THE UNCERTAIN LOGIC OF STANDARD-SETTING

Giandomenico Majone

September 1982
WP-82-89

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

THE UNCERTAIN LOGIC OF STANDARD-SETTING *

Giandomenico Majone
International Institute for Applied Systems Analysis
Laxenburg, Austria

INTRODUCTION

It is a truism, but one which is too often forgotten, that decision processes in general, and regulatory decision making in particular, vary greatly according to subject matter (e.g., the activity, process, or substance to be regulated), cognitive philosophy, ideological stance, available knowledge and skills, institutional setting, and so on. As Philip Selznick has observed in a different context, "[d]ecision-making is one of those fashionable phrases that may well obscure more than it illuminates ... The general features of all choices, or of all social choice, may some day be convincingly stated. But it will still be necessary to distinguish the more and the less trivial; and, if there is any order in this phenomenon, to identify some kinds of decisions, linking them to the distinctive problems or situations out of which they arise."[1]

As Selznick suggests, there may be too much variety in social choices to justify a single analytic approach or a single criterion of rationality. Yet, the tendency still prevailing in policy analysis is to force all kinds of decision problems into the Procrustean bed of "comprehensive rational analysis". The same stereotyped categories, the same models, the same evaluative criteria are applied to regulatory decisions regardless of specific differences and special circumstances.

Even conceding that some economy of thought may have been achieved, the cost in terms of understanding the standard-setting process in all its complexity has been, I suspect, too high. For example, differences in biological philosophy, conflicting

---

* forthcoming in Zeitschrift für Umweltpolitik.

views concerning the degree to which the human body can overcome the effects of toxic agents and pollutants, uncertainty about dose-response relationships, or the institutional context in which environmental inspectors operate, hardly play any role in most policy analyses of environmental regulation.

However, such factors have an enormous influence on the regulatory process. Differences in biological philosophy, for example, are the main reason for the striking differences between many Soviet and Western health standards.[2]

An oversimplified view of the environmental problem also pervades most current debates on the choice of regulatory tools. It is one thing to show that under certain idealized conditions pollution taxes are the most efficient (hence "rational"!) policy instruments. It is quite another thing to argue that such taxes should be used in practice, in spite of inadequate scientific and economic data, institutional problems, and the general reluctance of legislators, administrators, and the public to follow the economists' advice and accept economic efficiency as the basic criterion of social choice.

The analyst who evaluates environmental policies by the sole criterion of economic efficiency actually has something in common with the environmentalist who advocates regulation based exclusively on health criteria. For both of them, the important thing is outcome, not process; both are interested in the decisions that are made, not how they are made.

Evaluating social choices by their outcomes has a strong intuitive appeal, but presupposes the existence of some unambiguous measure of outcome. When the correctness or fairness of the outcome can be determined unambiguously, the manner in which the decision is taken is largely immaterial -- only results count. But when the factual and value premises are debatable, the consequences highly uncertain, when there is no consensus on evaluative criteria--then the process or procedure of decision making acquires special significance. This, as Niklas Luhmann has shown, is the fundamental insight on which the classical theories of judicial, legislative, and administrative procedures are based.

Regulators have traditionally sought legitimacy for their decisions by wrapping them in a cloak of scientific respectability. Their determinations (they claim) are firmly based on scientific analyses made by qualified experts. But the cognitive and institutional complexity of pollution control and risk evaluation has dispelled the initial faith in the power of such experts. As this paper attempts to show, the scientific and conceptual basis of environmental regulation is so precarious, the empirical evidence so ambiguous, that most regulatory decisions can only be evaluated and legitimated in terms of procedural, rather than substantive, rationality -- by process, not by outcome.

THE CONCEPTUAL BASIS OF REGULATION

Environmental and health standards are derived, and used, differently in different countries. A major source of variations lies in differences in the definition of what is a state of health, and conflicting views concerning the degree to which the defense mechanisms of the body can be safely drawn upon to offset the effects of toxic agents and pollutants.

Toxicological procedures used in the West rely on the idea that no threat to health exists so long as the exposure does not induce a disturbance that overloads the normal protective mechanisms of the body. On the other hand, Soviet toxicologists maintain that <u>any</u> change in the normal response to a stimulus represents an unacceptable deviation from normal conditions, and <u>any</u> concentration, however small, places an undesirable toxic or nuisance stress on the organism. Thus, a potential for ill-health is assumed to exist as soon as the organism undergoes the first detectable change of whatever kind from its normal state.[3]

To better visualize these conceptual differences, imagine the familiar dose-response curve (for example, curve A in Figure 1 below) as being subdivided into three zones: an upper zone corresponding to high doses of a toxic substance, where ill-effects due to exposure are clearly detectable; a compensatory zone where the body adjusts to the stresses imposed by lower levels of exposure, but at some cost; and, finally, a lower, homeostatic zone where the adjustments are automatic.

According to the biological philosophy prevailing in the West, the defense mechanisms in the compensatory zone, as well as the normal adaptive processes in the homeostatic zone, can be safely drawn upon, within limits, to offset the levels of stress imposed by minimum exposure to hazardous agents at the workplace and in the environment, just as they are called upon to counter the wear and tear of ordinary life. Hence, dose-response relationships are extrapolated downward from the zone of demonstrable health burdens to a point of "non-detectable" ill effects.[4]

Soviet toxicologists start at the other end of the dose-response curve, moving upwards from zero dose and a corresponding initial benchmark of normality in the test organism. The permissible level of exposure is established below the lowest dose needed to induce a statistically significant difference from the normal state, as revealed by highly sensitive measures of behavioral response. The assumption underlying this procedure is that the protective mechanisms in both homeostatic and compensatory zones should be kept in reserve to ward off unexpected toxic effects, and their effectiveness should not be weakened by the continuous demands of stress knowingly permitted in the environment or at the workplace.

Neither the Soviet nor the western position can be dismissed as being unreasonable or contrary to known biological laws, but the practical implications in terms of acceptable levels of exposure are vastly different in the two cases. The official goal of Soviet standard-setters is a zero level of exposure. By contrast, goals of zero exposure have not been seriously discussed in the United States or in other countries in the West, except for radiation protection and for carcinogens in the workplace.[5]

Conceptual differences concerning the nature of health and the adaptive capacity of the human organism are magnified by differences in research techniques. Soviet toxicologists place major emphasis on studying the effects of toxic agents on the nervous system. Central nervous system sensitivity (conditioned reflexes, electroencephalograms) and reflex responses (changes in

heart and respiratory rates, in blood pressure, and so on) play a central role in standard-setting. In the words of a Soviet expert

> We attach great significance to chronic changes in the higher nervous activity of animals under the influence of toxic substances in the air they breathe. We believe that changes in the functioning of the cortex of the cerebral hemispheres occur very early, even with small concentrations, since the cerebral cortex is highly sensitive to the effects of external factors in the environment ... One of the early manifestations of the influence of various chemical substances on the higher nervous system is the development of phasic states. Later, disinhibition of differentiation occurs, then individual reflexes begin to disappear and finally none of the reflex pattern is left. When the animal is more severely affected, the natural conditioned reflex to the sight and smell of food disappears.[6]

Because of this preoccupation with the role of the higher nervous system as controller of all bodily activity, considerably more importance is given to the pathology of this system than is the case in Western studies.[7] Incidentally, the interest of Soviet toxicologists in the nervous system and reflex behavior can be explained by the enormous influence of Pavlovian theories on all domains of Soviet medicine.[8] In particular, the insistence on testing the nervous system is justified by reference to Pavlov's theory that living organisms adapt to their environment by means of two nervous mechanisms: the unconditioned reflexes for the permanent features of the environment, and the conditioned reflexes for the temporary (conditional) features.

American and European scientists, while not fully convinced that tests of the nervous system necessarily provide more sensitive indicators of toxic action, agree that sophisticated measurements of nervous-system effects should be a more important part of toxicological testing in the West.[9]

Another interesting methodological difference is the limited role which epidemiology seems to play in standard setting in the Soviet Union. In the West, and particularly in the United States,

epidemiology has historically provided important, and sometimes decisive, evidence on which standards have been based, although there are indications that its role may be decreasing relative to toxicological testing. For the Soviets, on the other hand, epidemiological studies represent a form of human experimentation in which prior toxicological tests and subsequent prevention have failed. In short, epidemiological studies represent a reactive rather than a preventative approach. Moreover, epidemiological studies abroad, showing the effects on health of concentrations higher than those allowed in the Soviet Union, encourage continued faith in the value of the traditional toxicological approach that has led to the lower Soviet concentrations.[10]

SCIENTIFIC UNCERTAINTY IN STANDARD-SETTING

Extrapolation is a key step in the establishment of environmental and health standards, and a good part of the uncertainty inherent in standard setting originates in various types of extrapolation processes. Consider, first, the problem of extrapolating from animal experiments.

A major issue in toxicology is the determination of the animal species that best predicts the response of man. Would the same species be equally predictive for all pollutants being tested? Do species differ in the degree to which they can predict toxicity for specific organ systems -- kidney, liver, lungs, and so on? Which "animal model" best simulates the pregnant woman, the new-born child, or individuals with inadequate diet or genetic deficiencies?

There are no unequivocal answers to such questions. Thus, many researchers have criticized the excessive use of rodents as predictive models because rodents are phylogenetically further removed from humans than other species, such as the dog or the monkey. However, a scientific panel of the United States Food and Drug Administration on carcinogenesis has not recommended the general use of the dog in the testing of chemical carcinogenesis because of its large size and relatively long life-span.[11]

There is, in fact, little hope that one species could provide the broad range of predictive potential needed to assess the responses of a highly heterogeneous human population to different types of pollutants. Predictions could be improved by using multiple species in toxicological experiments. But heterogeneity in human populations is often social in origin, and social conditions cannot be reproduced in the toxicologist's laboratory.

The issue of human heterogeneity also arises in connection with the prediction of adverse health effects on individuals who are (or may be) at high risk with respect to certain pollutants. Once the toxic dose for the "normal healthy" population has been derived, consideration must be given to high-risk groups, i.e., "those individuals who experience toxic and/or carcinogenic effects significantly before the general population as a result of one or more biological factors, including developmental influences, genetic factors, nutritional inadequacies, disease conditions, and behavioral or life style characteristics."[12] Thus, children and adults with vitamin C deficiency are hypersensitive to ozone and to a number of heavy metals; pregnant women, to lead and carbon monoxide; people with asthmatic and chronic respiratory diseases, to respiratory irritants such as nitrogen dioxide, ozone, and sulfur dioxide.

Standards developed for statistically "normal" individuals should be adjusted (by downward extrapolation or some other means) in order to protect the sections of the population at high risk. Unfortunately, high-risk groups are seldom considered specifically and separately in setting environmental and health standards, except perhaps through the dubious device of "safety factors" (see below). There are several reasons for this neglect; for example, lack of detailed exposure information, and the widespread assumption that high-risk groups represent a negligible percentage of the population. But recent research indicates that the number of high-risk individuals is quite large in some cases, and can include significant percentages of the population of specific racial ancestries.[13]

Strictly speaking, each individual has a unique genetic composition and life history, and thus a unique response to environmental pollutants. This heterogeneity of human populations leaves public authorities with an almost impossible regulatory task. In an effort to find a way out of this dilemma, toxicologists and statisticians have developed several mathematical models expressing the probability of a lifetime response, P, as a function of dosage D: $P = f(D)$. This is the dose-response function; different models are obtained, with different choices of f.

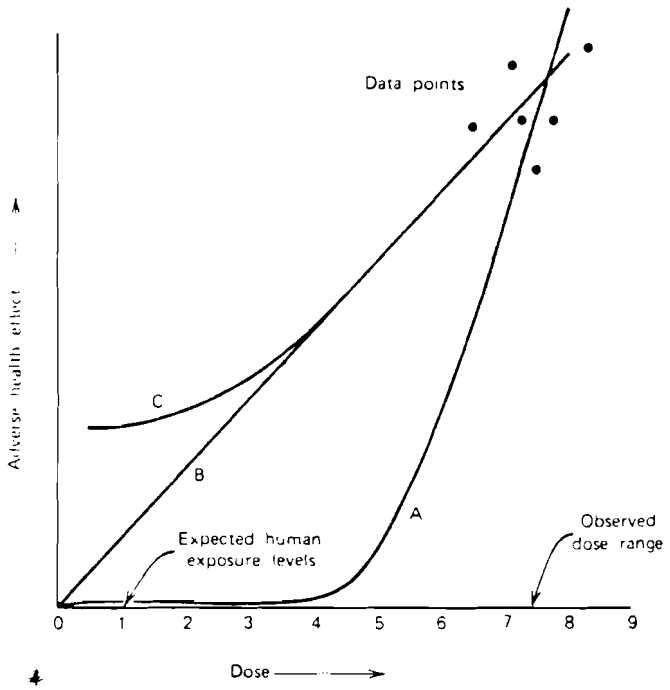

Figure 1: Downward extrapolation with different dose-response functions

Source: E.J. Calabrese, Methodological Approaches to Deriving Environmental and Occupational Health Standards, Wiley-Interscience, 1978, 131.

Figure 1 shows three choices of the extrapolating function from the many possible options. Although all three choices are consistent with the data points obtained at high dose levels, their policy

implications are quite different. Curve A represents the traditional threshold (nonlinear) dose-response model. Using A, it would be possible to establish a "virtually safe" level of exposure at dose 3, even though high doses produce adverse health effects. Curve B represents a linear (nonthreshold) dose-response relationship: adverse health effects occur at every level of pollutant exposure and there is no obvious point at which a reasonable standard could be set. Finally, the dose-response relationship expressed by curve C is linear at high and moderate doses, but at lower doses it indicates more serious health effects than the linear model would have predicted.

More sophisticated models consider the distribution of responses to different dose levels over the experimental population. Then $f(D)$ becomes the density of the probability distribution of responses, and the proportion of the population that will respond to a dose level D_0 is given by the cumulative function

$$P(D_0) = \int_0^{D_0} f(D)\,dD \quad .$$

Thus $P(0) = 0$ (i.e., there is no spontaneous occurrence of the particular response), and $P(\infty) = 1$ (i.e., no immune group exists within the population; all members will respond to sufficiently high doses).

A virtually safe dose (VSD) is now defined as a dose level D_0 such that $P(D_0) \leq P_0$, where P_0 is some preassigned small probability such as 10^{-8} (the value favored by many toxicologists) or 10^{-6} (the value used, for example, by the U.S. Food and Drug Administration). The VSD can be computed as soon as f is known.

But this is precisely the problem: how do we determine f? The usual procedure consists in fitting a curve (by one of many available methods) to the observations in the observable range, and then extrapolating downward to the unobservable response P_0 to determine the VSD.

There are three major problems with this procedure. First, the choice of function has a major effect on the value of the VSD -- more than 100,000-fold according to the Advisory Committee on Safety Evaluation of the Food and Drug Administration. Second, the different extrapolating functions often cannot be distinguished from each other in the range of the observable responses. Finally, no firm scientific basis now exists for choosing among the different possibilities.[14] An additional problem with downward extrapolation (an empirical rather than a mathematical problem) is that high-dose exposure to pollutants may totally swamp many protective mechanisms of the body that function at low-level exposures.

Why, it may be asked, are test animals exposed to levels of toxic substances far in excess of those to which humans would be exposed under normal circumstances -- thus making downward extrapolation necessary? This is done in order to compensate for the small samples of animals usually tested.
For example, if we assume that a chemical agent will cause cancer in 1 out of 10,000 people who are exposed to it, and that humans and test animals do not differ significantly in sensitivity with respect to the given agent, it would be necessary to test at least 10,000 animals (but preferably something like 30,000 animals) in order to detect one case of cancer.

With 1000 test animals and an unacceptably low confidence level of 90%, the upper confidence limit for a negative experiment (no cancer induced at the given dose level) is 2.3 cancers per 1000 tests. "No one could wish to introduce an agent into a human population for which no more could be said than that it would probably produce no more than 2 tumors per 1000. To reduce the upper limit of risk to 2 tumors per one million with confidence coefficient 0.999 would require a negative result in somewhat more than three million test animals."[15]

In practice, no more than 50 or so animals are usually available per dose level; hence the use of high doses on small samples of animals. To reduce the experimental doses, and thus the unreliability of extrapolations outside the experimental range, one could think of conducting experiments with extremely large numbers of animals. Such "megamouse" experiments have

in fact been proposed; but the costs would be prohibitive, and the validity of the conclusions still doubtful because of the problems connected with human heterogeneity and extrapolations from animal tests.

Unable to find clean theoretical solutions, standard-setters deal with the uncertainty of toxicological predictions by means of various rules of thumb -- safety factors, for example. Often a safety factor of 100 is used, meaning that test animals should show no adverse health effects from a given pollutant when exposed to doses at least 100 times greater than the likely human dose. This particular rule of thumb is sometimes justified by the reasoning that man may be ten times more sensitive than the experimental animals used, and that there may be in addition a tenfold variation in sensitivity among individuals.

This is all rather speculative and, besides, how does one justify the safety factors of 50 or 500 which are also in use?

COMPETING REGULATORY PHILOSOPHIES

It may be argued that if there is no firm scientific basis for choosing among the different mathematical models, then one should prefer the safest or most conservative procedure. One problem with the conservatism argument is that it is not clear where one should stop.[16] A no-threshold model is more conservative than one that admits the existence of thresholds for carcinogenic effects. But within the large class of no-threshold models many degrees of conservatism are possible. Again, in designing a toxicological experiment one could use the most sensitive species, the most sensitive strain within species, and so on down to the level of the most sensitive individual animal, thus obtaining 100 percent incidence at each dose level. In short, it is difficult to be conservative in a consistent manner, unless one is prepared to propose a zero level of exposure in each case.

As a decision rule, conservatism in the face of risk is as unsatisfactory as the minimax or the "most likely event" principles, or indeed any principle that does not balance expected risks against expected benefits. On the other hand, the only

consistent (Bayesian) decision procedure requires information -- prior probabilities of all possible scientific hypotheses, utilities for all possible consequences -- which no regulator is likely to supply.

What about determining acceptable levels of exposure on the basis of a cost-benefit or risk-benefit analysis? There are, of course, well-known difficulties in quantifying benefits, costs, and risks. The danger that the estimates represent little more than disguised value judgments is always present. Problems of quantification aside, a number of rather stringent conditions must be satisfied before the cost-benefit criterion may be meaningfully used as a decision rule in health regulation.

First, it has long been recognized that cost-benefit analysis is not applicable under extreme circumstances, for example, when potential health damage is so large that marginal tradeoffs between the risks and the benefits of certain kinds of activity become virtually meaningless. Also, price-based marginal calculations (and the tâtonnement procedures necessary to discover a correct set of prices) are hardly appropriate when immediate action is required. To use Martin Weitzman's example, suppose that a certain number of airplanes is required for an emergency operation. In strict economic terms, it would be inefficient to issue orders to different commercial airlines to supply a given number of airplanes, since marginal opportunity costs will typically vary from company to company. Yet, in practical terms, this approach would be preferable to the economically correct procedure of announcing a price for plane services and letting profit-maximizing companies decide on the number of planes they would be willing to commit to the rescue operation.[17]

A third case in which the relevance of the approach is doubtful is when the biological effect of the toxicant to be regulated becomes evident only after a long time (perhaps 20 to 30 years). If people are unaware of such long-run effects, no externalities are generated and the cost-benefit criterion would indicate a status quo policy -- a solution which most toxicologists would consider irresponsible. Moreover, pollutants

such as cadmium, mercury, or radioactive wastes, which deteriorate very slowly over time, pose particular problems, since the damage they cause arises mostly from an irreducible stock, while only incremental damage is caused by the flow of pollution. Now, it is a basic assumption of the marginal calculus that the relevant variables can be controlled in all directions. In the case of persistent pollutants the stock of pollution is, to all practical purposes, irreducible, so that an essential feature of the cost-benefit approach is missing in this toxicologically important situation.

Perhaps the most serious difficulties in using cost-benefit criteria in standard setting arise in connection with the dynamic aspects of pollutant exposure. It can be shown that in a biologically unstable situation, a standard set at the point where marginal net benefits equal external costs -- corresponding to a "Pareto-optimum" level of pollution -- may fail to prevent continuing environmental deterioration and eventual destruction of the ability of organisms to cope with the permissible level of pollution.[18]

Aside from the technical and conceptual limitations of cost-benefit analysis, a key issue of regulatory philosophy is the role that considerations other than health and the environment should play in the standard-setting process. Debate on this issue has been particularly intense in the United States, and American legislation shows quite clearly the difficulty of reaching a consensus on the basic principles of regulation.

Thus, while the Toxic Substances Control Act, the Federal Environmental Pesticide Control Act, the Safe Drinking Water Act, the Federal Food, Drug, and Cosmetics Act (with the exception of the Food Additives Amendment of 1958), and the Occupational Safety and Health Act (OSH Act) call for some balancing of costs and benefits, the 1970 Clean Air Act Amendments, the Federal Water Pollution Control Act, and the Resource Conservation and Recovery Act are silent on this issue.

Even when the law requires some balancing of costs and benefits, the language is often ambiguous. In the case of the

OSH Act, for example, the courts had to determine whether Section 6(b) (5) of the Act, which speaks only of feasibility ("The Secretary ... shall set the standard which most adequately assures, to the extent feasible, on the basis of the best available evidence, that no employee will suffer material impairment of health or functional capacity ..."), permits OSHA to consider economic as well as technological factors in setting standards for toxic substances.

Since Congress has set few coherent guidelines on the extent to which benefits, as well as costs and risks, must be considered, regulators have to rely on ad hoc procedures to somehow balance these incommensurable factors. A former director of the Office of Toxic Substances of the Environmental Protection Agency has described the process:

> Usually the procedure is to postulate a numerical standard for a toxic chemical or a specific type of limitation on the use of the chemical, with the restriction designed to reduce environmental levels to the point that concerns over health or environmental damage disappear. Then an assessment is carried out to see if the favorable environmental impact from the restriction warrants the concomitant economic costs. If the costs are too high, the level of control is adjusted until an appropriate balance is reached.[19]

It is interesting to compare this approach with the philosophy of Soviet regulators (and of many scientists in both East and West).

Health standards, Soviet authorities maintain, should be based on health effects alone, without regard to the availability of adequate control technology, to economic feasibility, or even to the ability to adequately measure the concentrations in practice. A currently unattainable standard can still represent a guideline for enforcement and an incentive for future research in control technology. Conversely, since technically or economically attainable concentrations will coincide with harmless concentrations only by chance, standards based on considerations of economic or technical feasibility "can act only as an obstacle

to the search for better techniques, ... they sanction what has already been achieved without stimulating new technical advances."[20]

Since "scientifically based" standards cannot always be achieved, the Soviets also use secondary ("sanitary") standards that may modify, for a limited period of time, the requirements set by the primary ("hygienic") standards. Professor V.A. Rjazanov, a leading toxicologist, distinguishes between the two types of standards (in the context of air pollution) in the following terms:

> Hygienic standards ... must in themselves reflect the scientifically based ideal towards which we must strive in order to ensure that the public is not subject to unfavorable effects from air pollution. This ideal cannot be achieved always and everywhere at a given time. Therefore, alongside the general hygienic standards for maximum permissible concentrations, there may be sanitary standards of a temporary character, serving the needs of the moment. They may modify for a defined period the requirements for cleanliness of the external atmosphere, taking into account economic and technological factors ... Such air pollution standards are permissible temporarily, but should be abandoned after a certain period, during which the condition of the air must be brought into conformity with the hygienic standards. If this approach is adopted, hygienic standards for the cleanliness of the atmosphere will not be used to sanction existing technical achievement, but will represent the goal towards which we must strive.[21]

The criticism that standards used in the West tend to codify existing economic and technical conditions, to the detriment of their normative character, has some validity. It is often said that one of the main goals of environmental or health standards is to channel growth away from hazardous industries and materials toward safer forms of production and employment. But it is hard to see how a "feasible" standard (in the sense in which this term has been recently used) can provide the necessary signals. Consider, for example, the history of OSHA regulation of the carcinogen vinyl chloride (VC).

In April 1974, OSHA promulgated an Emergency Temporary Standard (ETS) reducing the previous National Consensus Standard for vinylchloride from 500 parts per million (ppm) to 50 ppm. The National Consensus

Standard of pre-OSHA times was a standard proposed by the American Conference of Governmental Industrial Hygienists, and voluntarily accepted by industry, at a time when it was unknown that VC could induce cancer. The statement of reasons supporting the ETS reveals that the 50 ppm standard was an uneasy compromise between conflicting considerations and interests.

During the summer of 1974, OSHA held intensive hearings on the proposal for a permanent standard of 1 ppm (as a time-weighted average over an 8-hour work period, with permissible excursions up to 5 ppm averaged over any 15-minute period). Although the disagreement on the medical evidence was considerable, most of the debate concerned the "feasibility" of the proposal. Industry opposed the proposed level of 1 ppm on the grounds that OSHA lacked sufficient evidence on the harmfulness of VC at low doses; that it was technologically impossible to meet the 1 ppm ceiling; and that the cost of approaching the ceiling would force the companies out of business. Conceding industry's claims of infeasibility, OSHA finally promulgated a somewhat weakened permanent standard.

Subsequent experience was to show that meeting the 1 ppm level was neither as difficult nor as costly as industry had predicted. The permanent VC standard lacks explicit criteria of feasibility, but a careful case study comes to the conclusion that "OSHA's statements and actions suggest that it was following an unarticulated principle that a standard is not feasible if it would cause more than slight changes in the number of firms in an industry, or in an industry's profit and growth rates, its output, and competitive position."[22]

The case of the vinyl chloride standard is far from being unique. American regulators are constantly urged to treat economic and technical feasibility as important considerations in the derivation of health and environmental standards. The result of these pressures has been an increasing confusion of the conceptually distinct stages of standard-setting and standard-using. The notion of aggregating scientific, technical, economic, and political data into a single value is appealing, but in practice it has led to logically inconsistent conclusions.

The meaning of the numerical value chosen for a given standard becomes ambiguous, representing neither a policy goal, nor a scientific judgment of health risk, nor even (since the standards are supposed to be enforceable at the national level) a measure of the level of protection that can be reasonably achieved in specific local situations. Whatever reservations one might have about the logic of the distinction drawn by Soviet regulators between primary and secondary standards, it must be admitted that at least it allows a clear statement of objectives, while avoiding the danger of sanctioning existing technical and economic conditions.

REGULATORY REFORM: SOME SUGGESTIONS

It is time to draw some conclusions from the preceding discussion. Environmental and health standards are, and will long remain, basic instruments of regulatory policy. At the same time, the standard-setting process rests on precarious conceptual, scientific, and economic foundations. This contradiction poses almost insoluble problems of administrative rationality and legitimacy. For example, the demand for "conclusive" scientific evidence or thorough risk analyses before a standard is adopted is more likely to delay public action than to improve the quality of decision-making, and to generate dissension rather than consensus.

What is needed is a fundamental restructuring of procedures, institutions, and evaluative criteria along lines that explicitly recognize the uncertainty and complexity of regulatory decisions. Two directions of regulatory reform seem to be particularly important. First, statutory regulations should be replaced as much as possible by non-statutory codes and standards; in particular, standard-setting should be clearly distinguished from standard-using. Although emphasis on compulsory standards is more characteristic of recent American legislation, pressure for statutory regulation is also building up in several European countries (in part as a result of the activities of environmentalist groups), and at the level of the European Community.

Second, greater attention should be paid to the procedural aspects of standard-setting than has so far been the case. With the present state of knowledge, it is unrealistic to require that regulatory decisions be supported by "proof" in the strict sense of the word. But cognitive uncertainty, far from justifying carelessness in choosing among alternative data, theories, and methodologies, in fact demands strong procedural controls to make sure that the implications of these choices are explored from a variety of viewpoints, and to facilitate a detailed factual analysis of the intellectual merits of the conclusions.

Concerning the first point -- the need for greater regulatory flexibility -- it is clear that environmental and health standards should be revised as scientific knowledge improves, empirical evidence accumulates, and socioeconomic conditions change. However, frequent revisions are unlikely (or very costly) when standards are embedded in legal codes. Also, the more uncertain the scientific basis of regulation and the greater the need for flexibility and adaptability, the more discretion should be left to the regulatory agency. But statutory regulation sets narrow limits to administrative discretion.

The experience of a number of European countries, particularly in the area of occupational health, shows that an effective regulatory system can be operated without heavy reliance on legally enforceable standards. In the Federal Republic of Germany and in France, maximum acceptable concentrations (MACs) for toxic substances and other environmental limits are not embedded in legal codes but are used by the inspectors -- together with other information on the physical, chemical, and toxicological characteristics of different substances -- for giving preventive advice and monitoring working and environmental conditions. MAC values and numerical standards are typically based on health criteria only. Guidelines interpreting the standards in the light of technical and economic constraints are issued by separate governmental commissions, such as the German Committee for Dangerous Materials in the Workplace (Ausschuss für gefährliche Arbeitsstoffe) set up by the Federal Ministry of Labor and Social Security.

In the United Kingdom, too, occupational health standards have no specific legal status, but are used by the Factory Inspectorate of the Department of Employment for control and surveillance of working conditions. A strong case for greater reliance on voluntary standards and codes is presented in the official report of the Parliamentary Committee appointed in May 1970 under the chairmanship of Lord Robens. Although the report deals with occupational health and safety, many of its arguments have more wide-ranging validity. The following recommendations of the Robens Committee are particularly relevant to the present argument:

- Wherever practicable, regulations should be confined to statements of the broad objectives to be achieved.

- In future, no statutory regulation should be made before detailed consideration has been given to whether objectives might adequately be met by a non-statutory standard or code of practice.

- Greater emphasis should be placed on standard-setting by means of non-statutory codes and standards. As a general rule, statutory regulations should only be made when the non-statutory alternatives have been fully explored and found wanting.

- The whole regulatory system should be more flexible and more discriminating. Industry should be encouraged to deal with more of its own problems, thereby enabling official regulation to be more effectively concentrated on serious problems where strict official regulation is appropriate and necessary.[23]

These recommendations express the belief that statutory regulations are largely ineffective, intrinsically rigid, and have a built-in tendency to become obsolete quite rapidly. On the other hand, "standards and codes developed within industry and by independent bodies are, over a large part of the field, more practical and therefore potentially more effective instruments

of progress than statutory regulations."[24] The Report concludes that what is needed is "less law" and more provision for voluntary self-regulation at the plant level.

However, in order to provide credible sanctions when needed, inspectors should have the power, without reference to the courts, to issue formal Improvement Notices, i.e., orders to comply not only with any relevant statutory regulation, but also with any relevant voluntary code or standard that has been formally approved by the Authority for Safety and Health at Work. Voluntary codes and standards would also be admissible evidence in proceedings before tribunals (the Report suggests that appeals against improvement notices should be heard not in the criminal courts but by the industrial tribunals set up under the 1964 Industrial Training Act). In cases where serious hazards or imminent dangers exist, the inspector could issue a Prohibition Notice ordering that, in the event of non-compliance within the stated time limit, the use of specified plant, machinery, processes or premises must be discontinued, or continued only under specific conditions.

But ensuring compliance with minimum legal requirements is not the main task of the inspectorate. Rather, inspectors should be concerned with the broad aspects of safety and health at the workplaces they visit, as much as with those narrow aspects which may have been the subject of detailed statutory regulations. "We believe," the Report states, "that, as a matter of explicit policy, the provision of skilled and impartial advice and assistance should be the leading edge of the unified inspectorate."[25]

The second direction of reform is concerned with what Herbert Simon has called "procedural rationality". In situations characterized by great uncertainty and cognitive complexity, Simon argues, "we must give an account not only of substantive rationality -- the extent to which appropriate courses of action are chosen -- but also of procedural rationality -- the effectiveness, in light of human cognitive powers and limitations, of the procedures used to choose actions."[26]

Policy analysis has been traditionally concerned with the problem of choosing the best means to achieve given ends. The basic conceptual categories of the policy analyst -- goals, alternatives, impacts, effectiveness, choice -- clearly reveal his deep commitment to a teleological conception of policy making. According to this conception, rationality means maximizing something; it means choosing the best alternative, subject to a set of constraints. Hence the preoccupation with methods of analysis and evaluation that emphasize outcome rather than process, and the interest in what decisions are made, rather than in how they are made. As a result, policy analysis lacks the methodological equivalents of legal notions like reasoned decision, proper form, and rules of evidence.

This indifference toward procedures and the formal layout of arguments is justifiable under the assumption that there is "one best way" of making a decision or, if several methods are possible, that there is a well-defined rule for choosing among them. This is certainly not the situation in standard-setting. Here, Jerome Cornfield points out, "[a]ll present safety evaluation procedures ... must be regarded as mathematical formalisms whose correspondence with the realities of low dose effects is, and may long remain, largely conjectural."[27] Thus, the most important problem is not determining the "correct" value for a certain standard -- is it 5 or 2 ppm? -- but which criteria and procedures should be used to choose among competing models, approaches, and regulatory philosophies.

In other words, the main problem with many environmental policy decisions is not that they are, in some sense, suboptimal (we generally lack the scientific and medical knowledge to know what the correct decision should be), but that they leave much to be desired in terms of procedural rationality. Standard-setters often fail to probe deeply into the quality of the available evidence, or to test the sensitivity of the chosen model to uncertainty and alternative assumptions. Even more commonly, the methodology used in reasoning from the data to a proposed standard is so informal that it is impossible to retrace the steps of the agency's argument and its factual basis. Again,

the standard-setting process usually does not include any procedures specifically designed to bring out unstated assumptions, differing interpretations, and gaps in logic or in the empirical evidence.

An interesting example of procedural reform in the area of environmental regulation comes from the United States. Here the courts have recently developed "paper hearing" procedures that combine many of the advantages of a trial-type adversary process (without oral testimony and cross-examination), while avoiding undue costs and delays in decision-making. The procedural requirements imposed by the courts on the Environmental Protection Agency have been summarized by Pedersen:[28]

> First, both the essential factual data on which the rule is based and the methodology used in reasoning from the data to the proposed standard must be disclosed for comment at the time a rule is proposed ... Second, the agency's discussion of the basis and purpose of its rule -- generally contained in the "preambles" to the notices of proposed and final rule-making and in the accompanying technical support documents -- must detail the steps of the agency's reasoning and its factual basis. Third, significant comments received during the public comment period must be answered at the time of final promulgation. However, comments must meet a standard of detail equal to that required of the agency in promulgating its rule before they will be considered significant. Fourth, only objections to the regulations which were raised with some specificity during the public comment period, and to which the agency thus had an opportunity to respond, may be raised during judicial review.

Although these requirements are only a first step, and much remains to be done in reducing the ineffectiveness and rigidity of the present system and its built-in tendency to become obsolete, there is already some evidence of improvement in the quality of environmental decision making. Data and technical studies are collected and organized more systematically; external criticism is explicitly taken into account so that policies reflect a broader range of considerations

and interests; the various subunits of the regulatory agency are motivated to coordinate their assessments, methodologies, and conclusions. The new procedures should also increase the influence of the people who, because of their special knowledge, are more directly involved in standard-setting.

I would argue that the experience of the "paper hearing" procedures developed at EPA under the Clean Air Act has general relevance. The requirement of an open record that includes the factual and methodological bases of an agency's conclusions, as well as external criticism and responses to such criticism, is always a powerful incentive to more careful agency deliberations. The need to improve the intellectual quality of administrative deliberations is not, however, the only reason why procedural questions are so important today. In situations of great complexity and cognitive uncertainty it is essential that the groups affected should be willing to accept the outcome of the administrative process even before this has been determined. By ensuring adequate representation of conflicting opinions and examining a wide range of alternatives, well-designed procedures can greatly improve not only the rationality but also the legitimacy of regulatory decisions.

REFERENCES

1. Philip Selznick, Leadership in Administration, New York: Harper and Row, 1964, p. 56.

2. Giandomenico Majone, "The New Perspective on Health: Prevention and Health Standards", Laxenburg, Austria: International Institute for Applied Systems Analysis, PP-81-6, March 1981.

3. Roger I. Glass, "A perspective on environmental health in the USSR", Archives of Environmental Health, vol. 30, August 1975, pp. 391-395.

4. Theodore F. Hatch, "Permissible levels of exposure to hazardous agents in industry", Journal of Occupational Medicine, vol. 14, 1972, pp. 134-137.

5. Ib.. p. 135.

6. V.A. Rjazanov, "Criteria and methods for establishing maximum permissible concentrations of air pollution", Bulletin of the World Health Organization, vol. 32, 1965, p. 392.

7. Bertram D. Dinman, "Development of workplace environment standards in foreign countries, Part 1", Journal of Occupational Medicine, vol. 18, no. 6, 1976, pp. 409-417.

8. Mark G. Field, Soviet Socialized Medicine, New York: The Free Press, 1967, ch. 9.

9. Dinman, cit.

10. Glass, cit.

11. Food and Drug Administration Advisory Committee on Protocols for Safety Evaluation. Panel on Carcinogenesis: "Report on cancer testing in the safety evaluation of food additives and pesticides". Toxicology and Applied Pharmacology vol. 20, 1971, pp. 419-438.

12. Edward J. Calabrese, Methodological Approaches to Deriving Environmental and Occupational Health Standards, New York: John Wiley and Sons, 1978, p. 47.

13. Ib., pp. 48-59.

14. Jerome Cornfield, "Carcinogenic risk assessment", Science, vol. 198, 18 November 1977, pp. 693-699.

15. Food and Drug Administration, "Report on cancer testing", cit., p. 431.

16. Cornfield, cit.

17. Martin L. Weitzman, "Prices vs. quantities", Review of Economic Studies, vol. 41, October 1974, pp. 477-491.

18. David Pearce, "The limits of cost-benefit analysis as a guide to environmental policy", Kyklos, vol. 29, Fasc. 1, 1976, pp. 97-112.

19. Glenn E. Schweitzer, "Toxic chemicals and regulatory decision making: philosophy and practicality", in National Academy of Sciences, Decision Making for Regulating Chemicals in the Environment, Washington, D.C.: 1975, pp. 72-73.

20. Rjazanov, cit., p. 390.

21. Ib., p. 390.

22. David D. Doniger, The Law and Policy of Toxic Substances Control, Baltimore: Johns Hopkins University Press, p. 65.

23. Safety and Health at Work, Report of the Committee 1970-1972, Chairman Lord Robens, London: H.M. Stationery Office, Cm 5034, 1972, pp. 44-46.

24. Ib., p. 48.

25. Ib., p. 65.

26. Herbert A. Simon, "Rationality as process and as product of thought", American Economic Association Proceedings, vol. 68, May 1978, p. 90.

27. Cornfield, cit., p. 698.

28. W.F. Pedersen, Jr. "Formal records and informal rule making", Yale Law Journal, vol. 85, 1975, pp. 75, 76.