Failure Analysis Associates

ENGINEERING AND METALLURGICAL CONSULTANTS
750 WELCH ROAD, SUITE 116, PALO ALTO, CALIFORNIA 94304 (415) 321-6350

HOW SAFE IS SAFE ENOUGH?

Prepared by

Alan S. Tetelman
Professor of Engineering, UCLA
Head, Technical Staff, Failure Analysis Associates

and

Philip M. Besuner
Senior Analytical Engineer
Failure Analysis Associates

June 1977

(To be published in: Social Consequences of Engineering, Ed. L.S. Hager, published by, Dun-Donnelley Publishing Corporation.)

TABLE OF CONTENTS

ABSTRACT

This paper deals with the meaning of risk and safety and the various methods that society utilizes to assure that there is a reasonable balance between them. Systems are designed to run properly, without failing, and this requirement is usually met. Failures that lead to energy releases, which impact on persons or property and damage them, are called accidents. This paper describes the nature of accidents in technological systems, and their characterization in terms of failure frequency and severity. Following a description of the elements of an accident, the paper describes various methods by which the risk of an accident or hazard potential may be quantified. This provides a basis for comparing the safety of one system with that of another. Further sections describe parameters by which society has determined levels of acceptable risk, and the procedures that society utilizes to maintain a balance between risk, cost, and benefit.

"You're a Yankee boy, right? And you're a college boy, right? And you never worked on one of these rigs before, right? Okay, then if you listen real careful and pay close attention, you may not get killed your first week on this job. And if you make it through the first week, you'll probably make it fine through the Summer. But, if you work much longer than that, you're gonna lose a finger or an eye like me and the rest of the boys." The Failure Rate Curve, as first described to the senior author by F. C. Marshall, Driller, Tyler, Texas (1956).

1. INTRODUCTION

This paper deals with the meaning of risk and safety and the various methods that society utilizes to assure that there is a reasonable balance between them. Society, in the person of governments, corporations and individuals, builds and operates engineering systems to meet human needs. An engineering system is a process or device in which matter, energy, or information in one observable and definable condition is transformed into another definable condition, or is maintained in its original state when subjected to external forces. The total cost (in dollars) that society, in the aggregate or as an individual, spends on a given system is a measure of the worth or benefit of that system. The total cost is primarily that required to design, manufacture, operate, and maintain the system.

All systems are subject to possible malfunction and/or failure. Additional costs must therefore be allotted to system replacement (e.g., the value of an aircraft that has crashed) and to other systems and people who are damaged as a consequence of the failure (e.g., widows and children of wage earners who die in a plane crash). The hazard or risk that is associated with a given system is the cost (dollar effect) of failure. Since the effect of a failure can range from a minor incident to a major catastrophe, an entire spectrum of consequences can be associated with a given failure event. However, to simplify discussion, it is often convenient and informative to speak of the average consequence or average severity which results from a failure event.

Systems are designed to run properly, without failing, and this requirement is usually met. Those few failures that do occur in the life of an operational system are isolated instances, chance happenings, that

are called "accidents". An accident is defined as "one or a combination of unanticipated failure events (incidents) that release stored energy, which directly or indirectly causes personal injury, property damage, or malfunction of a system". An accident is an event that is described in terms of both a "cause" and an "effect".

Despite the tremendous increase in the safety awareness of our society in the last decade, and despite the various methods and technologies that are available to prevent accidents or reduce hazards, failures and accidents can never be completely eliminated from our society. In fact, a certain number of accidents are to be expected according to the Second Law of Thermodynamics. This law states that equilibrium in a total system is approached as the entropy of the system increases. Entropy has several meanings. The one that is germane to this paper is that entropy is a measure of <u>disorder</u> in a system. Chaos and disorder are the natural state of the physical world, and man's efforts to separate U_{238} and U_{235}, to organize a disordered corporation, and to lower the failure rate of a product are simply external energy inputs that combat entropy (randomness, disorder). The Second Law of Thermdynamics shows that entropy can never be completely overcome. Some disorder, some failures are inevitable. The real question society must address is not "How do we guarantee zero-defects?" or "How do we achieve complete technical security and freedom from failure?" but rather "How do we minimize the rate of failure?", "How do we reduce the consequences of a failure?" and "What level of total risk is acceptable to our society?"

When a consumer purchases a household product, or a military service purchases a new weapons system, a conscious decision has been made that the cost (to the purchaser) is less than or equal to the actual and perceived benefits that are associated with the product. Until recently, the

cost of failure and accidents was crudely accounted for in the total cost on the basis of past performance of the product. The number of expected failures would be estimated and multiplied by the average cost of a failure. This amount would be set aside (e.g., in the form of insurance premiums) to be used to cover the cost of repair and replacement and recompensation after an accident took place. Since failure and accidents were rate events and the cost of failure was minimal, little societal attention was given to questions of safety and acceptable risk.

In the last decade, society's interest in and its approach to safety has changed greatly. A full discussion of the reasons for the change are beyond the scope of the paper, but several major factors should be noted. First, with the development of larger systems, the magnitude of a catastrophe such as an airplane crash has greatly increased. Improved communications, primarily through the evening news on television, has brought these catastrophes close to a large body of the public, in a dramatic fashion that affects individual senses. The development of nuclear weapons and nuclear power has led to public concern for long term, genetic effects of accidents (e.g., due to leaks of radioactive waste), that cannot be quantified, nor accounted for in terms of insurance premiums paid by the manufacturer of the product. Finally, there is now a public demand for increased corporate concern for the welfare of the individual and his natural environment. This demand was originally associated with the efforts of Ralph Nader, but today a host of environmental action and public interest groups closely scrutinize governmental decisions that directly and indirectly affect public safety.

In developing answers to the question as to what constitutes an acceptable or reasonable risk, several factors must be kept in mind. First, that acceptable risk to one group may be unacceptable to another. Second, that there is often a vast difference between the actual, quantified level of risk and the level of risk that is "perceived" in the minds of individuals or groups of individuals. Third, as stated earlier, that zero-risk is an unreasonable, utopian ideal. Finally, that lower risk can be achieved by reducing benefits (e.g., there is no risk of air crashes if no aircraft are flying) or by increasing the product costs by improving basic design or safety systems.

The purpose of this paper is not to try and provide definitive solutions to risk problems, but rather to explain basic concepts of accidents and risk. With an understanding of some basic concepts, the reader should then be able to explore the more complex questions in those areas of direct interest to him. Since risk is directly related to the frequency and severity of accidents, it is appropriate to begin the body of this paper with a description of the basic elements of an accident in a technological system. A following section describes the means by which the risk of accidents can be quantified, so that there exists a method for comparing the safety of one system with that of another. The next section describes some parameters by which society has determined levels of acceptable risk, while the final section describes the means by which society maintains a balance between risk, cost and benefit.

2. ACCIDENTS IN TECHNOLOGICAL SYSTEMS

The conversion of matter and/or energy from one form to another in an engineering system is best described by a simple block diagram containing inputs and outputs. The inputs are combined at the transformation point (also called reaction point or node) into one or more outputs. The input/output process is observed and monitored by a controller (governor, regulator) which monitors the rate, direction and magnitude of the inputs and outputs. Let us consider, for example, a steam-turbine powered pump that operates in a refinery. Hydrocarbon liquid at pressure P_1 enters the pump which is driven by a steam turbine shaft. The turbine shaft is driven by steam at a rate that is controlled by a governor. The turbine shaft turns the compressor wheel in the pump such that the liquid exits at a higher pressure, P_2. The governor monitors the speed of the turbine shaft by operating a valve, such that if the turbine slows too much, steam enters the turbine and increases its speed and conversely, less steam is allowed to enter the turbine if the turbine starts to overspeed. A second governor is often added to the system to prevent gross overspeeding. In this sub-system of a refinery, the compressor, turbine and governor are called components; the shafts, bearings, casings and valves are principle parts of the pump and turbine.

The sub-system described above is typical of the multitude of sub-systems that perform a function in a manufacturing process, sporting event, or the home. These systems and components possess an accident potential or hazard that is related to 1) the mechanical, chemical, thermal or electrical energy that they store in normal operation, and 2) the energy that is released if they malfunction. For example, if the pipes leak, if the pump fractures, if the bearings are destroyed, then hydrocarbon

liquid will be released into the atmosphere and will vaporize. The gas is combustible, and if it contacts oxygen in the presence of a spark, an explosion and fire may occur. If the fire, in turn, ignites a larger body of stored hydrocarbon, then even more energy is released, and more property damage will occur, along with possible personal injuries and/or fatalities. The risk associated with this system is directly related to the probability of the different malfunctions, and to the probable spectrum of consequences associated with each malfunction. The term "risk" therefore implies a cause and effect relationship between input events that could lead to energy releases, and output events (physical and psychological damage) that are the result of released energy.

The cause/effect aspect of an accident implies that any quantitative description of risk contain a measure of probability that a given failure event will occur (e.g., a fracture of the pump casing). This probability is determined by the frequency, f, of such incidents, which is the number of such failures that occur in a given time period (typically in one year or in one hour), divided by the number of units that are operating in the given time period. If 10,000 compressors of a particular class operate in one year, and one compressor typically fails per year, then the probability of compressor failure is 10^{-4} per year. The second part of the cause/effect relation, or the consequence, is the severity, S, associated with a given failure. The severity is related to the energy released by the failure event, ΔE, and to the probability, p, that a sufficient portion of the energy impacts a person or object such that he/it is injured or malfunctions. If the second malfunction in turn leads to another energy release, which in turn leads to a third malfunction, then the accident chain has been created. In the case (Fig. 1) of the compressor/turbine

system, a typical accident chain might be the following:* vaporization of the liquid reduces drag on the turbine which causes the turbine shaft to rotate at a higher speed. In principle, the governor should sense this phenomenon and maintain steam flow rates such that the turbine does not overspeed. Suppose that the monitoring governor and the overspeed governor are both inoperative. Failure that the governors in turn allows turbine overspeed, which in turn leads to bearing failure in the compressor, which in turn allows air to enter the compressor. If the failure is detected at this point, the damage is minimal. However, if the shaft rotates eccentrically and at high speed for sufficient time that high frictional heat is generated in the compressor, the hydrocarbon/air mixture may explode causing compressor rupture. The hydrocarbon then can spill out of the compressor at a high rate. If the system is shut down and fire prevention measures are used, the hazard can still be contained. However, there is a probability that the hydrocarbon will ignite, causing a large fire and extreme property damage and personal injury to workers. In this case, vaporization and governor malfunctions, due to poor maintenance or wear out, have led to a spectrum of consequences. While the frequency of one governor failure, f_1, may be 10^{-3} per year, the probability of a large fire due to a governor failure is very low, perhaps as low as 10^{-15}, <u>if five separate and independent 10^{-3} events must occur in order to produce the one, major 10^{-15} catastrophe</u>. This factor is the basis for one of the fundamental concepts of risk analysis -- <u>that the more severe the accident, the less frequently it will occur</u> (c.f., Figs. 2 and 3). This concept will be explored more fully in the following section where some methods of quantifying risk are described.

*This is a description of an actual accident sequence which the author investigated.

3. THE QUANTIFICATION OF RISK

The fundamental notion that the risk associated with an event is a function of both the event's frequency and its severity is intuitively reasonable; that notion seems in keeping with the judgements we make as individuals, as well as the broader judgements society makes. This notion of risk also reflects modern concepts of reliability engineering. To the reliability engineer, there is no such thing as a perfectly safe product; for every product there is a non-zero probability of failure, (small, perhaps, but still non-zero) and a non-zero probability that such a failure will result in injury to somebody (not severe, perhaps, but still an injury). In order to make a product safer, the engineer can decrease the frequency (i.e., the probability) of failure and of the resulting injuries, or he can decrease the severity of such injuries, or he can do both.

The simplest (in terms of data gathering) method of evaluating personal risk is to determine the number of fatalities associated with a given process or activity. While helpful in certain areas such as aviation, where most major system malfunctions lead to non-survivable accidents, the method is too simple to account for most accidents where a spectrum of consequence (e.g., minor to major injuries) can occur. In these cases, it is more realistic to consider an average consequence, and to evaluate risk in these terms.

The most direct method for evaluating the risk or hazard from a process or activity is to multiply the frequency of failures, f, by the average severity, $\bar{S}$, that results from the use or exposure of the process. This product, I, is called the hazard index or risk index.

$$I = f \times \bar{S} \qquad (1)$$

The parameter S can be evaluated in terms of number and extent of injuries, lost man days, dollar cost and outage time for equipment.

A detailed method for describing personal injury accidents during the use of consumer products has been developed by the U. S. Consumer Product Safety Commission (CPSC)[1] and serves as a good example of the quantification of risk and hazard. The CPSC operates the National Electronic Injury Surveillance System (NEISS) to estimate the frequency and severity of injuries incurred by individuals who are treated at 2% of the emergency rooms of U.S. hospitals. Injuries resulting from 800 consumer products are indexed, and the number of injuries from each product is multiplied by 50 to estimate the total number of injuries per year, f, in the U. S. population from that product. The CPSC also assigns an average severity value S_i for each of eight injury categories, as shown in Table I. The severity value is related to the number of lost man days to society as a result of the accident. The average severity $\bar{S}$ of all injuries and fatalities due to a given product is obtained from the relation

$$\bar{S} = \frac{\sum_{i=0}^{i=8} f_i S_i}{\sum_{i=0}^{i=8} f_i} \qquad (2)$$

where f_i is the number of accidents that cause an injury of severity S_i, and f is the total number of accidents involving that product, including those "incidents" of zero severity (no injury). The CPSC quantifies "risk" in terms of a hazard index, I, by simply multiplying the accident frequency, f,

by the mean severity, $\bar{S}$, for that product. Thus, from Eq. (1)

$$\text{Risk Index } I = f \times \bar{S}$$

and hazardous products are ranked in order of decreasing value of I. Table II lists the 33 most hazardous consumer products in the U.S. in 1974, as evaluated by the CPSC.[2]

It should be noted that the CPSC system does not account for the time of exposure that a given product is used. The frequency of injuries is computed on an annual basis for the U.S., and thus represents the risk to society as a whole. The risk to an individual is different, since an individual will use a bed 2500 hours per year and a welding torch only 50 hours per year. Individual risk is therefore better quantified in terms of number of accidents per exposure hour, but this parameter is difficult to quantify since the usage spectrum of certain products (e.g., welding torches) varies from one individual to another.

The CPSC method can also be extended to quantify the risk due to individual sub-systems in, say, an automobile.[3,4] In certain states, police officers are required to complete a form for each accident that involves measurable property damage or personal injury. Cause categories are included, along with consequences, and the data are stored in a form that enables parametric analysis to be made. When the various injury classifications are normalized to the CPSC classifications (shown in Table I) the sources of automobile risk can be better defined. Figure 2 indicates the inverse frequency/severity relationship for auto accidents that was mentioned earlier. It should be noted that the relation is the same, irrespective of whether the accident was caused by an apparent component "defect" or resulted from driver or pedestrian error. Defective

vehicles account for only 2-5% of the total risk of driving. Figure 3 shows the risk breakdown for the major sub-systems of an automobile, along with the tabulated hazard indices. Again, the inverse frequency/ severity relation applies for the different sub-systems, although the slopes of the curves vary widely. The low slope for tires indicates that tire failures lead to higher severity accidents than, for example, failure of turn-signals. This fact, coupled with the high tire failure rate, combines to make tires almost as hazardous as all other automobile sub-systems combined (Table III).

4. THE MEANING OF ACCEPTABLE RISK (SOCIETAL VIEW)

Having described the physical factors that contribute to the hazard associated with processes and activities, and having discussed some methods of quantifying risk, we are now able to address the question as to what constitutes an "acceptable" risk. This question is basically at the core of the safety question. All products can probably be made "safer"; however, if the risk associated with their usage is "acceptable", there is an implication that the product is already safe enough.

Although the level of risk can be quantified in terms of frequency and severity as discussed earlier, the level of "acceptable" risk is a relative or comparative parameter, and varies from one individual to another and from one society to another. Certain individuals are more likely to undergo hazardous sporting activities, such as hang-gliding, than other individuals. The acceptable level of pollution, noise, and the speed limit on highways varies from one state to another. Consequently, acceptable risk is not an absolute parameter or engineering constant, but rather is a variable with respect to both geography, history, and personal philosophy. The United States has stronger occupational health and safety laws than non-western nations, and even in the United States, acceptable levels of risk to workers in critical occupations (e.g., miners) has decreased in the last century.

Since "acceptable risk" is a relative rather than invariant parameter, any discussion of acceptability is, by definition, comparative. Dr. Chauncey Starr, former Dean of Engineering at UCLA and currently President of the Electric Power Research Institute, has done the most original and extensive studies of what society has chosen as a level of acceptable risk. Staff[5,6] has compared the probability

of death from a variety of activities with variables such as natural risk, benefit, etc. His findings are summarized here. Briefly, the fatality rate in the United States from all causes (primarily old age and disease) is about 1% per year or about 10^{-6} per exposure hour. Hence, the natural probability of death is 10^{-6} per exposure hour. The probability of death from truly "pure chance" events such as lightening strikes is 10^{-10} per exposure hour. Figure 4 shows the risk of fatality due to motor vehicle accidents that occurred from 1900 to 1960, and Fig. 5 shows a similar curve for flying. The graphs indicate that when only a small fraction of the population are involved in the activity, the risk is high. As greater numbers of people become involved, the risk level decreases, and approaches an equilibrium value of 10^{-6}, the risk level due to natural causes. Many activities, such as driving or exploration, begin as sporting events or distractions for the wealthy. The main attraction of the activity is the participants' desire to achieve a direct personal gain or positive stimulus such as the enjoyment of the sensation of extending their natural habitat (e.g., flying), the releasing of energy (e.g., exploding gunpowder) or the extending of life (through medical advances). The initial attempts or products may be crude, and the risk is high, well above the 10^{-6} norm (Figs. 4 and 5). One of the attractions of the activity at this point may even be the "macho trip" that is associated with facing and conquering high risk situations so that high risk is perceived as part of the benefit. As time passes, increasing numbers of people find benefits from the activity or product, so its utilitarian function replaces the sportive function. Higher levels of reliability and lower risk are then desired, for both improved product efficiency and public safety. The improved "state-of-the-art" in safety results primarily from 1) increased attention to product liability, 2) increased knowledge

gained from past failure experience, and 3) from increased knowledge gained from engineering efforts, whose costs can be spread over the increasing population of users. While the risk decreases most significantly in the first 20 years of the product's usage, 50 years are generally required before the 10^{-6} norm is reached. This time period is required for development and implementation of improved technology, often an iterative, trial-and-error process, legislation of safety codes and standards, and the enforcement of these standards by government agencies.

The equilibrium level of acceptable risk to society appears to depend on whether the activity is voluntary or involuntary.[5,6] Sports and transportation modes are voluntary, in that the consumer has a choice as to whether he wishes to participate and by which mode (e.g., driving vs. air travel). Other activities on which society spends its resources are nonvoluntary. For example, the local utility company determines what type of fuel (coal, oil, nuclear) it will use in providing its customers with electric power, and the customers are involuntarily exposed to the risk associated with this power source. Figure 6 shows a plot of risk vs. benefit for a variety of activities, where benefit is measured in terms of dollars expended on the activity. Starr has found that the risk accepted is proportional to the third power of benefit. He also notes that the risk level of voluntary activities can be about 1000 times that of involuntary activities. In other words, individuals will accept higher risk levels, if they are free to choose them, than they will allow society to set for them.

These facts suggest some guidelines that society has subconsciously set for determing what constitutes "acceptable risk". If the activity is voluntary, an acceptable risk level would be the natural level set by disease and old age. While driving and commercial aviation have reached this point, general aviation currently poses a risk which is about 25

times greater; this higher risk level will probably remain unless and until more individuals become dependent on non-commercial aviation, an the activity becomes more integrated into our way of life. With respect to involuntary activities, it appears that society imposes a requirement that the activity be 1000 times safer. Consequently, in choosing to expose the population to an involuntary risk, some considerations need to be given to both the risk level and the benefits. Any activity whose risk is below that due to random ("Act-of-God") phenomena such as lightening strikes (10^{-10}) must be considered safe by any criteria, as society has little or no control over the probability that a random event will occur; society can only reduce this risk by minimizing the consequences of the event (e.g., through building codes designed to assure some earthquake resistance). The acceptable level of risk for other involuntary events should be related to the benefits of the event. As the benefit level increases, society should be willing to assume a higher risk level, although in only rare instances should the level exceed that set by disease and old age (10^{-6}).

While Starr's concepts have great merit in defining the scope of the "acceptable risk" question, systematic data dealing with naturally occurring levels of risk are often not available, and even when available, they may be either scanty or of questionable reliability. Moreover, the basic premise of the method -- that the level of risk associated with such naturally occurring events as disease constitutes a threshold of acceptability somehow recognized (though perhaps tacitly) by society -- is still as much of an hypotheses as a proven fact. In fact, people often appear to find levels of risk associated with natural occurrences to be too high and act to lower the levels of risk. Thus, people use

lightening rods on their homes, and they have themselves and their families immunized against many common diseases.

There is, however, another method that can, in an important class of situations, be used to assess whether the risk associated with a particular kind of event or product is "acceptable". Suppose we are studying a system -- an airplane or an automobile -- that contains a particular part or component P, and we wish to determine whether the risk associated with a defect in part P is acceptable. Suppose further that the system has numerous parts and components besides part P, and that people customarily use comparable systems -- comparable airplanes or automobiles -- and accept the overall level of risk as a whole. Inherent in the overall level of risk is some degree of natural variability, reflecting the fact that accidents can never be prevented entirely, and that their frequency and consequences can never be predicted with certainty. Different severity values will be associated with different accidents resulting from a given product, and the frequency of accidents will not be perfectly uniform; thus, there will always be some scatter in the overall risk about an average value. In such a situation, the level of risk associated with a defect in part P may be considered to be "acceptable" when that level of risk is not only smaller in magnitude than the average level of risk associated with the system as a whole, but is also as small as (or smaller) than the variations in the overall level of risk.[3,4]

This method of determining what constitutes an "acceptable" risk for designers to aim at seems intuitively reasonable. In voluntarily using the entire system,people unavoidably encounter and accept an average overall level of risk with an inherent variability. For instance, although there is an average level of risk associated with driving during the course

of a year, the precise level of risk varies from time to time depending on many factors, such as weather, road conditions, local population density, and whether it is a holiday weekend. If the risk associated with a defect in part P (a particular automotive component for instance) is about the same as or smaller than the unavoidable variations in the overall system risk, the risk associated with part P will look to the user of the system exactly like one of the unavoidable fluctuations in the overall risk. Since the user is willing to accept the overall risk together with its inherent variability, he should also be willing to accept the risk associated with the defect in part P. Table III shows an application of this method.[4] In this particular case, a determination of the risk of automobile accidents due to fatigue failure of a steering component (the Pitman Arm) was determined from analysis of failure frequency and severity. The data shown in Table III indicates that the risk due to Pitman Arm failure is 0.02% of the risk of driving and 0.7% of the risk of driving a defective vehicle. As these percentages are well below the natural variations in risk, the particular component was considered to pose no unreasonable risk to society if continued in service.

5. PROBLEMS OF RISK PREDICTION AND FORESEEABILITY

In the preceding section, we have considered three primary criteria by which society, either consciously or subconsciously, determines whether a risk is acceptable:

1) Whether the risk is low compared to naturally occurring risk or to pure chance,

2) Whether the risk is justified in terms of its benefit, and

3) Whether the risk contributes a negligible or undetectable degree compared to the total risk of an acceptable activity (comparable to peer group risk).

Individuals, like society, utilize these same general criteria for setting levels of acceptable risk. The chance of a failure, its consequences, and the benefit associated with the product, enters the thinking process of engineers, plant operators and consumers who design, build, maintain and use these products. In the ideal world, the cost of failure could be accurately predicted in terms of the failure rate and average cost per failure and, knowing all other costs of the system, (e.g., the cost of design, manufacture, quality control, etc.) it would be possible to optimize the system by selecting the highest benefit/cost option. However, in the real world it is often not possible to accurately calculate the risk based on evaluation of failure frequency, and to evaluate all possible consequences of a failure. Indeed, this "event tree" approach is only used to evaluate critical paths in large complex systems (e.g., Nuclear Power Plants) where the acceptable failure rate is extremely small. The cost of these efforts is then justified, even if they identify only a few hazards that have not already been noted by other means.

Consequently, past experience serves as a major guide for making decisions about the future and for setting levels of acceptable risk. Although imperfect, this may be the only reasonable guide that is available. In using past experience to evaluate future failure rates and consequences, certain factors need to be considered. Figure 7 illustrates the "bathtub" curve (so named because of its shape) that represents the rate of failure of a product, process, human being, etc. as a function of its age. When first introduced, unforeseen factors in design, errors in tooling and quality control, and lack of understanding of operators contribute to make the failure rate of a system or component relatively high. This period in the product's life is called the "infant mortality" period and is the analog to both a weeding out of highly defective individual units or to the period where the state-of-the-art in system reliability increases rapidly (Figs. 4 and 5). After proper adjustments are made, the failure rate decreases to a constant level. During the "wear-in" period, where the failure rate is low, constant and "acceptable", those failures that do occur are the result of abusive loads that are placed on the system, overloads that are far in excess of what the system was designed to handle. The period where the failure rate finally increases is known as the "wear-out" or "end-of-life" period. Here, failure occurs because time-dependent degradation processes such as fatigue, corrosion, wear (for metals), and old-age (for individuals) destroy the effectiveness of the components in the system. The mean life of the part, divided by some factor of safety, is then used to set an allowable operating lifetime.

Decisions relative to the safe wear-out period are complicated by different preceptions of acceptable failure rates. If service history, tests, and/or analysis indicate that the mean life of a safety related system on, say, a light aircraft is 4000 hours, the manufacturer may suggest (and/or the FAA may require) that the system be replaced at 1000

hours, the safety factor of 4 being introduced to account for scatter in fatigue life. This approach to a safety problem may bring on complaints from individual aircraft operators that the system life should be extended to 2000 hours since they (the operators who voluntarily use the system) have had no problems during the first 1000 hours. The inference is even made that a low allowable system or life had been set so that the manufacturer could sell more spare parts. It is no wonder that manufacturers, accused on one hand of producing products with too high a failure rate, and on the other hand of replacing parts with too low a failure rate, often feel squeezed by competing interests of the government, environmental groups and their customers.

Companies and others responsible for compliance with safety regulations need to monitor the failure rate curve to determine whether a defect causing infant mortality remains in the system, whether the defect can be repaired, or whether the component should be replaced. In large systems (e.g., autos, household appliances), evaluation of warranty complaints provide good indications of infant mortality, provided the data is collected and analyzed; but with smaller systems or individual components, there is often no method for systematically evaluating the infant mortality period and for screening out defective parts prior to a major accident. In these instances, it may be necessary to use literature reviews and/or data surveys to determine the failure rates of analog components, and then use engineering analysis to account for differences in stress level, environment, etc. to predict the failure rate of the subject system or component. This technique is called Combined Analysis.[7]

The biggest unknown in risk assessment deals with the severity or consequences of the failure of a part, usually inexpensive, that is placed in the interior of some component. The failure of a 50-cent accelerator

linkage spring on an automobile can be handled by switching off the ignition and bringing the vehicle to a stop at the side of the road. But, if the failure occurs at a busy intersection, and the driver is too flustered to turn off the ignition, pedestrians may be injured or killed before the vehicle can be stopped.

Because of the inverse relation between failure frequency and severity (Fig. 3) it is likely that thousands of springs will fail before one fatal accident takes place. It is even possible to find and review spring failure data and determine that an accident of this type was probable. This review might suggest that the event was so foreseeable that the auto manufacturer should have installed a second, parallel spring into the system, so that one spring failure would not lead to system failure. However, there are no simple methods available for predicting the frequency or the slope of the frequency/severity curve for spring failure, and, in the absence of this information, it may be difficult to make a quantitative analysis about the need for the second spring. Clearly, there is a need for a more detailed review, analysis, and feedback of failure incident reports before a major accident occurs, but even at this time few manufacturers and operators of large systems have set up reporting and analysis programs to perform this feedback service.

The risk associated with a given product or process can be reduced by reducing the failure frequency, f, the average severity, $\overline{S}$, or by doing both of these things. Figure 8 shows the $(f,\overline{S})$ graph divided into four regions. The upper right-hand corner (Region 1) represents high frequency, high severity events, where the hazard is high and can only be reduced by drastic reductions in f, $\overline{S}$ or both of these parameters. The lower left-hand corner (Region 3) describes low frequency, low severity events that pose little hazard and are "safe".

The failure of certain components (e.g., tires) lead to major accidents because the energy release associated with the failure event leads to loss of control of a system possessing large amounts of kinetic energy. The slope of the f/S curve is then shallow and the hazard I=f x S is large. Hazard reduction can be achieved by reducing the failure rate of the component, or by adding a back-up system (e.g., dual wheels) such that control is still maintained in the event of component failure. Alternately, if the severity $\overline{S}$ associated with a failure event is low, a small reduction in failure rate (Δf) will have little impact on system hazard, as $\Delta I = (\Delta f)(S)$ will also be small.

In the other two areas (Regions 4 and 2) the risk is acceptable, as described by a curve of constant hazard, I_a,

$$f = \frac{I_a}{S}$$

The curve is drawn as a band to account for natural variations in risk, as described previously. In Region 4, most of the risk to society results from a high failure rate, f. These accidents generally involve individuals engaged in activities where small amounts of stored energy are released by a failure (e.g.,a slip-and-fall, a burn from a cleaning solvent). In the absence of a protective covering for vulnerable parts of the body (e.g., helmets) it is unlikely that society can lower the risk by lowering the average severity, which is already low. Instead, improvements will come from lowering the failure rate (e.g., from A to B) through increased efficiency of warnings and placards and through increased public awareness of the need for individual responsibility in activities such as going down stairs, riding bikes, etc. Most of the most hazardous products listed by the CPSC (Table II) fall into the high frequency, low severity category.

Region 2 deals with risk associated with low frequency, high severity events such as explosions, air crashes, and derailments. The far lower right portion of the region describes the real or perceived high severity event where no failures have occurred over a limited amount of service (e.g., catastrophic failure of a nuclear pressure vessel). In these instances, f is so ill defined that part failure (and success) experience is not sufficient to make reliable predictions of risk. Risk assessments must then be based on the operational history of similar structures (e.g., non-nuclear pressure vessels) and Combined Analysis[7]. Primary reductions in risk in Region 2 can best be made through the use of fail-safe warning systems (e.g., leak before break), and through reduction in $\overline{S}$ (e.g., from C to D) through isolation of sub-systems and structures that store large amounts of energy; this will assure that failure in one sub-system does not trigger off a large accident chain.

If past experience or engineering analysis indicates that severe consequences could result from a failure, then a new product or process needs to be fully tested before it is placed on the market. The testing should be severe and in some cases the product should be proof tested at stress levels far in excess of operational levels, to determine its failure modes and its potential for failure due to abusive loading. In some cases, such as medical prosthetic devices, it is impossible to simulate the in-vivo environment with in-vitro testing. Continual evaluation of product performance in patients then provide the most significant "test" of a product's design adequacy.

One of the major questions that face a designer is "foreseeable" use, and a proper definition of the allowable service loads and the abusive overload. Suppose that a gas pipeline is designed to operate at a certain

pressure, P, without failing. Proper maintenance procedures can be set up to assure that cracks or corrosion do not lower the residual strength of the pipe. But if a scraper impacts the pipeline, which causes a large crack, which leads to a large scale brittle fracture, and then to an explosion from the escaping gas, a personal injury accident is likely. Is this form of loading "reasonably foreseeable" and should it be accounted for in design? Should a manufacturer of bearings assume that proper lubrication schedules, although well outlined in company literature, will probably not be followed, and that there is a moderate probability that bearing seizure and overload fracture will occur? There are no well defined answers to these questions, (in contrast to the question of whether a casting defect caused a failure) and only some general observations are in order.

The fact that an accident of the type described above has occurred once does not make it foreseeable (unless it has happened several times before). There is simply not enough engineering manpower available to postulate all of the 10^{-15} scenarios; even if an event is postulated, that does not mean it is "probable". Furthermore, even if a 10^{-15} event is identified, what should be done about it? Should additional safety systems be added to reduce the calculated failure probability to 10^{-20}, or is 10^{-15} already low enough? Four factors should be considered in defining the "state-of-the-art" and determining whether the design was adequate, or whether either the failure rate or the failure consequences are unacceptably high. First, does the product meet state or federal regulations? Second, do competitive products and systems guard against this accident? Third, does the risk associated with the "unforeseen" event contribute a significant fraction with the risk associated with similar activities? Finally, will introduction of a safety system to prevent the "foreseeable" event lead to little or no improvement in safety because a) the safety system itself may fail and cause other problems, b) because it gives an operator a false sense of security.

6. SOCIETAL CONTROLS ON RISK

This paper has considered risk in quasi-technical terms, without regard for the mechanisms that society has to control risk, and discussion of this topic is in order. Traditionally, the market place provided the control on risk; if one product is safer than the other, and if safety is of concern to the individual, then he will choose the safer of two similar products when given the choice. Many risks, however, are involuntary; and even in voluntary activities such as hunting and skiing, the risk level associated with the purchasable hardware (e.g., rifles, ski bindings) is very low compared to the risk resulting from user error. The major voluntary decision that is available to the consumer is whether or not to purchase insurance, but this decision provides no protection against the accident.

Decisions relative to risk reduction and accident prevention have, in the last decade, been made primarily by federal and state governments. Federal legislation has created a whole group of agencies, such as the Nuclear Regulatory Commission, the National Highway Traffic Safety Administration, and the Federal Aviation Administration, to name but a few. These agencies have the broad responsibility to determine whether drugs, automobiles, aircraft, and nuclear power plants are "safe". If these agencies decide otherwise, they have the authority to remove these products from service, and require repair or upgrading (e.g., auto recall campaigns, grounding of a fleet, etc.). Manufacturers and operators that are faced with these recall decisions have the option of complying with them or of challenging these decisions in court. In effect, this has brought the judicial system directly into the decision-making process of what constitutes an acceptable risk.

The courts have traditionally been concerned with the question of reasonable risk on or after-the-fact basis, as accidents invariably lead to civil law suits between injured parties and/or corporations. Engineers often question whether lay judges and juries have the ability to render correct decisions in complex technical matters, and corporations and insurance companies are concerned that a sympathetic jury will make awards that are grossly unfair.

This author has had the privilege of testifying as an expert witness in a variety of product liability litigations. It is his opinion that the present system is both fair and satisfactory. Somehow, after all the conflicting evidence is presented, the juries are able to arrive at reasonable decisions, even if they have not understood all the complex technical facts. Similarly, in seven states where critics of nuclear power have sought to severely restrict its use through passage of Initiatives, the voters have decided that the risk of a nuclear incident is more than justified by the benefit of having the necessary electric power, and have voted down the Initiatives. These observations suggest that engineers and scientists should, like politicians, respect the ability of the public to determine its own best interest.

The two major problems associated with risk assessment relate to "perceived" risk and to long-term effects of decisions made today. The major disaster gets major attention by news media and causes the public to consider the possibility of nuclear catastrophes, dangers from nuclear sabotage, and leaks of radiation and other highly toxic substances. These perceived risks may be actually much lower than more mundane risks due to radiation damage from excess sunbathing. The perception issue is particularly acute when there is a possibility that not only the present generation, but future generations as well, will be influenced by

regulatory decisions. This author takes the position that these decisions and risks are not different nor greater than those taken by individuals such as Queen Isabella and Columbus. The discovery of America led to profound changes in the lifestyle of Europe, changes that could never be allowed on the basis of the state of "risk acceptance" that existed in the 16th Century. Likewise, it is not possible for engineers in the 20th Century to predict all possible forms of hazard and guard against them. Since a refinery manager cannot prevent an explosion by assuring that all overspeed governor devices are functional on his turbines, it is unreasonable to expect that engineers, legislators and administrators can make meaningful, safety related decisions that will have a guaranteed, nondestructive impact on future generations. The best method of assuring reasonable risk levels is to assure that reasonable, well-trained individuals are involved in the decision-making and operation of safety related systems, that proper data be kept and maintained, and that the public be made more aware of its role and responsibility in preventing accidents.

REFERENCES

(1) NEISS News, National Electronic Injury Surveillance System, U. S. Consumer Product Safety Commission (1973).

(2) United States Consumer Product Safety Commission, Computer Tabulated Accident Data Sheets, U. S. Consumer Product Safety Commission Bureau of Epidemology (1973-74).

(3) Besuner, P. M., Tetelman, A. S., Egan, G. R. and Rau, C. A., "The Combined Use of Engineering and Reliability Analyses in Risk Assessment of Mechanical and Structural Systems," Proceedings of Risk Benefit Methodology and Application Conference, Asilomar Conference Grounds, pp. 353-389, (December 1975) (UCLA ENG 7598).

(4) Tetelman, A. S. and Burack, M. L., "An Introduction to the Use of Risk Analysis Methodology in Accident Litigation," Journal of Air, Law and Commerce, Southern Methodist University School of Law, Dallas, pp. 133-164 (1976).

(5) Starr, C., "Benefit-Cost Studies in Socio-Technical Systems," Colloquium on Benefit-Risk Relationships for Decision-Making, Washington, D.C. (April 1971).

(6) Starr, C., "General Philosophy of Risk-Benefit Analysis," presented at EPRI/Stanford IES Seminar, Stanford, California (September 30, 1974).

(7) Tetelman, A. S. and P. M. Besuner, "The Application of Risk Analysis to the Brittle Fracture and Fatigue of Steel Structures," submitted for presentation at the Fourth International Conference on Fracture, Waterloo, Canada, June 19-24, 1977 and published in Conference Proceedings.

TABLE I

CPSC - NEISS VALUES FOR ACCIDENT SEVERITY CATEGORIES (1974)

SEVERITY CATEGORY	REPRESENTATIVE DIAGNOSIS	SEVERITY VALUE
0	Incomplete or otherwise not acceptable data	0
1	Mild injuries/small areas, dermatitis and sprains	10
2	Punctures - fractures	12
3	Contusions - scalds	17
4	Internal organ injury	31
5	Concussions - cell and nerve damage	81
6	Amputations - crushing and anoxia	340
7	All hospitalized category sixes	2,516
8	All deaths	34,721

TABLE II

MOST HAZARDOUS CONSUMER PRODUCTS IN 1974,
ACCORDING TO CPSC METHOD

PRODUCT DESCRIPTION	CPSC RANK
Bicycles	1
Stairs	2
Doors	3
Drain Cleaners	4
Tables	5
Beds	6
Football	7
Swings	8
Gasoline and Kerosene	9
Home Structures	10
Power Motors	11
Baseball	12
Nails	13
Bathtub and Shower Structures	14
Gas Space Heaters	15
Swimming Pools	16
Gas Ranges	17
Basketball	18
Non-Upholstered Chairs	19
Storage Furniture	20
Unpowered Cutlery	21
Clothing	22

TABLE II (Continued)

PRODUCT DESCRIPTION	CPSC RANK
Paints	23
Household Chemicals (other than caustic)	24
Money	25
Floors	26
Glass Bottles	27
Washing Machines	28
Matches	29
Ladders	30
Sun Lamps	31
Home Workshop Saws	32
Fences, Non-electric	33

TABLE III

Calculated Values of Frequency, f,
and CPSC-Defined, Mean Severity, $\bar{S}$, and Total Severity I
for Several Vehicle Systems, Major Sub-Systems, and Components

	f	$\bar{S}$	$I=f\bar{S}$ (1)	% of Driving Risk	% of Risk Due to Defective Vehicles
All Vehicles	6.65×10^{5}	197	1.3×10^{8}	100	
All Defective Vehicles	1.39×10^{4}	254	3.78×10^{6}	3.4	100
Tires	2455	695	1.7×10^{6}	1.6	45
Other Defects	1341	509	6.82×10^{5}	0.6	18
Brakes	7065	92	6.5×10^{5}	0.6	17
Lights	410	552	2.27×10^{5}	0.2	6
All Steering	587	280	1.64×10^{5}	0.14	4
Trailers	951	36	3.42×10^{4}	0.03	0.9
Wheels Come Off	585	47	2.75×10^{4}	0.02	0.7
Turn Signals	428	32	1.37×10^{4}	0.01	0.3
Wipers	8	5	40	3×10^{-7}	10^{-5}
Pitman Arm Separations (Incidents)	8270	3.24	2.68×10^{4}	0.02	0.7
Pitman Arm Accidents	258	104	2.68×10^{4}	0.01	0.7

(1) All data have been normalized to the total vehicle usage, 4.45×10^{6} vehicle-years, in the State of Texas in CY 1971.

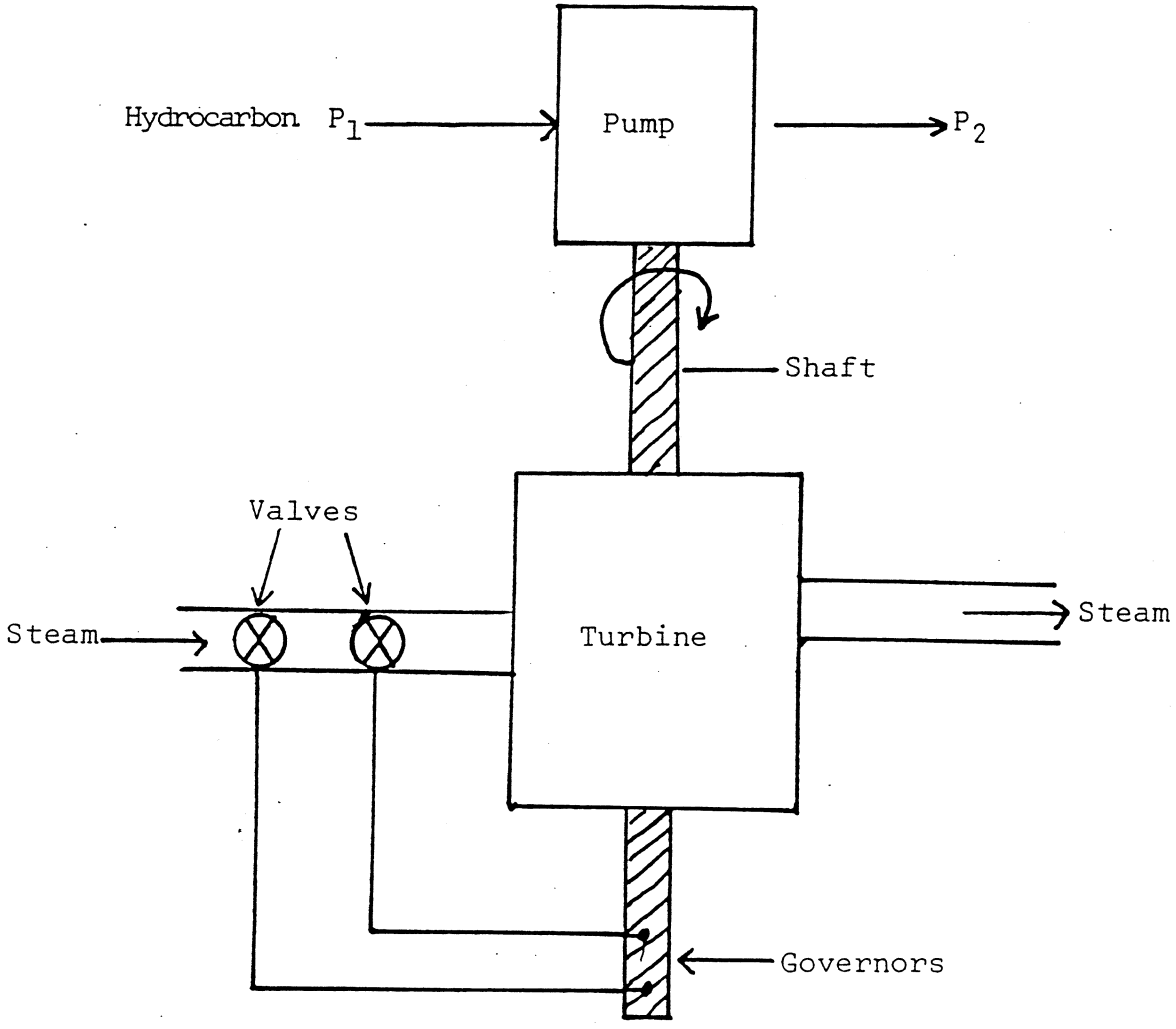
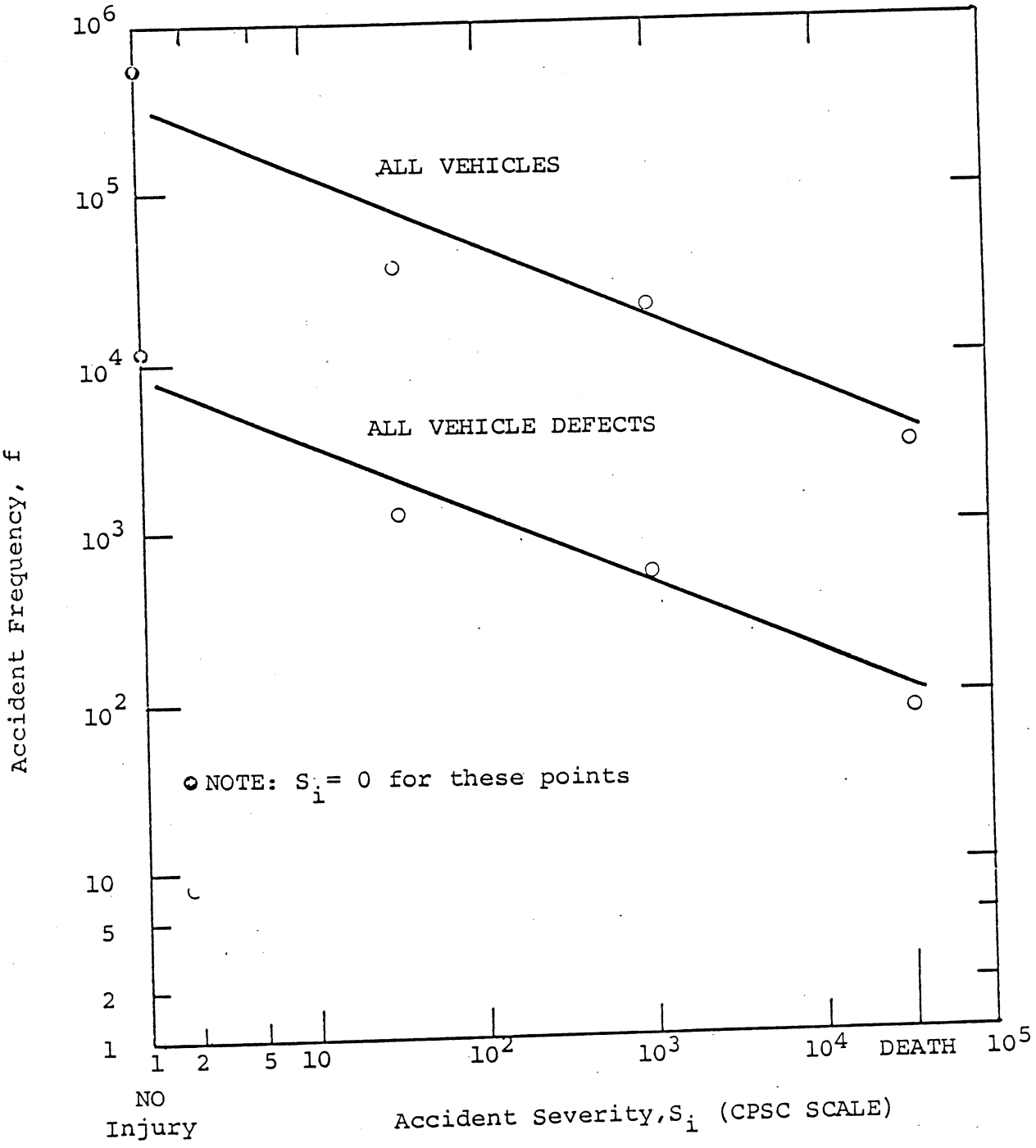


Figure 1: Simple pump/turbine system.



Automobile Accident Frequency-Severity Distribution (State of Texas 1971)[4]

Figure 2

Frequency - Severity for Automobile Accidents Caused by Various Defective Vehicle Subsystems and Components (From All Data Categories in Table III Listing One or More Fatalities).. (4)

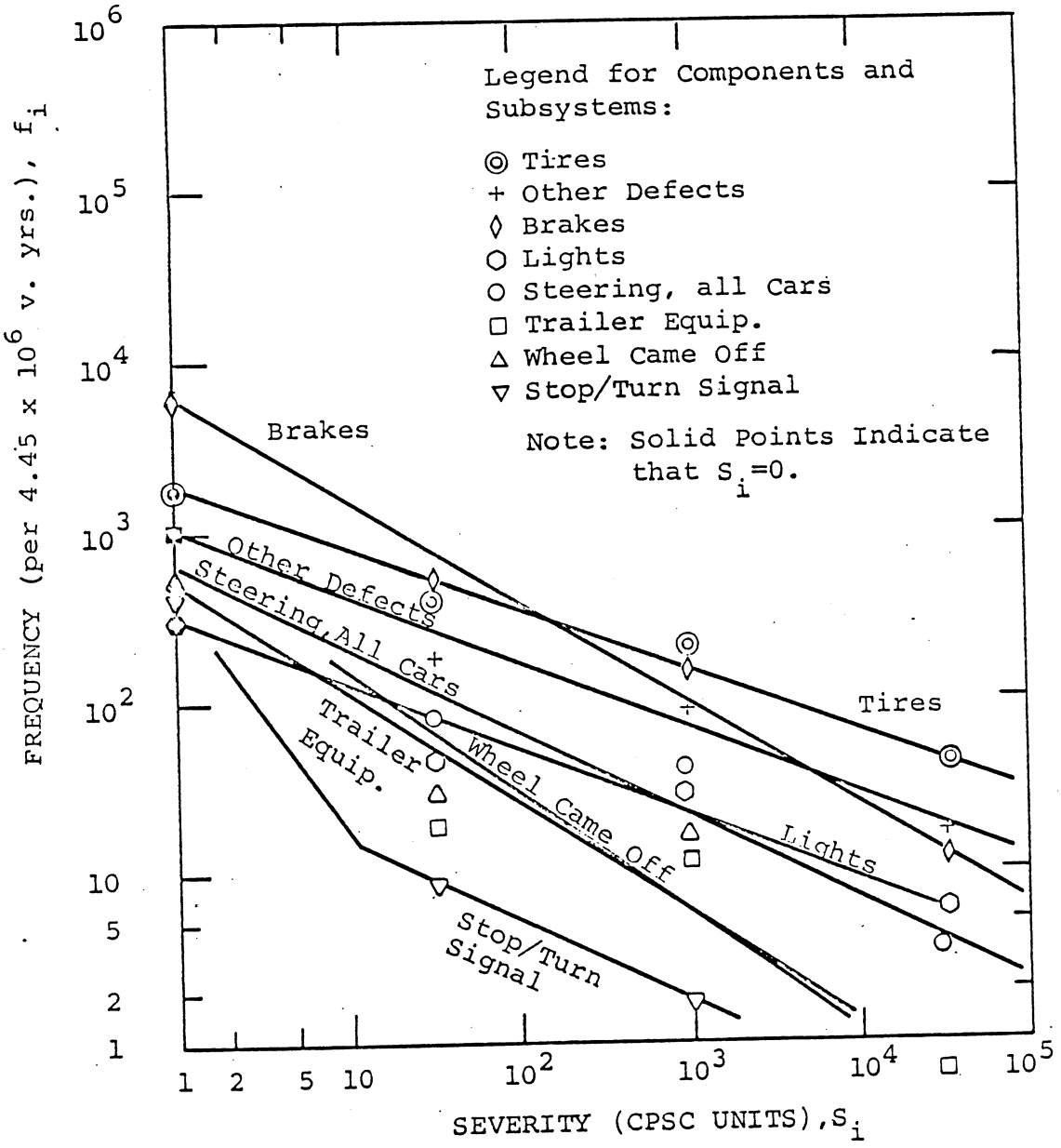


Figure 3

RISK AND PARTICIPATION TRENDS
FOR
MOTOR VEHICLES [6]

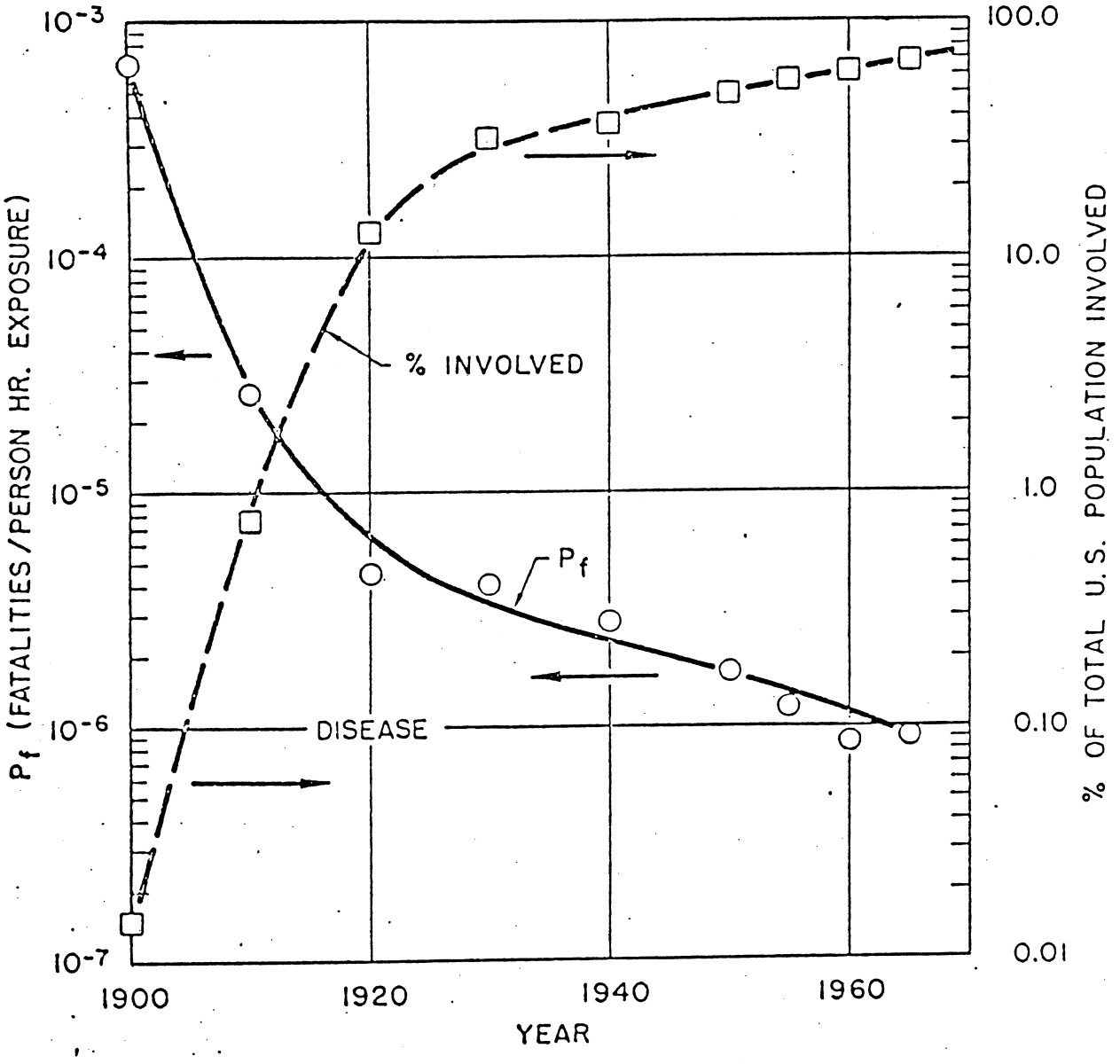


Figure 4.

RISK AND PARTICIPATION TRENDS
FOR
CERTIFIED AIR CARRIERS[6]

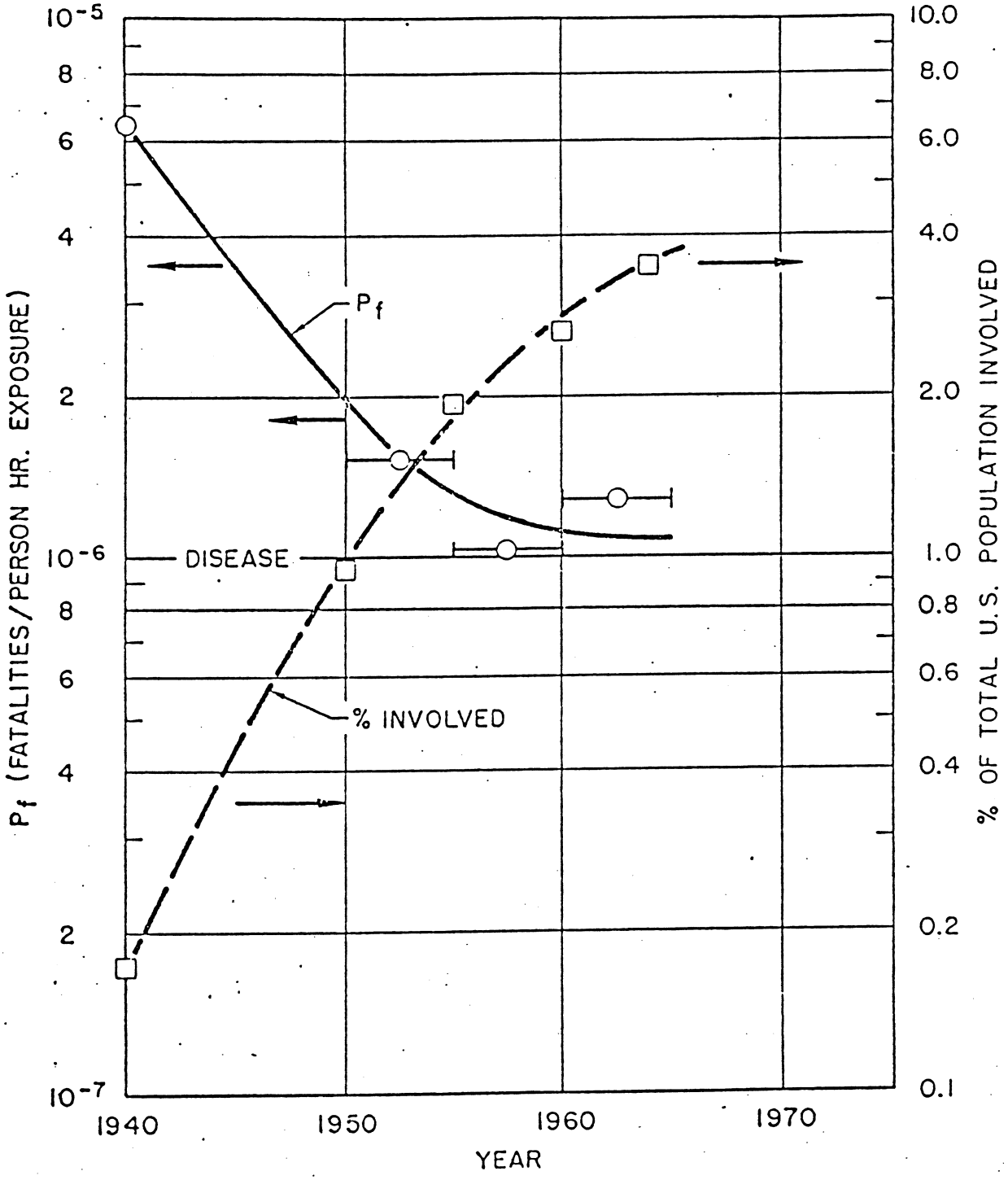


Figure 5

RISK VS. BENEFIT
VOLUNTARY AND INVOLUNTARY EXPOSURE[6]

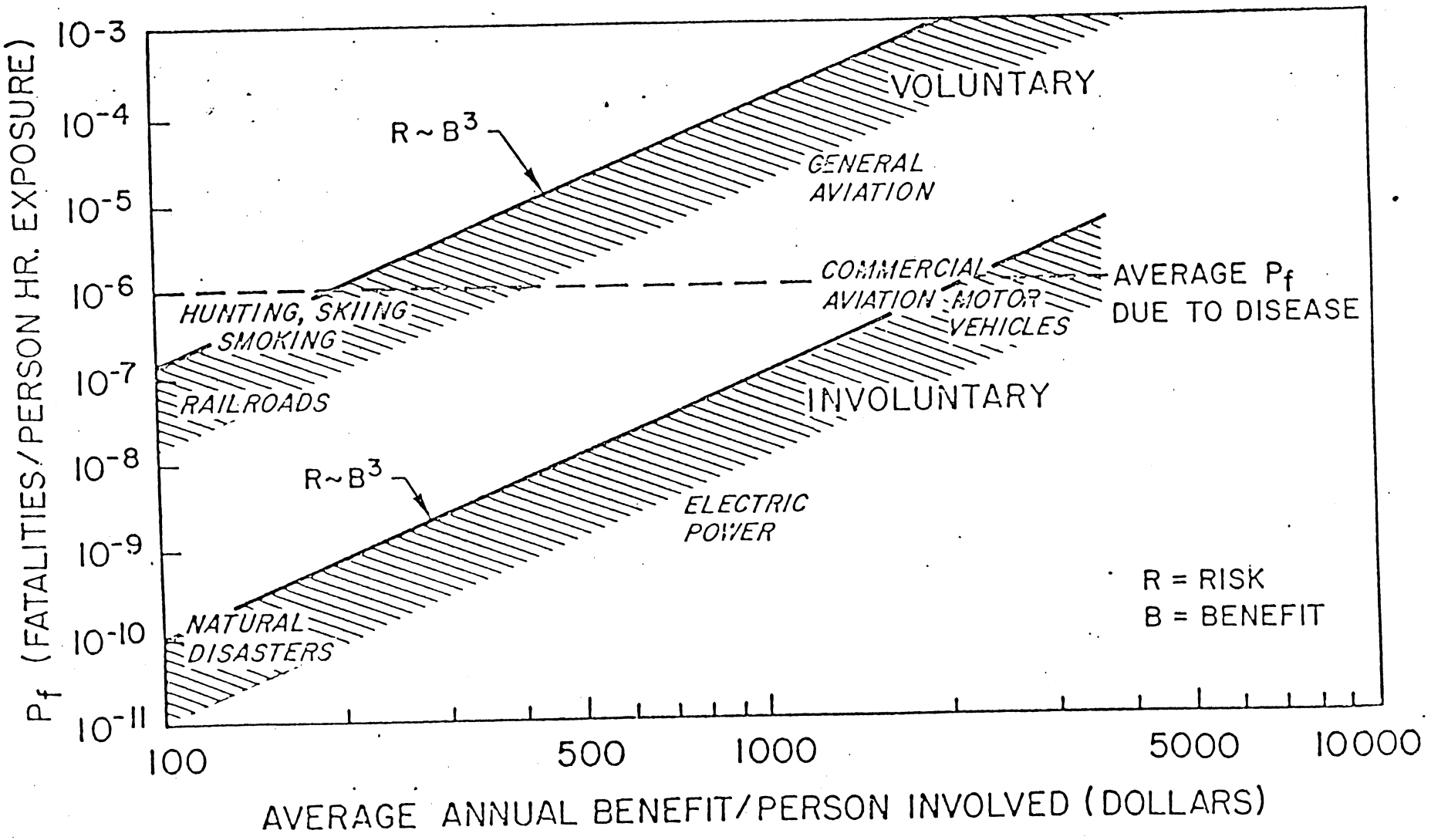


Figure 6

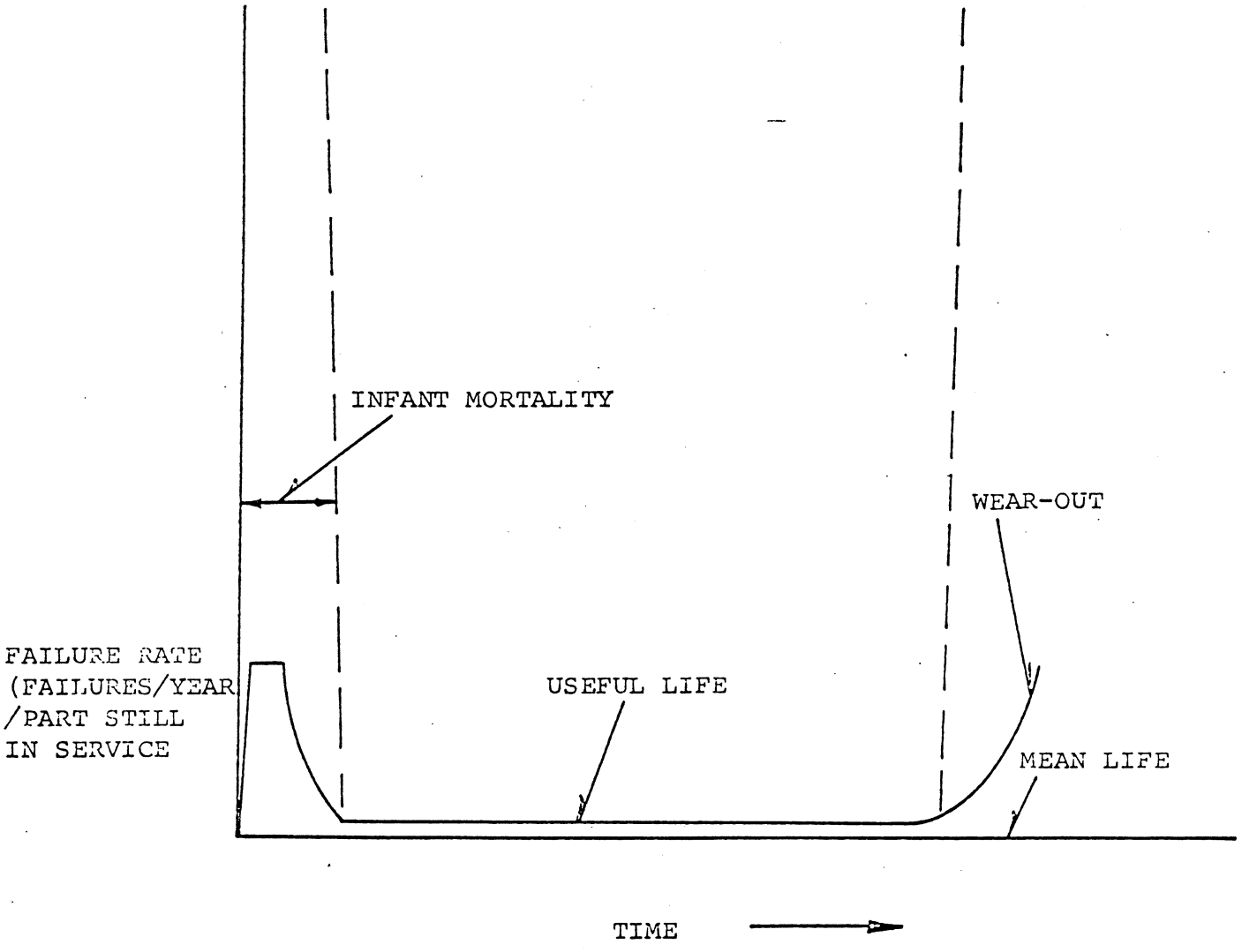


Figure 7

Typical Failure Rate Curve

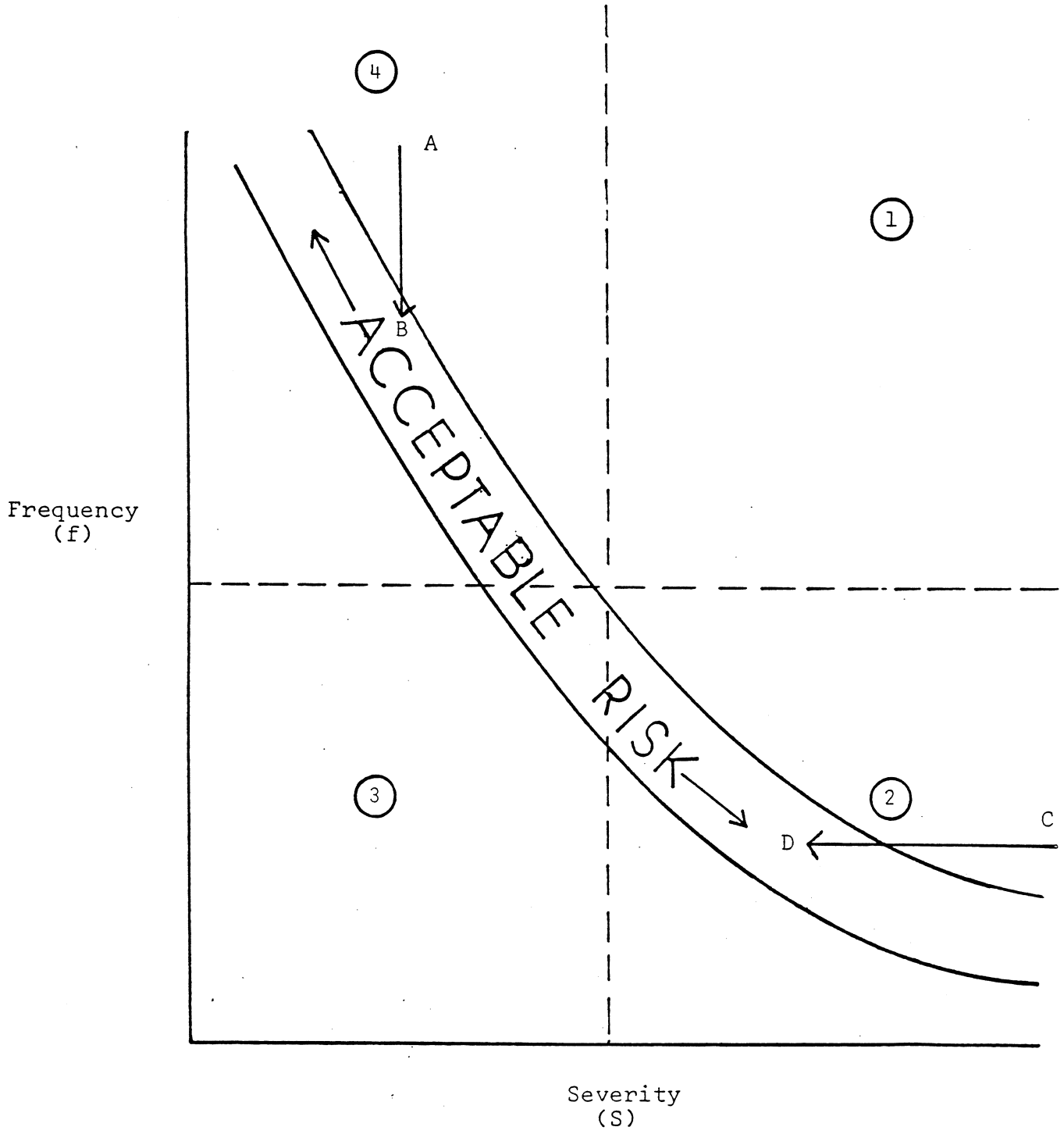


Figure 8

Frequency/severity curve for accidents. See text for explanation.