MANAGEMENT SCIENCE
Vol. 30, No. 7, July 1984
Printed in U.S.A.

SETTING STANDARDS: A SYSTEMATIC APPROACH TO MANAGING PUBLIC HEALTH AND SAFETY RISKS*

BARUCH FISCHHOFF
Decision Research, 1201 *Oak Street*, *Eugene*, *Oregon* 97401

Standards are an effective means for managing hazardous technologies only if three conditions are satisfied: (a) setting general standards is preferable to case-by-case decision making; (b) some general safety philosophy, balancing risk and other factors, can be justified on normative grounds; (c) that philosophy is faithfully translated into operational terms. In practice, standards are rarely developed and enforced in an integrated systematic way. As a result, they often miss their mark. This guide presents a general framework for the design, development, and implementation of safety standards. That framework is derived from the logical character of the standard setters' task and from experience with actual standards. It first identifies the conditions under which standards are an appropriate management tool. Second, it presents four generic methods that may be used to develop safety policy. Third, it characterizes the design issues that arise in making that policy operational. At each step, it suggests particular strategies along with their inherent strengths and weaknesses. In particular, it shows the sensitivity of a standard's effectiveness to seemingly technical aspects of the way it is drafted.
(GOVERNMENT REGULATION; HAZARD MANAGEMENT; COST-BENEFIT ANALYSIS; DECISION ANALYSIS; RISK)

1. Introduction

Standards are a pervasive feature of modern society. They are set by governments to protect public safety and welfare; by professional organizations to assure consumers of the quality of their work; by manufacturers to protect themselves against liability claims. Familiar standards include the Code of Federal Regulations, national ambient air quality standards, auto emission limits, the Good Housekeeping Seal of Approval, NCAA helmet specifications, the Geneva Conventions, and Underwriters' Laboratories requirements. The form of these standards is as varied as their source and subject matter. Some refer to allowable effluents; others specify the physical properties of machinery; still others define permissible impacts on the environment. Their one common feature is that each separates activities or technologies into two categories: those that meet the standard are acceptable; those that do not are unacceptable.[1]

The present analysis proposes a systematic, general approach to setting standards, focussed on those standards used to manage risks to public health and safety. It views standard setting as involving three distinct stages. First, there must be a decision to rely on standards. Second, a safety philosophy, balancing risk and other factors, must be determined. Finally, that philosophy must be translated into operational terms. The paper describes problems that arise at each stage and offers guidelines to their resolution. Descriptively, it can be used to predict how successful particular standards will be. Prescriptively, it can be used to design sounder standards (and to identify the limits of standards as a managerial tool).

In practice, standards are often applied quite flexibly. Here, they are treated as fixed, explicit rules. Such literal interpretation is essential for determining whether the

* Accepted by Warren E. Walker, received July 1982. This paper has been with the author 4 months for 2 revisions.

[1] Some pertinent references include: Driver (1980), Portney (1979), American Enterprise Institute (1980), Lave (1981).

0025-1909/84/3007/0823$01.25

operational form of a standard is faithful to the intent of the standard setters. The translation process is shown to involve many apparently technical specifications that can make the effective tradeoffs realized by the operational standard quite different from the intended ones.

2. When to Set Standards

There are two archetypal ways of deliberatively managing hazardous activities and technologies—decision making and standard setting. By considering the differences between these two procedures, it is possible to identify the conditions that are most (and least) conducive to standard setting.

A decision is a choice between alternative courses of action. Standards, on the other hand, pass judgment on individual courses of action. Decision-making procedures attempt to order options according to relative attractiveness; standards simply categorize them as "pass/no pass." Applied to a set of proposed actions, decision making attempts to identify the one action that is most desirable; reliance on standards may lead to all actions being accepted or to none being judged adequate. If an option has been judged acceptable, it will remain so until either the option or the standard is changed. The judgment rendered by a standard is, thus, irreversible. Decision making is conditional; if a better option comes along, then the selected option may be replaced. With decision making, options must beat the competition; with standard setting, they must beat the standard.

Once set, a standard is meant to be applied repeatedly. The same fixed rule, expressing a particular set of values, is used to judge all proposals falling within its jurisdiction. Decision making is tailored to the particulars of individual choice problems, deriving the implications of the decision maker's values for each. The tradeoffs between benefits, risks, and nonrisk costs are considered only when standards are set. Their application is an administrative act, unlike decision making which is always a political act.

Identification of these differences provides a basis for circumscribing the conditions in which it makes sense to set and apply a standard rather than to make a decision. Conditions that favor standard setting are listed in Table 1 and are discussed below.[2] In particular, standards are appropriate in the following situations:

(1) *When no choice among options is possible*. Many hazard managers do not have the opportunity to weigh the relative merits of competing options. Rather, they must judge each on its own merits, without explicitly considering what will come in its stead if it is rejected or what will be foregone if it is accepted. For example, industrial trade associations are seldom empowered to decide which product is best; government regulators typically must act independently upon each option presented to them. A general standard will allow them to treat all options in a consistent manner.

(2) *When no choice among options is required*. Much of the effort required by decision-making schemes is invested in ordering all available options so as to identify the best one. That effort will be wasted if it is possible to choose several options, or none at all. For example, society need not restrict itself to one energy source, nor need it accept any member of a new family of prescription drugs. In such cases, all that is required is an accept/reject rule.

(3) *When predictability is important*. Standards can simplify life for hazards as well as for their managers. They offer a clear and fixed target that a technology's designers

[2] Although the remainder of this discussion considers actions by policy-making or regulatory bodies, a comparable choice of procedure faces individuals in managing the hazards in their own lives. That is, they can deal with each hazard in isolation or establish a general rule for dealing with many hazards. A parallel analysis generates some interesting hypotheses regarding individual choice behavior.

TABLE 1

Feature	Standards	Decisions	Conditions Favoring Standards
Number of options chosen	None at all	One	(1) No choice possible (2) No selection needed
Task facing options	Satisfy standard	Beat opposition	(3) Predictability important (4) Future options need shaping
Range of application	Category	Single case	(5) Competing technologies in same jurisdiction (6) Category members homogeneous
Expression	Rule	Choice	(7) Explicit policy attractive (8) Value issues sensitive
Application	Technical	Political	(9) Political resources limited (10) Process unimportant
Flexibility	Little	Great	(11) Awkward applications avoidable

can strive to meet. Once the standard has been met, approval is irreversible, allowing the designers to concentrate on factors other than safety (e.g., cost containment, market share). The desire for predictability has led the nuclear industry to propose that the U.S. Nuclear Regulatory Commission (1982) adopt an overall safety standard. Once a power plant has met this standard, it would no longer be subject to the Commission's case-by-case decisions regarding what designs and procedures to adopt. Although meeting the overall standard might be more expensive than abiding by the local decisions, doing so would eliminate worry about retrofitting and withdrawal of operating permission.[3]

(4) *When regulators hope to shape future options.* When regulators do not like the choice of options offered to them (or to the public), they may set a standard that is clear, predictable, but out of reach. Such "technology-forcing" standards are intended either to require a particular product or to give a competitive advantage to those who produce it. U.S. automobile fuel economy goals, for example, were intended to do the former, but ended up doing the latter as well. Designed to reduce the environmental and national security risks from excessive gasoline consumption by coercing manufacturers to produce more efficient cars, they inadvertently helped the industry meet foreign competition. On the other hand, when standards mandate a particular technical solution, they can stifle innovation, perhaps even increase monopoly pressures. For example, requiring the catalytic converter as a solution to pollution control greatly benefitted General Motors, the leader in its development (*Wall Street Journal* 1979). Standards generally favor large corporations which have greater resources to track, fight, shape, and adjust to them. Anything that decreases competitive forces will also tend to reduce innovation.

(5) *When competing technologies fall in the same jurisdiction.* Standards are meant to apply to a category of options. If that category includes one method of accomplishing a particular task, it should include its competitors as well (e.g., all ways to generate electricity, all modes of urban transport). If competing options are subject to stronger or weaker standards (or to none at all), then these inconsistencies may be exploited. For example, tough standards on hazards with acute health effects will, over time, encourage the development of competing products and processes with primarily chronic health effects. If a stricter overall safety standard is adopted for nuclear power

[3] An application of this theory to the Commission effort can be found in Fischhoff (1983). As currently proposed, the goals would be an addition to case-by-case decisions, at least for the foreseeable future.

than for other energy technologies, then it will be at a disadvantage relative to its competitors (which have at the moment no such standard). The range of a standard may change over its working life. Each technology affected by a standard may try to be released from its constraints or to have them extended to its competitors.

(6) *When category members are homogeneous*. Having one standard for all category members inevitably means that it will be more suited for some than for others. For example, a uniform standard for atmospheric asbestos levels will allow quite different risks in workplaces with high and low proportions of smokers (Selikoff and Lee 1978). It may exact quite different costs of compliance in different settings (and, hence, implicitly assign different values to human life and health). A standard that allows a level of risk that is commensurate with the benefits derived from the average member will seem unduly restrictive when applied to a member with particularly high benefits. The extent of such "injustices" should depend upon the heterogeneity of the technologies within its jurisdiction. It is possible, in principle, to develop a standard with correction factors to cope with special cases. However, doing so produces a standard that is more complicated, more discretionary, more situation-specific—and less a standard.

(7) *When an explicit policy statement is attractive*. When a decision is made, observers may see just the act of choice, which reveals little about its underlying rationale. When that rationale is explicated, it may involve a complex tradeoff between diverse considerations. By contrast, many standards offer a tidy statement of principle saying, "This is what we allow." That statement may tell the whole story, or it may just summarize the results of more complex deliberations. For example, after weighing many factors in deciding to allow fishing in Chesapeake Bay but not in the James River, the Environmental Protection Agency announced a standard for Kepone in fish that lay between the levels found in those two bodies of water (Johnson 1980). That statement may be meant literally or figuratively, as an expression of policy maker's concerns. For example, Eastern European countries want strong standards for worker protection on their books, even if economic problems limit their realization (Derr et al. 1981). The Delaney Amendment made a similarly strong, if impractical, statement of American values toward health and food. Whatever their original motivation, simplifications may develop lives of their own. They create both precedents and expectations (on the part of producers, workers, and consumers).

(8) *When value issues are sensitive*. Because they always tell some story, but need not tell the whole story, standards can be used to blur value issues (as well as to highlight them). Unlike formal decision-making methods, such as cost-benefit analysis or decision analysis, standard setting need not leave an audit trail explaining its rationale. Such vagueness may cover confusion or it may be designed to obscure what the standard was walling in and walling out. It might, then, serve vested interests who want a high-sounding rule to legitimate their actions, or regulators who must make political decisions, but lack the mandate to do so, or warring parties who can negotiate a compromise within the context of a standard that hides what each has conceded.

(9) *When political resources are limited*. Every decision is a political act, requiring a statement of values regarding the appropriate tradeoffs between conflicting objectives. Particularly where the tradeoffs are difficult and society is divided, it may be hard to muster the political resources needed to reach a decision. When many tough decisions must be made, one way to keep the political system from being overwhelmed is to replace some "small decisions" with one big one, by choosing a general standard. That standard can then be applied administratively by technical experts. Three conditions for this strategy to work are: (a) the standard setters tackle and resolve the value issues; (b) the operational standard leaves little room for discretion; (c) the technicians receive adequate resources to apply the standard faithfully. For example, the Toxic

Substances Control Act made a strong value statement, but failed to provide sufficiently precise guidelines or adequate funds for the Environmental Protection Agency to devise a defensible plan, ensuring that the tens of thousands of existing chemicals receive equal justice under the standard (Shaw and Canape 1980, US National Academy of Sciences 1981, 1982).

(10) *When process is unimportant*. By concentrating political attention on the standard-setting process, standards do away with the repeated discussion of value issues that comes with case-by-case decision making. To realize these potential increases in efficiency, the standard-setting process must be able to replace the political and intellectual functions of repeated decision making. Part of the political function can be served by ensuring that varied parties interested in an issue are represented in setting the standard for it. Concentrated participation in a standard is not, however, the same as continued participation in individual decisions, each of which offers an opportunity to educate oneself, monitor the regulators and generate public discussion. The intellectual functions are most likely to be filled when the issues are well understood in advance, so that standard setting centers on the resolution of known conflicts. The very acts of grouping options into categories and searching for a consistent rule may even generate new insights. However, when problems are too new or complex to be understood in advance, they defy the setting of informed, lasting standards. Although it is possible to have a standard evolve through trial-and-error learning, doing so would be contrary to the notion of a general rule. Thus, standards are inappropriate where a political or learning process is essential.

(11) *When awkward applications can be avoided*. Once any policy has been made (in the form of a decision or standard), those affected adversely may try to overturn it. Standards are inherently vulnerable to attacks that focus on specific cases in which the general rule seems ill suited. For example, standards look bad when they reject otherwise attractive alternatives whose risk levels are just beyond the threshold of acceptability, especially when it would be very expensive to gain the small decrease in risk needed to achieve compliance (Derby and Keeney 1981). Even if they cannot have the standard overturned on the basis of one sticky example, critics may still obtain a precedent-setting relaxation. One way to defend a standard is to avoid awkward applications, perhaps by specifically excluding them or showing how to bend the standard deliberately (to keep it from being broken).[4]

Summary. Reliance on standards differs from case-by-case decision making in certain fundamental ways. From these, it is possible to derive a set of conditions that are particularly conducive to standards. When standard setting is mandated, consideration of these conditions can help one to anticipate and perhaps avoid some of the problems that will be encountered. When a choice can be made between a general standard and specific decisions, then this kind of analysis can point to the tradeoffs that should underlie that choice. A typical tradeoff question would be whether the timeliness and predictability of a standard's application would compensate for its occasional injustices.

3. Justifying Standards

Deciding to rely on a standard sets into motion one big decision, choosing a safety philosophy, followed by many small ones, enountered in translating it into operational terms. Often that translation process leads to an operational standard that is expressed

[4] Another part of the defense is being alert to bogus claims, such as that the money required to bring a particular technology in compliance with a safety standard would be better spent on high-sounding and far-flung alternatives (e.g., mobile trauma units), even though there is no chance of funds being transferred from one realm to the other.

in terms of some measure of risk, such as a level of toxic emissions, a failure rate, or expected casualty toll. Because that level expresses the maximum risk that will be tolerated in a technology it is often termed the "acceptable level of risk." Although perhaps technically correct, that term obscures the fact that benefits and nonrisk costs are considered in the setting of most standards (otherwise the tolerated risk level would be near zero). If the standard is not simply a political statement about safety, the focus on risk may reflect the feeling that the other consequences are relatively fixed (hence, "here are the risks that those net benefits[5] can justify").

The term "acceptable" can be misleading in other ways as well. It may connote that society is satisfied with that risk level, rather than just resigned to it in the light of current economic and technical realities. It may also connote that everyone exposed to that risk should be content with risk levels that reflect a compromise among the particular interest involved in setting the standard, or are the result of weighing the interests of society as a whole (Otway and von Winterfeldt 1982).

However it is labeled, the product of the standard setters' deliberations is a safety philosophy, expressing their view of an appropriate tradeoff between risks and other consequences. Four generic methods can be identified for justifying such tradeoffs. They differ both in the information that they consider and the composition rule that they use to integrate that information. The choice among them is in part an empirical question: Just how well does each fulfill its promise of eliciting and accommodating a particular range of information? The choice is also a political one: Each method embodies a somewhat different perspective on whose views count most and how political processes should be conducted, hence the method chosen can affect society beyond the context of the particular standard.[6]

3.1. *Formal Analysis*

Applying standards rather than making case-by-case decisions means that one cannot apply to those decisions the sophisticated techniques of decision analysis, cost-benefit analysis and related procedures (e.g., Raiffa 1968, Stokey and Zeckhauser 1978). Those techniques might, however, be used to guide the choice among alternative standards. When so applied, they would try to estimate the consequences following from implementation of each possible standard, in order to identify the candidate offering society the best deal. Those consequences would be estimated in terms such as the costs of compliance, the beneficial health effects expected, and the cost-benefit offerings of substandard projects (Keeney 1981).

Here, as elsewhere, formal analysis is more readily proposed than accomplished. It is hard to anticipate all significant consequences of a standard; it is hard to calculate the magnitude of many consequences that can be anticipated (e.g., reduced incentives to innovation or increased respiratory disease); it is hard to put a consensual (or societal) value on many consequences, such as health effects or loss of life (e.g., Crandall and Lave 1981, Lave 1981, Salem et al. 1980). These difficulties are, of course, inherent in the problem. By highlighting them, formal analysis may increase the chances of their being solved, or at least of being recognized when the standard is operationalized and applied.

Formal analysis often shapes standards indirectly through various parties' analyses of how different possible standards affect them personally (May 1982). The adopted standard is often the one that looks best on the private balance sheets of the most

[5] Net benefits will be used to represent the difference between benefits and nonrisk costs.

[6] A fuller general discussion of these arguments, without the focus on their usage for standard setting, may be found in Stokey and Zeckhauser (1978). That source also provides extensive references.

powerful parties. Modeling these hedonic perspectives is one way to design standards that will be well received.[7]

3.2. *Professional Judgment*

The technical community that creates technologies has always been the primary source of standards for governing them. Builders determine building codes; engineers specify acceptable metal fabrication techniques; medical personnel prescribe the conditions for licensing new drugs. In addition to such explicit expressions, other standards are implicit in the norms of accepted professional practice. The putative advantage of such standards comes from professionals' skill at identifying people's desires and reconciling them with one another and with technical realities. Devising creative solutions lies at the heart of the professionals' craft. Furthermore, their intimate involvement with the technology should ensure that the standard will be translated faithfully into operational standards and applied appropriately.

The trouble with letting professionals set standards is that they are a vested interest. They must balance the public's welfare with their own, that of their immediate employers, that of their industry, and that of the national economy. Depending upon the situation, there might be temptations to allow too much risk (in order to ensure a technology's survival), to require too much safety (in order to reduce professionals' legal liability), or to design highly technical standards (in order to provide enforcement work for professionals). Professionals' ability to achieve a judicious balance may be greatest when the other vested interests are equally capable of censuring them: that is, when manufacturers may refuse to employ them, consumers may refuse to buy their creations, legislators may disband (or disregard) their licensing boards, and other professionals can monitor their actions.

3.3. *Political Processes*

Reliance on either formal analysis or professional judgment means entrusting some group of technical experts with setting public policy. The justification for this transfer of power is that they are capable of making decisions in the public's best interest, perhaps even more capable than the public itself which is often accused of having a flitting involvement with problems, unstable preferences, and limited understanding of risk issues (Starr 1981, Fischhoff et al. 1983). Whether the experts can produce thoughtful, balanced solutions and whether the public really needs their help are empirical questions. If the answers are negative, then the obvious alternative is to leave (or return) primary responsibility for standard setting to political processes such as referenda, legislatures, commissions, and, by some accounts, courts (Mazur 1981, Nelkin and Pollak 1980).

The laypeople taking part in these processes can solicit expert opinion, but need not be bound by it. For example, they may choose to ignore, override, or adjust the experts' recommendations when the experts have ignored important consequences that are hard to quantify, been insensitive to the public's full risk burden, represented a vested interest, or exaggerated the definitiveness of their analyses (Fischhoff et al. 1983). Although political processes cannot promise to identify the optimal standard directly, the ruthless criticism of the political arena should reveal the deficiencies of suboptimal ones. Proposals that survive may have accumulated a considerable constituency.

[7] An even more indirect way to rely on formal analysis is possible when the set of future technologies can be anticipated. Analysis could be used to identify the best options. The standard could then be drafted so as to permit those options and no others (i.e., by finding some dimension(s) on which they had the most extreme scores).

The credibility of a particular political process depends upon how representative it is, how well it exploits existing technical knowledge, how well its lay participants resist bullying by expert opinion, and how committed participants are to the issue at hand (as opposed to using it as a venue for addressing other problems).

3.4. *Revealed Preferences*

The preceding methods might be called "expressed preferences," the standard that emerges is what some people say it should be. The revealed preference method substitutes action for words by codifying the tradeoffs implicit in past decisions. A weak revealed preference claim is that no new product should be riskier than the riskiest product on the market (unless it had unprecedented benefits).[8] A stronger one is that a new technology can be as risky as existing technologies with comparable benefits. Even more specific would be a safety philosophy that required the adoption of a safeguard only if its cost[9] were less than the market value of its expected reduction in mortality and morbidity.

As the argument goes, the historical record reveals not only real behavior, but also appropriate behavior. Through trial and error, the desires of all citizens involved with existing technologies have been incorporated in its current cost-risk-benefit tradeoffs. If a technology is too risky, then people will stop buying it until it has been made safe. If they have suffered from a technology, then they will turn to the courts for redress, forcing safer products in the future. If a technology is too safe, then either people will use it more dangerously so as to get more benefit out of it (Wilde 1982) or they will turn to riskier, but cheaper alternatives.

In order for a particular tradeoff to reveal society's preferences, it must have been achieved in conditions like those needed for efficient market operation. These include: (a) information is shared equally by all actors; (b) all actors optimize their hedonic preferences in each action; (c) a full range of choices, varying in risks and benefits, has been offered; (d) the risky consequences of technologies can be unambiguously assigned to them.

The preferences so revealed are necessarily those of the past, reflecting its tastes and mores, as well as its distribution of the power to influence social and economic realities. If the past appears to have been unjust, or insensitive to risk issues, then so will be its tradeoffs. Changes in taste should either invalidate past experience or require the application of some correction factor to it.

3.5. *Summary*

The four generic approaches to setting standards that were discussed above are summarized in Table 2. They differ in the perspectives that they consider and the method used for integrating them. The choice among them in a particular case is in part an empirical question: To what extent will the potential advantages and disadvantages of each be realized? The choice is also political: How important are these advantages and disadvantages? Hybrids are also possible. For example, the U.S. Nuclear Regulatory Commission (1982) has combined the political and professional approaches in its overall safety goals for nuclear power; although prepared by the Commission's technical staff, the proposed goals were presented periodically to a panel representing diverse constituencies for shaping.

[8] A nonrisk variant of this approach may be found in Dawes' (1971) description of an attempt to discourage unpromising candidates to graduate school by suggesting that they apply to themselves a rule specifying the profiles of candidates that had never been chosen in the past.

[9] Or, its net cost, being the cost minus any benefits other than risk reduction (e.g., increased productivity due to better working conditions).

TABLE 2
Methods for Setting Standards
(*How Can Standards Be Justified?*)

Approach	Locus of Wisdom	Description	Potential Advantages	Potential Disadvantages
Formal analysis	Formalized intellectual processes	Choose standard offering highest utility (or best cost-benefit tradeoff)	Systematic explicit sophisticated techniques	Impractical oversold centralizes power
Professional judgment	Intuitive intellectual processes	Let technical experts identify best standard	Realistic implementable creative compromises	Vested interests incomplete perspectives inscrutable
Political processes	Body politic	Have lay groups set standards, informed by technical advice	Broad perspective legitimacy open to criticism	uninformed unrealistic unstable
Revealed preferences	Past social processes	Adopt standard implicitly emerging in actual decisions	Reflects deeds shaped through experience influenced by whole society	Inefficient unfair insensitive to risk

4. Making Standards Operational

Once a safety philosopy has been determined, it must be translated into operational terms. That act of translation requires answering the questions of formulation discussed below. The answers will determine how readily the standard can be understood and monitored, how likely it is to be overturned by courts or legislatures, how porous it will prove to the search for loopholes. If the answers are poor, the tradeoff emerging from the operational standard may be quite different from that intended by the standard setters. Thus, although their task is ostensibly technical, the translators may deliberately or inadvertently set social policy. The standard formulators must address these issues, or let them be resolved ad hoc.

4.1. *What Should Be the Range of Application?*

The formulators' first task is to define the technologies to which the standard applies. That definition has two aspects: (a) which technologies are in its jurisdiction, and (b) what aspects of each technology are covered by the standard.

4.4.1. *Which Technologies Are to Be Included?* Whenever compliance with a standard imposes a serious burden, technologies and those responsible for them will be motivated to seek a change of venue. They may hope to be treated by other, more lenient standards, or to avoid regulation altogether. If standards are to work as intended, then their formulators must specify their jurisdiction in such a way that they are not imposed upon technologies for which they create too great a burden, that they do not become a refuge for technologies escaping more onerous standards, that they do not allow exemptions for all but those technologies for which compliance is trivial, and that they do not inadvertently subject technologies to multiple regulation, whose cumulative burden is much greater than what was intended by the developers of the individual standards.

There are three basic ways to define categories. One is to enumerate all members; that leaves nothing to chance, as long as all potential members are known or can be anticipated. The second way is to specify analytically the properties of category members, in terms of structure (e.g., all polymers), function (e.g., all modes of transportation), or consequences (e.g., all carcinogens). Such a definition can be applied to any present or future technology providing the designers can anticipate jurisdictional conflicts and ambiguities. The third way is to identify the archetypal category member, and then assign technologies to the category whose archetype is most similar to them. The standard itself will be set in terms of that archetype with instructions on how to adjust it in the light of possible deviations. For this strategy to work, competing jurisdictions must also specify archetypes, and a metric is needed for assessing similarity.[10]

It is beyond the scope of the present analysis to provide a more detailed discussion of the problems faced by these three procedures (or, indeed, for many of the other design issues raised below). As but one example of the kind of issue that can arise on the next level of detail, consider McNamara's (1980) description of the esoteric, but financially consequential, deliberations regarding whether cosmetic products, such as lip balms, toothpastes, and suntan lotions, are also drugs (hence, subject to the more stringent drug licensing standards).

> *Wrinkle removers/smoothers.* While it may seem incredible, there have been five published judicial decisions, arising in three different FDA enforcement actions, concerning this class of product. Speaking generally, all of this litigation may be summarized as leading to the following conclusion: If you represent a product as 'removing' wrinkles, you are in danger of having it regulated as a drug by the FDA; if, on the other hand, you represent the product solely as a wrinkle 'smoother,' it may be subject to the regulation solely as cosmetic. Note that two such products identical in chemical composition may nevertheless have different regulatory status depending upon their claims (pp. 470–471).

In this light, the seemingly frivolous debate over whether Certs is a candy mint or a breath mint may have serious legal ramifications.[11]

4.1.2. *What Aspects Should Be Covered?* Technologies typically have both spatial and temporal extent. At any point in time, they may be used in various methods and places. Any given use evolves over time through its life, material, or fuel cycle. In both regards, the formulator must specify what constitutes a technology for the purposes of the standard.

Spatially, the standard could pass judgment on a technology such as asbestos on the basis of each usage, each point of manufacture, each manufacturer, or the entire industry. As the definition broadens, the consequences of failing to meet the standard increase. The probability of that happening may decrease because of the political pressure to keep large economic units alive (even if it means overturning a standard). As the definition narrows, the responsibility for compliance becomes more focused, increasing the chances of individual users being held accountable for excesses. On the other hand, the absolute magnitude of the risk arising from small units may make them too inconsequential to regulate seriously (American Public Health Association 1980).

Temporally, a technology such as asbestos goes from mining to milling to manufacture to usage to intermediate and final disposal. Each stage produces some risks; some

[10]The actual standard set here constitutes a decision regarding the disposition of the archetypal member. In some cases, the detailed analysis of that member will improve understanding of it, thereby increasing its safety (by uncovering design problems that are addressed voluntarily); hence, the standard may have to be made more restrictive to ensure that other members are as safe. Alternatively, such close scrutiny will lead to an exaggerated perception of all the things that could go wrong with the archetype. The result would be unduly tight regulations, which should be relaxed for other members.

[11]The *Food, Drug, and Cosmetic Law Journal* is a rich source of such issues. See, for example, Shaw and Canape (1980), McNamara (1982), Anastasio (1981).

also produce benefit. An extensively defined standard would consider the tradeoffs between cradle-to-grave risks and benefits from a technology. Even though the benefits are most apparent at the stage at which a technology is used, they reflect the cumulative value accruing from the entire production process, hence they should be compared with its cumulative risk. The comprehensive view would give manufacturers the freedom to look through the entire fuel cycle for the most efficient way of achieving the requisite safety (Whipple 1981).

The comprehensive view is less acceptable when different people receive the risks and benefits accruing from the different stages. Those who bear risks may want risk reduction or compensation and be unmoved by the bargains to be realized by cost-effective expenditure of safety resources elsewhere; those who enjoy benefits may be reluctant to pay for or be restricted because of risks occurring far away. If a technology is required to balance risks and benefits at every stage, then it is no more acceptable than its weakest link. Whereas the comprehensive definition allows excess safety at one stage to compensate for inadequate safety at another, the stage-wise definition would require compensation actually to be paid, to sweeten the burden of the riskier stages. This way of defining technologies would likely result in greater expense or less risk.

4.2. *What Should Be the Locus of Standard?*

Just as technologies go through a number of stages from inception to demise, so do hazardous events go through several stages en route from constituting to realizing a threat (Hohenemser et al. 1983). Figure 1 shows one conceptualization of how automobile accidents go through each stage. Similar figures could trace the path of air pollutants up the smoke stack, through the air, into people's lungs and into health effects, or the path of infectious diseases into and through public water supplies. It is possible, in principle, to block the event by interventions at each stage, designed to

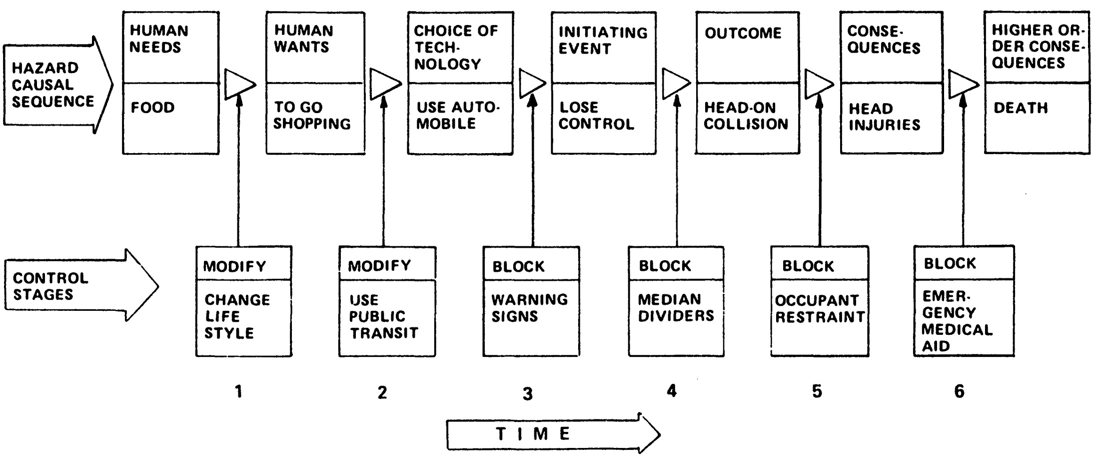

FIGURE 1. Illustration of the causal chain of hazard evolution. The top line indicates seven stages of hazard development, from the earliest (left) to the final stage (right). These stages are expressed generically in the top of each box and in terms of a simple motor-vehicle accident in the bottom. The stages are linked by causal pathways denoted by triangles. Six control stages are linked to pathways between hazard states by vertical arrows. Each is described generically as well as by specific control actions. Thus, control stage 2 would read: "You can modify technology choice by substituting public transit for automobile use and thus block the further evolution of the motor-vehicle accident sequence arising out of automobile use." The time dimension refers to the ordering of a specific hazard sequence; it does not necessarily indicate the time scale of managerial action. Thus, from a managerial point of view, the occurrence of certain hazard consequences may trigger control actions that affect events earlier in the hazard sequence (Source: Bick, Hohenemser and Kates).

TABLE 3
Locus of Standard

Typical Advantages of Placing Standard Close to: Technology	Consequences
Responsibility for problems more readily attributed	Greater relevance
Costs of compliance more easily assessed	Allowable risks more readily compared with benefits
Greater predictability	Easier access for measurement
Measurement easier	Creativity in compliance encouraged

keep the incipient threat posed there from developing further. It is also possible, in principle, to measure the magnitude of the threat at each stage. As a result, standards can be set in terms of each stage, measuring there whether technologies exceed the maximum of risk alloted to them by the standard setters (considering their benefits, alternatives, etc.). Although the details vary from case-to-case, each locus has particular problems, prejudices, and loopholes. These are discussed below and summarized in Table 3, which shows the advantages that accrue as the locus of the standard comes closer to the extremes of being set on the technology's hardware or on its averse consequences.

4.2.1. *Apply to the Technology*. Standards that prescribe the specific hardware to be used in the acceptably safe version of a technology are commonly called *design* standards. Their strongest advantage is their potential enforceability. Compliance may be checked simply by observing the hardware. When problems are detected, it is relatively clear whom to hold responsible. Whenever people must operate and maintain the technology or can frustrate its safety features (e.g., disconnecting seat belt interlocks or emission controls), then specification of their behavior is part of the design. Unlike hardware, which might be evaluated once and for all, behavior requires continual monitoring, greatly increasing the costs of applying a standard (or the opportunities for circumventing it).

Perhaps the weakest point of design standards is the hypotheticality of the risks that they permit. It may be hard for either risk analysts or lay observers to understand the link between the design and actual health effects. As a result, it may be hard to convince the producers of a technology that particular (expensive) design features are needed or to convince those potentially affected by it that their welfare has been protected. Such frustrations have motivated the Nuclear Regulatory Commission's (1982) plan to supplement its design standards with standards defining allowable consequences. They are also behind worries about what safety has been achieved by the building codes intended to protect the residents of earthquake country (Wiggins 1972).

Design standards effectively ban a technology when they cannot be met with current techniques and within current economic constraints. Such standards are often called "technology forcing" in the hopes that they will prompt the development of better techniques. In practice, however, it is hard to make impossible demands when the link between design and consequences is unclear. The desire for stringency is more likely to express itself in requirements for costly, but feasible safety devices and procedures. When these design requirements are tacked on to an operating technology, they may destroy the integrity of its original design without allowing for (let alone encouraging) the creation of new, more effective designs (Hammer 1980).

4.2.2. *Apply to the Initiating Event.* Standards set on the remaining loci specify different aspects of a technology's *performance*. In doing so, they should permit industry technicians to use their creativity in designing cost-effective ways to comply. A logical place to impose performance standards is at the source, constraining the likelihood or extent of initiating events and, thereby, keeping things from getting out of hand. Exemplary initiating-event standards include limits on blood alcohol levels (for automobile accidents), on smokestack emissions (for power plants), or on concentration levels (in pesticide sprays). Standards placed close to the source not only are imposed before any damage is done, but apply to events that are relatively easily measured and traced to their source. Proximity to the source means distance from the (risky) consequence itself, subjecting such standards to attenuated versions of the difficulties facing design standards (e.g., knowing what risk remains).

Two inherent problems of all performance standards emerge with initiating event standards. One is that measuring the performance of a technology is akin to measuring the behavior of an individual; it may fluctuate considerably, requiring either constant monitoring or some sampling scheme (topics discussed below). The second problem is that some of the technicians' creativity may be devoted to detecting and exploiting the standard's loopholes. Harriss (1978) suggests that emission limits on mercury stimulated the development of new production processes using (and emitting) mercury in smaller quantities. As a result, the yearly increment in mercury emissions has remained constant. The environmental problem may even have become worse, insofar as it is harder to control the more numerous current sources.

4.2.3. *Apply to the Outcome.* Once the initiating event has occurred, the next step toward realizing a risk is to damage the integrity of the environment, creating an imminent threat to those in it. Outcome standards limit the amount of environmental insult. Examples include ambient air quality standards, motor vehicle rollover requirements, and E. coli limits for drinking water (Calabrese 1978). Even more than initiating event standards, outcome standards are set in politically relevant terms and allow freedom in reaching compliance. For example, one option that opens up here is trading in pollution rights, whereby a technology is allowed to pollute an overburdened airshed by reducing the emissions of an existing industry (Mahoney and Yandle 1980).

The debates over pollution rights illustrate one general problem with standards set far from the source, assigning responsibility for violations. The medium that the monitors must measure is directly accessible, unlike hardware inside an industrial facility where prior warning is needed. However, the diffusion of that medium may make it hard to trace excesses to their source (e.g., Brombaugh and Lee 1981). When tracing is possible, the question remains of who must cut back (or pay fines or offer compensation) when a threshold is exceeded: The source that pushed the medium over the limit? The largest source? The most recent source? The source with the lowest ratio of societal benefits to societal costs? The source that is best able to bear the costs? The source with the cheapest control options? All sources in proportion to their contribution to pollution? All sources in proportion to their ability to pay? Without an attributional rule, an outcome standard is meaningless.

4.2.4. *Apply to the Consequences.* Setting standards in terms of allowable consequences brings to their logical conclusion the advantages and disadvantages that emerge as the locus moves away from the source technology. On the positive side, the standard is set in the most politically relevant terms; it allows maximum latitude in compliance strategies; the relationship between risks and benefits is clearest. On the negative side, it can be most difficult to respond to violations. In some cases (e.g., cancer), it is hard to trace the consequences to a source; in others (e.g., lead poisoning), there are many identifiable sources; in others (e.g., auto accidents), producers of the technology involved can attribute responsibility to intervening events (e.g., poor drivers, bad roads). Where violations cannot be linked to action, consequence standards are not advisable.

The controversy over hazardous waste dumps illustrates some of the issues associated with choosing the locus of standard. Although interest has been stimulated by actual consequences, the situation may be out of hand before health effects can be attributed confidently to a particular dump (Lennett 1980). As a result, a mixture of outcome standards (e.g., polluted water supplies) and consequence standards for less important effects (e.g., chromosomal damage) is used to judge existing dumps, whereas design standards have been devised for new dumps. The superfund for cleaning up abandoned dumps, which punishes the present for the excesses of the past, reflects these attributional problems (Trauberman 1981).

4.3. *How Should Compliance Be Measured?*

Once the locus has been selected, a scheme must be created for measuring the technology's design or performance. Although each standard raises special problems in this stage of operationalization, many of those can be grouped under the rubrics of a few basic statistical concepts.

4.3.1. *Selecting the Features to Be Measured.* Whatever locus is selected, many hazards have several features that might be measured. For example, there are many dimensions to most hardware and operating procedures, each might be specified in a design standard; industrial processes often emit several pollutants, each could be controlled as the initiator of a hazardous event chain; a chemical may have several sized particles, each could have its own outcome standard; separate consequence standards might be created for each of the adverse physiological and psychological effects potentially caused by a technology.

All of these features could be measured, or only a subset of representative features, or a single feature. Features might be chosen because they are easily observed, politically sensitive, well-understood, or associated with large health effects. Assuming a reasonable selection criterion, as the number of features increase, so should a standard's comprehensiveness, credibility, explicitness, and expense of application. Depending upon the rule used to aggregate the measurements from the different features, increasing their number may also tend to increase the difficulty and cost of complying with the standard.

4.3.2. *Aggregation of Measures over Features.* When several features are measured, the result may be a complex pattern of compliance and noncompliance on different features. An aggregation rule is needed to translate this pattern into a overall assessment of acceptability. *Conjunctive rules* set a separate cutoff criterion on each feature, all of which must be met for a technology to be acceptable. Strictly interpreted, such rules would fail a generally acceptable technology with a single excess but pass a technology that barely passed on each feature. They might even favor the development of technologies with diverse (but just tolerable) problems. Because each feature provides another opportunity to fail, when a feature is added to the standard, the cutoff criteria on existing features should be relaxed (if the standard's demands are to remain constant). Conjunctive rules are particularly appropriate with design standards in cases where one weak link can threaten the viability of a system and with consequence standards where each affected group's safety must be protected.

Compensatory rules allow more-than-adequate performance on one feature to compensate for the less-than-adequate performance on others. These rules are most appropriate with performance standards concerned with the total risk burden and not with upon whom it falls. Here, increasing the number of features may reduce the standard's stringency insofar as additional problems that a technology does not cause are allowed to compensate for those that it does cause.

4.3.3. *Handling Statistical Variation.* Any empirical measure will show some variability, due to unreliability in the measurement procedure or to fluctuations in the measured phenomenon. Descriptive statistics summarize distributions of measurements, in order to guide actions based upon them. One obvious set of candidate summaries is the statistical moments of a distribution. Figure 2 suggests how each of the first four moments (mean, variance, skew, and kurtosis) might provide the most appropriate way to characterize a risk. The choice between them depends on society's values and the actions it contemplates. Of course, other statistics are possible. One candidate with some uses is an extreme fractile of the distribution, setting a limit on the probability (or relative frequency) of extreme observations.

4.3.4. *Determining the Unit of Observation.* Each point in the distribution of observations may take in larger or smaller amounts of experience. An elementary

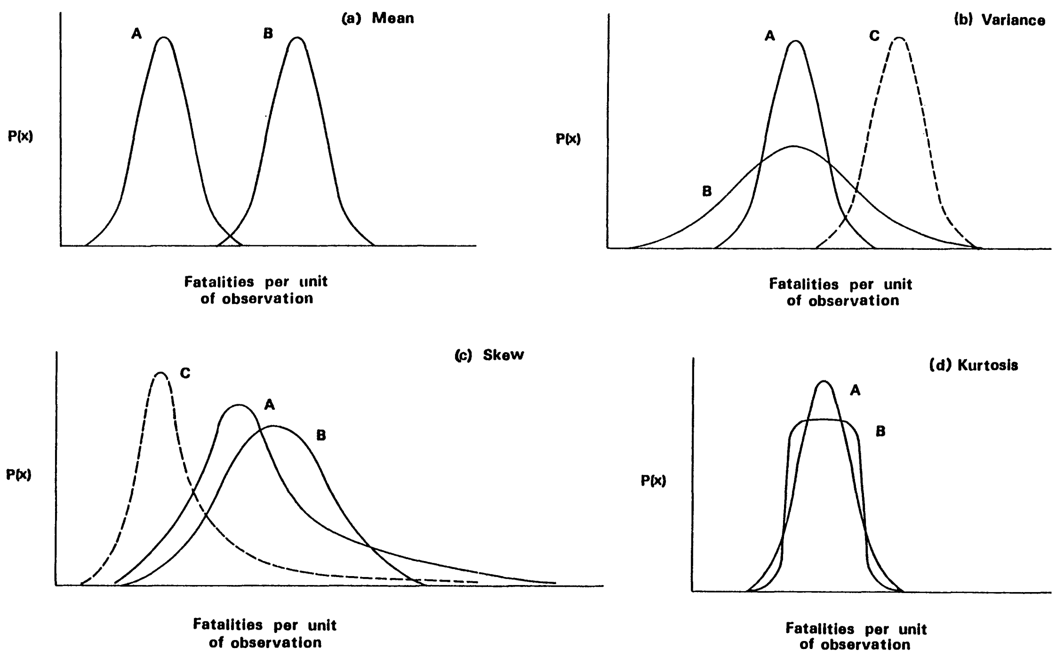

FIGURE 2. Characterizing risks by their statistical moments (of casualties). How one might be concerned primarily with different moments and how choice of moment may affect relative acceptability of hazards.

(a) *Reliance on mean.* Hazards *A* and *B* differ only in means. All other things being equal, one would prefer Risk *A*.

(b) *Reliance on variance.* Hazards *A* and *B* have the same mean, but differ in variance. One might prefer *B* because it offers some chance of low losses or because its high variance suggests the presence of a variable factor that might be understood and controlled. One might prefer *A* because its greater predictability allows a standard allocation of resources to compensation or adjustment. Indeed, such reasoning might lead one to prefer hazard *C*, with a higher mean but reduced variability, over *B*.

(c) *Reliance on skew. A* and *B* have similar means and variances, but differ in skew. That is, *A* has a long, flat tail, reflecting the somewhat remote possibility of a large number of fatalities. This catastrophic potential may induce a preference for *B* over *A*, or even a preference for *B* over *C*, which has a lower mean and variance but a longer tail.

(d) *Reliance on kurtosis. A* and *B* are similar in mean, variance, and skew, but differ in kurtosis (i.e., their flatness or homogeneity of density). One might prefer *A* because one standard coping behavior is adequate for a larger proportion of problems. Or one might prefer *B* because its restricted range means that one can more readily prepare for all feasible contingencies. Indeed, the fact that there is hardly a modal consequence with *B* may force one to be more imaginative and flexible.

result of sampling theory is that as the unit of observation increases, so does the stability of the measure. Conversely, the probability of exceeding any threshold decreases. For example, with fixed limit values, violations are more likely if a design standard is applied to the thickness of each tube, rather than the mean thickness for batches of 100; violations are more likely if an initiating event standard is applied to instantaneous smokestack emissions, rather than daily averages; and so on. Thus, standards that sound the same when expressed in terms of their limit values may require quite different degrees of control. Failure to specify the unit of observation can lead to a standard becoming more rigorous in mid-life simply because finer measurements become possible (Norcross 1980). Smaller units are appropriate when fluctuations are large and the environment is sensitive to local peaks. For example, abundant rain keeps the mean air pollution in the Willamette Valley low; however, making a month or a year the unit of observation would obscure the occasional atmospheric inversions that create local pollution maxima which pass the threshold for issuing air stagnation advisories or alerts. On the other hand, for pollutants with cumulative effects, long-term averages may be adequate.

4.3.5. *Defining a Sampling Scheme.* When it is impossible to look at all aspects of the regulated technology (and its byproducts) constantly, a sample must be taken. The definition of the sampling frame must consider the dimensions along which a technology may vary. For example, some hardware may be evaluated when it leaves the plant, with the assurance that it will not change over time; on the other hand, its manufacturing process may be unreliable enough to require intensive sampling of items as they are produced. Other kinds of hardware may be manufactured to precise specification, but be subject to changes as they are used (e.g., disconnecting safety devices); they should be measured over time. Air pollution may vary significantly over altitude, but if ground level concentrations pose the only threat to health, then only spatial variations need be sampled.

Once the frame has been defined, a representative sample must be drawn, through application of statistical theory informed by a substantive knowledge of the phenomenon. The random sampling needed to ensure a representative sample may have the added advantage of being hard to circumvent by those who might try to look especially good at observation time (as might happen with periodic inspections). A key parameter of the sampling scheme is the number of observations to be drawn. Within a fixed budget, the number of observations is inversely related to their intensity. Even when there is reason to get many observations, some resources might be allocated to intensive studies that might uncover biases in the ordinary observations (Brooks and Bailar 1978). For example, such studies might show that exotic cancers, which are unfamilar to many physicians, are underestimated in public health statistics that supposedly incorporate the entire population.[12]

4.3.6. *Using Surrogate Measures.* Selecting a locus was to some extent a choice between monitoring what is important (risk) and monitoring what can be more readily observed (hardware). Design standards are, then, a special case of regulating a surrogate for the actual topic of interest. With surrogates, it is harder to prove the standards' faithfulness to the intent of the standard setter. Such standards' defense is practicality and predictability.

The possibility (and problems) of using surrogates may also be found within loci. When a problem affects the whole population in a weak form, it may be easier to monitor the health of a particularly sensitive subgroup (e.g., air quality standards defined in terms of allowable health effects among people with respiratory problems).

[12] Analogously, anthropological studies of inner-city neighborhoods have identified classes of individuals that are underrepresented in census records of the same neighborhoods (Martin 1980).

When a health effect has measurable precursors, a standard defined on them might make control possible, as well as censure (e.g., using interferon levels as a predictor of resilience to medical problems). When accidents are too rare for their rate to be measured, near accidents may allow more stable estimates.[13] When problems have long latency periods, it may be useful to judge current affliction rates even though they reflect the exposure problem of many years ago.

To adopt surrogates, there must be some plausible link between them and the actual effect. If they are adopted, it is important to keep their surrogacy in mind when considering a standard. The expense of air quality standards may seem too high to "save a few asthmatics" but not if their health reflects that of the whole population.[14]

4.4. *How Should the Standard Be Enforced?*

To be more than just an academic exercise, standards must be linked to action. For action to follow in a timely and predictable manner, the standard formulators must anticipate ambiguities in interpreting the standard, provide guidelines for the usage of uncertain evidence, and specify the consequences of noncompliance. Finally, they must review their efforts and estimate how well the resultant standard represents the safety philosophy of the standard setters. If operationalization has introduced some bias, then it may be redressed by incorporating a correction factor in the enforcement scheme, making it especially stringent or lax.

4.4.1. *Anticipating Problems of Interpretation.* The essence of the standard formulators' job is to eliminate any need for interpretation when the standard is applied. Although attention to the details described above should reduce ambiguities, it should be supplemented by a special effort to anticipate surprises. One way is to assume that every ruling of noncompliance will initiate a search for a loophole that will exempt the technology involved. Advance consideration of the validity of such claims can help ensure that the standard is relaxed responsibly.[15] A second strategy is to anticipate novel applications of the standard. That can be done both by trying to stretch the standard to new problems and by identifying possibly related technologies that have less satisfactory (or no) standards. A third strategy is to consider the ramifications of the precedents set by possible applications. For example, when a particular piece of equipment is mandated for a new technology, should it be extrapolated literally to other existing technologies (even if retrofitting is more expensive and less effective than with new designs) or should it be interpreted as embodying a principle (e.g., take any step that will reduce risk by at least amount x for less than $\$y$)? The economic consequences of accidents such as Three Mile Island may vary greatly depending upon how widely the lessons learned are generalized (Evans and Hope 1982).

Finally, because it is impossible to anticipate all contingencies, the standard should include a procedure for resolving interpretative questions, along with a criterion for when recourse to that process is needed. It might specify which parties should participate in resolution and when to go back to the standard setters for guidance.

4.4.2. *Anticipating Problems of Incomplete Evidence.* The measures chosen to determine compliance with a standard inevitably involve random variables. The variation in these variables creates the risk of Type I and Type II errors: erroneously

[13] With a highly politicized technology, like nuclear power, considerable economic stakes may actually ride on the rate of casualtyless near-misses, making a near-miss standard important for nonsafety reasons (Slovic et al. 1980, Starr 1981).

[14] Some of the more facetious criticisms of the Delaney Amendment's treatment of saccharin made it seem like a standard protecting laboratory rats from cancer. In principle, a standard might be written to protect the health of street rats if they were particularly sensitive to something that eventually affected humans.

[15] Just as property owners often hire ex-criminals to diagnose their vulnerability to theft, standard formulators might hire potential opposing attorneys to expose standards' points of vulnerability.

determining that a technology is in compliance or erroneously determining that it is not. Statistical decision theory shows how to formulate decision rules to achieve different balances between these error rates. Its application requries a value judgment, regarding the relative social cost of the two kinds of errors, being unfair to technologies or to risk bearers (Page 1978). The decision rule establishes whether the burden of proof is on the technology to prove its innocence or on the standard applier to demonstrate its guilt.

The application of statistical procedures requires assessment of the probability of the measured feature having different values. This presents no problem when a large sample of observations has been drawn with a definitive measurement procedure. When the sample is small or the measurement procedure faulty, some extrapolation procedure is needed. When data have been collected from several sources, perhaps using different research paradigms, an aggregation procedure is needed, such as counting "votes" for or against compliance, pooling the probabilities of compliance (assuming the independence of tests), or computing the mean of best guesses, weighted by the quality of the studies (Hedges and Olkin 1980). Different extrapolation and aggregation procedures can produce quite different summaries of available evidence. These can affect determinations of compliance and, hence, should be specified in advance.

Specification is also needed regarding two other issues that follow from incomplete evidence. One is how far to exploit the opportunities available to reduce uncertainty at the cost of time and investigatory resources. Value-of-information analysis (Raiffa 1968, Weinstein 1979) provides the conceptual framework for addressing these issues systematically. Its application requires (and reveals the role of) assigning explicit costs to information and to misclassification (where inaction implies some classification, depending upon how the burden of proof is defined). With uncertainty, it is inevitable that subsequent evidence will show some classifications to have been in error. The standard must specify whether classifications are irreversible (and whether damages are due). If disapproval prevents a technology from being used, then it may also prevent the collection of information showing that that classification was wrong.

4.4.3. *Determining the Rigor of Enforcement*. The response to violations has two dimensions, speed and severity. The time allotted to an unacceptable technology to get into compliance (or suffer the consequences) should depend upon the rate at which unacceptable risks (or actual damages) accumulate and the chances that the time will be used successfully to improve the technology (rather than to gain market share or unfair profits by selling a cheaper, riskier product). A wild card with waiting is that political moods may change, weakening or strengthening all regulation.

The severity of the response to a noncomplying technology (like its speed) is a continuum, along which a number of marker positions may be identified. The most salient of these responses is to ban the technology until it comes into compliance. At times, however, this is impractical; although the technology's benefits are insufficient to justify its risk, they are too great to be forfeited. If these risks are intolerable in part because they are borne by people who do not benefit from the technology, then it may be possible to redress the imbalance by compensating them. That transfer could be viewed as a penalty for violations or as a redesign of the technology to conform to a standard with an equity component. More extreme penalties are also possible, such as accompanying the ban with punitive damages. These could be justified by their signal value, warning potential violators to take the standard seriously; or they could be designed to compensate for known defects in the standard's monitoring system. If only a certain percentage of violations will be detected and proven, then the penalty might be set so that violators' expected penalty would be the same as it would be with perfect detection and nonpunitive enforcement.

Whatever enforcement policy is written into the standard must be commensurate with the standard appliers' ability to make their determinations stick. If they lack power, then it may be preferable to design an enforcement plan that focuses on encouraging voluntary compliance, eliciting public displeasure, and attacking the clearest excesses.

4.4.4. *Including a Correction Factor*. The idea of using punitive damages to compensate for imperfect monitoring is one example of the kind of correction factor that standard formulators might use to bring an operational standard into compliance with its underlying safety philosophy. A second is to increase all estimates describing the uncertainty of knowledge about the technology by a fixed factor in order to counter people's tendency to be overconfident in their own knowledge (Fischhoff 1982, Lichtenstein et al. 1982). A third is to adjust estimates of risk or benefit upward or downward by a fixed factor to nullify biases introduced to the data at their source. A fourth way would be to disregard mild violations when a standard has been formulated too restrictively. An analogous activity from another domain was the attempt to adjust the results of the 1980 U.S. Census to correct a suspected undercount of certain socio-economic and ethnic groups.

Although a logical possibility, such corrections may be political impossibilities. Especially when they change the fate of a technology, these adjustments may be a lightning rod for criticism. They can be defended only by a technical analysis of the standard-safety philosophy match for which political arenas may have little patience. Moreover, the correction itself undermines the credibility of the standard.

The defensibility of a correction can be improved by specifying it in advance, so that it emerges from the standard formulators' deliberations rather than from the standard appliers' disquiet. It may even be requested when the overall bias of a standard is common knowledge to those who provide inputs to it. Rather than participate passively in such an exercise, they may spontaneously adjust their own inputs in order to introduce correction factors from the bottom. For example, scientists may give high-side or low-side summaries of studies if they believe that they will be taken too lightly or too heavily, or inspectors may take their job more or less seriously depending upon the anticipated response to their reports. The cumulative impact of such tinkering may be so hard to assess that an explicit overall adjustment factor may be accepted in return for a promise of candor in local estimates.

4.5. *Summary*

The acceptability of any injustices associated with a standard depends upon the alternatives. When the costs of waiting and uncertainty run high, even those treated unfairly may view occasional injustices as a legitimate price to pay for a timely, predictable system. Translating a safety philosophy into an operational standard raises many design issues, whose resolution will determine the practicality and stringency of the resultant standard. If not faced explicitly, these issues will be decided by default as the technicians do their work. Resolving them in the standard-setting process can help ensure that they draw the open controversy they deserve. Moderate ignorance about the specific cases to which a standard will be applied may encourage the parties involved to adopt a balanced position and try to design it fairly.[16]

[16]This research was supported by the Nuclear Regulatory Commission under subcontract 7656 from Union Carbide Corporation, Nuclear Division, Inc., to Perceptronics, Inc.

I would like to thank S. A. Carnes, T. E. Cole, E. D. Copenhaver, H. Inhaber, H. Kunreuther, G. Morgan, D. C. Parzyck, P. Rathbun, P. Slovic, K. A. Solomon, S. Watson, and two anonymous reviewers for their many helpful comments on previous drafts. Correspondence should be addressed to the author at Decision Research, A Branch of Perceptronics, 1201 Oak Street, Eugene, Oregon 97401.

References

AMERICAN ENTERPRISE INSTITUTE, "Proposals for Regulatory Reform," Washington, D.C., Author, 1980.

AMERICAL PUBLIC HEALTH ASSOCIATION, "Statement on S.2153 Occupational Safety and Health Improvement Act of 1980," Washington, D.C., Author, 1980.

ANASTASIO, E. L., "Restraints on Use of the Term 'Natural' in Advertising and Labeling of Food, Drugs, and Cosmetics," *Food, Drug, and Cosmetic Law J.*, 36 (1981), 424–430.

BICK, T., C. HOHENEMSER AND R. KATES, "Target: Highway Risks," *Environment*, 21, 2 (1979), 7–15, 29–38.

BROMBAUGH, K. J., AND K. W. LEE, "Finger-Printing Organic Emissions," *Environmental Sci. and Technology*, 18 (1981), 1142–1149.

BROOKS, A., AND B. A. BAILAR, "An Error Profile: Employment as Measured by the Current Population Survey," (Statistical Policy Working Paper 3), U.S. Department of Commerce, Washington, D.C., 1978.

CALABRESE, E. J., *Methodological Approaches to Deriving Environmental and Occupational Health Standards*, Wiley, New York, 1978.

CRANDALL, R. W. AND L. B. LAVE, *The Scientific Basis of Health and Safety Regulation*, Brookings Institution, Washington, D.C., 1981.

DAWES, R. M., "A Case Study of Graduate Admissions: Applications of Three Principles of Human Decision Making, *Amer. Psychologist*, 26 (1971), 180–181.

DERBY, S. L. AND R. L. KEENEY, "Risk Analysis: Understanding 'How Safe is Safe Enough?'" *Risk Anal.* 1, 3 (1981), 217–224.

DERR, P., R. GOBLE, R. E. KASPERSON AND R. W. KATES, "Worker/Public Protection: The Double Standards," *Environment*, 23, 7 (1981), 6–15, 31–36.

DRIVER, C., "A Theory of Regulatory Enforcement," *Public Policy* 28 (1980), 258–291.

EVANS, N. AND C. HOPE, "Costs of Nuclear Accidents: Implications for Reactor Choice," *Energy Policy*, 10, 4 (1982).

FISCHHOFF, B., "Debiasing," in D. Kahneman, P. Slovic, and A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, New York, 1982.

———, "Safety Goals for Nuclear Power," *J. Policy Anal. and Management*, 3 (1983), 559–575.

———, S. LICHTENSTEIN, P. SLOVIC, S. L. DERBY AND R. L. KEENEY, *Acceptable Risk*, Cambridge University Press, New York, 1981.

———, P. SLOVIC AND S. LICHTENSTEIN, "The 'Public' vs. the 'Experts': Perceived vs. Actual Disagreements about the Risks of Nuclear Power," In V. Covello, G. Flamm, J. Rodericks, and R. Tardiff (Eds.), *Analysis of Actual vs. Perceived Risks*, Plenum Press, New York, 1983.

FLORMAN, S. C., *Blaming Technology*, St. Martin's Press, New York, 1981.

HARRISS, R., "Mercury: Measuring and Managing the Hazard," *Environment*, 20, 8 (1978).

HAMMER, W., *Product Safety and Management Engineering*, Prentice-Hall, Englewood Cliffs, N.J., 1980.

HEDGES, L. V. AND I. OLKIN, "Vote-Counting Methods in Research Synthesis," *Psychological Bull.* 88 (1980), 359–369.

HOHENEMSER, C., R. W. KATES AND P. SLOVIC, "The Nature of Technological Hazard," *Science*, 220 (1983), 378–384.

JOHNSON, E. L., "Risk Assessment in an Administrative Agency," *Sixth Symposium on Statistics and the Environment*, National Academy of Sciences, Washington, D.C., 1980.

KEENEY, R. L., *Issues in Evaluating Alternative Standards*, Woodward-Clyde Consultants, San Francisco, 1981.

LAVE, L. B., *The Strategy of Social Regulation*, Brookings Institution, Washington, D.C., 1981.

LENNETT, D. J., "Handling Hazardous Wastes," *Environment*, 22, 8 (1980), 6–15.

LICHTENSTEIN, S., B. FISCHHOFF AND L. D. PHILLIPS, "Calibration of Probabilities: State of the Art to 1980, in D. Kahneman, P. Slovic and A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, New York, 1982.

MCNAMARA, S. H., "When Is a Cosmetic Also a Drug—What You Need to Know and Why," *Food, Drug, and Cosmetic Law J.* 35 (1980), 467–478.

———, "When Does a Food Become a Drug?" *Food, Drug, and Cosmetic Law J.* 37 (1982), 222–231.

MAHONEY, M. T. AND B. YANDLE, "Bubbles and Efficiency," *Regulation*, (May/June 1980), 49–52.

MARTIN, E., *Surveys as Social Indicators: Problems in Monitoring Trends*, Institute for Research in Social Science, University of North Carolina, Chapel Hill, N.C., 1980.

MAY, W. W., "Dollars for Lives: Ethical Considerations in the Use of Cost/Benefit Analysis by For-Profit Firms," *Risk Anal.*, 2 (1982), 35–46.

MAZUR, A., *The Dynamics of Technical Controversy*, Communication Press, Washington, D.C., 1981.

NELKIN, D. AND M. POLLAK, *The Atom Besieged*, MIT Press, Cambridge, Mass., 1980.

NORCROSS, M. A., "Sensitivity of Method—A Wave of the Future," *Food, Drug, and Cosmetic Law J.*, 35 (1980), 342–348.

Otway, H. J. and D. von Winterfeldt, "Beyond Acceptable Risk: On the Social Acceptability of Technologies," *Policy Sci.*, 14 (1982), 247–256.

Page, T., "A Generic View of Toxic Chemicals and Similar Risks," *Ecology Law Rev.* 7 (1978), 207–243.

Portney, P., "Efficient Use of Standards and Enforcement: The Case of Pollution Controls," *Policy Anal.*, 5 (1979), 512–524.

Raiffa, H., *Decision Analysis*, Addison-Wesley, Reading, Mass., 1968.

Salem, S. L., K. A. Solomon and M. S. Yesley, "Issues and Problems of Inferring a Level of Acceptable Risk," R-2561-DOE, RAND Corp., Santa Monica, Cal., 1980.

Selikoff, I. J. and D. H. K. Lee, *Asbestos and Disease*, Academic Press, New York, 1978.

Shaw, J. S. and C. Canape, "Is the EPA Saddled with More than It Can Handle?" *Business Week*, (May 26, 1980), 148.

Shindler, J. R., "A Common Sense Approach to Common or Usual Names if There Be Any," *Food, Drug, and Cosmetic Law J.*, 36 (1981), 284–292.

Slovic, P., B. Fischhoff and S. Lichtenstein, "Perceived Risk and Quantitative Safety Goals for Nuclear Power, *Trans. Amer. Nuclear Soc.*, 35 (1980), 400–401.

Starr, C., "Risk Criteria for Nuclear Power Plants," *Risk Anal.* 1 (1981), 113–120.

——— and C. Whipple "Risks of Risk Decisions," *Science*, 208 (1980), 1114–1119.

Stokey, E. and R. Zeckhauser, *A Primer for Policy Analysis*, Norton, New York, 1978.

Trauberman, J., "Superfund: Legal Update," *Environment*, 23, 2 (1981), 25–29.

U.S. National Academy of Sciences, *Strategies to Determine Needs and Priorities for Toxicity Testing*, Washington, D.C., Author, Vol. 1, 1981; Vol. 2, 1982.

U.S. Nuclear Regulatory Commission, "Safety Goals for Nuclear Power Plants: A Discussion Paper," NUREG-0880, Washington, D.C., Author, 1982.

Wall Street Journal, "Down with Big Business," Editorial, 1979.

Weinstein, M. C., "Decision Making for Toxic Substances Control: Cost Effective Information for the Control of Environmental Carcinogens," *Public Policy*, 27 (1979), 333–383.

Whipple, C., "Energy Production Risks: What Perspective Should We Take?" *Risk Anal.* 1 (1981), 29–36.

Wiggins, J. H., "Earthquake Safety in the City of Long Beach Based on the Concept of Balanced Risk," in *Perspectives on Benefit-Risk-Decision Making*, National Academy of Engineering, Washington, D.C., 1972.

Wilde, G. J. S., "The Theory of Risk Homeostasis: Implications for Safety and Health," *Risk Anal.*, 2 (1982), 209–225.