

The essays in this book were commissioned with an aim of advancing discussion beyond foot-stamping criticism of prominent indicators of quality in science. The essayists were challenged to scrutinize what Harvey Brooks calls the “inherent tension between simplicity and reality,” to propose new approaches to the problem, and to introduce into policy discussions the ethical and value issues that affect the invention, selection, and use of particular indicators of quality in science.

The essayists and seminar commentators approach these topics from several different perspectives. Some, like Lewis Branscomb and Harvey Brooks, report on what actually takes place when indicators have been or are being used for decision-making. Branscomb’s description of the decision-points in IBM Corporation’s evaluation of researchers, research, and product development affords a window onto a process that could have applications in other settings. Brooks’ retrospective assessment of decision-making in four research areas has similar usefulness, for it gives exposure to the processes in university-based work, where the values of the individual investigator or field can influence powerfully the choice of topics, methods, or hypotheses.

Restraints on research also may affect the validity of certain indicators. In her essay on secrecy in science, Sissela Bok provokes many complex questions about the applicability of indicators to newly-classified fields such as cryptography, where the results are important for both military and civilian users of computers and high-speed data processors, but the control of publication (and, hence, of widespread recognition) rests with only one group. Brooks’ description of how scientists may regard certain topics as “low-grade,” “routine,” and “scientifically not very challenging,” and therefore “not likely to attract the interest or attention of independent inves-

tigators,” provides examples of internal restraints that affect the validity of indicators.

Many of the contributors emphasize the importance of the dichotomy between internal and external indicators—for example, “intrinsic and extrinsic value” of research (Branscomb), “inside and outside measures of quality” (Mazlish), or “intrinsic vs. extrinsic indicators” (Brooks). As Mazlish points out, quantitative indicators tend to be concerned with the “inside” image of science, but their predominance should not obscure the importance or significance of “alternative modes of understanding the quality of science, especially its ‘outside.’ ” The three essays by Yankelovich, Prewitt, and Weingart provide the backdrop to these discussions through their often sharply disparate views of (1) how science should “seek thoughtful public judgments about priorities” and (2) how it reacts to social challenges to the legitimacy and direction of research. The writers paint pictures of “science increasingly politicized” (Weingart), of science no longer allowed to “stand aloof” from changing political and social values (Yankelovich), and of the increasing impingement on science policy of “deeply held political values of democratic accountability and public scrutiny” (Prewitt).

Examples of such values underpin the statements by Senator Hatch and Representatives Brown, Fuqua, and Walgren. Because they serve on committees that oversee government-funded or government-regulated scientific and technical research, these members of Congress are representative not only of their voting constituencies but also of the “consumers” of indicators of quality, that is, the decision-makers who look for quantitative and qualitative measures for help in allocating resources or directing research policy at the national level. Brooks’ question “Will the science we are generating today be the science wanted and needed to attack the salient problems two decades hence?” puts into words the uncertainty feeding the Congressmen’s underlying tone of impatience with scientists. For, despite the fact that policy for science and technology is, by and large, directed by *policy-makers* (who, even if they are working scientists, perceive their role as a double one), the non-scientist’s perception is that *scientists* maintain control and, as Hatch forcefully states, must therefore acknowledge a “special responsibility” for directing science.

One thing that these essays—and the seminar discussions preceding them—clearly indicate is that this is a difficult problem. Why? Three

reasons spring immediately to mind—although they are by no means the only ones.

First, as a practical matter, measures of “quality” are, by their very nature, difficult to devise. The insufficiency of quantitative measures pushes the development effort; but, when discussion gets to the point of concrete suggestions for a real indicator for real-time use, inevitably we ask what such an indicator would look like. How would it be used? Are there data to feed into it? As the Comptroller General’s critique of the *Science Indicators* series states, at the practical level “it is how well the *data* fit and measure a relevant concept that determines the indicator’s validity” (emphasis supplied).* Consideration of indicators of quality must be a process in which proposals for new approaches or new measures proceed hand in hand with comprehensive information about existing data sets and/or sources and with proposals for the generation of new data (such as the public attitudes research described in the Yankelovich seminar).

Second, there are a host of “political” reasons, some having to do with real or perceived attitudes within the community of social scientists concerned with science indicators and some having to do with the sensitivity of both the choice of measures and the real or potential conclusions based upon them. Each of the writers refers to this situation, some enigmatically but all cognizant of the significance of the influence.

A third, mitigating factor is the effect of the tugs and pulls of the pertinent constituencies’ differing value systems. Conflicting social, economic, or political pressures can make the requirements for a usable indicator at least appear different at the different decision-points. Let me name just a few such constituencies. (a) *Scientists* resist the “unscientific” self-scrutiny implied by formal indicators of quality; they are much more comfortable with informal, intuitive, “intrinsic” indicators guided by the norms of science and closely fitted to the history, demands, sociology, current “success,” and morale of each field or research area. If you ask scientists about their research agendas or why they do what they do, one colleague observed, they get defensive because they feel obliged to explain; but if you ask them

* Report by the Comptroller General of the United States, *Science Indicators: Improvements Needed in Design, Construction, and Interpretation*, PAD-79-35 (U.S. General Accounting Office, Washington, DC, 25 September 1979), p. 9.

to describe how they made a certain decision, they may talk for hours. (b) *The public* may have its own special set of indicators of quality. Certainly, one criterion is the potential economic and social “value” (positive and negative) of the research product. Another criterion may be the timing of the proposed investment—the public may not always evaluate the same research in the same way. The size of the investment required may also be important—although it is not simply that big is always better or that small is always considered more economical. A small research project that meets few other criteria could be evaluated as a waste of time and resources. For positive evaluation by the public, scientific research also must not conflict with contemporary cultural ideals, myths, or values, and must match the public’s expectations of how and when those ideals may be achieved. Positive public assessment may also be significantly affected by the political or economic strength of the groups benefited or harmed by the research. (c) The value systems of *the users* of indicators—in government, in industry management, in Congressional committees—play important (and occasionally conflicting) roles in determining the parameters for new indicators and in directing their use. The user’s choices can be governed, among other factors, by political considerations, by availability of resources, and by the “results” the user desires. One of the questions for consideration in the future might be the differing requirements of the various users.

Given that this is a difficult problem (and I have touched on only a few of the difficulties), in what direction do these essays point? Each reader will have her or his own agenda arising from questions raised here, but let me make two observations about some threads apparent in the essays, just to get the discussion started.

(1) The ethical aspects of indicators

The essays clearly show that considerations of the ethical and value aspects of science are germane to the measurement of its quality. Here are just a few questions raised by the essays and seminars: Is one indicator of quality of science the openness of the enterprise to all promising science students, male and female, black and white, Native American and naturalized refugee? Or the availability of research funds to young scientists just starting their careers? How can these be measured? What about measures of scientific freedom—at the level both of the individual researcher and of the field? What

about the economic status or job security of young scientists? or the effect of rapid “industrialization” in fields such as genetic engineering? Much of the science indicators discussion has focused on measures for choosing the “right” proposals, the upper 10%, the research with the “greatest potential.” The other side of that coin, however, may be an approach that asks how we can assure that proposals (or people) are not rejected for the wrong reasons.

This type of search for underlying ethical assumptions can help to assure the fit of science indicators to both scientific and social reality. As Fuqua and Walgren summarize, “only when the society as a whole can participate [in these choices] in an informed manner, can we expect wise decisions concerning the quality and direction of science.”

(2) Indicators and the public

Public opinions about science are usually responded to only when they are hostile, only when there is a crisis or controversy. This is unfortunate. The Public Agenda Foundation research described by Yankelovich*—which, as of this writing, is in the data collection stage—represents one of many possible experimental approaches to measuring the non-scientific public’s assessment of scientific quality. Upon observing some of the survey research groups (composed of ordinary citizens discussing scientific research proposals), I was surprised at the participants’ apparent unquestioning assumption that they *can* talk about research priorities in a positive way, and that they have a right to discuss such things. The scientific community—and the social scientists and policy-makers concerned with science—have, I believe, long felt that public attitudes represent primarily a negative pressure on science, especially when scientists argue over the public’s “right” to limit science or to have certain technical information. The Public Agenda Foundation survey does not introduce false or “let’s pretend” controversy into the presentations to the participants; rather they are presented with examples of real dilemmas in resource allocation and research priorities as positive, serious questions. These “average” Americans do not *appear* (and I use that word deliberately

* The research is being undertaken by the Public Agenda Foundation, under subcontract to Harvard University. Persons involved with the survey include Daniel Yankelovich, John Doble of the Public Agenda Foundation, Gerald Holton of Harvard University, and the author.

because of the early stage of the survey research) overtly hostile to science *per se*, although they seem to recognize intuitively the politicized nature of science, even of basic science.

The results from the experimental survey—as from the Miller, Prewitt, and Pearson survey for the National Science Foundation*—will undoubtedly raise a host of new research questions. But one observation that I might make now from our preliminary data and from the essays in this book is that serious consideration can and must be given to how to give “weight to the social functions of the scientific enterprise” (Brooks), and how to construct a “more effective decision-making partnership between the public and the scientific community” (Fuqua and Walgren)—and that these are *not* extraneous matters.

* Jon D. Miller, Kenneth Prewitt, and Robert Pearson, *The Attitudes of the U.S. Public Toward Science and Technology* (Chicago, IL: National Opinion Research Center, University of Chicago, July 1980).