Lessons Learned from the History of Social Indicators

Article · January 1998

CITATIONS
119

READS
409

3 authors, including:

Clifford Cobb
American Journal of Economics and Sociology
13 PUBLICATIONS **462** CITATIONS

SEE PROFILE

REDEFINING PROGRESS

CLIFFORD W. COBB AND CRAIG RIXFORD

LESSONS LEARNED FROM THE HISTORY OF SOCIAL INDICATORS

NOVEMBER 1998

REDEFINING PROGRESS
ONE KEARNY STREET, FOURTH FLOOR
SAN FRANCISCO, CA 94108

1.800.896.2100

TEL 415.781.1191
FAX 415.781.1198
E-MAIL info@rprogress.org
WEBSITE www.rprogress.org

ACKNOWLEDGEMENTS

The authors would like to thank Rob Allen, Monty Hempel, Tony Hodge, Maureen Kennedy, and Greg Oliver for providing useful comments on an earlier version of this paper.

The Community Indicators Project is grateful to the Ford Foundation and the James Irvine Foundation for their support of this publication.

Design: ReVerb, Los Angeles

CLIFFORD W. COBB

Clifford Cobb has been at Redefining Progress (RP) since its founding in 1994. He pioneered an alternative to the Gross National Product and was one of the key researchers and co-authors for the cover story in the *Atlantic Monthly*, "If the Economy Is Up, Why Is America Down?", published in October 1995. He has also conducted research on an alternative tax system for the state of California, and will be a contribution to RP's upcoming report on environmental tax reform in California.

Mr. Cobb's work has been published widely. He contributed to the ground-breaking book, *For the Common Good: Redirecting the Economy Toward Community, the Environment, and a Sustainable Future*, written by Herman Daly and John Cobb, Jr., published in 1989. He is also co-author of *The Green National Product* and author of *Responsive Schools, Renewed Communities*, an analysis of education vouchers from a communitarian perspective. Mr. Cobb served in 1995 as a nonvoting member of President Clinton's Council on Sustainable Development. He is a graduate of the University of California at Berkeley, where he earned a Master's degree in Public Policy.

CLIFFORD W. COBB
SENIOR FELLOW
REDEFINING PROGRESS
cobb@jps.net

CRAIG RIXFORD

Craig Rixford has been at Redefining Progress since its inception in 1994. As a research associate for the Indicators Program, he contributed research to the *Genuine Progress Indicator,* co-authored the *Community Indicators Handbook*, and conducted research on the history of social indicators. He currently manages Redefining Progress's computer and information systems. Mr. Rixford holds a B.A. in French with a minor in anthropology from San Francisco State University.

CRAIG RIXFORD
REDEFINING PROGRESS
rixford@rprogress.org

TABLE OF CONTENTS

I. INTRODUCTION

Over the past two centuries, nations and communities have become increasingly reliant on statistical measures to assess their status and to set policies. Successive generations have endeavored to gather and use social statistics, or indicators, in a more systematic and purposeful manner. Technically speaking, an indicator refers to a set of statistics that can serve as a proxy or metaphor for phenomena that are not directly measurable. However, the term is often used less precisely to mean any data pertaining to social conditions.[1]

We hope that by looking at the history of social indicators, we can avoid earlier mistakes and promote better strategies for indicators' development.[2] Indicator movements often begin with extreme optimism, aspiring to guide society rationally. Later, when social change does not seem to flow clearly from the effort, practitioners can become unduly pessimistic.

The realistic hope is that social indicators will enable nations and communities to face their condition more effectively. An annual report on the well-being of a community or nation has the potential to focus energy on problems, analyze their interrelationships, and attack them in a systematic manner. But based on past experience, these goals are not likely to be reached if the production of social indicators is confined to collecting statistical data, publishing chartbooks and reports, and publicizing findings.

In the United States, there have been presidential commissions to set national goals, the publication of social trend reports, and efforts to institutionalize a social reporting system. Those activities largely failed to make social indicators practical tools, but the lessons provided by those failures can lead to success.

1. Although the term "social indicators" can be used narrowly, to distinguish them from economic and environmental indicators, we intend that it be construed broadly to encompass all types of measures, except where we explicitly focus on the distinction between social and economic indicators.

2. We have focused primarily on the history in United States, but this does not imply that significant developments have not occurred elsewhere. In fact, the impetus for the renewed interest in indicators in the U.S. is due in large part to continuing efforts in other countries.

II. FROM DESCRIPTION TO ANALYSIS: Making Indicators More Outcome Oriented

Indicators are invariably developed with the intent of changing some outcome in the world. Even academic researchers, who often work with abstract models and who seem the least interested in their application to practical problems, usually believe that their work will effect change by improving our understanding of problems.

At a minimum, indicators can serve to inform the public and policymakers by providing statistical information about a condition that might otherwise be overlooked. Through this process, indicators can influence policy outcomes by expanding awareness and focusing attention.[3]

However, indicator practitioners are often frustrated by the lack of change brought about by the production of indicator reports. Raising awareness is not their ultimate goal; they want to achieve a more direct link between indicators and outcomes. For that, we believe they need an *analytic* approach to indicators, one that involves developing a theory, not simply conveying social trends.

Whereas the descriptive approach asks, "What social conditions exist?" the analytic approach raises the underlying question, "Why do those conditions exist?" Analyzing the reasons for a problem is, of course, more difficult than merely pointing out that the problem exists. Yet, without a theory about the cause of a problem, a solution can rarely be found.

This report is intended to improve outcomes by outlining some general principles gleaned from the past that continue to bear on the development of indicators. Our strongest recommendation will be to make a transition from *description* to *analysis*. In order to move from indicators to action, projects must examine the causes behind the symptoms, a process that could lead the indicators movement in a new direction.

3. This report does not look at the benefits associated with the process of developing indicators. Some point to the indirect benefits of using indicators, such as the creation of new social networks in the group developing the indicators. However important this is as an outcome, it does not necessarily result from the indicators per se.

Judith Innes (1990) points to the importance of the "social construction" of indicators. According to this perspective, knowledge, in the form of indicators, needs to be relevant to the community it is intended to influence. The process of developing indicators can lead to a new shared understanding of a problem among those involved in the effort. Since the tacit assumptions of such an understanding often frame a problem and set bounds for possible solutions, this is a crucial outcome (35). Innes provides a provocative introduction to this topic (1–47). However, these positive results will not necessarily flow from the process of doing indicators. It is our contention that it is more often an analytical approach, rather than a descriptive one, that leads to a new understanding.

III. CONFLICTING PRINCIPLES IN THE HISTORY OF INDICATORS

In the following section, we sketch out the chronology from which the principles are drawn. Doing so provides a context for the examples that illustrate the "lessons learned" in the third section. More important, it illustrates the inevitable conflicts that emerge in indicators work. Although there is no easy resolution to the conflicts, which remain unresolved in social science generally, it is helpful to bear them in mind.

The proponents of descriptive indicators claim that motivating decisionmakers or the public to do something is a key part of creating change. While this is true, it remains insufficient. Providing evidence about which policies may actually work is perhaps the most crucial step to create change.

This brings up the first and most basic conflict about the nature and purpose of indicators. Should they be aim to be *prescriptive*, providing guidance about what to do, or descriptive, simply highlighting conditions that might otherwise be overlooked?

A second dispute concerns the means used to develop social indicators. Those who advocate a *deductive* approach believe that social indicators should be based on abstract models that produce testable hypotheses. Thus, for example, indicators might be useful in testing various theories about the causes of poverty or homelessness rather than simply measuring their extent. Those who favor an *inductive* approach tend to compile data about social conditions before making generalizations. Even some of the strongest proponents of this approach have been frustrated by its failure to yield many insights useful for problem solving (e.g., Sheldon et al. 1983).

A third issue, largely resolved by 1900, is whether or not the process of collecting statistics should be impartial or nonideological. The conclusion, not surprisingly, was that objectivity is desirable. The issue that is not resolved was how any institution (or individual for that matter) can be truly impartial. Thus, the real conflict has always been between partisanship and pseudo-objectivity.[4]

A fourth conflict is about whether indicators should be primarily considered academic tools for understanding or rough guides for practical action. Historically, the academics hoped to keep indicators aloof from politics. They wanted the luxury of having

4. The absence of objectivity stems primarily from inevitable biases in the selection of topics on which to gather data. There are also hidden biases in techniques of gathering and publication of data. The pretense of objectivity stands in the way of public appreciation of those biases.

several decades of compiling data and conducting pure research, believing that publishing conclusions before that time would be premature. Practitioners, by contrast, could not wait for final answers. They needed to make judgments on the basis of models and information available in the present.

IV. A BRIEF HISTORY OF THE SOCIAL INDICATORS MOVEMENT

Although humans have been using indicators since the dawn of history, the self-conscious use of indicators to judge social conditions is of much more recent origin. In the 1830s, social reformers in Belgium, France, England, and the United States began using statistical indicators to improve public health and social conditions. In Europe, physicians and statisticians led the way in the development of social indicators. They looked for ways to understand the nature of epidemics in industrial cities. Using social components of census data, which were collected for the first time during this period, they formulated causal models that showed how disease was linked to poverty and other social conditions.

THE BEGINNINGS OF SOCIAL MEASUREMENT IN THE UNITED STATES

In the United States, religious groups and other social reformers were the leaders in the development of social indicators. The earliest known use of statistics to champion social change was for prison reform in Philadelphia in the 1810s (Cohen 1982, 170). The reformers produced tables of figures showing the number of people in jail still awaiting trial for each of the five previous years. Presumably their aim was to show that the justice system had broken down and that defendants were not receiving speedy trials. During this early period, however, there was little understanding of how to apply statistics to a debate over policy. The prison reformers mostly presented tables of figures on categories of crime and the race and sex of prisoners and assumed the numbers would speak for themselves.

By the 1830s, the temperance movement had begun to learn how to make inferences from statistics. They used statistics not only to show the extent of drinking problems, but they also hoped to prove that alcohol was the *cause* of crime, moral depravity, poverty, and economic wastefulness. Gathering data from poorhouses and jails in four states, Samuel Chipman sought to show that intemperance promoted poverty and criminal behavior (Cohen 1982, 212). Other temperance advocates calculated the number of acres of grain devoted to alcohol production—like a modern Ecological Footprint analysis—in order to show that alcohol was economically wasteful. But this latter sort of argument laid the temperance movement open to the counterargument from the United

States Brewers' Association that liquor taxes paid for much of the care of the poor and thus were socially beneficial (220). Thus, the two-sided nature of statistics was discovered at an early stage.

In the last third of the nineteenth century, political life in the U.S. and Europe was dominated by conflicts over wages, unemployment, and the conditions of the working class. The Massachusetts Bureau of Statistics of Labor was created in 1869 (Leiby 1960, 54), and the U.S. Bureau of Labor was established in 1884 (69). Labor statistics were among the first social statistics gathered on an official basis. The way the labor agencies defined their roles was instrumental in shaping social indicators for generations afterwards.

The first directors of the Massachusetts Bureau were labor sympathizers Henry Oliver and George McNeill. They sought information that would not only demonstrate the terrible conditions of workers but also show how entrenched interests perpetuated those conditions. Oliver and McNeill began with a vague, socialist hypothesis—that wages were set below the true value of labor—and set out to show its validity. They lacked the requisite methodology to prove their theory, but their efforts to use statistics to prove a point were plain. Propertied interests objected to the "strongly preconceived theories . . . with which the officers are incessantly forcing their facts to conform" (Leiby 1960, 61). Oliver and McNeill saw their role in managing a statistical agency not as neutral fact-finders, but as agents of social change. This partisan approach to data-gathering offended real estate, financial, and manufacturing interests. After the Labor Reform party collapsed in 1872, the governor replaced the first directors of the bureau.

The man who took over in 1873 was Carroll Wright.[5] Wright's desire to achieve objectivity in the gathering of statistics and to avoid theorizing put an indelible stamp on social indicators. In contrast to his predecessors, Wright saw his role as a purely neutral technician (Leiby 1960, 68). By appealing to "facts," Wright assumed that he could rise above partisan controversy. Yet, there was a subtle anti-labor ideology imbedded in his method. He was aware that his approach would thwart structural change. He argued publicly that the facts would prevent labor agitation. Although his impartiality was more apparent than real, he planted the seeds of pseudo-objectivity that continue to affect social indicators.

AN EARLY COMMUNITY INDICATORS MOVEMENT

Around 1910, the Russell Sage Foundation initiated the development of what are now called "community indicators" using processes that are remarkably like the ones that

5. Wright's reputation for evenhandedness in Massachusetts won him, in 1885, the position as the first U.S. Commissioner of Labor, a position he held until 1905.

have been re-established in the 1990s. Sage provided a grant to the Charity Organization Society (of New York) to survey industrial conditions in Pittsburgh (Smith 1991, 40–41). After the study was released in 1914, the Russell Sage Foundation was besieged with requests to fund similar studies in other cities. Since it did not have the funds to do that, the foundation provided technical advice instead. Partly as a result of this initiative, over two thousand local surveys were taken on education, recreation, public health, crime, and general social conditions.

Just as community indicators projects today are often associated with some existing organization, the surveys of this period were "conducted under the supervision of citizens' committees, church federations, chambers of commerce, or civic improvement associations" (Smith 1991, 41). The process by which this information was expected to affect decisionmaking was not much different from today either: "These groups then relayed the findings of the technical experts to the public who, enlightened by the facts, were expected to mobilize public opinion and press for appropriate reforms" (41).

Yet the collection of factual information could only yield observations about symptoms. In the absence of theories, the accumulation of facts could not offer any explanation of causes. "The surveys usually explained much less than met the eye. In reality, they were less an instrument for testing hypotheses and designing reforms than for arousing a community's conscience and 'quickening community forces' for reform, as one staff member of the foundation put it" (Smith 1991, 42). That is why "the actual political results [of the surveys] seldom lived up to the organizers' expectations" (41).

OGBURN'S *RECENT SOCIAL TRENDS*

The next major period in the shaping of social indicators was the 1920s and 1930s. As Secretary of Commerce from 1921 to 1929, Herbert Hoover improved national statistics on business and the economy, ultimately commissioning a report entitled *Recent Economic Changes in the United States.* As president from 1929 to 1933, he set up the Research Committee on Social Trends. The director of research, William F. Ogburn, a prominent sociologist from the University of Chicago, was charged with overseeing a comparable volume on recent social changes.[6] Other committee members included Wesley Mitchell and Charles E. Merriam.[7] All of these social scientists favored an inductive approach to understanding social problems (Bulmer 1983, 111).

6. Ogburn hoped to discern the causes of social change. He believed that statistical time-series would lead to more objective and reliable theories of change than would qualitative assessments. In particular, Ogburn was searching for underlying laws. For example, he held that technological change precedes cultural change, making it not only a driving force of social change but also a leading indicator of it.

7. Mitchell was a pioneer in the development of business cycle indicators at the National Bureau of Economic Research. Merriam was the founder of the Social Science Research Council (which was later to house the Center for Coordination of Social Indicator Research in the 1970s).

In 1933, the committee issued its mammoth report, *Recent Social Trends*. This report was hailed as a crowning achievement of social science by some and considered practically useless by others. Over 1,500 pages in length, it was the first official document devoted to social measurement, covering numerous social conditions such as demographics, health, and education. Although it served as an encyclopedic tome of social trends, it offered little insight into how to understand or solve the huge problems brought on by the Great Depression.

Ogburn, like Wright, believed that social reports should present facts, not opinions; they should present the data and trends, but refrain from interpretation and certainly from policy recommendations (Bulmer 1983, 114). His approach to studying social events—descriptive, inductive, pseudo-objective—came to dominate indicators work in subsequent years. A number of his students, including Albert D. Biderman, Otis Dudley Duncan, Albert J. Reiss Jr., and Eleanor Bernert Sheldon, were later active in the social indicators movement of the 1960s.

As the nation focused on the Depression, and then the Second World War, much more attention was paid to economic indicators. The development of a standardized measurement of unemployment and the use of surveys to gather the data were begun during the Depression as the government struggled to assess living conditions. Work continued on business cycle indicators as many sought to devise systems to predict and possibly thwart economic downturns.[8] The GDP was developed largely as a means of analyzing and organizing wartime production. For the time, the study of social trends was on the back burner.

THE SOCIAL INDICATORS MOVEMENT OF THE 1960s AND 1970s

By the early 1960s, as a result of the increasingly successful management of economic policy, the deductive or analytic approach to policymaking had effectively triumphed over the inductive or descriptive method favored by the students of Ogburn. The success of the Kennedy tax cut of 1964 and the apparent accuracy of econometric predictions of its effect on the economy made economists quite influential forces in guiding public policy.

The work of professional economists in government (such as the Council of Economic Advisers) and in policy institutes (such as Brookings) became a model for

8. The story of business cycle indicators illustrates the conflict between the inductivists and deductivists. On one hand, empiricists sought to look for patterns in the data and attempt to identify leading, coincident, and lagging indicators as a way of forecasting and to understand the nature of the business cycle. Two pioneers in the field, Wesley Mitchell and Arthur Burns, used this approach in their work. On the other hand, deductivists thought that it was futile to "measure without theory" and next to impossible devise policy solutions without a theory about what causes the business cycle. To a certain extent, the debate replayed itself in the social indicators movement of the 1960s.

applying social science to government policy. However, critics charged that economic considerations were given undue priority in policy decisions. If social theory and planning were given equivalent institutional support, they surmised that social policy could be rationalized in the way economic policy had been. Thus, the success of economic indicators was one spur to the birth of the social indicators movement in the 1960s and early 1970s.

The perceived limitations of economic thought and economic indicators also led to a call for greater support for social indicators. There was a "growing perception by policy makers and their advisors that the nation's rich array of economic statistics and related measures were simply inadequate indicators of emerging developments and issues under prevailing conditions of rapid social change and severe social strains" (Johnston 1989, 433). As these crises mounted during the 1960s, some politicians and social scientists began to champion the development of systems of social statistics comparable to the existing economic ones.

The event that signified the launching of the social indicators movement in the United States was the publication in 1966 of *Social Indicators*, a project sponsored by NASA (Bauer 1966).[9] Raymond Bauer, Albert Biderman, and Bertram Gross, the primary authors and influential forces in the budding social indicators movement, argued for increased collection of statistics that would be published as a social report. They also advocated the development of a system of social accounts which could help guide policy decisions.

A second influential publication, *Toward a Social Report*, was issued by the U.S. Department of Health and Welfare (HEW) in early 1969, on the last day of the Johnson Administration. It called for the establishment of an annual social report of the type advocated earlier by Bertram Gross in *Social Indicators*. This report was representative of the view that by definition social indicators should tell us if we are moving in the right direction, be relevant to setting policy, and help evaluate the effectiveness of social programs.

In the meantime, Senator Walter Mondale and others put forth legislation from 1967 to 1973 calling for the creation of a Council of Social Advisers (CSA), comparable to the Council of Economic Advisers (CEA) (Booth 1992, 380–85). The CSA was to issue an annual social report like the *Economic Report of the President*. Underlying this effort was the belief that the creation of the CEA had institutionalized the use of economic information and the power of economists. Creating a comparable institution to address social problems seemed like a logical next step.

Although the authors of *Social Indicators and Toward a Social Report* tended to use a descriptive approach, they ultimately saw indicators as tools for guiding public policy.

9. NASA was interested in determining the second-order consequences of the space program for American society.

The most profoundly descriptive approach was the work sponsored by the Russell Sage Foundation and the Social Science Research Council (SSRC). The foundation, which had funded the early community indicators movement described above, was also instrumental in promoting an inductive, descriptive approach to developing social indicators. In 1968 it published *Indicators of Social Change*, a volume edited by Wilbert Moore and Eleanor Bernert Sheldon (who was a program officer at the foundation). This volume was a successor to the *Recent Social Trends* volume of the 1930s in both content and spirit.

Sheldon and others opposed the creation of Mondale's Council of Social Advisers or any other premature application of social indicators to social policy. Instead, they argued, the pressing needs were basic research and better data series (Sheldon et al. 1983, 79). Since Sheldon was president of the SSRC (from 1972 to 1979), her views had an impact. Although SSRC established the Center for the Coordination on Social Indicators, it promoted basic research and advocated against the establishment of even an annual social report (83).

Following the academic approach she had inherited from William Ogburn, Sheldon believed that the development of a theoretical framework for indicators was premature. Social indicators could not follow in the footsteps of economic modeling since: 1) social goals were more ambiguous than economic ones, 2) social problems were less clearly understood than economic ones, and 3) the theoretical foundations of economics were much clearer than those underlying the analysis of social problems (Sheldon and Freeman 1970). Sheldon argued that an inductive approach was needed: First gather descriptive data, then develop the categories that would allow meaningful generalization and eventually work towards analysis of social change (103–5).

Another significant approach to improve social measurement was pioneered at this time: the use of perceptual indicators as an alternative way to understand changes in quality of life. The Russell Sage Foundation was a strong supporter of this research and sponsored the publication of a pioneering volume on the subjective measurement of well-being (Campbell and Converse 1972). This approach, which grew out of social psychology research, measures welfare by assessing an individual's personal interpretations of their own well-being. Subjective measurement showed a different dimension of quality of life than objective measures, such as housing conditions or income levels. As Kenneth Land summarizes: "The principle that the link between changes in objective conditions and psychological states is both indeterminate and sometimes paradoxical and therefore that subjective as well as objective states should be monitored is well-established . . . " (1992, 1846). Work on measuring perceptions has continued to blossom in the last few decades, much of it appearing in the journal *Social Indicators Research,* founded by Alex Michalos in 1974.

The social indicators movement entered the 1970s generally united in its use of a descriptive approach, but conflicted over the immediate goals of the movement. Work

on social indicators flourished in the United States in 1970s with thousands of relevant articles and books being published.[10] The idea was also quickly picked up by other nations and some international organizations, which created the sense of a social movement by the mid-1970s.

In the United States, three volumes of *Social Indicators* were published (U.S. Office of Management and Budget 1974; U.S. Census Bureau 1977, 1981). Although some had envisioned these as the beginning of institutionalized social reporting, their hopes were quickly dashed as political pressure within the Nixon Administration turned them into neutral chartbooks, replete with facts but void of interpretation. Viewed in retrospect, the publication of these volumes seems somewhat anticlimactic. According to Denis F. Johnston, who oversaw the development of the second two volumes of *Social Indicators,* "The federal effort to produce comprehensive social indicator reports was terminated after the publication of *Social Indicators III . . .*" (Johnston 1987, 299). Among the reasons he cites for this failure is "the inherent weakness of descriptive statistics in explaining social phenomena . . ."(299).

Thus the social indicators movement in the United States was effectively over by the early 1980s, although it continued in the form of annual statistical reports issued by government agencies (for example, *Health United States, and The Condition of Education*). Some data series that were started in part because of the movement (for example, the National Criminal Victimization Survey, and the American Housing Survey) also continued to be collected. Numerous explanations have been offered that account for the early demise (or hibernation) of the movement (based on Andrews 1989; Bulmer 1989; and Noll and Zapf 1994):

- Economic anxieties, which pushed aside concerns about social issues.
- The ideological shift toward conservatism and distrust of government (and a decrease in government support for social indicators research).
- The limited usefulness of social indicators to policymakers.
- The lack of a theoretical framework comparable to economic theory.
- The lack of an agreed-upon method of making normative judgments (about whether trends are good or bad).
- The lack of a common unit of measurement to permit aggregation, comparable to the use of money in economics.
- The declining faith in econometric modeling, which failed to avert rising inflation and unemployment.

10. Thousands of articles, reports, and books were published by 1972 when the first major bibliography of social indicators was published (Wilcox, et al. 1972). Thousands more had been written by the time a second major bibliography on social indicators was published in 1979 (Gilmartin et al.).

While the U.S. government was working on social indicators sporadically during the 1970s, several European nations, notably Great Britain, France, Germany, and the Netherlands, went further than the U.S. and institutionalized social reporting. Examples of European reports include *Données Sociales* in France, the *Social and Cultural Report* in the Netherlands, and *Social Trends* in the U.K. Among the preconditions that enabled this were "an articulated welfare-state program of social policy, an interventionist orientation of government, innovative statistical agencies and geographical centrality" (Noll and Zapf 1994, 5). Although support for social indicators has waxed and waned in Europe and Canada, it has been more constant than in the U.S.

RECENT DEVELOPMENTS

The excitement generated by the social indicators movement had effects on international agencies, who began to develop indicators as a part of their mission (Rothenbacher 1993). In the early 1970s, the Organization of Economic Co-operation and Development (OECD) established a program in social indicators which continued work until the publication of *Living Conditions in OECD Countries* in 1986. In the 1980s, many agencies developed annual reports that focused on social conditions. At the United Nations, the relationship between economic development and human development was explored via the Human Development Index. At the World Bank, the *World Development Report* and *Social Indicators of Development* reflected similar concerns, albeit from a much different vantage point. The World Health Organization (WHO) emphasis on human health led to the healthy cities movement, which developed indicators to assess improvements in public health broadly construed (Waddell 1995, 213–5).

The social indicators movement also inspired the development of environmental indicators. In the United States, the Council on Environmental Quality (CEQ) and the Environmental Protection Agency both began to develop indicators to monitor and publicize the state of the environment. The CEQ's *Environmental Trends,* first published in 1981, is an early example of this work. A similar effort began at OECD in the 1980s. In addition, policy institutes such as WorldWatch and World Resources Institute began producing annual books describing and analyzing environmental trends. Their regularly appearing volumes, *State of the World and World Resources* respectively, helped fill the void of environmental reporting by government agencies.

In the 1980s, work by researchers in Canada and the Netherlands led to the development of new approaches to environmental indicators, including the pressure-state-response framework. The notion of sustainable development, as highlighted by the Brundtland report and later by the Rio conference, brought a new framework for developing indicators in the early 1990s. These sustainability indicators emphasize the implications of current trends on the future. They also provide conceptual models that

illustrate the interrelationships among social, environmental, and economic concerns, although they don't necessarily offer an analytical understanding of them.

The social indicators movement has experienced a revival in the United States in the 1990s. A number of different frameworks guide the development of indicator reports. The quality of life and healthy cities approaches, like sustainability indicators, are also concerned with economic, social, and environmental conditions.[11] The sources of the renewed movement are somewhat different than in previous decades. There is far less impetus from federal agencies or major national institutions like the SSRC. Instead, the major focus in the 1990s has been on community indicators, similar to the work catalyzed by the Russell Sage Foundation around 1910.

11. For an introduction to some of these approaches, see Waddell 1995, Walters 1994, and Maclaren 1996.

V. THE LESSONS OF HISTORY FOR PRACTITIONERS TODAY

In the absence of an understanding of what has happened in the past, the new indicators movement may once again meet obstacles that have blocked progress of previous efforts. In the following section, we outline some of the lessons that can be learned from a study of how indicators have functioned historically.

In the anecdotes and stories we have drawn on, we have not limited ourselves to the experience of formal work in social indicators. We have instead considered the ways in which social statistics have been used and abused in the past. These are offered as examples that might help those who are engaged in the nitty-gritty problems of defining purposes, gathering data, and formulating appropriate indicators for their communities. It is our hope that the successes and failures of the past can be used to improve outcomes in the present and future.

LESSON 1

Having a number does not necessarily mean that you have a good indicator.

A common mistake in working on social indicators is to believe that because some official agency has measured something, an indicator based on that measure is likely to be valuable.

One way to think of indicators (and there are many ways) is to consider them quantities that reveal qualities. It is easy to find numbers that tell us the magnitude of something: the number of inhabitants of a city, vehicle miles traveled, the acres planted in various crops, the number of children vaccinated in a given year.

What is much harder to develop are numbers that tell us about quality. The best that can be done is to devise numerical measures from which a quality can be inferred. It is as if what we most want to measure is something that we cannot see if we look directly at it; we can see it only out of the corner of the eye.

Thus, if you want to know the health status of a segment of the population, you cannot measure that quality directly. But you might be able to infer something about it by surveying people to ask how they feel, by using public health data about the incidence of certain diseases, or by determining the amount of fat consumed, tobacco smoked, and exercise engaged in by that group.

Recognizing the elusive nature of quality is a far cry from saying that quality is subjective and that nothing can be said about it. It would be better to say that quality is always ambiguous and that any statements about it are provisional rather than final. This problem is not new. It has arisen repeatedly as people have tried to capture elusive qualities—such as the health of the economy or human intelligence—with quantitative measures.

Example 1 | Trying to measure the health of a nation's economy with a single number has always given misleading results. In the seventeenth century, foreign trade was becoming an important part of England's economy. It was also relatively easily measured: customs laws required that all ships be unloaded during the daytime so inspectors could inspect the cargo. As a result, a lot of information began to accumulate about trade (Cohen 1982, 29). (It actually took at least 50 years before someone realized it would be a good idea to centralize that information so that government could use it.) Since the amount of goods going in and out of England were almost the only economic statistics that were even partially available during much of the seventeenth century, it became common for writers to treat the "balance of trade" as if it represented the health of the economy as a whole. If exports were greater than imports, gold accumulated in the treasury. As a result, an increase in the stock of gold became viewed as a measure of the nation's well-being.[12]

That habit of mind has still not worn off after two centuries. Economists today may think production and income are better indicators of national economic health than gold inflows. But they are still thinking that quantity measures quality. In another two centuries, citizens will look back in amazement that we were so dazzled by production figures in our day. As better quality indicators are developed, older quantity measures seem out of place.

Example 2 | Educational testing also creates the illusion that the quality known as educational potential can be captured with a numerical score. Despite a stream of criticism, Scholastic Assessment Test (SAT) scores continue to be used as an indicator of the readiness of individual students to succeed in college. From 1926 until recently, the College Board and then the Educational Testing Service (ETS) created the misleading impression that the SAT was a disguised intelligence test (Crouse and Thusheim 1988, chapter 2). Nowadays, ETS no longer talks about aptitude or intelligence, but the SAT still maintains a public aura of being able to measure ability. Using it in that way perpetuates the image

12. The mercantilist philosophy measured economic health according to stockpiles of gold. When Spain used that measure in the seventeenth and eighteenth century, its economy plummeted from being the strongest in Europe to one of the weakest. From the perspective of a sustainable economy, this was the wrong indicator to use.

that college admissions are based on individual "merit" rather than family background. In effect, the supposedly neutral SAT has become the fig leaf behind which the effects of social class on educational opportunity are hidden. The problem is summarized by David Owen (1985) in his book on the SAT:

> If one decides to structure society according to the results of a test, that test becomes the blueprint for society . . . . The meritocratic impulse can be quite unmeritocratic. Far from being egalitarian, the new "fair" system perpetuates old injustices by making them look like the neutral workings of the merit market. The testing industry is the mighty engine of the status quo. The meritocracy, as interpreted by ETS, is eugenics by other means. (198–99)

LESSON 2
Effective indicators require a clear conceptual basis.

If you set out to create a good indicator, you need to spend time clarifying exactly what you are trying to measure. If you don't, you may end up with an indicator that measures something other that what you intended.

This may seem like obvious advice, but it is not easy to follow in practice. There is an understandable tendency for groups intent on developing indicators to start compiling data right away without a clear understanding of what needs to be measured. Taking the time to develop conceptual clarity seems to many people a kind of useless intellectual exercise; however, as the following examples show, a lack of clarity can lead to endless problems. Although measurement can help clarify a concept, the concept itself will not simply emerge from the data.

Example 1 | When Samuel Chipman visited New England prisons and poorhouses in the 1830s gathering information on the effects of alcohol consumption, he presumed the difference between a "temperate" and an "intemperate" person was obvious (Cohen 1982, 212). He asked jailers and wardens to apply those labels to inmates as if they were precise, neutral terms. Yet, in retrospect, it is clear that those terms were not true measures of the quantity of alcohol consumed, but generalized measures of approval or contempt. The same problem would apply today if one tried to gather data on the extent of addiction. In the minds of some people, anyone who uses an illegal drug would be classified as an "addict." Careless definitions can lead to misleading statistics and bad policies.

Example 2 | The 1930 census treatment of unemployment provides an example of an indicator being developed without a clear definition to guide it (Innes 1990, 126–127). Classifying someone as employed or unemployed is not as simple as it would at first seem. Should people be classified as unemployed if they worked at some time during the

previous year? If they work occasionally for a family-owned business? If they are on strike? Should they be included in the labor force if they are looking for work for the first time? If they "retired" early because they lost a job and couldn't find another one?

Without clear definitions, counting is impossible. As the nation drifted into its worst depression, the census takers were forced to formulate questions without a concept of what they were trying to measure. They placed people into seven categories with regard to employment without any prior definition of unemployment. The designers of the categories had hoped that the data could be aggregated after consensus was reached on the definition of unemployment (Innes 1990, 126,127).

After the data were collected, it became apparent that some of the categories contained both unemployed and employed persons. It had been hoped that the 1930 census would provide a benchmark against which to judge progress in lowering unemployment levels, but the ambiguity of the data prevented this. By failing to clarify their concept to begin with, they could never decipher the numbers they generated.

LESSON 3
There's no such thing as a value-free indicator.

Whatever anyone tells you to the contrary, all serious indicators work is political. The very act of deciding what to count and how to count it involves making value judgments. Because all indicators are laden with values or carry implicit messages, indicator reports really can not be neutral. Consideration of the values or concept underlying each indicator can lead to a more balanced presentation. But omitting analysis or interpretation does not make an indicators report neutral.

The formulation of survey questions is a complex process that incorporates value judgments in subtle nuances. Often the data from which indicators are derived are seriously biased because the surveyors have failed to consider the values of the people answering questionnaires. When you fill out a survey that asks your family income, do you sometimes feel suspicious or resentful about being asked for that sort of information? Most Americans do. If people know how information is likely to be used, that might also bias their responses. Professional survey designers know that. They have developed complex methods to deal with bias, but on some sensitive matters, there are no means of correcting for it.

Example 1 | The idea of a census now seems politically neutral to most citizens, but the term "census" comes from the same Latin root as "censure." In the Roman Empire, the office of censor was responsible not only for counting people and property and distributing food but for maintaining public morals (Duncan 1984, 51). In other words, when the census office required you to enroll, they were checking up on you. Later in Europe,

the same was true. Free cities in Renaissance Italy examined the status of households to determine who was contributing to society and who was likely to be a vagrant or thief. The former could be taxed, the latter expelled. Counting and judging were seen as one and the same thing. There was never a time when counting was regarded as a purely technical task. It always carried moral overtones by which the ruling class was imposing its values on the rest of society.

Example 2 | In the 1970s, some researchers talked about quality of life indicators as if they could be devised as purely technical measures (e.g., Liu 1975). Yet a quality of life measure cannot be carried out in a value-free manner. The choice of the components and their weighting are the most obvious ways in which such measures are governed by the values of the people who develop them. There is nothing wrong with value-laden measures of progress. It is essential, however, that the values and methods that go into constructing indicators be open for inspection so that others can determine if they share those values. Thus, for example, the Index of Social Health, developed by Marc Miringoff at Fordham University, describes in detail the components that are used in computing the index (Miringoff 1997). Similarly, most of the sustainability measures that have been adopted at the community level in the 1990s have sought to be transparent.

LESSON 4
Comprehensiveness may be the enemy of effectiveness.

A narrow range of indicators is more powerful than a laundry list. Historically, the most powerful indicators work has focused on a single issue. It has moved people to look beyond the most obvious features of a situation and to ask deeper questions than before. If an indicators project emphasizes more than two or three indicator categories, that is unlikely to happen. It is natural to explore all of the facets of society by using many indicators to paint a detailed picture. However, it is more effective to find a few insightful and compelling indicators that represent that complex whole.

A corollary to this is that the story told by the indicators is probably more important than the indicators themselves. A chartbook of indicators with little interpretation does not clearly depict the state of a community's health or otherwise hint at what needs to be done. The indicators may be of interest to specific audiences, but probably not to the general public. Perhaps a large number of indicators would be of interest if the geographic focus of the report is narrowed to a small area, such as a city block or a neighborhood. But even then, fewer is probably better than more.

Example 1 | The *Social Indicators* volumes published in the 1970s contained hundreds of indicators and graphs (U.S. Office of Management and Budget 1974; U.S. Bureau of the

Census 1977, 1980). But the reports had little effect partially because they contained so much information with no interpretation or analysis of overall trends. According to Denis F. Johnston, ". . . the original idea of a social report was abandoned in favor of a more 'factual' and 'neutral' mode of data presentation" (1989, 286). Instead, the Nixon and Carter Administrations authorized the publication of hundreds of pages of data, arranged solely by subject matter. The "significance of the data, [and] possible policy implications" were left up to the reader (286). To interpret would have been political. They correctly understood that the chartbooks would have no political impact.

Example 2 | Another case of a superabundance of indicators is the Oregon Benchmarks project. Beginning with 158 indicators, organized according to government program, the number increase to 272 over several years. To begin to make the benchmarks more meaningful and effective, the Oregon Progress Board shrank the number down to 92 and reorganized them according to the three primary goals of Oregon's Strategic Plan. Over the next two years the group plans to explore the cause and effect relationships between the indicators with an eye towards identifying key relationships and the indicators that best represent overall status of system (Oregon Progress Board 1997).

LESSON 5
The symbolic value of an indicator may outweigh its value as a literal measure.

Most of us are accustomed to thinking of accounting as pretty cut-and-dried. Numbers satisfy our desire for understanding when we are being literal-minded about information. "What's the bottom line?" is the question people often ask when they want just the straight facts.

Although numerical data often serve the purpose of reporting literal facts, it is important to keep in mind that numbers can also act as metaphors. That is especially true of index numbers which combine a variety of factors. In those cases, it is often not clear what the index number measures in concrete terms, so it can only function as a metaphor. This metaphorical function of an indicator may be what distinguishes it most from a statistic.

Example 1 | Efforts to measure the health of a whole economy or society invariably lead to numbers that are more metaphor than literal truth. The debate is then over whether the metaphor is a good one or not.

Gross domestic product (GDP) may be an appropriate measure in some contexts, such as when the Federal Reserve is trying to estimate the growth of the money supply in relation to market production. As a technical tool, GDP has its place. However, when GDP is used as a metaphor of well-being it fails utterly. It does not distinguish between

constructive expenditures and those that merely reflect spending to avoid the damage caused elsewhere in the economy. To draw attention to this defect in the metaphoric use of GDP, Redefining Progress created an alternative measure called the Genuine Progress Indicator (GPI). The GPI is not intended as a literal measure of well-being either. It is best understood as a better metaphor of progress than the GDP.

Example 2 | The use of the spotted owl as an indicator of biodiversity in old-growth forest should also be understood as a metaphor. This does not mean that the owls do not exist as literal animals or that they do not deserve protection as a species. The point is rather that that single species is symbolic of a wide range of ecological values and the overall health of a complex ecosystem. Like any good indicator, the vitality of the spotted owl population points beyond itself to a broader concern that is not easily measured.

LESSON 6
Don't conflate indicators with reality.

Indicators can help us understand complex situations by condensing an array of information into a simple number or graph. Yet there is always a danger that the indicator itself will be taken for the underlying reality. It is easy to become intellectually lazy and forget the complex process that created the number that serves as a proxy for the concept. This "hardening" of concepts then stands in the way of a supple understanding; the number becomes a barrier to the truth.

Every indicator is a flawed representative of a complex set of events. Confusing the statistic with the reality is all too common, but it should be avoided by those who care about creating high-quality indicators. Even the best indicator is only a fractional measurement of the underlying reality. One of the best ways to guard against this solidification of ideas is to try to develop multiple indicators for the same phenomenon. In this way, it is possible to remain constantly clear that no single indicator completely represents reality.

Example 1 | The modern history of the concept of intelligence has been beset by the failure to distinguish test results from the underlying reality. Educational policymakers have treated IQ scores as if they represented an indelible characteristic of each individual. Racial differences in test scores have been used to implement "tracking" (channeling by ability), which has amounted to a resegregation within many public schools.

The use of faulty measures of intelligence has been used for over a century to justify racism. In the 1840s, for example, Samuel George Morton provided what he regarded as compelling evidence of the intellectual inferiority of various people of color by demonstrating (with a flawed methodology) that their cranial capacity was smaller than

that of Indo-Europeans (Gould 1981, 69). Robert Bennett Bean published similar brain-size measurements in 1906. His mentor at Johns Hopkins, Franklin P. Mall, demonstrated the error in Bean's procedures, "but not before a leading journal had recommended that blacks be barred from voting as a consequence of their innate stupidity" (82).

Stephen Jay Gould also shows how a confused hereditary theory, combined with misuse of test scores as true measures of innate intelligence, led to continued pseudo-scientific validation of racist and nativist theories in the 1920s (1981, 192–233).

More recently, the controversial book *The Bell Curve,* which asserted the significance of measured IQ, sold 500,000 copies (Herrnstein & Murray 1996). Specifically, it charged that innate differences in intelligence are the cause of social disparities and that egalitarian policies cannot overcome those differences. Since the authors believe that racial differences in intelligence are also innate, their logic effectively endorses a two-tier society.

Our entire society has been shaped by the belief that standardized intelligence tests provide true measures of human intellectual capacity. This has not only affected the ways schools are organized, it has profoundly affected people's self-perceptions. One of the most valuable outcomes of the indicators movement would be a rejection of single measures of any quality. In the case of intelligence, our society would be far better served by taking seriously the multiplicity of kinds of intelligence (Gardner 1983, 1993).

Example 2 | Prior to the 1970s, the FBI's Uniform Crime Reports (UCR) were the only source available for crime statistics. These administrative statistics were highly unreliable as an indicator of crime trends (Innes 1990, 153–170). They were compiled from police reports which contained a variety of biases. Although they were authoritative they weren't always accurate. Some crimes such as rape tend to be underreported to police and therefore underrepresented in the UCR. Even homicide, seemingly the most unambiguous crime, is counted differently in different jurisdictions. Others, such as property theft, tend to be reported more because a police report is necessary for insurance. (Interestingly, in this case, an increase in the number of people who have homeowners' insurance may lead to an appearance of increase in property theft simply because the reporting rate is higher.)

The National Criminal Victimization Survey conducted by the Bureau of Justice Statistics was developed to address some of the deficiencies of the UCR (Biderman and Lynch 1991, 101–104). By surveying victims, a very different picture of crime emerges. Crimes that go unreported to the police show up, raising statistics for crimes that are embarrassing to the victims. On the other hand, asking people to remember events that took place over the previous six months leads to gaps in the record.

Neither one of the crime indicators measures the "true" incidence of crime. Each is likely to underrepresent or overrepresent different kinds of crime. Thus, it is inappropriate to use either of them as a standard.

LESSON 7

A democratic indicators program requires more than good public participation processes.

There is a tendency for people engaged in any community activity to believe that public review and stakeholder meetings are the best way to make the process democratic and just. Yet democracy is more than enabling large numbers of people to influence outcomes, and justice is more than procedures.

Many indicators groups seem to start with the implicit and tenuous assumption that procedural justice will automatically bring about substantive justice. According to this view, social indicators will lead to better social outcomes if they are developed by a broad representation of community members. In practice, however, an insistence on achieving a consensus of stakeholders or citizens usually produces a set of indicators that do little to challenge prevailing practices.

Rather than focusing exclusively on issues of procedural justice (ensuring that all groups are represented at stakeholder meetings), it would be useful for indicator projects to make substantive justice a priority. Charges of elitism are unlikely to arise if a project seeks to address issues of fair treatment in schools, job sites, housing, and the criminal justice system. That does not mean that procedural issues should be ignored. It simply means that widespread participation may not be the best "indicator" of whether an indicator project is really democratic.

Example 1 | Social reports have often had a political edge only when the authors did not try to reach a consensus. Reports on working conditions or public health in nineteenth century Europe and the United States were hard-hitting only when a commission was dominated by an individual willing to put forth an uncomfortable perspective that did not satisfy all sides (Cullen 1975, 35–40). The Massachusetts Labor Bureau, for example, advocated a pro-labor position for its first few years until the governor replaced its first directors with someone who could provide a more "balanced" presentation of the facts (Leiby 1960, 55–60).

Example 2 | The recent indicators project by Working Partnerships in conjunction with the Economic Policy Institute provides an example of how a local group can raise similar issues (Benner 1998). In its report "Growing Together or Drifting Apart?" the conditions of the "other side" of Silicon Valley are described in detail: the decline in wages of the

bottom 25 percent of the workforce and the continuing problems of workplace safety. Working Partnerships solicited community input and created an advisory board to assist in drafting the report, but it did not insist on gaining a consensus before publication.

Example 3 | The economic benchmarks program at the Center for Social and Urban Research of the University of Pittsburgh was able to analyze the racial disparities in that city forthrightly because it was not primarily concerned with consensus-building. Their report assesses the overall economic well-being of the whole population, but since their study revealed glaring inequities between blacks and whites in Pittsburgh they published a report focusing on that issue (Bangs and Hong 1995).

LESSON 8

Measurement does not necessarily induce appropriate action.

Indicators make sense as a tool only to the extent that they are part of a larger plan of action. It is possible that new information contained in indicators may change perceptions, but the connections to actions are not automatic.

After an indicators project produces and promotes its final report, the natural question arises: What next? The authors of the report want some positive results to flow from their work, but it is often not clear whether the indicators themselves will accomplish the larger goals of the project. In short, the linkage between indicators and action is often tenuous.

There is a saying in the field of social indicators that "What gets measured gets done." That makes a nice slogan, but it is not entirely true. On the one hand, governments have historically compiled lots of statistics that did not affect policy. Health statistics have not necessarily saved lives. If a city decided to measure the total length of its sidewalks as an indicator of how "pedestrian-friendly" it was, that might have no effect on public policy at all. (It would simply be viewed as an interesting statistic.)

On the other hand, people have demanded action on air pollution when their only measuring devices were watery eyes and asthmatic children. Knowing ozone levels in various parts of a city or metropolis is of value when determining precisely what policies to enact, but those indicators may have little to do with creating the political momentum for action. In fact, public bodies often act initially on the basis of anecdotal information and only develop indicators later to improve, modify, or justify the policy.

Sometimes indicators don't lead to action; other times action precedes the development of indicators. There are a number of actual cases that fit both of these reactions.

Example 1 | In the 1830s, early public health physicians in France and England were able to show a statistical relationship between poverty and serious illness. Again in the 1870s,

studies by the Massachusetts Labor Bureau developed statistical indicators showing the poor working and living conditions of the working class. None of those studies had the same impact on the political scene as the writings of Victor Hugo or Charles Dickens, who appealed to the hearts as well as minds of readers.

Example 2 | In the 1970s scientific noise indicators were developed, and studies were conducted showing a connection between noise and stress, but little was done to counteract this threat to health. There has been little connection between scientific measures of noise and the amount of political organizing to curb it. This does not mean the issue is not politically significant. Complaints about noisy neighbors, leafblowers, and barking dogs are among the most common received by mayors and city council members.

Example 3 | Studies were conducted from the 1950s onward which showed a statistical connection between smoking and various illnesses. There was no significant action to prevent smoking deaths, however, until the anti-smoking campaign got caught up in the general attack on all drugs.

Example 4 | Water pollution laws in the 1960s and 1970s were enacted largely on the basis of photographs showing dead fish in rivers and the like, not refined measurements of water quality. Even today there is little evidence demonstrating the degree of water pollution in most of the nation's rivers and lakes. The statistical monitoring system lags far behind efforts to improve water quality.

Example 5 | The U.S. military build-up during the Cold War took place prior to any evidence of a "bomber gap" in the 1950s or a "missile gap" in the 1960s. In both cases, the public support for defense spending occurred not because of the availability of statistical indicators of Soviet superiority (which was the claim), but because of a generalized "red scare." In fact, the U.S. led in every facet of the arms race throughout the entire Cold War.

LESSON 9

Better information may lead to better decisions and improved outcomes, but not as easily as it might seem.

The policymaking function of indicators is always indirect and roundabout. It is never as simple as textbook charts would have us believe. Seldom do problems go unsolved merely as a result of missing information. Of course, if government officials operate on the basis of incorrect statistics, they are likely to make bad decisions. But that doesn't mean that better statistics will always lead to better decisions. Think of the situa-

tion this way: If you are a parent and someone provides you with new information on childhood development patterns, will that automatically make you a better parent? Not likely. To change behavior, information needs to affect motives or perceptions of how the world works. Indicators, which are one form of information, can only be a piece in a larger puzzle.

Example 1 | An interesting case of the ambiguity of information is how competing groups interpreted the energy crisis of the 1970s differently (Tenenbaum and Wildavsky 1981). Consumer groups wanted lower prices for oil, whereas environmentalists tended to favor higher oil prices to encourage conservation. A great deal hinged on the question of whether the U.S. was "running out of oil" or not. Different groups used different indicators to make their judgments. Those who claimed the U.S. (or the world) was running out of oil pointed to figures on recoverable reserves and showed there were only about 20 years worth of oil left (the estimates varied). Consumer groups pointed to the high profits of oil companies and the decline in drilling rates and said there was plenty of oil and that the oil companies were rigging the market to create the illusion of a shortage.

The situation was inevitably fraught with ambiguity. Oil reserves are like a sponge soaked with water. How much oil there is depends on how much effort you apply to squeezing it. More information was not likely to change either side's view of the situation. One side used indicators of oil in the ground, whereas the other side used indicators of oil company behavior. Government officials were caught in the middle. What they needed was not more information or indicators, but new models to help them think about the situation.

LESSON 10
Challenging prevailing wisdom about what causes a problem is often the first step to fixing it.

If indicator reports are to do more than take up shelf space, they need to address problems that people care about. That could involve drawing attention to a condition that had previously been ignored. It might also mean showing that a widely shared idea is wrong, so that money won't be wasted pursuing misguided policies. Sometimes it might mean demonstrating a connection between two factors, so a new approach to a problem might be tried.

The greatest power in public policy debates lies in being able to change the definition of a problem. This is the first step in changing a policy and perhaps one of the most effective uses of indicators work. This function of indicators is often referred to as the enlightenment function. Based on our understanding, it is the power of indicators to alter the common understanding of a problem—not merely to point out the problem—

that leads to a change in problem definition. By offering a convincing analysis of why a problem exists, a group can gain support for a new solution.

Example 1 | During the early stages of the Industrial Revolution, a number of commissions were set up to study the conditions of the poor who were working in factories. Careful research by Louis-René Villermé in France and William Farr in England showed that death rates were consistently higher in the poorest neighborhoods than in the richer ones, when all other factors were held constant (Cullen 1975, chapter 2; Eyler 1979, chapter 6; Coleman 1982, chapter 6). This was a profound conclusion. It challenged the prevailing view that "bad air" (mal aria) or living in proximity to a river were the primary threats to health. Their work suggested that improving public health depended on economic reform. These findings did not immediately change society, but they were a first step toward justifying social security, minimum wage laws, and other elements of the safety net.

Example 2 | In recent years, indicators have been used to undermine the credibility of the welfare state. One influential author (Murray 1984) used statistics to convince many people that efforts to help the poor were not only costly to taxpayers (the traditional conservative argument), but also counterproductive to the poor themselves (the new conservative argument). Other scholars challenged the analysis, but politically the new argument took hold. The groundwork had been laid for the dismantling of the welfare state, a process that began in the 1990s. Murray and others had succeeded in shifting the terms of the debate and reframing perceptions about the nature of poverty. Once the question became *how* we should reform welfare and not *whether* we should reform it, the battle was over.

LESSON 11
To take action, look for indicators that reveal causes, not symptoms.

Indicators that focus only on symptoms can rarely solve the actual problem. In order to alter a symptom, it is necessary to have a theory about what is causing it and to test that theory repeatedly.

Some indicators projects collect a lot of statistics about a city, a region, a state, or a country, and then try to discern how they can affect the trends. If one starts with a vague idea, it is easy to end up with ambiguous conclusions. This can happen if an indicators project does not have a theory or hunch about what causes problems and what enables them to be solved. If indicators just tell about existing conditions without adding some insight into how they got to be that way, then the reports will not easily lead to action.

Thus, a group might theorize that certain kinds of education will raise employment levels, then it can try to measure both education programs and unemployment to see if they are inversely related (so that when one goes up the other goes down). Alternatively, it might think that employment opportunities are tied more to local economic conditions than to the qualifications of potential job seekers. That hypothesis could also be tested. By specifying one or more hypotheses, it becomes possible to gain some clarity about what one expects an indicator to "do" once it has been developed.

Example 1 | When the temperance movement turned to indicators in the early nineteenth century, it did more than just compile statistics about the horrors of alcohol. It also began developing hypotheses about the effect of alcohol on the quality of life that could be tested with statistics. The Massachusetts Society for the Suppression of Intemperance gathered statistics on the amount of money spent annually on liquor in order to show a connection to family debt (Cohen 1982, 170). It also used statistics to correlate poverty and premature death with intemperate drinking. By using indicators to illustrate the relationship between drinking and other social problems, they were able to transcend simple moralism and appeal to a broader audience.

Example 2 | In the 1960s and 1970s, concern about the rising costs of the welfare system led to a number of proposals to reform welfare. One proposal, the negative income tax, was favored by some within the Nixon Administration.[13] A negative income tax essentially allows those on welfare to work and retain a certain percentage of their welfare benefits based on their earned income. Rather than focusing only on overall trends of inputs (such as welfare costs) and outputs (such as changes in household income or poverty levels) to assess the overall effects of the welfare system, the demonstration research project used indicators to measure the effectiveness of a particular policy proposal.

The research project designed to test the idea of the negative income tax was formulated during the Johnson Administration and carried out during Nixon's. Advocates of the negative income tax believed that the policy would increase the incentive to work for those on welfare and decrease the overall cost of welfare. By allowing people to retain a percentage of their welfare benefits while they worked to earn additional income, they would increase their total income and decrease their welfare receipts. The catch was that it was possible for someone to maintain their current income level by working less and obtaining an additional part of their income from welfare payments. The question, therefore, was whether there would be a net increase or decrease of total welfare payments. Although there is some controversy over the interpretation of results, the general

13. Meanwhile, in California, Governor Ronald Reagan was developing the "conservative" vision of welfare reform (simply eliminate benefits) that culminated in the recent welfare reform in the United States.

consensus is that the negative income tax does not work to reduce the total cost of welfare (Nathan 1988, 49–60). This surprised many advocates of this approach who had expected the results to be positive.

Based on his involvement as a user and producer of public policy research, Nathan advocates the use of large-scale demonstration research projects that are interdisciplinary in design, use both quantitative and qualitative methods, and address *how* to do something rather than *what* is to be done.[14] In his view, applied social science research should move away from the description of conditions and trends (with its implied goal setting) and onto testing whether a policy (based on some causal hypothesis) works to ameliorate the problem (Nathan 1988, 16–17).

Example 3 | The community indicators project in Honolulu provides another good example of the iterative process of testing a hypothesis. Living in a community that depends heavily on tourism, local community members wondered if they could increase employment by increasing the tourism industry. Rather than assume that this was the case, they tested the hypothesis by looking at the relationship between the size of the tourism industry and the number of jobs in the region. To their surprise they found that a growing tourism industry does not necessarily bring more jobs. A subsequent examination of trends tried to determine whether a) increasing worker productivity had made the difference, or b) the base number of jobs created by the industry can support any number of tourists (Hart 1998). By testing their hypothesis, the project discovered that the key to creating jobs does not lie in attracting more tourists, but in diversifying the economy.

Example 4 | One caveat: A correlation between two trends does not always mean there is a causal relation between them. It is clear that economically depressed areas tend to experience greater crime rates. In the mid-1990s California passed their "three strikes" law which some, such as California's Attorney General Dan Lungren, claim caused the recent decrease in crime in the state. During the same period there was a dramatic economic turnaround with increasing wages and decreasing unemployment. Who can say which led to the decrease in crime based on just that evidence? New York, Houston, and other major cities experienced a comparable decrease in crime with no "three strikes" law. A basic lesson in indicators work as in statistics, correlation does not prove cause.

14. Demonstration research amounts to a social experiment to test a hypothesis about whether or not a policy will work. Ideally the projects use random assignment to create a control group to allow comparisons with the test group. The many technical, administrative and ethical difficulties in conducting large scale experiments are discussed at length in Nathan's account of his experience with the Manpower Demonstration Research Corporation.

LESSON 12

You are more likely to move from indicators to outcomes if you have control over resources.

Indicators are not an end in themselves. Their purpose is to alert the public and policymakers about the existence and cause of problems so that they might be solved. This is only possible when the groups responsible for indicator development have a connection to those with the power to make substantive changes. Otherwise, indicators may not influence outcomes at all.

As far as national economic policy is concerned, the connection between indicators and action poses few problems. The President's Council of Economic Advisers publishes the monthly report *Economic Indicators,* which supplies information on general economic activity, inflation, and unemployment. The audiences—the U.S. Congress, business leaders, and the Federal Reserve Board—do have an ability to influence these macroeconomic trends.

It is helpful to think about what your audience could do to change the value of the indicator. What causes the problem? Who has the power to change it? Perhaps your audience or your community does. Perhaps they only have the power to pressure someone else to act. Determining who has the power to take action can help clarify what outcomes you expect from your indicators report.

Example 1 | The city of Santa Monica provides an example of making use of indicators to change outcomes. The city developed a set of sustainability indicators to monitor its resource use. One of their indicators showed that only 15 percent of the municipal fleet used reduced-emissions fuels. In response the city instituted a plan to increase that percentage to 75 percent by the year 2000. So in Santa Monica, the agency was able to use indicators to monitor and change its own behavior. They had the ability to act on the information and create a successful outcome.

By contrast, when a group publishes local unemployment rates, it may not have much power to change those rates directly. This isn't to imply that indicator groups should choose only indicators over which they have some influence. It is just to point out that the chances are greater that groups will see results from their indicators work if they have control over the causes.

Example 2 | The Truckee Meadows Regional Transportation Agency and Truckee Meadows Tomorrow developed an indicators program as part of a regional planning process. As in many communities, two major concerns were the unemployment rate and average real wages. Although the project couldn't control the conditions themselves, a regional economic development agency did have the power to attract new industry. One

of the indicators, "Median Weekly Wage by Industry," was chosen by a regional economic development agency as a benchmark. By focusing on sectoral characteristics, they have been able to target industries that have higher wages, thereby exerting some influence over both the employment and income trends. Although the indicators group themselves did not have the power to change local economic conditions, they were able to reach a group that did (Truckee Meadows Tomorrow 1997).

VI. CONCLUSION

If any single lesson emerges from the long history of social statistics and indicators, it is that there are many blind alleys that can be followed. While it might seem easy at first to develop indicators that define the condition of a society and point it in the right direction, the missteps of the past suggest that only careful forethought will lead to success.

The biggest challenge facing current indicators practitioners is how to make their efforts meaningful—how to make sure they matter. There is a constant possibility that groups will collect statistics and publish reports that will have no visible impact on social processes and outcomes.

Solving this problem requires a perspective on indicators that is at variance with the common or popular view of them. If they are seen simply as descriptive statistics that point to the existence of problems, they will likely be of little use in solving those same problems. Merely describing a situation does little to reveal how it arose or what sustains it. A different approach to indicators is needed to move beyond that stalemate.

The approach that has most often worked in the past has been an analytic method, one that looks for causal relationships between events rather than simply at the events themselves. The success of this approach has been based on the development of models and hypotheses about how the world works. The purpose of indicators is then to help develop and test the validity of the models. If the indicators chosen illustrate the validity of a theory, then the indicators project has a handle on solving real problems.

Unfortunately, the indicators movement has been heavily influenced by its association with a school of academic sociology that emphasized the seemingly endless collection of numbers and resulted in provisional conclusions at best. No hypotheses were formulated, no potential solutions to social problems were found. Social indicators came to be viewed as an end in themselves. Skeptics understandably questioned the value of that sort of exercise.

There is a danger that community indicators in the 1990s could slip into the same sort of doldrums unless an effort is made to change course. In order to develop indicators that will have an impact on society, it will be necessary to learn from the lessons of the past. We have discussed the dozen that we find to be the most helpful:

- Having a number doesn't necessarily mean that you have a good indicator.
- Effective indicators require a clear conceptual basis.
- There's no such thing as a value-free indicator.
- Comprehensiveness may be the enemy of effectiveness.
- The symbolic value of an indicator may outweigh its value as a literal measure.
- Don't conflate indicators with reality.
- A democratic indicators program requires more than good public participation processes.
- Measurement does not necessarily induce appropriate action.
- Better information may lead to better decisions and improved outcomes, but not as easily as it might seem.
- Challenging prevailing wisdom about what causes problem is often the first step to fixing it.
- To take action, look for indicators that reveal causes, not symptoms.
- You are more likely to move from indicators to outcomes if you have control over resources.

REFERENCES

ANDREWS, FRANK M. 1989.
– "The Evolution of a Movement." *Journal of Public Policy* 9(4):401–405.

BANGS, RALPH L., AND JUN HYUN HONG. 1995.
– *Black and White Conditions in the City of Pittsburgh.* Pittsburgh, Penn.: University Center for Social and Urban Research, University of Pittsburgh.

BAUER, RAYMOND A., ED. 1966.
– *Social Indicators.* Cambridge, Mass.: MIT Press.

BENNER, CHRIS. 1998.
– *Growing Together or Drifting Apart: Working Families and Business in the New Economy.* San Jose, Calif.: Working Partnerships USA.

BIDERMAN, ALBERT D., AND J. P. LYNCH. 1991.
– *Understanding Crime Incidence Statistics: Why the UCR Diverges from the NCS.* New York: Springer Verlag.

BOOTH, TIM. 1992.
– "Social Indicators and the Mondale Initiative." *Knowledge: Creation, Diffusion, Utilization* 13(4):371–98.

BULMER, MARTIN. 1983.
– "The Methodology of Early Social Indicator Research: William Fielding Ogburn and 'Recent Social Trends', 1933." *Social Indicators Research* 13:109–130.

_____. 1989.
– "Problems of Theory and Measurement." *Journal of Public Policy* 9(4):407–412.

CAMPBELL, ANGUS, AND PHILIP E. CONVERSE, EDS. 1972.
– *The Human Meaning of Social Change.* New York: Russell Sage Foundation.

COHEN, PATRICIA CLINE. 1982.
– *A Calculating People: The Spread of Numeracy in Early America.* Chicago: University of Chicago Press.

COLEMAN, WILLIAM. 1982.
– *Death Is a Social Disease: Public Health and Political Economy in Early Industrial France.* Madison, Wisc.: University of Wisconsin Press.

CROUSE, JAMES, AND DALE TRUSHEIM. 1988.
– *The Case against the SAT.* Chicago: University of Chicago Press.

CULLEN, M. J. 1975.
– *The Statistical Movement in Early Victorian Britain: The Foundations of Empirical Social Research.* New York: Barnes and Noble.

DUNCAN, OTIS DUDLEY. 1984.
– *Notes on Social Measurement: Historical and Critical.* New York: Russell Sage Foundation.

EYLER, JOHN M. 1979.
– *Victorian Social Medicine: The Ideas and Methods of William Farr.* Baltimore, Md.: Johns Hopkins University Press.

GARDNER, HOWARD. 1983.
– *Frames of Mind: The Theory of Multiple Intelligences.* New York: Basic Books.

_____. 1993.
– *Multiple Intelligences: The Theory in Practice.* New York: Basic Books.

GILMARTIN, KEVIN J., ROBERT J. ROSSI, LEONARD S. LUTOMSKI, AND DONALD F. B. REED. 1979.
– *Social Indicators: An Annotated Bibliography of Current Literature.* New York: Garland Publishing.

GOULD, STEPHEN JAY. 1981.
– *The Mismeasure of Man.* New York: W. W. Norton.

HART, MAUREEN. 1998.
– "Indicator of the Month–March 1998." [WWW document]. Accessed September 21, 1998: <http://www.subjectmatters.com/indicators/HTMLSrc/Indicator_199803.html>.

HERRNSTEIN, RICHARD J., AND CHARLES MURRAY. 1994.
– *The Bell Curve: Intelligence and Class Structure in American Life.* New York: Free Press.

INNES, JUDITH E. 1989.
– "Disappointments and Legacies of the Social Indicators." *Journal of Public Policy* 9(4):429–432.

_____. 1990.
– *Knowledge and Public Policy: The Search for Meaningful Indicators.* New Brunswick, N.J.: Transaction Publishers.

JOHNSTON, DENIS F. 1987.
– "The Federal Effort in Developing Social Indicators and Social Reporting in the United States during the 1970s." In *Social Science Research and Government: Comparative Essays on Britain and the United States,* ed. Martin Bulmer. Oxford: Cambridge University Press.

_____. 1989.
– "Some Reflections on the United States." *Journal of Public Policy* 9(4):433–436.

LAND, KENNETH. 1983.
– "Social Indicators." *Annual Review of Sociology* 9:1–26.

_____. 1992.
– "Social Indicators." In *The Encyclopedia of Sociology,* ed. Edgar F. Borgatta and Marie L. Borgatta, 1844–1850. New York: Macmillan Publishing Company.

LEIBY, JAMES. 1960.
– *Carroll Wright and Labor Reform: The Origin of Labor Statistics.* Cambridge, Mass.: Harvard University Press.

LIU, BEN-CHIEH. 1975.
– "Quality of Life: Concept, Measure, and Results." *The American Journal of Economics and Sociology* 34(1):1–12.

MACLAREN, VIRGINIA W. 1996.
– "Urban Sustainability Reporting." *Journal of the American Planning Association* 62(2):184–202.

MIRINGOFF, MARK L. 1997.
– *The Index of Social Health 1996: Monitoring the Social Well-Being of the Nation.* Tarrytown, New York: Fordham Institute for Innovation in Public Policy.

MURRAY, CHARLES. 1984.
– *Losing Ground: American Social Policy 1950–1980.* New York: Basic Books.

NATHAN, RICHARD P. 1988.
– *Social Science in Government: Uses and Misuses.* New York: Basic Books.

NOLL, HEINZ-HERBERT, AND WOLFGANG ZAPF. 1994.
– "Social Indicators Research: Societal Monitoring and Social Reporting." In *Trends and Perspectives in Empirical Social Research,* ed. Ingwer Borg and Peter Ph. Mohler. Berlin: Walter de Gruyter.

OREGON PROGRESS BOARD. 1997, JANUARY 21.
– "Oregon Shines II: Updating Oregon's Strategic Plan, Highlights." [WWW document]. Accessed September 21, 1998: <http://www.econ.state.or.us/opb/highlights/highlights.html>.

OWEN, DAVID. 1985.
– *None of the Above: Behind the Myth of Scholastic Aptitude (How the Educational Testing Service Controls the Gates to Higher Education and Success in American Society).* New York: Houghton-Mifflin.

PRESIDENT'S RESEARCH COMMITTEE ON SOCIAL TRENDS. 1933.
– *Recent Social Trends in the United States.* New York: McGraw Hill.

ROSSI, ROBERT J., AND KEVIN J. GILMARTIN. 1980.
– *The Handbook of Social Indicators: Sources, Characteristics, and Analysis.* New York: Garland STPM Press.

ROTHENBACHER, FRANZ. 1993.
– "National and International Approaches in Social Reporting." *Social Indicators Research* 29:1–62.

SHELDON, ELEANOR BERNERT, AND WILBERT E. MOORE, EDS. 1968.
– *Indicators of Social Change.* New York: Russell Sage Foundation.

SHELDON, ELEANOR BERNERT, AND HOWARD E. FREEMAN. 1970.
– "Notes on Social Indicators." *Policy Sciences* 1:97–111.

SHELDON, ELEANOR BERNERT, ROBERT PARKE, AND MURRAY ABORN. 1983.
– "Recollections and Views of Key Figures in the Social Indicators Program." *Items* 37(4):78–89.

SMITH, JAMES ALLEN. 1991.
– *The Idea Brokers: Think Tanks and the Rise of the New Policy Elite.* New York: Free Press.

TENENBAUM, ELLEN, AND AARON B. WILDAVSKY. 1981.
– *The Politics of Mistrust: Estimating American Oil and Gas Resources.* Beverly Hills, Calif.: Sage Publications.

TRUCKEE MEADOWS TOMORROW. 1997.
– *Quality of Life in the Truckee Meadows.* Reno, Nev.: Truckee Meadows Tomorrow.

U.S. BUREAU OF THE CENSUS. OFFICE OF FEDERAL STATISTICAL POLICY AND STANDARDS. 1977.
– *Social Indicators, 1976: Selected Data on Social Conditions and Trends in the United States.* Washington, D.C.: U.S. Government Printing Office.

U.S. BUREAU OF THE CENSUS. 1981.
– *Social Indicators III: Selected Data on Social Conditions and Trends in the United States.* Washington, D.C.: U.S. Government Printing Office.

U.S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE (HEW). 1969.
– *Toward a Social Report.* Washington, D.C.: U.S. Government Printing Office.

U.S. OFFICE OF MANAGEMENT AND BUDGET. STATISTICAL POLICY DIVISION. 1974.
– *Social Indicators, 1973: Selected Data on Social Conditions and Trends in the United States.* Washington, D.C.: U.S. Government Printing Office.

WADDELL, STEVE. 1995.
– "Lessons from the Healthy Cities Movement for Social Indicator Development." *Social Indicators Research* 34:213–235.

WALTERS, JONATHAN. 1994.
– "The Benchmarking Craze." *Governing* (April):33–37.

WILCOX, L. D., R. M. BROOKS, G. M. BEAL, AND G.E. LONGAN. 1972.
– *Social Indicators and Societal Monitoring: An Annotated Bibliography.* San Francisco, Calif.: Jossey-Bass.