with that I take here: A satisfactory social theory cannot remain at the macro level, with only Type 4 relations; and when it descends to the individual level, it cannot return to the macro level through simple aggregation, which is a pseudo–Type 3 relation. Like interests do not simply aggregate to common interests. But in Olson's application of the point, it is unclear just what are the Type 3 relations that allow collective action in the cases which he discusses.

From Micro to Macro: Markets, Organizations, Social Choice, and Collective Behavior

Despite the misnomer, "aggregation," that economists have given to the problem of moving from individual to macro level, economists may have made the most progress in addressing it. Their principal tool is the conception of rational action carried out in a competitive market.

In neoclassical economics, this produces a perfect market for exchange of private divisible goods, and such a market is the conceptual apparatus that enables economists to move from individual tastes and endowments to a set of prices and a distribution of goods.[4] Modification of that theory to include monopolistic and oligopolistic markets, and to include markets with imperfections, such as transaction costs, expands the scope of this theoretical orientation.

Among the papers of this conference four have explicitly used some variant of this paradigm: Becker, Eccles and White, Oberschall, and Olson. With each I would have some disagreement: Oberschall never examines the individual-level actions which combine to bring about a set of rules; Olson neglects to show the conditions under which his distributional coalitions would overcome the freerider problem and through Type 3 relations would come into being; Becker uses not only the economic paradigm, but also restricts himself to economic variables as independent variables; and Eccles and White focus only on economic dependent variables. Yet these papers taken together show something of the capability of this paradigm for moving beyond the confines of economic markets. In addition to these papers that explicitly use the paradigm, the work by Ben-David and Zloczower reported in Ben-David's paper makes a somewhat less explicit use of it as well, when they show how the market structure of the university system in Germany generated a new discipline.

There are, however, ways other than competitive markets in which a social product results from individual actions. Organizations constitute structures explicitly designed to bring social products into being through the combination of individual actions. Again, the paradigm of economic rationality has been used to provide a theory for certain aspects of this process of combination, in the so-called theory of the firm, a theory to which Eccles and White's paper is designed to contribute. Such questions as the amounts of different factors of production to be used (answered by the marginal productivity of the factor) are addressed by this body of work, although it is far less developed than the theory of markets.

In the theory of the firm, the orientation used in the theory of markets, with the actions of the organization being the result of the interplay of various individuals with differing resources and differing interests, is largely abandoned by economists. In its place is the simpler, but less satisfactory, device of the single decision-maker, making decisions which

[4]While this is true in principle, in fact macroeconomics is not well grounded in microeconomics, and the combination of differing individuals' tastes and endowments, as well as differing firms' production functions, to produce macroeconomic phenomena are activities largely missing from economic theory construction.

are rational for the goals of the firm. Far more satisfactory are those approaches exemplified by Crozier's *The Bureaucratic Phenomenon* (1964), by some of Peter Blau's work, and by some recent work in economics in the problem of agency, which sees the organization as a structure of incentives within which individuals, who have differing resources depending on their positions, pursue their interests.

Another set of phenomena, collective decisions or social choice, constitutes a mode of combining individual actions that is, on its face, quite different from a competitive market or a productive organization. Here, the resources with which individual actors begin are *rights*, most prominently the right to cast a vote. A large body of work carried out by economists and political scientists has focussed on the processes of combining individuals' actions to produce a social choice (see Riker & Ordeshook 1973). This work, too, is less developed than the theory of private goods markets and explicitly neglects some of the most important problems of social choice: conflicts, revolts and revolutions, social movements.[5]

Besides markets, organizations, and social choice, there are other less institutionalized social processes through which the micro-macro transformation takes place. There is a broad class of phenomena termed "collective behavior": panics, fads, crowds, riots, audience behavior. Theoretical work here is probably least developed. It involves such elusive and fluid concepts as "trust," and the conception that persons will voluntarily give up temporary control over their actions to others.

These four institutions by which individual actions are translated into macrosocial products are not exhaustive. I have said nothing, for example, about the development of social norms. Nevertheless, these give a sense of the problem.

It is important to recognize the correspondence between social reality and the existing or potential social theory. What is necessary for reality is to have *social institutions* (such as markets or elections) which translate individual tastes and endowments into a set of prices and a distribution of

---

[5]Part of the reason for the lesser development than in neoclassical economics is the greater theoretical intractability of the phenomenon. An illustration of this intractability is the fact that if an economic market involving exchange of private divisible goods is conceived as a game of strategy, it always has a nonempty *core* (the set of potential outcomes that cannot be blocked by a coalition of players, each of whom could benefit by doing so). An election, however, with three or more alternatives, again considered as a game, will always have an empty core, for certain distributions of preferences: Any proposed social choice can be blocked by a coalition of players who would each benefit by doing so.

goods or into a collective decision. What is necessary for social theory is to have conceptual devices to describe that translation.

The reality may be simple or complex; and a more complex reality creates greater difficulties for the social theory designed to describe it. For example, an early form of rule of a society was rule by king, and the question of where the right to rule originated was decided by fiat: The rights were of divine origin, and the king was endowed by a divinity with these rights. In such a system the translation of individual tastes into a collective action is not problematic, for the collective action merely follows the will of the king. The conceptual apparatus necessary to describe this is, as a result, quite simple.

The evolution of forms of societal rule, involving multiple branches of government, various forms of representation, and a variety of procedures for the registering of citizen preferences, has made social reality extraordinarily complex. This in turn means that the conceptual apparatus for describing the way these tastes are translated into collective action must itself be complex. Thus, the evolution of complexity in the social institutions which create macrosocial outcomes from individual behavior forces a similar evolution in social theory. There is no fixed and invariant, ultimately correct theory of this process of combining individual behavior into macrosocial behavior.

The question then arises whether, despite the necessary variability in specific theory, there is a paradigm that is generally applicable to these processes through which individual actions combine to produce a social product. To pursue that question, however, would distract attention from the central thesis of this paper: that satisfactory social theory must attempt to describe behavior of social units, not merely that of individuals; that it must nevertheless be grounded in the behavior of individuals; and that the central theoretical challenge is to show how individual actions combine to produce a social outcome.

General Discussion

Mancur Olson: While it would have some other shortcomings, wouldn't a Walrasian general equilibrium theory meet all of your standards?

James Coleman: I am glad you asked me that question. I believe that the appropriate paradigm for sociology is one which is derivative from Walrasian equilibrium theory, though one which deviates from that theory in part because not all social goods are divisible, without externalities, and obey properties of conservation; and in part because of social structure, which a Walrasian system ignores.

Siegwart Lindenberg: Sometimes it helps to know that something that is described and deemed important already has a name. What you are describing as an adequate Type 3 effort in sociology has been called "the problem of transformation" for many years now in Europe. The problem of transformation means: How is the concert of individual action transformed into a social or collective phenomenon. All sorts of institutional and social structural things are needed to make that step, so that the *minimal* sociological explanation would always fall into two parts; one would be the explanation of action or behavior on the basis of social conditions and the second would be the solution of the problem of transformation. In this connection I years ago criticized Olson's theory of collective action because it really does only the first and not the second. It explains apathy as a function of group size, using the microeconomic apparatus. It does not explain the underproduction of public goods as a function of apathy. It is very useful, I think, to pay particular attention to this *two*-step kind of explanation in the social sciences.

The other point that I wanted to raise is connected to that. The adequate solution to the problem of transformation depends to a large extent on the quality of the explanation of the behavior, i.e., the quality of the first step. The more psychological the explanation of behavior, the more difficult it will be to relate it to the production of social phenomena. Conversely, to the degree to which you can bring in institutions and structural aspects already on the individual level, to that degree it will be much easier to make the second step. For example, if you model "motives" of bureaucrats (or better: utility arguments) already on the basis of institutions and social structure, say, by assuming that bureaucrats attempt to maximize budgets, then it is much easier to relate this behavior to some social consequences, in this case the budgetary process and social costs.

James Coleman: This is what I illustrated by the example of social stratification research. If the model or theory explaining individual job choice appropriately took into account the institutional structure of jobs (for example, the matching market), then the Type 3 relation or the transformation process would already be obtained; the distribution would automatically result.

Edward Laumann: There may be an instructive error either in my understanding or in your coding. Your discussion of my paper was to suggest that the proposition about similarities in status, attitudes, and beliefs facilitating the formation of intimate of consensual relationships among the incumbents of social positions should be coded as a step from the individual to the social. I think this raises a fundamental question about what is the nature of a structural or a social entity. One could not measure an attribute of the individual as "a similarity"; you must know how it combines with attributes of relevant others. It requires the joint observation of the pair, whether the person is a Protestant or a Catholic, to determine that it is a similar or different religious characteristic.

James Coleman: I agree. Any paper in network theory would fit less well into the scheme for exactly that reason. It is only a highly simplified scheme, but it helps in certain cases to focus on problems of moving from micro to macro levels. This is all I claim for it.

I also might comment about theories composed of Type 4 relations. Suppose the social unit is at the level of a nation. Any theory that can be empirically tested must, if it is based on Type 4 relations, be a very simple theory, since the empirical data consist of a small number of cases. With an analysis which goes down to an individual level and then back up again, one can get a stronger test of most components of the theory, since at the lower level there are many more cases to test the theory.

William Labov: Often there is no other way than to be in a Type 4 situation. Let's take theories of language change. Some people look to slips of the tongue as explanations for language change. But the response we have always made to that is that the language does not change when one person changes his way of speech, only when a group of people agree to accept that counter as a way of conveying social information.

Surely there must be many other social phenomena that don't exist on the individual level.

James Coleman: I agree with you. That is one point I am trying to make. Often relations are mistakenly seen as Type 1 or 2 when they should be conceived of as social products. And the explanatory task is to find out how this social product arises from individual actions. Changes in speech begin with individual actions, but the diffusion of that action through a population is collective behavior, and the social product is language change.

Terry Clark: It may be useful to look for the importance of different types of independent and dependent variables. For example, a state which is providing public goods may override the normal importance of selective incentives to individuals which, in turn, are much more critical when you are dealing with a marketlike phenomenon, such as the job market. A Type 4 relation may be more reasonable when selective incentives are not so important and it may not lead to an adequate analysis when they are important. Have you worked on that possible distinction?

James Coleman: No, I have not done that but I agree with you; it would be useful, although a formidable task.

Gary Becker: I would like to return to the Walrasian general equilibrium example. It is a great achievement in economics, but 99 percent of the time economists do not deal with the complete system. They break this system into parts and study supply, occupational choices, etc., in a manner that is (and should be) consistent with the total picture.

It would be a monumental achievement to integrate economic, social, and political phenomena into one general system. Practically all the time, however, social scientists will continue to analyze in detail particular parts.

James Coleman: The question is what kind of chunks one takes. When one is confronted with, say, individual job change, the piece of this overall Walrasian system that one should look at is that particular job market and not only one fragment of that market. So I would agree with your general point but I think not with the specifics.

Mancur Olson: As Alfred Marshall said, you have to get both blades of the scissors—the preference or demand blade and the resource constraint or supply blade—together, and many studies don't even do that.

Harrison White: I wonder if I might introduce a lighter note. Mary Douglas could do this better than I, but in honor of so many guests we have here from history and economics and linguistics, I thought I might venture my amateur hand using four of the papers to characterize what has just been going on with Jim Coleman. Labov's work alerts us to the possibility that, in addition to the manifest communication Jim Coleman has been giving us, there may be subliminal features that are distinctive of a particular locale or a tribe. And I would like to suggest that indeed Jim Coleman as an old Columbia sociology hand may have been engaged in an activity known to him as classifying everyone into a 2 x 2 table. Then I would like to refer to Hannan's work to say that he indeed wouldn't have survived Columbia unless he learned to do that. At this point I suggest we put our hands in the hand of David Heise and say that the name of the game is to find out if Jim can persuade us that indeed this 2 x 2 table exercise is "potent," "good," and "lively." And as always, the last word must go to Joe Gusfield because it is up to him to come to an agreement whether there is enough coming together that that has happened.

James Coleman: I would like to ask for a second hand of applause for the man who was initially responsible for the whole conference and that is Stefan Nowak who is not here . . . Ladies and Gentlemen, there are drinks outside.

References

Current Issues in Social Theory

Janowitz, Morris. *The Last Half Century.* Chicago: University of Chicago Press, 1978.

Lindenberg, Siegwart. "Sharing Groups: Theory and Suggested Applications." *Journal of Mathematical Sociology* 9(1982):33–62.

———. "An Assessment of the New Political Economy: Its Potential for the Social Sciences and for Sociology in Particular." Paper presented at the annual meeting of the American Sociological Association, Detroit, 1983 [1983a].

———. "Utility and Morality." *Kyklos* 36(1983):450–68 [1983b].

———. "Normen und die Allokation sozialer Wertschätzung." In *Normengeleitetes Verhalten in den Sozialwissenschaflen,* H. Todt. Berlin: Dunker & Humblot, 1984.

The Emergence of Sociology as a Discipline

Cohen, Michael D.; March, James G.; and Olsen, Johan P. "A Garbage Can Model of Organizational Choice." *Administrative Science Quarterly* 1(1972):1–25.

Goodman, Leo A. "How to Ransack Social Mobility Tables and Other Kinds of Cross-Classification Tables." *American Journal of Sociology* 75(1969):1–40.

Parsons, Talcott. *The Social System.* New York: Free Press, 1951.

Stinchcombe, Arthur L.; Jones, Calvin; and Sheatsley, Paul. "Nonresponse Bias for Attitude Questions." *Public Opinion Quarterly* 45(1981):359–75.

Thomas, W. I., and Znaniecki, Florian. *The Polish Peasant in Europe and America.* 1918. Reprinted. New York: Knopf, 1927.

Sociology of Knowledge

Barnes, Barry. *Interests and Growths of Knowledge.* London: Routledge & Kegan Paul, 1977.

Ben-David, Joseph. "Scientific Productivity and Academic Organization in 19th Century Medicine." *American Sociological Review* 25(1960):828–43.

———. *The Scientist's Role in Society: A Comparative Study.* Englewood Cliffs, N.J.: Prentice-Hall, 1971.

———. "Science, Scientism and Anti-Scientism." Paper presented at the Progress in Science and Its Social Conditions Symposium, Stockholm, August 15–19, 1983.

———, and Collins, Randall. "Social Factors in the Origins of a New Science: The Case of Psychology." *American Sociological Review* 31(1966):451–564.

Ben-David, Joseph, and Zloczower, A. "The Idea of the University and the Academic Market Place." *European Journal of Sociology* 2(1961):303–14.

Bloor, David. *Knowledge and Social Imagery.* London: Routledge & Kegan Paul, 1976.

Bourdieu, Pierre. "The Specificity of the Scientific Field." In *French Sociology: Rupture and Renewal Since 1968*, edited by C. Lemert. New York: Columbia University Press, 1981 [1975].

Cranefield, Paul F. "The Organic Physics of 1847 and the Biophysics of Today." *Journal of the History of Medicine and Allied Sciences* 12(1957):407–23.

Curtis, James, and Petras, John W., eds. *The Sociology of Knowledge: A Reader.* New York: Praeger, 1970.

Fuhrman, Ellsworth. *The Sociology of Knowledge in America, 1883–1915.* Charlottesville: University of Virginia Press, 1980.

Hamilton, Peter. *Knowledge and Social Structure: An Introduction to the Classical Argument in the Sociology of Knowledge.* London: Routledge & Kegan Paul, 1974.

Hesse, Mary. *Revolutions and Reconstructions in the Philosophy of Science.* Bloomington: Indiana University Press, 1980.

Jacob, J. R. *Robert Boyle and the English Revolution.* New York: Burt Franklin, 1977.

———, and Jacob, M. C. "The Anglican Origins of Modern Science: The Metaphysical Foundation of the Whig Constitution." *Isis* 71(1980):251–67.

Jacob, M. C. *The Newtonians and the English Revolution.* Ithaca, N.Y.: Cornell University Press, 1976.

Kuhn, Thomas. *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press, 1970.

Lenoir, Timothy. *The Strategy of Life: Teleology and Mechanics in Nineteenth Century German Biology.* Dordrecht: Reidel, 1982.

———. "Social Interests and the Organic Physics of 1847." Paper presented at a colloquium, Van Leer Foundation, Jerusalem, 1983.

Merton, Robert K. "Paradigm for a Sociology of Knowledge." In *The Sociology of Science: Theoretical and Empirical Investigations*, edited and introduced by N. W. Storer. Chicago: University of Chicago Press, 1973a [1945].

———. "Science and the Social Order." In *The Sociology of Science.* 1973b [1938].

———. "Normative Structure of Science." In *The Sociology of Science.* 1973c [1942].

Mulkay, Michael. *Science and the Sociology of Knowledge.* London: George Allen & Unwin, 1979.

Mulkay, Michael, and Milič, V. "The Sociology of Science in East and West." *Current Sociology* 28(1980):23–42.

Polanyi, Michael. *The Logic of Liberty.* London: Routledge & Keegan Paul, 1951.

Remmling, Gunter. *Road to Suspicion.* Chicago: University of Chicago Press, 1967.

Stark, Werner. *The Sociology of Knowledge.* London: Routledge & Kegan Paul, 1958.

Turner, R. Stephen. "The Growth of Professional Research in Prussia, 1818–1848, Causes and Contexts." *Historical Studies in the Physical Sciences* 3(1971):137–82.

Webster, Charles. *The Great Instauration: Science, Medicine and Reform, 1626–1660.* London: Duckworth, 1975.

Zloczower, A. *Career Opportunities and the Growth of Scientific Discovery in Nineteenth Century Germany, with Special Reference to the Development of Physiology.* New York: Arno Press, 1981.

Network Theory

Allison, Graham T. *Essence of Decision: Explaining the Cuban Missile Crisis.* Boston: Little, Brown, 1971.

Bailey, Kenneth D. "Cluster Analysis." In *Sociological Methodology*, edited by David R. Heise. San Francisco: Jossey-Bass, 1975.

Blau, Peter. *Exchange and Power in Social Life.* New York: Wiley, 1964.

———. "Parameters of Social Structure." *American Sociological Review* 39(1974):615–35.

———. *Approaches to the Study of Social Structure.* New York: Free Press, 1975.

Burt, Ronald S. *Toward a Structural Theory of Action: Network Models of Social Structure, Perception and Action.* New York: Academic Press, 1982.

Coleman, James S. "Loss of Power." *American Sociological Review* 38(1973):1–17.

Deutsch, Karl W. *The Nerves of Government: Models of Political Communication and Control.* New York: Free Press, 1966.

Galaskiewicz, Joseph. *Exchange Networks and Community Politics.* Beverly Hills, Calif.: Sage, 1979.

Heinz, John P., and Laumann, Edward O. *Chicago Lawyers: The Social Structure of the Bar.* New York: Russell Sage Foundation/American Bar Foundation, 1982.

Homans, George C. *The Human Group.* New York: Harcourt, Brace, 1950.

Knoke, David, and Laumann, Edward O. "The Social Organization of National Policy Domains: An Exploration of Some Structural Hypotheses." In *Social Structure and Network Analysis*, by Peter V. Marsden and Nan Lin. Beverly Hills, Calif.: Sage, 1982.

———. "Issue Publics in National Policy Domains." Paper presented at the annual meeting of the American Sociological Association. Detroit, Michigan, August 31–September 4, 1983.

Kruskal, Joseph B., and Wish, Myron. *Multidimensional Scaling.* Beverly Hills, Calif.: Sage, 1978.

Laumann, Edward O. *Prestige and Association in an Urban Community.* Indianapolis: Bobbs-Merrill, 1966.

———. *Bonds of Pluralism: The Form and Substance of Urban Social Networks.* New York: Wiley, 1973.

———, and Knoke, David. "On the Nature of Influence in Large Systems: The Cases of the National Policy Domains of Energy and Health." Paper presented at the third annual Sunbelt Social Network Conference, San Diego, California, February 10–14, 1983.

Laumann, Edward O.; Knoke, David; and Kim, Yong-Hak. "An Organizational Approach to State Policy Formation: A Comparative Study of Energy and Health Domains." *American Sociological Review* 50(1985):1–19.

Laumann, Edward O.; Marsden, Peter V.; and Prensky, David. "The Boundary Specification Problem in Network Analysis." *Applied Network Analysis: A Methodological Introduction*, edited by R. S. Burt and M. J. Minor. Beverly Hills, Calif.: Sage, 1983.

Laumann, Edward O., and Pappi, Franz U. *Networks of Collective Action: A Perspective on Community Influence Systems.* New York: Academic Press, 1976.

Lazarsfeld, Paul F.; Berelson, Bernard; and Gaudet, Hazel. *People's Choice.* New York: Columbia University Press, 1944.

Levine, Joel H. "The Sphere of Influence." *American Sociological Review* 37 (1971):14–27.

Lipset, Seymour M. *The First New Nation.* New York: Basic Books, 1963 [1963a].

———. "The Sources of the 'Radical Right,' Three Decades of the Radical Right: Coughlinites, McCarthyites, and Birchers." In *The Radical Right*, edited by Daniel Bell. Garden City, N.Y.: Doubleday, 1963 [1963b].

Lorrain, François, and White, Harrison C. "Structural Equivalence of Individuals in Social Networks." *Journal of Mathematical Sociology* 1(1971):49–80.

Marsden, Peter V., and Laumann, Edward O. "Collective Action in a Community Elite: Exchange, Influence Resources, and Issue Resolution." In *Power, Paradigms, and Community Research*, edited by R. J. Liebert and A. W. Imersheim. London: Sage, 1977.

———. "Mathematical Ideas in Social Structural Analysis." *Journal of Mathematical Sociology* 10(1984):271–94.

Mayhew, Leon. *Society, Institutions and Activity.* Glenview, Ill.: Scott, Foresman, 1971.

Nadel, S. F. *The Theory of Social Structure.* London: Cohen and West, 1957.

Parsons, Talcott. *The Social System.* Glencoe, Ill.: Free Press, 1951.

Pool, Ithiel de Sola. "Tracking the Flow Information." *Science* 221(1983): 609–13.

Prensky, David. *Interorganizational Structure and Organizational Participation in a National Policy Domain.* Unpublished doctoral dissertation, University of Chicago, 1985.

Rado, Leslie Ann. "Cultural Elites and the Institutionalization of Ideas: A Case Study." Paper presented at the annual meeting of the American Sociological Association, San Francisco, California, 1982.

Sheingold, Carl A. "Social Networks and Voting: The Resurrection of a Research Agenda." *American Sociological Review* 38(1973):712–20.

White, Harrison C.; Boorman, Scott A.; and Brieger, Ronald L. "Social Structure from Multiple Networks: I. Block-Models of Roles and Positions." *American Journal of Sociology* 81(1976):730–80.

Structural Theory

Bancroft, H. H. *History of California, 1848–1855.* Works, vol. 23. San Francisco, 1888 [1888a].

———. *Popular Tribunals.* Works, vol. 36. San Francisco, 1888 [1888b].

Ben-Porath, Yoram. "The F-connection: Families, Friends, and Firms, and the Organization of Exchange." *Population and Development Review* 6(1980):1–29.

De Alessi, Louis. "The Economics of Property Rights: A Review of the Evidence." *Research in Law and Economics* 2(1980):1–47.

Hamilton, Gary. "The Structural Sources of Adventurism: The Case of the California Gold Rush." *American Journal of Sociology* 83(1978):1466–90.

Shinn, Charles H. "Land Laws and Mining Districts." *Johns Hopkins University Series in Historical and Political Science.* 2nd ser., vol. 12, December 1884.

Shinn, Charles H. *Mining Camps,* New York, 1885.

Trimble, William J. "The Mining Advance into the Inland Empire." *Bulletin of the University of Wisconsin.* History Series, vol. 3, no. 2, 1914.

Umbeck, John R. *A Theory of Property Rights (with Applications to the California Gold Rush).* Ames: Iowa State University Press, 1981.

Williamson, Oliver E. "The Economics of Organization: The Transaction Cost Approach." *American Journal of Sociology* 87(1981):548–77.

Purposive Action Theory

Atkinson, A. B. *Social Justice and Public Policy.* Cambridge, Mass.: MIT Press, 1983.

Becker, Gary S. "Human Capital and the Personal Distribution of Income: An Analytical Approach." Woytinsky Lecture no. 1. Ann Arbor: Institute of Public Administration, University of Michigan, 1967.

———. "A Theory of Social Interactions." *Journal of Political Economy* 82(1974):1063–93.

———. *Human Capital.* 2nd ed. New York: Columbia University Press, for the National Bureau of Economic Research, 1975.

———. *A Treatise on the Family.* Cambridge, Mass.: Harvard University Press, 1981.

———, and Barro, Robert. "A Reformulation of the Economic Theory of Fertility," presented at the Conference on Causes and Consequences of Non-Replacement Fertility, Hoover Institution, Stanford, California, November 7–9, 1985.

Becker, Gary S., and Tomes, Nigel. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy* 84(1976):S143–162.

———. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy* 87(1979):1153–89.

———. "Human Capital and the Rise and Fall of Families." Conference on the Family and the Distribution of Economic Rewards, September 21–22, 1984, Snowbird, Utah.

Behrman, Jere, and Taubman, Paul. "Interrelationship of Education Across Three Generations." Center for Household and Family Economics, University of Pennsylvania, 1983.

Bevan, D. L. "Inheritance and the Distribution of Wealth." *Economica* 46(1979):381–402.

———, and Stiglitz, J. E. "Intergenerational Transfers and Inequality." *Greek Economic Review* 1(1979):6–26.

Bloom, Benjamin S. *Human Characteristics and School Learning.* New York: McGraw-Hill, 1976.

Bowles, Samuel. "Schooling and Inequality from Generation to Generation." *Journal of Political Economy* 80(1972):S219–51.

Ehrlich, Isaac, and Ben-Zion, Uri. "Asset Management, Allocation of Time, and Returns to Saving." *Economic Inquiry* 14(1976):558–86.

Goldberger, Arthur S. "Family Data Analysis: Assortment, Selection, and Transmission." Proposal to the National Science Foundation, 1979.

Jensen, Arthur R. "How Much Can We Boost IQ and Scholastic Achievement?" *Harvard Educational Review* 39(1969):1–123.

Mandelbrot, Benoit. "Paretian Distributions and Income Maximization." *Quarterly Journal of Economics* 76(1962):57–85.

Meade, J. E. *The Just Economy.* Albany: State University of New York Press, 1976.

Roy, A. D. "The Distribution of Earnings and of Individual Output." *Economic Journal* 60(1950):489–505.

Tinbergen, Jan. "A Positive and a Normative Theory of Income Distribution." *Review of Income and Wealth* 16(1970):221–34.

Tomes, Nigel. "The Family, Inheritance, and the Intergenerational Transmission of Inequality." *Journal of Political Economy* 89(1981):928–58.

Ecological Theory

Aldrich, Howard E. *Organizations and Environments.* Englewood Cliffs, N.J.: Prentice-Hall, 1979.

Arrow, Kenneth J. *The Limits of Organization.* New York: Norton, 1974.

Astley, W. Graham, and Van de Ven, Andrew. "Central Perspectives and Debates in Organization Theory." *Administrative Science Quarterly* 28(1983):245–73.

Blau, Peter M., and Scott, W. Richard. *Formal Organizations.* San Francisco: Chandler, 1962.

Brittain, Jack, and Freeman, John. "Organizational Proliferation and Density-Dependent Selection." In *Organizational Life Cycles*, edited by John Kimberly and Robert Miles. San Francisco: Jossey-Bass, 1980.

Burawoy, Michael. *Manufacturing Consent: Changes in the Labor Process Under Monopoly Capitalism.* Chicago: University of Chicago Press, 1979.

Carroll, Glenn R. "Concentration and Specialization: Dynamics of Niche Width in Populations of Organizations." *American Journal of Sociology* 90(1985):1262–83.

———, and Delacroix, Jacques. "Organizational Mortality in the Newspaper Industries of Argentina and Ireland: An Ecological Approach." *Administrative Science Quarterly* 27(1982):169–98.

Chandler, Alfred D. *The Visible Hand: The Managerial Revolution in American Business.* Cambridge, Mass.: Belknap, 1977.

Child, John, and Kieser, Alfred. "Development of Organizations over Time." In *Handbook of Organizational Design*, vol. 1, edited by William H. Starbuck and Paul C. Nystrom. Oxford: Oxford University Press, 1981.

DiMaggio, Paul J., and Powell, Walter W. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48(1983):147–60.

Downs, Anthony. *Inside Bureaucracy.* Boston: Little, Brown, 1967.

Edwards, Richard. *Contested Terrain: The Transformation of the Work Place in the Twentieth Century.* New York: Basic Books, 1979.

Etzioni, Amitai. *The Comparative Analysis of Complex Organizations.* 2nd ed. New York: Free Press, 1975.

Freeman, John. "Organizational Life Cycles and Natural Selection Processes." In *Research in Organizational Behavior,* vol. 4, edited by Barry M. Staw and Lawrence L. Cummings. Greenwich, Conn.: JAI Press, 1982.

———; Carroll, Glenn R.; and Hannan, Michael T. "The Liability of Newness: Age-Dependence in Organizational Death Rates." *American Sociological Review* 48(1983):692–710.

Freeman, John, and Hannan, Michael T. "Niche Width and the Dynamics of Organizational Populations." *American Journal of Sociology* 88(1983):1116–45.

Hannan, Michael T., and Freeman, John. "The Population Ecology of Organizations." *American Journal of Sociology* 82(1977):929–64.

Hawley, Amos H. "Human Ecology." In *International Encyclopedia of the Social Sciences*, edited by David L. Sills. New York: Macmillan, 1968.

Hermann, C. F. "Some Consequences of Crisis Which Limit the Viability of Organizations." *Administrative Science Quarterly* 8(1963):61–82.

Janowitz, Morris. *The Professional Soldier.* Glencoe, Ill.: Free Press, 1960.

Ladde, G. S., and Siljak, D. D. "Stability of Multispecies Communities in Randomly Varying Environments." *Journal of Mathematical Biology* 2(1976):165–78.

Lawrence, Paul, and Lorsch, Jay. *Organization and Environment.* Cambridge, Mass.: Harvard University Press, 1967.

McKelvey, Bill. *Organizational Systematics.* Berkeley: University of California Press, 1982.

March, James G. "Footnotes on Organizational Change." *Administrative Science Quarterly* 26(1981):563–97.

———, and Olsen, Johan P. *Ambiguity and Choice in Organizations.* Bergen: Universitetsforlaget, 1976.

May, Robert M. *Stability and Complexity in Model Ecosystems.* 2nd ed. Princeton, N.J.: Princeton University Press, 1974.

Mayhew, L. B. *Surviving the Eighties.* San Francisco: Jossey-Bass, 1979.

Meyer, John W., and Rowan, Brian. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83(1977): 340–63.

Nelson, Richard R., and Winter, Sidney G. *An Evolutionary Theory of Economic Change.* Cambridge, Mass.: Belknap, 1982.

Perrow, Charles. *Complex Organizations: A Critical Essay.* 2nd ed. Glencoe, Ill.: Scott, Foresman, 1979.

Pfeffer, Jeffrey, and Salancik, Gerald. *The External Control of Organizations: A Resource Dependence Perspective.* New York: Harper & Row, 1978.

Scott, W. Richard. *Organizations: Rational, Natural, and Open Systems.* Englewood Cliffs, N.J.: Prentice-Hall, 1981.

Selznick, Philip. "Foundations of the Theory of Organizations." *American Sociological Review* 13(1948):25–35.

Siljak, D. D. "When Is a Complex Ecosystem Stable?" *Mathematical Bioscience* 25(1975):25–50.

Simon, Herbert A. "The Architecture of Complexity." *Proceedings of the American Philosophical Society* 106(1962):67–82.

Stinchcombe, Arthur S. "Social Structure and Organizations." In *Handbook of Organizations*, edited by James G. March. Chicago: Rand McNally, 1965.

Thompson, James D. *Organizations in Action.* New York: McGraw-Hill, 1967.

Williamson, Oliver E. *Markets and Hierarchies: Analysis and Antitrust Implications.* New York: Free Press, 1975.

Weber, Max. *Economy and Society: An Outline of Interpretive Sociology.* 3 vols. New York: Bedminster Press, 1968.

Weick, Karl. "Educational Organizations as Loosely Coupled Systems." *Administrative Science Quarterly* 21(1976):1–19.

Interpretive Sociology

Aries, Phillipe. *Centuries of Childhood.* New York: Vintage Books, 1962.

Barnes, Barry, and Edge, David, eds. *Science in Context.* Cambridge, Mass.: MIT Press, 1982.

Beauchamp, Dan. *Beyond Alcoholism.* Philadelphia: Temple University Press, 1981.

Becker, Howard S. "Becoming a Marihuana User." *American Journal of Sociology* 59(1953):235–42.

Berger, Peter, and Luckmann, Thomas. *The Social Construction of Reality.* Garden City, N.Y.: Doubleday, 1966.

Bloor, David. *Knowledge and Social Imagery.* London: Routledge & Kegan Paul, 1976.

Burke, Kenneth. *A Grammar of Motives.* New York: Prentice-Hall, 1945.

Cameron, Tracy. "Alcohol and Traffic." In M. Arens et al., *Alcohol, Casualties and Crime.* Report C-18. Berkeley, Calif.: Alcohol Research Group, 1977.

Cicourel, Aaron. *Method and Measurement in Sociology.* Glencoe, Ill.: Free Press, 1964.

———. *The Social Organization of Juvenile Justice.* New York: Wiley, 1968.

———. "Interpretive Procedures and Normative Roles in the Negotiation of Status and Role." In A. Cicourel, *Cognitive Sociology.* Hammondsworth, England: Penguin Education, 1973.

Conrad, Peter, and Schneider, Joseph. *Deviance and Medicalization.* St. Louis: Mosby, 1980.

Davis, Murray. *Smut: Erotic Reality/Obscene Ideology.* Chicago: University of Chicago Press, 1983.

Douglas, Jack. *The Social Meanings of Suicide.* Princeton, N.J.: Princeton University Press, 1967.

Eliade, Mircea. *Myth and Reality.* New York: Harper & Row, 1963.

Emerson, Joan. "Behavior in Private Places: Sustaining Definitions of Reality in Gynecological Examinations." In *Recent Sociology 2: Patterns of Communicative Behavior*, edited by H. Dreitzel. London: Macmillan, 1970.

Foucault, Michael. *Madness and Civilization.* New York: Random House, 1965.

———. *Birth of the Clinic.* New York: Vintage Books, 1973.

Garfinkel, Harold. *Studies in Ethnomethodology.* Englewood Cliffs, N.J.: Prentice-Hall, 1967.

Geertz, Clifford. "From the Native's Point of View: On the Nature of Anthropological Understanding." In *Interpretive Social Science*, edited by P. Rabinow and W. Sullivan. Berkeley: University of California Press, 1979.

Goffman, Erving. *Frame Analysis: An Essay on the Organization of Experience.* Cambridge, Mass.: Harvard University Press, 1974.

Gusfield, Joseph. *The Culture of Public Problems: Drinking-Driving and the Symbolic Order.* Chicago: University of Chicago Press, 1981.

———. "Prevention: Rise, Decline and Renaissance." In *Alcohol, Science and Society Revisited*, edited by E. Gomberg, H. White, and J. Carpenter. Ann Arbor: University of Michigan Press, 1982.

———; Kotarba, Joseph; and Rasmussen, Paul. "The Social Control of Drinking-Driving: An Ethnographic Study of Bar Settings." *Law and Policy* 6(January 1984):45–66.

Hoy, David. *The Critical Circle.* Berkeley: University of California Press, 1978.

Latour, Bruno, and Woolgar, Stephen. *Laboratory Life: The Social Construction of Scientific Fact.* Los Angeles: Sage Books, 1979.

Levine, Harry G. "The Discovery of Addiction." *Journal of Studies on Alcohol* 39(1978):143–74.

Lindesmith, Alfred. *Opiate Addiction.* Bloomington, Ind.: Principia Press, 1947.

Mauss, Armand. *Social Problems as Social Movements.* Philadelphia: Lippincott, 1975.

Mead, George H. *Mind, Self and Society.* Chicago: University of Chicago Press, 1934.

Mehan, Hugh. *Learning Lessons.* Cambridge, Mass.: Harvard University Press, 1978.

Ricoeur, Paul. "Metaphor and the Main Problem of Hermeneutics." In *The Philos-*

ophy of Paul Ricoeur, edited by C. Reagan and D. Stewart. Boston: Beacon Press, 1978.

———. "The Model of the Text: Meaningful Action Considered as a Text." In *Interpretive Social Science*, edited by P. Rabinow and W. Sullivan. Berkeley: University of California Press, 1979.

Room, Robin. "Governing Images of Alcohol and Drug Problems." Doctoral dissertation, Department of Sociology, University of California, Berkeley, 1978.

Schutz, Alfred. *Selected Writings.* Edited by H. Wagner. Chicago: University of Chicago Press, 1970.

Spector, Malcolm, and Kitsuse, John. *Constructing Social Problems.* Menlo Park, Calif.: Cummings, 1977.

Weinberg, Ian. "Social Problems That Are No More." In *Handbook on the Study of Social Problems*, edited by E. Smigel. Chicago: Rand McNally, 1974.

Wolfe, A. C. "Characteristics of Late Night, Week-end Drivers." In *Alcohol, Drugs, and Traffic Safety*, edited by S. Israelstam and S. Lambert. Toronto: Addiction Research Foundation, 1975.

Wolfe, Kurt. "Phenomenology and Sociology." In *A History of Sociological Analysis*, edited by T. Bottomore and R. Nisbet. London: Heinemann, 1979.

Organization Theory

Abell, Derek F. *Defining the Business: The Starting Point of Strategic Planning.* Englewood Cliffs, N.J.: Prentice-Hall, 1980.

———, and Hammond, John S. *Strategic Market Planning.* Englewood Cliffs, N.J.: Prentice-Hall, 1979.

Anthony, Robert N. *Planning and Control Systems: A Framework for Analysis.* Cambridge, Mass.: Harvard University, Division of Research, Graduate School of Business Administration, 1965.

———, and Dearden, John. *Management Control Systems.* 4th ed. Homewood, Ill.: Irwin, 1980.

Arrow, Kenneth J. *The Limits of Organization.* New York: Norton, 1974.

Bailey, Elizabeth E., and Friedlander, Ann F. "Market Structure and Multiproduct Industries." *Journal of Economic Literature* 20(1982):1034–48.

Braudel, Fernand. *The Wheels of Commerce.* New York: Harper & Row, 1982.

Chandler, Alfred D., Jr. *Strategy and Structure.* Cambridge, Mass.: M.I.T. Press, 1962.

———. *The Visible Hand.* Cambridge, Mass.: Harvard University Press, 1977.

Corey, E. Raymond. *Industrial Marketing: Cases and Concepts.* 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall, 1976.

———. *Procurement Management: Strategy, Organization and Decision-Making.* Boston: CBI Publishing, 1978.

Cuff, Robert D. *The War Industries Board.* Baltimore: Johns Hopkins University Press, 1973.

Davis, Stanley M., and Lawrence, Paul R. *Matrix.* Reading, Mass.: Addison-Wesley, 1977.

deRoover, Raymond. *The Rise and Decline of the Medici Bank 1397–1494.* New York: Norton, 1966.

Eccles, Robert G. "Control with Fairness in Transfer Pricing." *Harvard Business Review*, November–December 1983, pp. 149–61.

Eccles, Robert G., and Carley, Kathleen M. "The Distribution and Exchange of Resources in the Multi-Profit Center Firm." Working paper, 1983.

———, and Dougherty, J. Stewart. *VISA International: The Management of Change.* Case 0-482-022, HBS Case Services. Cambridge, Mass.: Harvard University, Graduate School of Business Administration, 1981.

Grant, John H., and King, William R. "Strategy Formulation: Analytical and Normative Models." In *Strategic Management: A New View of Business Policy and Planning*, edited by Dan E. Schendel and Charles W. Hofer. Boston: Little, Brown, 1979.

Hamermesh, Richard G. *The Administration of Strategy: A Study of the Impact of Portfolio Planning.* Unpublished draft, n.d.

Haspeslagh, Philippe. "Portfolio Planning: Uses and Limits." In *Strategic Management*, edited by Richard G. Hamermesh.

Henderson, Bruce. *Henderson on Corporate Strategy.* Cambridge, Mass.: Abt Books, 1980.

Hirshleifer, Jack. "On the Economics of Transfer Pricing." *Journal of Business* 29(1956):172–84.

National Association of Cost Accountants. *Proceedings of the Sixth International Cost Conferences.* New York: Little and Ives, 1925.

Porter, Michael. *Competitive Strategies: Techniques for Analyzing Industries and Competitors.* New York: Free Press, 1980.

Reece, James S., and Cool, William R. "Measuring Investment Center Performance." *Harvard Business Review*, May-June 1978, pp. 28–49.

Salter, M. A., and Weinhold, W. A. *Diversification through Acquisition.* New York: Free Press, 1979.

Stinchcombe, Arthur L. "Contracts as Hierarchical Documents." Draft manuscript, Stanford Graduate School of Business, November 1982.

Udy, Stanley. *Work in Traditional and Modern Society.* Englewood Cliffs, N.J.: Prentice-Hall, 1970.

Vancil, Richard F. *Decentralization: Managerial Ambiguity by Design.* Homewood, Ill.: Dow Jones-Irwin, 1979.

Wallerstein, Immanuel. *The Modern World System: Vol. II, 1600–1750.* New York: Academic Press, 1980.

White, Harrison C. "Production Markets as Induced Role Structures." In *Sociological Methodology*, edited by Samuel L. Leinhardt. New York: Jossey-Bass, 1981 [1981a].

———. "Where Do Markets Come From?" *American Journal of Sociology*, 87(1981):517–47 [1981b].

Williamson, Oliver E. *Markets and Hierarchies.* New York: Macmillan, 1975.

Woodward, Joan. *Industrial Organization: Theory and Practice.* London: Oxford University Press, 1965.

Theory of Social Change

Addams, Jane. *Twenty Years at Hull-House with Autobiographical Notes.* New York: Macmillan, 1910.

Asher, Robert. "Business and Workers' Welfare in the Progressive Era: Workmen's Compensation Reform in Massachusetts, 1880–1911." *Business History Review* 43(1969):452–75.

Baltes, Paul B., Reese, H. W.; and Lipsitt, L. P. "Life-Span Developmental Psychology." *Annual Review of Psychology* 31(1980):65–110.

Berkowitz, Edward, and McQuaid, Kim. *Creating the Welfare State.* New York: Praeger, 1980.

Birch, A. H. *Federalism, Finance and Social Legislation in Canada, Australia and the United States.* London: Oxford University Press, 1955.

Brandeis, Elizabeth. "Labor Legislation." In *History of Labor in the United States, 1896–1932*, vol. 3, by Don Lescohier and Elizabeth Brandeis. New York: Macmillan, 1935.

Brandes, Stuart. *American Welfare Capitalism, 1880–1940.* Chicago: University of Chicago Press, 1976.

Brody, David. "The Rise and Decline of Welfare Capitalism." In *Workers in Industrial America: Essays on the Twentieth Century Struggle.* New York: Oxford University Press, 1980.

Buenker, John D. *Urban Liberalism and Progressive Reform.* New York: Norton, 1978.

Cameron, David. "The Expansion of the Public Economy: A Comparative Analysis." *American Political Science Review* 72(1978):1243–61.

Castles, Francis G. *The Social Democratic Image of Society.* London: Routledge & Kegan Paul, 1978.

———. "The Impact of Parties on Public Expenditures." In *The Impact of Parties*, edited by Francis Castles. Beverly Hills, Calif.: Sage, 1982.

Cohen, Emmeline. *The Growth of the British Civil Service, 1970–1939.* London: George Allen and Unwin, 1941.

Collier, David, and Messick, Richard. "Prerequisites Versus Diffusion: Testing Alternative Explanations of Social Security Adoption." *American Political Science Review* 69(1975):1299–1315.

Collins, D. "The Introduction of Old Age Pensions in Great Britain." *Historical Journal* 8(1965):246–49.

Cutright, Phillips. "Political Structure, Economic Development and National Social Security Programs." *American Journal of Sociology* 70(1965):537–50.

Douglas, Roy. *The History of the Liberal Party, 1895–1970.* London: Sidgwick and Jackson, 1971.

Featherman, David L. "The Life-Span Perspective in Social Science Research." In *Life-Span Development and Behavior*, vol. 5, edited by P. B. Baltes and O. G. Brim, Jr. New York: Academic Press, 1983.

———. "Individual Development and Aging as a Population Process." In *Individual Development and Social Change: Explanatory Analysis*, edited by J. Nesselroade and A. von Eye. New York: Academic Press, 1985.

———. "Biography, Society and History." In A. B. Sørensen, F. Weinert and L. Sherrod (eds.), *Human Development and the Life Course: Multidisciplinary Perspectives*, Hillsdale, N.J.: Erlbaum, in press.

———, and Lerner, Richard M. "Ontogenesis and Sociogenesis: Problematics for Theory and Research About Development and Socialization over the Lifespan." Unpublished manuscript. University of Wisconsin, Madison, 1985.

Featherman, David L., and Petersen, Trond. "Markers of Aging: Modelling the Clocks That Time Us." Unpublished manuscript. University of Wisconsin, Madison, 1985.

Finer, Herman. *The British Civil Service.* London: George Allen and Unwin, 1937.

Flora, Peter, and Alber, Jens. "Modernization, Democratization and the Development of Welfare States in Western Europe." In *The Development of Welfare States in Europe and America*, edited by Peter Flora and Arnold Heidenheimer. New Brunswick, N.J.: Transaction Books, 1981.

Gilbert, Bentley. *The Evolution of National Insurance in Great Britain: The Origins of the Welfare State.* London: Michael Joseph, 1966.

Greaves, H. R. *The Civil Service in the Changing State.* London: Harrap, 1947.

Gronbjerg, Kirsten; Street, David; and Suttles, Gerald D. *Poverty and Social Change.* Chicago: University of Chicago Press, 1978.

Hanham, H. J. *Elections and Party Management: Politics in the Time of Disraeli and Gladstone.* London: Longmans, Green, 1959.

Harris, Jose. *Unemployment and Politics.* London: Oxford University Press, 1972.

Heclo, Hugh. *Modern Social Politics in Britain and Sweden.* New Haven: Yale University Press, 1974.

Henderson, Charles. *Industrial Insurance.* Chicago: University of Chicago Press, 1909.

Hennock, E. Peter. "The Origins of British National Insurance and the German Precedent, 1880–1914." In *The Emergence of Welfare States in Britain and Germany, 1850–1950*, edited by W. J. Mommsen. London: Croom Helm, 1981.

Jackman, Robert W. *Politics and Social Equality: A Comparative Analysis.* New York: Wiley, 1975.

Kaim-Caudle, P. *Comparative Social Policy and Social Security.* London: Martin Robertson, 1973.

Katzenstein, Peter. "The Determinants of Policy in Switzerland and Austria." Paper presented at the SSRC Conference on States and Social Structures, Mt. Kisco, New York, 1982.

———. "Economic Dependence and Political Autonomy: The Small European States in the International Economy." In *The Antimonies of Interdependence*, edited by John Ruggie. New York: Columbia University Press, 1983.

Katznelson, Ira. *City Trenches.* New York: Pantheon Books, 1981.

Keller, Morton. *Affairs of State: Public Life in Nineteenth Century America.* Cambridge, Mass.: Harvard University Press, 1977.

Kerr, Clark; Dunlop, John; Harbison, Frederick; and Myers, Charles. *Industrialism and Industrial Man.* New York: Oxford University Press, 1964.

Korpi, Walter. *The Working Class in Welfare Capitalism.* London: Routledge & Kegan Paul, 1978.

———, and Shalev, Michael. "Strikes, Power and Politics in the Western Nations, 1900–1976." *Political Power and Social Theory* 1(1980):301–34.

Leff, Mark. "Consensus for Reform: The Mothers'-Pension Movement in the Progressive Era." *Social Service Review* 47(1973):397–417.

Lubove, Roy. *The Struggle for Social Security, 1900–1935.* Cambridge, Mass.: Harvard University Press, 1968.

McCormick, Richard. "The Party Period and Public Policy: An Exploratory Hypothesis." *Journal of American History* 66(1979):279–98.

———. "The Discovery that Business Corrupts Politics: A Reappraisal of the Origins of Progressivism." *American Historical Review* 86(1981):247–74.

McGill, Barry. "Francis Schnadhorst and the Liberal Party Organization." *Journal of Modern History* 34(1962):19–39.

McMurry, Donald M. "The Political Significance of the Pension Question, 1885–1897." *Mississippi Valley Historical Review* 9(1922):19–36.

Malloy, James M. *The Politics of Social Security in Brazil.* Pittsburgh: University of Pittsburgh Press, 1979.

Mann, Arthur. "British Social Thought and American Reformers of the Progressive Era." *Mississippi Valley Historical Review* 42(1956):672–92.

Martin, A. *The Politics of Economic Policy in the United States.* Beverly Hills, Calif.: Sage, 1973.

Namier, Sir Louis. *The Structure of Politics at the Accession of George III.* London: Macmillan, 1961.

Nisbet, Robert. *Social Change and History.* New York: Basic Books, 1969.

Ogus, A. I. "Great Britain." In *The Evolution of Social Insurance, 1881–1981*, edited by Peter A. Kohler, Hans F. Zacher, and Martin Partington. New York: St. Martin's Press, 1982.

Orloff, Ann, and Skocpol, Theda. "Why Not Equal Protection?: Explaining the Politics of Public Social Spending in Britain, 1900–1911, and the United States, 1880s–1920." *American Sociological Review* 49(1984):726–50.

Pryor, F. L. *Public Expenditure in Communist and Capitalist Nations.* London: George Allen and Unwin, 1968.

Report of the Commission on the Support of Dependent Minor Children of Widowed Mothers. Massachusetts House No. 2075. Boston: Wright and Potter, 1913.

Rimlinger, Gaston. *Welfare Policy and Industrialization in Europe, America and Russia.* New York: Wiley, 1971.

Roberts, David. *Victorian Origins of the Welfare State.* New Haven: Yale University Press, 1960.

Rogers, Daniel T. "In Search of Progressivism." *Reviews in American History* 10(1982):113–32.

Schiesl, Martin. *The Politics of Efficiency: Municipal Administration and Reform in America: 1880–1920.* Berkeley: University of California Press, 1977.

Seager, Henry. *Social Insurance.* New York: Macmillan, 1910.

Shalev, Michael. "The Social Democratic Model and Beyond: Two Generations of

Comparative Research on the Welfare State." *Comparative Social Research* 6(1983):315–51 [1983a].

———. "Social Protection, the Labor Movement and the State." Unpublished paper, Departments of Sociology and Political Science, Hebrew University of Jerusalem, 1983 [1983b].

Shefter, Martin. "Party and Patronage: Germany, England and Italy." *Politics and Society* 7(1977):403–51.

———. "Party Bureaucracy and Political Change in the United States." In *Political Parties: Development and Decay*, edited by Louis Maisel and Joseph Cooper. Beverly Hills, Calif.: Sage, 1978.

———. "Trade Unions and Political Machines: The Organization and Disorganization of the American Working Class in the Late Nineteenth Century." In *Working Class Formation: Nineteenth Century Patterns in Western Europe and the United States*, edited by Ira Katznelson and Aristide Zolberg. Princeton, N.J.: Princeton University Press, in press.

Skocpol, Theda. "Bringing the State Back In: False Leads and Promising Starts in Current Theories and Research." In *Bringing the State Back In*, edited by Peter Evans, Theda Skocpol, and Dietrich Rueschemeyer. Cambridge: Cambridge University Press, in press.

———, and Ikenberry, John. "The Political Formation of the American Welfare State in Historical and Comparative Perspective." *Comparative Social Research* 6(1983):87–148.

Skowronek, Stephen. *Building a New American State: The Expansion of National Administrative Capacities, 1877–1920.* New York: Cambridge University Press, 1982.

Smellie, K. B. *A Hundred Years of English Government.* 2nd ed. London: Duckworth, 1950.

Spalding, Rose J. "Welfare Policymaking: Theoretical Implications of a Mexican Case Study." *Comparative Politics* 12(1980):419–38.

Stephens, John D. *The Transition from Capitalism to Socialism.* London: Macmillan, 1979.

Thomson, David. *England in the Nineteenth Century, 1815–1914.* Baltimore: Penguin Books, 1950.

Tilly, Charles, "Big Structures, Large Processes, Huge Comparisons." Working paper no. 295. Ann Arbor: Center for Research on Social Organization, 1983.

Tishler, H. D. *Self-Reliance and Social Security, 1870–1917.* Port Washington, N.Y.: Kennikat Press, 1971.

Treble, James H. "The Attitudes of Friendly Societies Towards the Movement in Great Britain for State Pensions." *International Review of Social History* 15(1970):266–99.

Tuma, Nancy, and Hannan, M. T. *Social Dynamics: Models and Methods.* New York: Academic Press, 1984.

Webb, Beatrice. *My Apprenticeship.* London: Longmans, Green, 1926.

Webb, R. K. *Modern England: From the 18th Century to the Present.* New York: Dodd, Mead, 1970.

Werner, H. "The Concept of Development from a Comparative and Organismic

Point of View." In *The Concept of Development*, edited by D. B. Harris. Minneapolis: University of Minnesota Press, 1957.

Wiebe, Robert. *The Search for Order, 1877–1920.* New York: Hill and Wang, 1967.

Wilensky, Harold. *The Welfare State and Equality: Structural and Ideological Roots of Public Expenditures.* Berkeley: University of California Press, 1975.

———. *The "New Corporatism," Centralization and the Welfare State.* Beverly Hills, Calif.: Sage, 1976.

———. "Leftism, Catholicism, and Democratic Corporatism: The Role of Political Parties in Recent Welfare State Development." In *The Development of Welfare States in Europe and America*, edited by Peter Flora and Arnold J. Heidenheimer. New Brunswick, N.J.: Transaction Books, 1981.

Sociolinguistics

Abdel-Jawad, Hassan. "Phonological and Social Variation in Arabic in Amman." Unpublished dissertation, University of Pennsylvania, 1981.

Allen, Harold B. "The Use of Atlas Informants of Foreign Parentage." In *Lexicography and Dialect Geography: Festgake for Hans Kurath,* edited by H. Scholler and J. Reidy. Wiesbaden: F. Steiner, 1973.

Bailey, Charles-James N. *Variation and Linguistic Theory.* Washington, D.C.: Center for Applied Linguistics, 1973.

Baugh, John. "Linguistic Style-Shifting in Black English." Unpublished dissertation, University of Pennsylvania, 1979.

Cheshire, Jenny. *Variation in an English Dialect.* Cambridge: Cambridge University Press, 1982.

Chomsky, Noam. *Lectures on Government and Binding.* Dordrecht: Foris, 1981.

Cloward, Richard A., and Ohlin, Lloyd E. *Delinquency and Opportunity: A Theory of Delinquent Gangs.* Glencoe, Ill.: Free Press, 1960.

Cofer, Thomas. "Linguistic, Variability in a Philadelphia Speech Community." Unpublished dissertation, University of Pennsylvania, 1972.

Deutsch, Karl W. *Nationalism and Social Communication: An Inquiry into the Foundations of Nationality.* 2nd ed. Cambridge, Mass.: M.I.T. Press, 1966.

Fischer, John L. "Social Influences on the Choice of a Linguistic Variant." *Word* 14(1958):47–56.

Fishman, Joshua. "Some Contrasts Between Linguistically Homogeneous and Linguistically Heterogeneous Policies." *Language Problems of Developing Nations*, edited by J. Fishman, C. Ferguson, and J. Das Gupta. New York: Wiley, 1968.

Gal, Susan. *Language Shift: Social Determinants of Linguistic Change in Bilingual Austria.* New York: Academic Press, 1980.

Goffman, Erving. *Forms of Talk.* Philadelphia: University of Pennsylvania Press, 1981.

Graff, David; Labov, William; and Harris, Wendell A. "Testing Listeners' Reactions to Phonological Markers." Paper presented at the 12th Conference on New Ways of Analyzing Variation, Montreal, 1983.

Grimshaw, Allen. "Talk and Social Control." In *Social Psychology: Sociological Perspectives*, edited by M. Rosenberg and R. Turner. New York: Basic Books, 1981.

Gumperz, John, and Wilson, Robert. "Convergence and Creolization: A Case from the Indo-Aryan/Dravidian Border." In *Language in Social Groups*, by J. Gumperz. Stanford, Calif.: Stanford University Press, 1971.

Holmquist, Jonathan. "Language Loyalty and Linguistic Variation in Spanish Cantabria." Doctoral dissertation, Princeton University, 1982.

Labov, William. "The Social Motivation of a Sound Change." *Word* 19(1963):273–309.

———. *The Social Stratification of English in New York City.* Washington, D.C.: Center for Applied Linguistics, 1966.

———. "Crossing the Gulf Between Sociology and Sociolinguistics." *American Sociologist* 13(1978):93–103.

———. "The Social Origins of Sound Change." In *Locating Language in Time and Space*, edited by W. Labov. New York: Academic Press, 1980.

———. "Objectivity and Commitment in Linguistic Science: The Case of the Black English Trial in Ann Arbor." *Language in Society* 11(1982):165–202.

———; Cohen, P.; Robins, C.; and Lewis, J. "A Study of the Non-standard English of Negro and Puerto Rican Speakers in New York City." Cooperative Research Report 3288. Vols. 1 and 2. Philadelphia: U.S. Regional Survey, 1968.

Labov, William; Yaeger Malcah; and Steiner, Richard. "A Quantitative Study of Sound Change in Progress." Philadelphia: U.S. Regional Survey, 1972.

Labov, William, et al. "The Social Determinants of Sound Change in Philadelphia." Philadelphia: U.S. Regional Survey, 1982.

Lambert, Wallace. "A Social Psychology of Bilingualism." *Journal of Social Issues* 23(1967):91–109.

Lieberson, Stanley. "Bilingualism in Montreal: A Demographic Analysis." *American Journal of Sociology* 71(1965):10–25.

Mock, Carol C. "The Social Maturation of Pronunciation: A Family Case Study." *Learner in the Process* 1,3(1979):23–37.

Modaressi, Yahya. "A Sociolinguistic Investigation of Modern Persia." Unpublished dissertation, University of Kansas, 1978.

Myhill, John, and Harris, Wendell A. "The Use of the Verbal -s Inflection in BEV." Paper presented at the 12th Conference on New Ways of Analyzing Variation, Montreal, 1983.

Sacks, Harvey; Schegloff, Emanuel A.; and Jefferson, Gail, "A Simplest Systematics for the Organization of Turn-Taking in Conversation." *Language* 50(1974): 696–735.

Sankoff, David, and Labov, William. "On the Uses of Variable Rules." *Language in Society* 8(1979):189–222.

Shopen, Tim, and Wald, Benji. "The Use of (ing) in Australian English." Manuscript, 1982.

Spears, Arthur. "Come: A Modal-like Form in Black English." *Language* 58(1980):850–72.

Torrey, Jane. "Black Children's Knowledge of Standard English." *American Educational Research Journal* 20(1983):627–43.

Trudgill, Peter. "Sex, Covert Prestige and Linguistic Change in Urban British English." *Language in Society* 1(1972):179–95.

———. "Linguistic Change and Diffusion: Description and Explanation in Sociolinguistic Dialect Geography." *Language in Society* 3(1974):215–46.

Weinreich, Uriel; Labov, William; and Herzog, Marvin. "Empirical Foundations for a Theory of Language Change." In *Directions for Historical Linguistics*, edited by W. Lehmann and Y. Malkiel. Austin: University of Texas Press, 1968.

Winfold, Donald. "Phonological Hypercorrection in the Process of Decreolization—the Case of Trinidadian English." *Journal of Linguistics* 14(1978):277–91.

Social Psychology

Averett, Christine P. "Attribution of Interpersonal Qualities: An Affect Control Theory Approach." Doctoral dissertation, Department of Sociology, University of North Carolina, 1981.

Gollob, Harry F. "Impression Formation and Word Combination in Sentences." *Journal of Personality and Social Psychology* 10(1968):341–53.

Heise, D. R. "Affectual Dynamics in Simple Sentences." *Journal of Personality and Social Psychology* 11(1969):204–13.

———. "Potency Dynamics in Simple Sentences." *Journal of Personality and Social Psychology* 16(1970):48–54.

———. "Social Action as the Control of Affect." *Behavioral Science* 22(1977):163–77.

———. *Computer-Assisted Analysis of Social Action.* Chapel Hill, N.C.: Institute for Research in Social Science, 1978.

———. *Understanding Events: Affect and the Construction of Social Action.* New York: Cambridge University Press, 1979.

———. "Growing Knowledge Trees." *Micro: The 6502/6508 Journal* 48(1982):67–82.

———, and Roberts, E. P. "The Development of Role Knowledge." *Genetic Psychology Monographs* 82(1970):83–115.

Heise, D. R., and Smith-Lovin, Lynn. "Impressions of Goodness, Powerfulness, and Liveliness from Discerned Social Events." *Social Psychology Quarterly* 44(1981):93–106.

MacKinnon, Neil J. "Canadian Dictionaries of Social Identities and Interpersonal Behaviors." Research report, Department of Sociology and Anthropology, University of Guelph, 1983.

McCall, George J., and Simmons, J. L. *Identities and Interactions: An Examination of Human Associations in Everyday Life.* Rev. ed. New York: Free Press, 1978.

Marwell, G., and Hage, J. "The Organization of Role Relationships: A Systematic Description." *American Sociological Review* 35(1970):884–900.

Osgood, Charles E. "Studies of the Generality of Affective Meaning Systems." *American Psychologist* 17(1962):10–28.

———, May, W. H.; and Miron, M. S. *Cross-Cultural Universals of Affective Meaning.* Urbana: University of Illinois Press, 1975.

Parsons, Talcott. *The Social System.* New York: Free Press, 1951.

Smith-Lovin, Lynn, and Heise, D. R. "A Structural Equation Model of Impression Formation." In *Multivariate Applications in the Social Sciences*, edited by Nancy Hirschberg and L. G. Humphreys. Hillsdale, N.J.: Lawrence Erlbaum, 1982.

———. "Impressions, Expectations, and Behavior." Unpublished manuscript, Department of Sociology, University of South Carolina, 1983.

———, and Willigan, D. "Belfast, Northern Ireland, Dictionaries of Social Identities and Interpersonal Behaviors." Research report, Department of Sociology, University of South Carolina, 1983.

Spradley, James P. *The Ethnographic Interview.* New York: Holt, Rinehart, & Winston, 1979.

Stryker, Sheldon. *Symbolic Interactionism.* Menlo Park, CA: Benjamin/Cummings, 1980.

Thomas, W. I. "Situational Analysis: The Behavior Pattern and the Situation." In *W. I. Thomas on Social Organization and Social Personality*, edited by M. Janowitz. Chicago: University of Chicago Press, 1927.

Triandis, Harry C. *Interpersonal Behavior.* Monterey, Calif.: Brooks/Cole, 1977.

Werner, Oswald, and Fenton, J. "Method and Theory in Ethnoscience or Ethnoepistemology." In *A Handbook of Method in Cultural Anthropology*, edited by R. Naroll and R. Cohen. Garden City, N.Y.: Natural History Press, 1970.

Wiggins, Beverly A. "Intentions, Expectations, and Behavior: A Test of Affect Control Theory." Doctoral dissertation, Department of Sociology, University of North Carolina, 1980.

Social Movements

Erikson, Robert; Goldthorpe, John: Portocarero, Luciene. "International Class Mobility in Three Western European Societies." *British Journal of Sociology* 30(1979):415–41.

———. "Social Fluidity in Industrial Nations. England, France and Sweden." *British Journal of Sociology* 33(1982):1–34.

Hutt, W. H. *The Economics of the Colour Bar;* a study of the economic origins and consequences of social segregation in South Africa. London: A. Deutsch, 1964.

Landes, David. *The Unbound Prometheus: Technical Change and Industrial Development in Western Europe from 1750 to the Present.* London, Cambridge University Press, 1969.

Nehru, Jawarharlal. *The Discovery of India.* New York, Doubleday, Inc., 1946.

Olson, Mancur. *The Logic of Collective Action.* Cambridge: Harvard University Press, 1965.

———. *The Rise and Decline of Nations.* New Haven: Yale University Press, 1982.

Treiman, Donald J., and Terrill, Kermit. "The Process of Status Attainment in the United States and Great Britain." *American Journal of Sociology* 81(1975):563–83.

Vanneman, Reeve D. "U.S. and British Perceptions of Class." *American Journal of Sociology* 85(1980):769–90.

Wiener, Martin. *English Culture and the Decline of the Industrial Spirit, 1850–1980.* London, Cambridge University Press, 1981.

Closing Address

Becker, Gary. *The Economic Approach to Human Behavior.* Chicago: University of Chicago Press, 1976.

Blau, Peter M., and Duncan, O. *The American Occupational Structure.* New York: Wiley, 1967.

Coleman, James M. *Individual Interests and Collective Action.* Cambridge: Cambridge University Press, 1985.

———; Katz, E.; and Menzel, H. *Medical Innovation.* Indianapolis: Bobbs-Merrill, 1966.

Crozier, M. *The Bureaucratic Phenomenon.* Chicago: University of Chicago Press, 1964.

Durkheim, Emile. *Suicide.* New York: Free Press, 1951.

Gale, David, and Shapley, Lloyd. "College Admissions and the Stability of Marriage." *American Mathematical Monthly* 69(1962):9–15.

Jencks, Christopher, et al. *Inequality.* New York: Basic Books, 1972.

Katz, G., and Lazarsfeld, Paul. *Personal Influence.* New York: Free Press, 1955.

Olson, Mancur. *The Rise and Decline of Nations.* New Haven: Yale University Press, 1982.

Riker, W. H., and Ordeshook, P. C. *An Introduction to Positive Political Theory.* Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Rogers, E. *Diffusion of Innovations.* New York: Free Press, 1962.

Roth, Alvin E. "The Evaluation of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory." Mimeographed. University of Pittsburgh, November, 1983.

Sanderson, Warren. "Economic-Demographic Simulation Models: A Review of Their Usefulness for Policy Analysis." In *Laxenburg, Austria: IIASA Reports Vol. 1, No. 2."* Laxenburg, Austria, 1980.

Schoen, Robert. "Measuring the Tightness of a Marriage Squeeze." *Demography* 20(1983):61–78.

Weber, Max. *The Protestant Ethic and the Spirit of Capitalism.* New York: Scribner, 1958.

Contributors

Robert McCormick Adams, *Smithsonian Institution*

Howard Aldrich, *University of North Carolina*

Gary S. Becker, *University of Chicago*

Joseph Ben-David, *University of Chicago and Hebrew University**

Peter M. Blau, *Columbia University*

Ronald S. Burt, *Columbia University*

James S. Coleman, *University of Chicago*

Randall Collins, *University of California, Riverside*

Robert G. Eccles, *Harvard University*

David L. Featherman, *University of Wisconsin*

John Freeman, *University of California, Berkeley*

Allen Grimshaw, *Indiana University*

Joseph R. Gusfield, *University of California, San Diego*

Michael T. Hannan, *Cornell University*

Russell Hardin, *University of Chicago*

Michael Hechter, *University of Arizona*

David R. Heise, *Indiana University*

Morris Janowitz, *University of Chicago*

* Deceased

John Kitsuse, *University of California, Santa Cruz*

David Knoke, *University of Minnesota*

William Labov, *University of Pennsylvania*

Edward O. Laumann, *University of Chicago*

Siegwart Lindenberg, *University of Gröningen*

Arthur Mann, *University of Chicago*

Anthony Oberschall, *University of North Carolina*

Mancur Olson, *University of Maryland*

Ann Shola Orloff, *University of Wisconsin*

John Padgett, *University of Chicago*

Edward Shils, *University of Chicago*

Theda Skocpol, *Harvard University*

Arthur L. Stinchcombe, *Northwestern University*

Fred L. Strodtbeck, *University of Chicago*

Nigel Tomes, *University of Western Ontario*

Immanuel Wallerstein, *State University of New York, Binghamton*

Harrison C. White, *Harvard University*

Name Index

Numbers in italics indicate statements by conference participants.

Subject Index