Science & Risk: How Safe is Safe Enough

by David Kane

Published in Harpur Academic Review, 1992

Science and Risk: How Safe is Safe Enough

Abstract

Risk is a difficult concept to measure and effectively use in decision making. This paper seeks to present the problems of risk decisions in order to present models that can be used to answer the question, *"How safe is safe enough?"*.

First the basic terms in the study of risk are outlined. This is important, because many of the terms are used interchangeably, and the distinctions between them are not always clear.

The problems confronting those studying risk are discussed. There are three basic disparities that exist in the way risk is dealt with. There are (1) differences in the way risk is perceived, (2) differences in the way risk is measured, and (3) differences between these perceptions and measurements. These variations will each be discussed in turn, in order to understand the dilemmas they create for policy makers.

The resolution of these dilemmas is the subject of the last section of the paper. A model is presented that incorporates the perspective of both science and the public. It calls for science to provide risk assessment along with confidence intervals for their assessments. This puts the value decisions of safety into the hands of policy makers. Further, an interface is established, to provide policy makers with the tools to make effective risk decisions.

Introduction

In Managing the Unknown, William C. Clark presents a contemporary version of an ancient myth. It is a story that epitomizes our society's attempts to come to grips with the risks of science and technology:

> A young man came to two doors, and he could open either as he pleased. If he opened one, there came out of it a hungry tiger, the fiercest and most cruel that could be procured, which would immediately tear him to pieces. But if he opened the other door, there came a lady; the most suitable to his years and station that His Majesty could select among his fair subjects.
>
> He refused to take the chance of opening either door. He lived safe and died chaste.
>
> Then another young man came to the door, and he hired risk assessment consultants. He collected all the available data on lady and tiger populations. He brought in sophisticated technology to listen for growling and to detect the faintest whiff of perfume. He completed checklists. He developed a utility function and assessed his risk averseness. Finally, sensing that in a few more years he would be in no condition to enjoy the lady anyway, he opened the optimal door. He was then eaten by a low probability tiger.
>
> A third man took a course in tiger taming. He opened a door at random and was eaten by the lady. (1977)

This tale shows many of the dilemmas facing society. The potential rewards of science and technology are great, but so are the possible dangers. Our society must make choices, even if it chooses to do nothing. As we are try to make our technology safe, we wonder "when is it safe enough to open the door?".

Newspapers often carry stories about the latest studies connecting our eating habits to the risk of cancer or other diseases. Yet in these and other areas of our life, we find a close connection between risk, our habits, and public policy. Consequently, both government leaders and the general public need to make decisions based on risk, and try to answer the question *”how safe is safe enough?”*.

The answer is not always readily found. The already murky waters of risk are further muddied by three fundamental disparities: (1) The first of these concerns how people perceive the same risks differently. (2) The second disparity is the difference between risks calculated for the same event using different models. (3) Even if we can agree on a measurement of a risk and come to a consensus on the appropriate perception of that risk, we find that our assessments are rarely in agreement. These phenomena cloud the issue of what is safe science and technology.

Upon what basis then can we decide how safe is safe? Do we consider how the public perceives a risk, or how science measures that risk? Perhaps both ought to have a role in formulating public policy, but how can we combine these often divergent assessments?

The government must answer the question of what a sufficient level of safety is. In order for that decision to be the best possible choice for society, it must first be based on a sound understanding of the principles of risk. Then, it must allow science to provide a clinical assessment of the issue, while reserving the right to make value decisions for the society’s leaders.

This paper will begin its analysis by considering the definition of risk and other key terms. Next, it will examine the three basic disparities that occur when evaluating risk. With this groundwork established, it will then be possible to construct a model by which society can effectively answer the question, *“how safe is safe enough?”*.

Definitions

The study of risk, and its applications in science, brings many different ideas together. Unfortunately, as we try to answer the question of how safe is safe enough, the distinction between these concepts blurs. What is the difference between risk and uncertainty? What distinguishes uncertainty from ignorance? How are ignorance of risk and perceived risk different? (Zeckhauser & Viscusi, 1990) To clarify these differences and provide solid groundwork upon which the discussion of science and risk can be based, several definitions are needed.

•Risk

Risk is a frequently used term. However, it is also one of the most widely defined; what is risk to some is the perception of risk to others. For the purposes of this paper, risk will be as defined by Richard Zeckhauser and Kip Viscusi, “the states of the world that may prevail, and the precise probability of each state.” (1990) Risk is the collection of things that can happen, and how likely it is that each will occur. Associated with each of these states is the severity of the damage that each event can bring. By this definition, risk can then be expressed as:

$$\text{Risk} = \sum (\text{Probability} * \text{Damage})$$

This interpretation makes risk an objective quantity, but one that introduces subjective qualities when used in practice.

Perception of risk is another important, albeit distinct, concept. It is expressed by a population’s choice to commit to or avoid an action. Two actions may have the same measured risk, but the public may deem one risk to be acceptable, while dismissing the other. To some experts, the perception of risk is molded out of ignorance. To others, perceived risk reflects a more intricate understanding of the world than can be expressed in a simple equation. There is little argument, though, that disparities exist between how we view the world and how we measure it.

•Uncertainty

Uncertainty is also a common term used in discussions about risk and safety. There are two main factors that create uncertainty: (1) the level of precision, and (2) the level of accuracy. Precision is an indicator of the specificity of a piece of information. Accuracy refers to the correctness of the information. In Figure 1, we can see an illustration of accuracy and precision as shown on a target. (Halliday & Resnick, 1988)

Graphic Interpretation of Precision and Accuracy

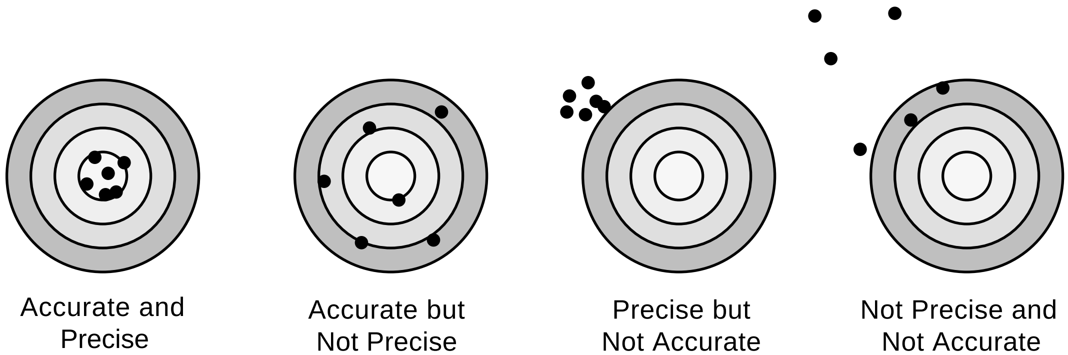

Figure 1

Ignorance also contributes to uncertainty. Zeckhauser and Viscusi write, "with ignorance we may not even be able to define what states of the world are possible." (1990) Ignorance is a result of an inability to define (1) the complete list of consequences resulting from an event, (2) the probability of the event, and (3) the severity of the damage involved. As we will see, failure to understand the magnitude of the consequences of an action can have a profound impact on our assessment of its risk.

•<u>Benefits</u>

In discussions concerning risk, we must consider what is meant by benefit; however, this too is a difficult concept to define. As with risk, there is a distinct difference between benefits and perceived benefits. This difference can be shown in the context of the Persian Gulf War when Iraq launched Scud missiles against Israel. The Israelis wanted Patriot missiles in order to better defend their country, yet the Patriots failed to minimize the damage from the Iraqi attacks.[1] (Conway, 1991) Within this paper, benefits will be defined as solid measurable goods, services or activities. In defining perceived benefits, psychological benefits will be considered as well. Should we weigh psychological benefits, even if there are no measurable benefits? From a scientific point of

1 Even though fewer than 10% of the Scuds launched after Patriot installation passed though the Patriot defensive shield, the damage estimates per missile before and after the defenses were installed were basically the same. That is, when a Patriot hit a Scud, the Scud would fall to the ground and do as much damage as if it had been allowed to travel unimpaired. (Conway, 1991)

view, the answer is probably no. Still, in the political arena, our leaders are obligated to do so. Pressure from constituents forces politicians to consider psychological benefits to some extent.

•**Product Z: A Hypothetical Example**

These definitions provide a good starting point for a discussion on science and risk. To see how these terms are interrelated and fundamental to this study, consider the hypothetical creation of *Product Z*. *Product Z* is a new chemical designed to enhance plant growth. There may be some risk that *Product Z* is harmful to humans or to the environment. Two studies are conducted to assess this potential danger: one experiment is conducted on 2,000 guinea pigs in a controlled laboratory; a second study follows the health of the 20 scientists and engineers exposed to the substance during its development. There is some risk associated with *Product Z*. However, the exact magnitude and probability of that risk, in terms of severity, or population size is unknown.

The first study offers a precise but potentially inaccurate assessment of the risk of *Product Z*. With great detail we could ascertain the effects of *Product Z* on guinea pigs, but its effects on people would be based on conjecture. Errors of significant magnitude can be introduced when applying results obtained from animal experiments to human populations.

The second study offers an accurate, but imprecise indicator of the risk of *Product Z*. Although using human subjects removes much of the uncertainty of the previous study, its small sample size makes precise assessments difficult. If some of the scientists contract health problems, it might indicate problems with the substance, but extrapolating them to the entire public would destroy any precision in the experiment. It would be analogous to predicting a Presidential election by asking just two dozen people.

Ignorance of the environment also plays a role. Though studies are being conducted on humans and animals, their interaction with the environment is never considered. Because of this, the full range of possible reactions to *Product Z* is not even known, let alone the risk potential of these responses.

In traditional regulatory practices, once the risk of *Product Z* is assessed, it will be weighed against the benefits. Economic, political and sociological impacts must be considered before any action will be taken on the product. That is, a risk-benefit comparison would be executed. Here is where the process moves from the domain of science to the political arena.

Perceptions of Risk

The study of risk does not occur in a vacuum. In fact, tremendous social, political and scientific pressures often confront those who study risk. These same kinds of pressures also face those who experience risk. These powerful, often conflicting forces can cause people in various situations to perceive the same risk in different fashions.

There are several factors that affect risk perception, over and above the measurement of the risk. First, we'll consider how the public compares new and unfamiliar risks with old and familiar risks. Next, we'll see how the level of control over an event influences how the risk is perceived. Lastly, we shall determine if the perceived benefit of an activity makes its risks palatable.

•New Risks versus Old Risks

The public's view of risk from an event is affected by a multitude of factors. Instead of a naive opinion, the public's perception of risk is based upon an intricate set of characteristics. How are "new risks" and "old risks" perceived? (Crouch & Wilson, 1982) Consider the case of energy; we have used coal power for more than a century. The risks associated with such an energy source are well understood, and have been documented with a high level of accuracy and precision.[2] Nuclear power, on the other hand, might be safer than coal power, but it may also be much more dangerous. Our understanding is neither necessarily precise nor accurate.

The relative youth of nuclear power influences the public's view of its risk. First, a risk that has always existed becomes a base by which other risks are judged. Generally, a risk that has always existed is regarded as an acceptable risk, while newer risks that are brought to the public eye receive greater scrutiny. This intense examination makes the new risk appear to be more dangerous.

[2] This argument does not consider the effects of global warming.

Ignorance of the new risk furthers this tendency. There is a great amount of uncertainty surrounding nuclear power, making it difficult for the public to view nuclear energy as safe. Otway and Simms describe this idea more simply as “fear of the unknown”. If the risk is difficult to conceptualize or comprehend, this fear of the unknown amplifies the magnitude of the the apparent danger. (1987)

When people contrast comparable risks, they feel safer with the ones they better understand. Unless the public has experienced an activity and has grown to trust it, they will be skeptical of its safety. Just as we eye a stranger on a dark corner with wariness, so the public views new risks.

•**Control of Risk**

The characteristic of control is important. An individual is more willing to accept the risk of an activity of which he or she is in direct control. For example, an airline passenger has little control over the flight of an airplane. On the other hand, a pilot of a small, single-engined airplane has a much higher degree of control. Although the risks of civil aviation both are much greater than those of commercial aviation, both are considered acceptable. If commercial aviation had the same failure rate as that of civil aviation, much of the public would not deem airline flight to be safe enough. Otway & Simms write that the general feeling of helplessness generated by lack of control, heightens the sense of risk. (1987) This underlying factor explains why poorly understood, or indirectly controlled activities have a high perceived risk.

A key type of control is the decision to be exposed to a risk. Voluntary risks involve this kind of choice, while involuntary risks lack this element of control. In 1969, Chauncey Starr presented a model of the public’s perception of risks and benefits to determine which risks the public had accepted in the past. He argued that there was a distinction between the perception of voluntary and involuntary risks. Because voluntary risks involve a choice based on an individual’s own set of values, the acceptance levels are a thousand times greater than those of involuntary risks. Sociologist Gary Wyatt’s study of risk behavior reinforces the importance of values in decision-making. His argument states that the values of an individual are key factors in how that individual will choose between risk and risk-averse alternatives. (1989)

Starr’s model also explains why the same activity may be safe enough for some, but not for others, even if they agree on the probable consequences of the act. If an activity is voluntary for one group but not for another, the first group might deem the risk of the activity to be acceptable, while the second may not. For example, a community might consider building a nuclear power plant. To the affluent and mobile residents, the choice of living near the proposed facility would be a voluntary one. However, economically disadvantaged citizens might not have the resources to relocate. For them, the risk would be involuntary and perhaps unacceptable.

•Risks versus Benefits

Chauncey Starr points out that the public’s acceptance is roughly based upon the perceived benefit of a risk taken to the third power. This model, which was built upon previous experience as the normative basis, explicitly seeks to answer the question, “how safe is safe enough?” (Starr, 1969) The following graph of what the public finds as acceptable and unacceptable, Figure 2, demonstrates the relationship between risk and benefit. The public is only willing to tolerate a minute level of risk for activities which it considers to be of little value, such as constructing nuclear power plants. Conversely, for those events which the public perceives great benefit, the acceptance level is quite high.

Risk Benefit Awareness[3]

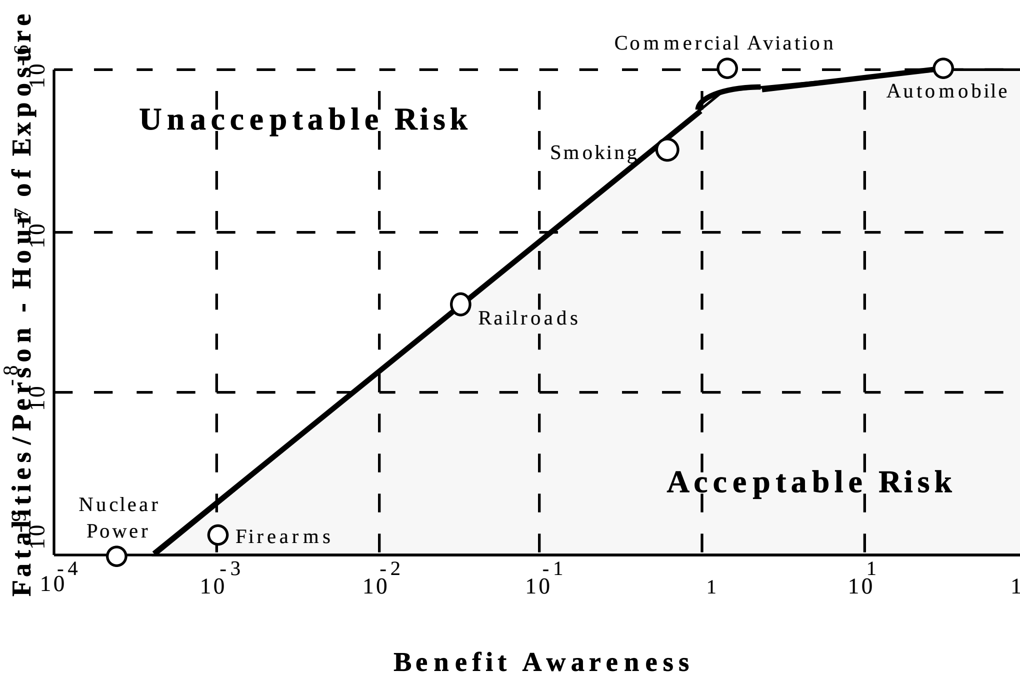

Figure 2

Another disparity explained by Starr's model is the receipt of difference benefits by distinct groups. In the above example, the citizens near the facility might not find the risk benefit ratio acceptable, for they are subject to greater risk. (Otway & Simms, 1987) Citizens some distance upwind might enjoy the ready availability of electricity enough to warrant the diminished risk that they would face. In order for these benefits to be persuasive, they must be visible. (Otway & Simms, 1987) Even if all parties concerned agreed on the likelihood of an accident at the plant, each could come to a different conclusion regarding the acceptability of that risk.

We come to the question of not only what the risks and benefits are, but of who experiences them. In his book <u>Normal Accidents: Living With High Risk Technologies</u>, Charles Perrow attempts to understand the risks of a high-technology society by classifying the victims of accidents into four categories. (1984) This classification system is based on the amount of

[3] Starr (1969).

knowledge and control that a person has over a system and can be broadened to include those exposed to the risks of any technological hazard. First Party victims are those who run the facility, and would therefore have the best information for making decisions, as they are in direct control of the system. Second Party victims are those who use or supply the facility. Third Party victims are innocent bystanders, and Fourth Party victims are the intergenerational victims of an accident.

Classification of Victims

First Party	Operators: first-level supervisor maintenance personnel, low-level engineering personne laborers, and assisting personnel.
Second Party	Associates: suppliers, passengers, users, and others knowledgeable of the risk, who "choose" to participate.
Third Party	Bystanders: Innocent bystanders, found in proximity to the accident, but with no direct connection.
Fourth Party	Descendents: Those who may n even have been alive, but are affected by an accident. Include both physical and social damage

Figure 3

We must understand the limitations of the Third and Fourth Party victims. If we only consider the risk’s perception, we may neglect the interests of those who might be affected. The next generation may perceive the same risk in an entirely different fashion, yet they do not have a voice with which to argue their perspective. For a conscionable decision concerning safety, we must be able to consider the interests of potential Third and Fourth Party victims.

The issue of Third and Fourth Party victims is not cut and dried. By seeking to make the world safe for future generations, we may still threaten their safety. Aaron Wildavsky writes, “The future will not be allowed to make mistakes because the present will use up its excess resources in prevention.” (1979) We can draw an analogy to the previously stated tale of the young man who chooses neither door and died chaste. If we become incapacitated by an inability to take risks, we will be unable to ensure that future generations will have opportunities of their own. A middle ground must be found, even when considering these Fourth Party victims.

•**Summary**

Upon reflection, it is easy to see why disparities can exist in the public’s perception of risk. In general, people view an activity as riskier if they feel helpless to understand, control, or avoid the consequences of that activity. How the public perceives benefit is also important. Not only must there be a risk-benefit balance, but the risks and benefits must be shared within a given group. That is, one group will not find risks acceptable if the benefits are received by another group, which presents a policy dilemma. While considering the interests of the public, which has a voice, we also must be able to consider the interests of future generations, which lack a say in our policy choices. Their perception of risk is potentially important, yet it is also easy to ignore in the panorama of public opinion.

Measurement of Risk

Just as the public does not view risk homogeneously, experts are seldom in accord. One reason for the experts’ disagreement is their use of different models to measure risk. Each of these models makes different assumptions and introduces different biases into its measurements. Now, we’ll look at these different processes, so we can better understand the way they affect our description of the world.

We will start with the most traditional method to measure risk, the event-tree, and will consider its limitations. Then we shall move onto the basic regulatory methods used to decide how safe is safe enough. By considering alternate models and methodologies, we can see how even the experts can come to such varied conclusions from the same basic set of data.

•**Event-Tree Analysis**

Traditionally, risk analysis has used simple probability models to assess the danger of a given system. One such model is called the "Event-Tree." In a nuclear power plant, for instance, the probability of any system failing is computed, based upon the experience with that device and its design specifications. A serious accident occurs when a given set of systems fail simultaneously. Such an accident would be unlikely since each system is independent, and the chance of such a path being chosen is small. Only in these cases, would serious consequences result. Figure 4 depicts a simplified example of such a tree. Surprisingly, proponents of nuclear power argue that Three Mile Island (TMI) and Chernobyl do not tarnish their logic.

Simplified Event-Tree

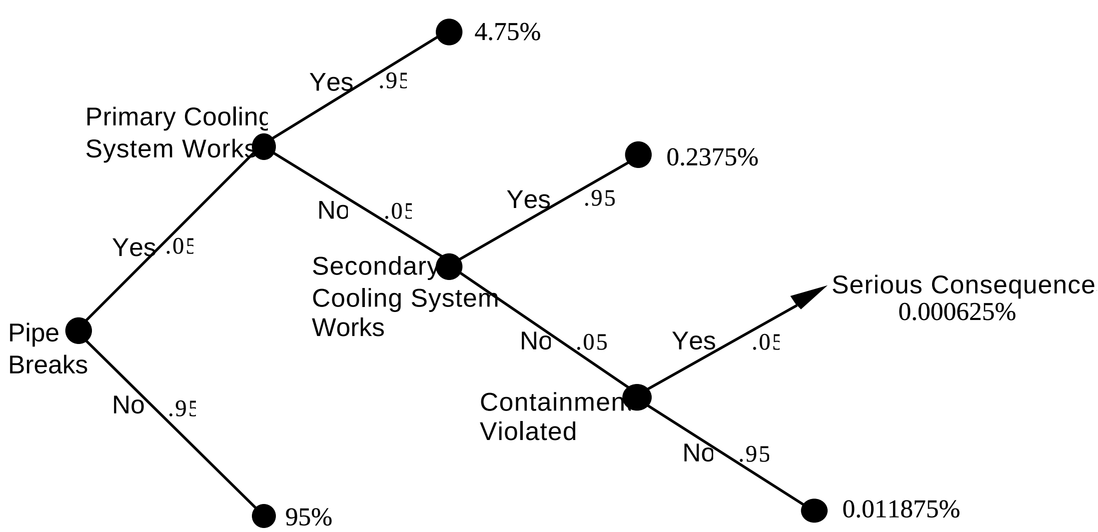

Figure 4

Chernobyl Event-Tree

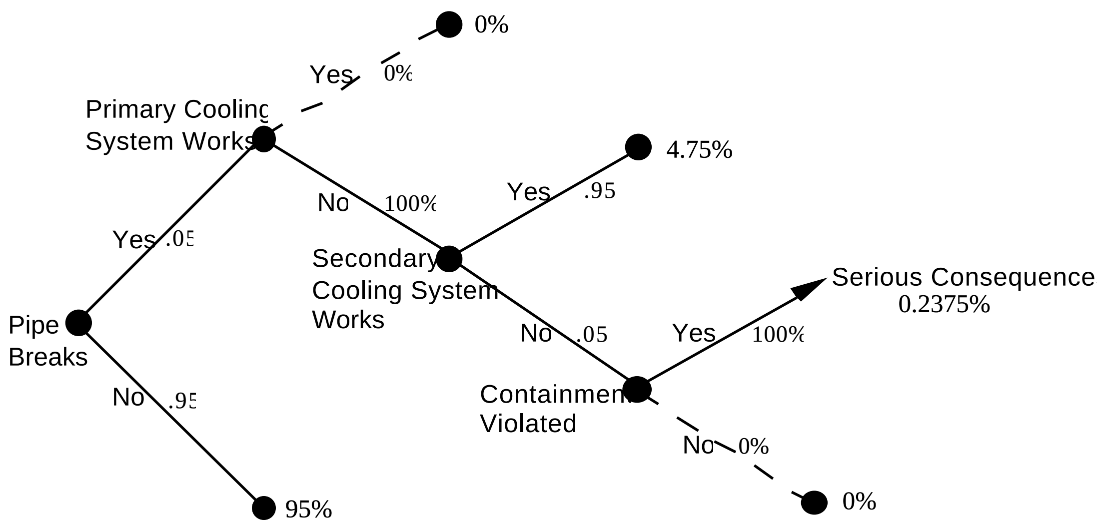

Figure 5

At TMI, multiple human and technical errors occurred, yet there was relatively little human damage. A minimum of radiation exposure resulted, and the vast majority of the damage was limited to the facility and not the surrounding area. (Ford, 1981) With such a major accident causing so little damage, relative to the disaster of a containment breech, experts believe that the event-tree was vindicated. Even after multiple failures, the odds were still against a destructive outcome.

At Chernobyl, several major safety systems were turned off to run an experiment on the system. When applying the event tree, we find that most of the paths leading to safe resolutions were systematically turned off. The probabilities were significantly altered by the conscious shut down of the plant's systems, as can be seen in the modified event tree, Figure 5. Such unnatural operation of the plant does not mean the tree failed as a predictor of reactor safety. The Chernobyl event-tree does not have much bearing on the predictive value of a normally operated reactor. On the contrary, it took a concerted effort to destroy the reactor. One analogy that could be drawn involves driving your car through a red light to decide if the car is safe. Such non-standard operation of the vehicle would not invalidate previous safety studies. Instead the case would be dismissed as a bad piece of data.

•**Limitations**

Such an analysis may inadvertently simplify the operation of complex systems. In complex systems such as aircraft and nuclear power facilities, the interaction of the various systems may create inadvertent side effects. As a result, a given safety system may not work independently of the system. In such a case, the event-tree evolves into a twisted, and incomprehensible web, Figure 6. If a system fails, the environment created might make it more likely for subsequent systems to fail. In any case, accurate predictions of performance can become impossible to compute.

This interactiveness is inherent in immature technical systems. Such interactions between non-linear systems are anticipated, to a degree. That is why different sections in a ship are compartmentalized. Nonetheless, all interactions cannot be anticipated. A fire on a ship which disables its electrical systems, could cripple the crew’s ability to respond by disabling the communications systems as well. So emergency responses would be unable to be implemented because of an unforeseen connection between the initial accident and the ship’s emergency response procedures, and the ships’s electrical and communication systems. Only with more experience with complex systems and en eye toward increased effectiveness can this interaction be simplified. (Perrow, 1984)

Diagram of Complex Interactions

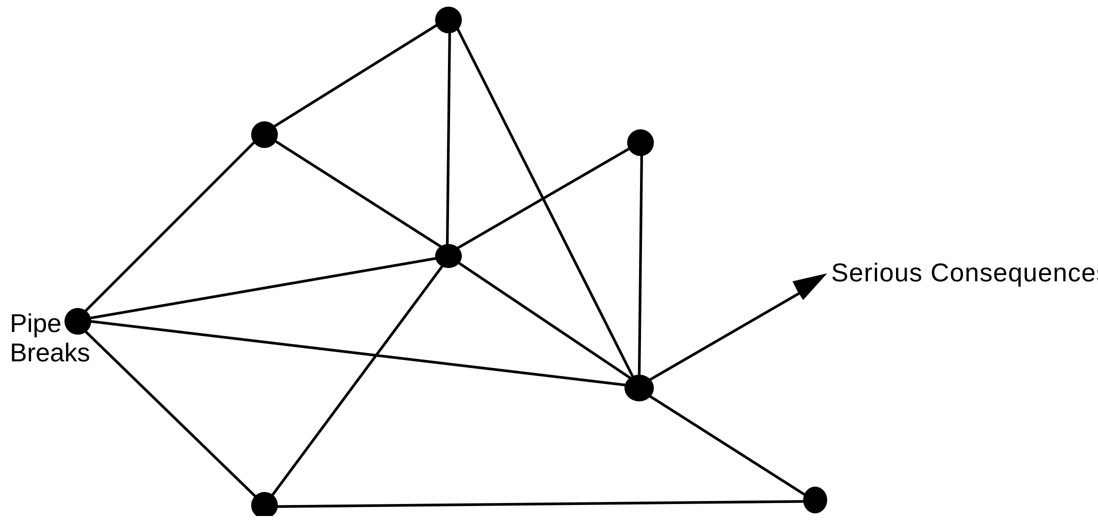

Figure 6

The difficulties of complex systems also apply to human operators. As systems begin to fail, the operators must handle increasingly complex counter-measures. The likelihood of further mistakes increases because more is being demanded of the operators. Poor human interface in a system can compound such failure.

Ironically, the solution to poor human interface introduces a new problem to the system. By automating reactions, and therefore making the operator work from a higher level position, the flexibility of the system is limited. The operator is no longer able to leave the high-level control environment to fix low-level difficulties. (Perrow, 1984) For example, a car with an automatic transmission gives the driver less direct control over the automobile than one with a standard transmission.

Another failing is the inability of the model to predict multiple failures caused by a single incident. Many aircraft have redundant controls, but these controls can be centralized in one part of the airplane's structure. An explosion or collision at one of these key points could eliminate many systems, and again the chances of failure are no longer independent.

Many important technologies, like nuclear power, suffer from an additional problem. Precision is especially difficult to achieve when evaluating the low-probability high-consequence events associated with these technologies. Obtaining objective data for these events is rare.

Therefore subjective methods are used including: “expert opinion, extrapolated data, deductive analysis, etc...” (Martz & Bryson, 1984) Further, this reliance on “expert opinions” clouds the issue, because of “frequent disagreements within the expert community about interpreting and evaluating these data.” (Regens, Dietz, & Rycroft, 1987) Since contemporary issues affecting risk assessment involve low-probability high-consequence systems, finding a way to resolve these conflicts becomes more critical.

Even if the experts can agree on how likely it is that the separate components of a system will fail, they may still disagree on the safety of the system as a whole. In reference to the kind of basic model used, experts come to different conclusions based on different assumptions about the function and interaction of the system. With updated technology, this trend becomes even more prevalent. Because of a lack of hard data, scientists are forced to rely more on secondary sources of information, making a consensus difficult to obtain.

•Regulatory Models

Unfortunately, We often need to regulate an activity before we can come to any agreement about the safety of that activity. For example, Rieth and Starr (1984) write, “It should be recognized that U.S. regulatory agencies are often in the undesirable position of having been mandated to perform carcinogen risk assessments for chemicals most of which do not have a adequate database for the task.” The problem of regulating risk has been brought to the front page by the controversy over dioxin.

Dioxin, a byproduct of pesticide production, was labeled the most dangerous substance known to man. In animal studies, it was found to be highly carcinogenic. However, there were limited direct data suggesting a threat to humans. As a matter of public policy, the Environmental Protection Agency (EPA) set very strict guidelines on the allowable exposure levels for the chemical. In some cases, it was more strict than European and Canadian standards by factors of 70 to 1700. (Roberts, 1991)

Their argument was that without a better model, a worst case model ought to be assumed. This linearized multistage (LMS) model is designed to predict a “plausible upper limit,” but it “does not necessarily give a realistic prediction of risk.” (Dowd, 1988) Others criticize the model for failing to consider the mechanism for action. In the case of dioxin, the LMS model does not account for how dioxin causes cancer, and in addition, the dose range for a given substance is very narrow. Consequently, the resulting regulations are usually very strict. (Rieth & Starr, 1984) Because of the basic biases built into the methodology of the EPA, they will almost certainly estimate that a substance has a relatively high level of risk.

The model also fails to incorporate a realistic threshold of risk. Any detectable amount of dioxin was considered too high. This policy originated in 1958 with Delaney clause in the Food, Drug and Cosmetic Act. This legislation banned any food additive found to be carcinogenic. It was thought that only a handful of substances would be affected. (Smith, 1990) With new technological developments, even lower levels of potentially harmful substances could be detected. As the government's regulatory restrictions are tightened, substances are increasingly restricted. These changes occur without any further evaluation by our policy makers, even when such changes are not justified. (Rulis, 1987) Aaron Wildavsky describes the situation colorfully; "How extraordinary! The richest, longest-lived, best-protected, most resourceful civilization, with the highest degree of insight into its own technology, is on its way to becoming one of the most frightened." (1979) Our ability to gain new insights into the world has done much to heighten our sense of fear. This ability has unquestionably affected our regulatory process as well.

•**Alachlor**

For dioxin, alternative pesticide and disposal techniques ensured that the affect of the restrictions would not severely impact society. George Hoberg Jr. studied the regulatory responses of the United States and Canada to the pesticide Alachlor. Agriculture Canada banned the substance, while the EPA allowed for its continued use. The divergent regulations resulted from the same basic scientific data, for both agencies agreed on the basic effects of the substance. (Hoberg, 1990)

In the United States Alachlor is one of the more widely used pesticides, leading the EPA to conclude that "Upper bound levels of this order of magnitude do not outweigh the substantial benefits of Alachlor use." In Canada, the substance was only the tenth most widely used pesticide of its kind. This relative lack of utility makes alternative solutions easier and more practical, thus making regulation more likely. (Hoberg, 1990) Even if both were found to cause the same amount of environmental damage, mopeds would be easier to bad than automobiles. Because they are used in much smaller numbers they would be much easier less expensive to replace.

What about substances for which there are no other available choices? Many treatments for Acquired Immune Deficiency Syndrome (AIDS) have been brought through the regulatory process with unheard of speed. For the patients there are few alternative treatments other than the experimental drug AZT. However, the "Upper bound levels of risk" compare favorably with the potential benefits because of the desperate situation of the disease's victims.

•**Status Quo Model**

Some have proposed other models for regulatory guidelines based upon previous risk acceptability levels. Curtis Travis and Holly Hatterner-Frey surveyed 132 federal exposure standards for various carcinogens. What they discovered was that the general lifetime exposure risk levels for small populations was 10^{-3}, and 10^{-4}[4] for larger populations. By relying on these established norms, questions of "acceptability" become much simpler by maintaining these standards. (Dwyer, 1989)

This model presents an ethical question of why larger groups ought to be afforded greater safety than smaller groups. The source of this disparity may simply come from greater political power. If that is the case, the government clearly has an obligation to protect the smaller groups from environmental hazards. At the same time, the model fails to consider the benefits resulting from these risks. A 10^{-3} risk for a vital part of the economy might compare well with a 10^{-4} risk to make our fruit a brighter color, regardless of the size of the affected population.

•**Toxicological Model**

Richard Dowd, the president of R.M. Dowd and Company, a scientific and environmental policy consulting firm, has suggested two possible alternatives. The first is a "traditional" toxicological approach. The other uses a "best estimate" model. (1988)

The traditional toxicological approach seeks to find a safe level for exposure. Unlike the

[4] One death per thousand and one death per ten thousand, respectively.

EPA's LMS[5] model, it assumes that there is a threshold level, under which the dangerous effects of a substance do not materialize. The toxicological approach determines from animal studies the levels under which there is no effect. These numbers are then extrapolated to human populations with a margin for uncertainty. "This is the least conservative model of the three[6] ". (Dowd, 1988)

To generally describe its application to the EPA's dioxin study, consider the test animals that did not develop cancer after dioxin exposure. The mice with the greatest dioxin exposure that did not develop the disease would be used to generate a threshold model. The threshold is the amount of exposure under which no deleterious effects occur.

•<u>Best Estimate Model</u>

The best estimate model is designed to approximate the actual risk to humans as accurately as possible. Part of this process involves an attempt to closely represent risk as a function, instead of seeking to bound the function. Secondly, the processes that make a substance risky need to be considered. A two-stage process could be used to determine not only the physiological affects of a substance, but also the processes involved. For example, is the chemical a promoter or an initiator? These specific hypotheses are more readily testable, eliminating much of the ambiguity that plagues the LMS model. At this time however, no regulatory agency has adopted either of these models. (Dowd, 1988)

---

[5] Linearized Multistage

[6] LMS, toxicological and best estimate models.

LMS verses Best Estimate

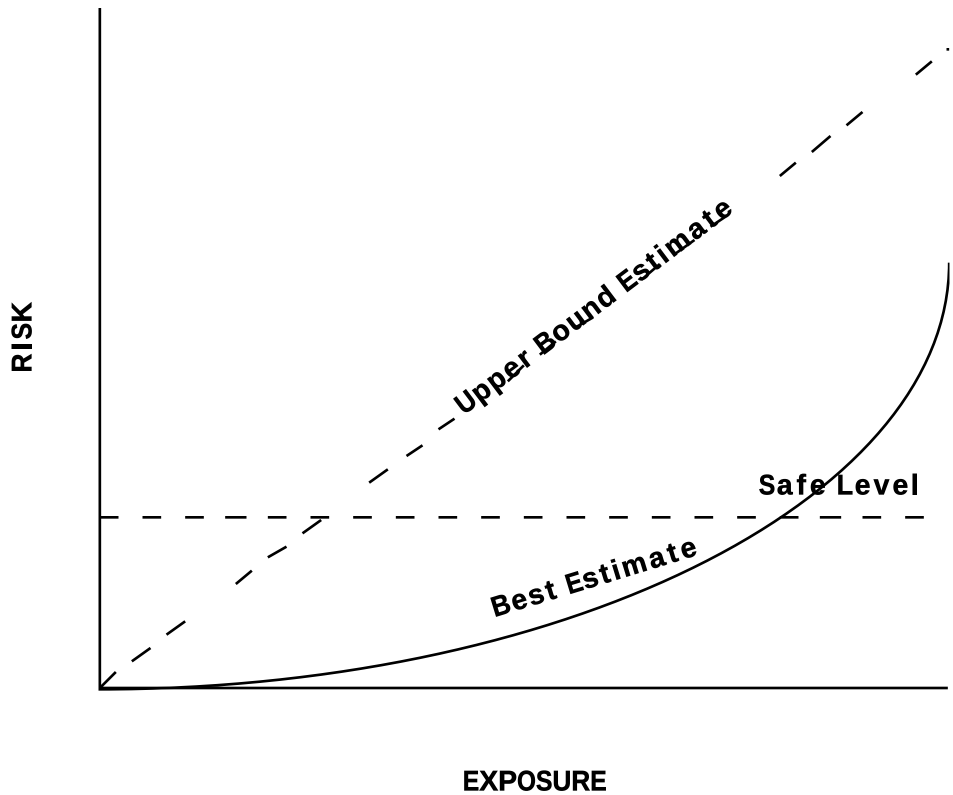

Figure 7

A comparison of the best estimate model and the LMS model is shown in the Figure 7. Notice how the LMS graph serves as an upper bound for the system. Also note where the two graphs intersect the safety level. The gap between these two intersection points is the difference between the two models' answers for how safe is safe enough.

The best estimate model can be used to incorporate complexity and coupling in its measurement. Because the model specifically looks at the processes in question, the underlying weakness or strength of a system can be revealed. If nuclear power plants were the subject of study, the tight coupling and dizzying complexity could be used in addition to the many statistics concerning general reactor safety. We could incorporate these ideas by testing specific hypotheses about how the systems components react with each other. These tools ought to create an accurate picture of a substance's or facility's safety.

As applied to dioxin, the LMS model establishes a safe exposure level of 0.006 pg/kg/day.[7] . The best estimate approach yields a safety level between .6 and 1.0 pg/kg/day, and the traditional toxicological approach indicates between 1.0 and 10.0 pg/kg/day as a safe level. (Dowd, 1988) Clearly, we can see how from the original set of data, the experts can draw three entirely different conclusions. This confusion carries itself into policy making processes too.

•Summary

There are many different ways to measure risk. Some models, such as the event-tree, are deceptively simple. They can generate false impressions of an activity's safety. Still, necessity requires us to make choices concerning risk before we have a solid foundation of information. Many assumptions are then made in an attempt to create a coherent policy. The result is chaos, as the experts even disagree on what is safe and what is not.

We find many different models available to scientists. Each model of measurement carries its own biases as well, and can lead different investigators to wildly divergent conclusions. There are no "right" models. Instead, we need to be keenly aware of each model's limitations in making policy decisions.

Risk Versus Perceived Risk

Perhaps the most obvious conflict in risk assessment is the one between science and the public. There seems to be a tremendous gap between our perception and our measurement of risk. After having looked at both, one difference is obvious: the two use separate criteria. Yet both have a valid claim in the policy making process.

•Priorities

[7] pg/kg/day: picograms per kilogram per day

The distinction between the expert's and the public's view of risk can be seen as we look at the issues that are important to both groups. Emergency chemical releases, hazardous wastes, point-source air and water pollution have received much notice. Events such as Bhopal, Love Canal, and the Cuyahoga River have drawn the public's attention. Currently the EPA has deemed radon, indoor pollution, and global warming to be the most important environmental issues. (Fiorino, 1990) The public is motivated by the rhetorical impact of these events, whereas the EPA is moved by the statistical indicators of danger. This disparity makes resource allocation difficult for public officials, because each group has an important and valid role in making policy choices.

•Relative and Absolute Risk

A problem that often occurs is how to compare risks of different magnitudes and frequencies. Is it better to have a risk that would kill a million people, but only had a one in a million chance of occurrence, or is it better to have a risk that has a near unity chance to kill one person?

Some argue that the public overestimates risks of low-probability high-consequence dangers, but that they underestimate risks of high-probability low- consequence events. (Zeckhauser & Viscusi, 1990) Should we discount the public as uninformed, or do they have a keener understanding of risk than is reflected by conventional scientific thought? Is it possible that the public uses a different, but still valid set of criteria?

Perhaps the public implicitly recognizes the differences between high probability-low cost and low probability-high cost risks. While a society could easily absorb the loss of a single life every year for a million years. The loss of a million lives once every million years would tear the culture's social fabric.

We often think of the risk of doing something, but what about the risk of not doing something? The perceived risks of omission are even more difficult to calculate. However, science can usually generate data for the consequences of omission. This discrepancy creates yet another disparity.

•**Arrogance**

There is a tendency by experts to “overvalue” their estimates. Studies have found that instead of 5% of true values falling outside expert-assessed nominal probability bounds, from 20% to 50% fell outside these bounds. (Martz & Bryson, 1984) To put this in the terms defined earlier, their estimates are precise, but inaccurate. write, “The existence of these systematic methods for estimating risks and for measuring public perceptions of them has perhaps led to over-optimistic expectations on the part of some members of the technical community about their application to policy decisions.” Otway and Simms (1987) This can lead to sharp differences within the public’s perceptions, which are intuitive and are formed by deeply held beliefs concerning institutions and values. (Fiorino, 1990) Thus, it might be wise to consider expert opinion with a grain of salt, based on past performance.

•**Summary**

Considering these two powerful visions of risk, we discover the source of their conflict. The public and science have different criteria from which they draw disparate conclusions. The priorities of the two groups are distinct and their assessment of various kinds of risks reflects divergent values. Lastly, the traits of confidence and skepticism, or arrogance and prudence lead to further divisions. The impasse is not easily resolved, for neither viewpoint is necessarily invalid. Instead, policy makers need to find a way to weigh these two important perspectives.

Conclusions and Recommendations

It is difficult to see through the muck to discern, “how safe is safe enough?”. We quickly realize that there is not one simple answer. As the Royal Society in a group study report notes, “there is not, and cannot be a single acceptable level of risk.” (Otway & Simms, 1987) Instead, there are many important considerations that must be weighed before a choice is ultimately made. The road chosen must separate the decisions of science from the decisions of policy. Although such a distinction will be fuzzy at best, it is only by making such a distinction that we can preserve the importance of both science and public opinion.

•<u>Subjectivity</u>

We must realize that the question of “how safe is safe enough?” is not an objective one, nor is its solution to be found exclusively in the realm of science. Instead “the myth of absolute scientific objectivity should be discarded and expert advice seen as a support in clarifying issues and evolving pragmatic solutions to problems rather than as a clear an unambiguous basis for decisions.” (Otway & Simms, 1987) In fact at one point, the National Academy of Science was commissioned to determine Recommended Maximum Contaminant Levels for a variety of substances, and they concluded that it could not be done. (Regens, Dietz & Rycroft, 1983)

What is needed is a role for science to be a provider of information, not judgments. In addition, the information that is provided needs to be presented in a manner in which its inherent biases are also revealed. It is not simply enough that some separation exists between policy and science. A working interface needs to be established as well, otherwise, the answers that are generated become warped.

•<u>Guidance</u>

The role of science should be that of a guide. Carl F. Cranor presents a model that allows science to take just such a role. Instead of withholding scientific judgment until a high level of certainty is reached or using an arbitrary regulatory model, he suggests that scientists should present their data with appropriate confidence intervals. This would allow political leaders to make the judgments of acceptable risks and leave the hands of science clean. (1990)

This model also provides an opportunity to account for a largely unstudied source of risk assessment error: the scientists. Evidence is accumulating that scientists can introduce their own biases as a group to the study of risk. (Freudenburg, 1988) By limiting the number of scientific assumptions, it would be possible to reveal this source of error. Scientists provide accurate information qualified by the level of precision. However, the level of precision is reported as well, according to Cranor’s model.

Such a model introduces values α and β. α measures the likelihood of a false positive error. β marks the probability of a false negative error. Both errors present pitfalls for regulators. A false negative could lead to lives being endangered by continued exposure to a harmful substance, while a false positive could destroy the fabric of a community or an economy. α, β, and 1- β represent measures for "risk of error" and "standards of proof." Cranor (1990) argues that rather than having science deciding what chance to take in making policy decision, the choice should be left to the politicians.

Figure 8 uses Cranor's model to present the findings in the case of dioxin.

α and β Error Model

	Null hypothesis is actually true **Dioxin is associated with cancer**	**Null hypothesis is actually false** **Dioxin isn't associated with cancer**
Null hypothesis is accepted	No error $1-\alpha$	Type 2 Error β
Null hypothesis is rejected	Type 1 Error α	No Error $1-\beta$

Figure 8

If scientists seek an extremely high level of precision, they are likely to be inaccurate. Conversely, accurate hypotheses are less likely to be precise. This model allows science to escape the paradox: the issues that most need risk estimates are the ones about which we know the least. Under the public's eye, policy makers can make the decisions once they've been provided with this basic information.

It is at this point that the concerns and considerations of the public can be met. Science is not capable of being a representative of the public, but allowing science to circumvent the representative process would be elitist. Therefore, it falls upon the shoulders of the government to carry out this task. In essence, this model reaffirms the principle of representative government. These elected leaders are the ones equipped to evaluate and respond to the public's perceptions of risk.

Cranor's model provides a mechanism to resolve conflicts surrounding our views of risk, in an attempt to answer "how safe is safe enough?". Conflicts between different groups' perceptions of risk can be resolved just as other policy disputes are. Disparities amongst different scientific models can be weighed, by using Cranor's model. The resolution of the conflict between the measurement and perception of risk then takes on a new light. The sharp contrast between the different kinds of choices: cold and objective science and hot and impassioned public opinion, is smoothed. The myths that science is completely objective, and that public opinion is uninformed or irrational, are both shattered. By overcoming these two obstacles, it is then possible for policy makers to view the two in comparable terms.

The answers to the question "how safe is safe enough?" are then complete. These resolutions reflect the many different interests of the science, the public and society as a whole. Although perfect solutions can never be found, our answers offer us the best hope to avoid our own "low-probability tigers."

Reference List

Clark, William C. (1977) Managing the unknown. Kates, Robert W. (Ed.) Managing technological hazard:Research needs and opportunities. Boulder CO:Institute of Behavioral Science, pp. 111-127.

Conway, J., (1991) Scud's impact?. The New York Times vol.141 p. A15

Cranor, C. F., (1990) Some moral issues in risk assessment. Ethics, vol. 101, 123-143

Crouch, E. A. C., & Wilson, R., (1982) Risk/Benefit Analysis (pp. 51-74)

Dowd, R. M., (1988) EPA revisits dioxin risk. Environmental Science and Technology vol. 22, p. 373

Dwyer, J. P. (1989) Coming to terms with acceptable risk. Environmental Science and Technology vol. 23, pp. 145-146.

Ford, Daniel F. (1981) Three mile island: Thirty minutes to meltdown. New York:Penguin Books.

Fiorino, Daniel J. (1990) Can problems shape priorities? The case of risk-based environmental planning. Public Administration Review vol. 50 pp. 82-90.

Freudenburg, W. R., (1988) Perceived risk, real risk: Social science and the art of probabilistic risk assessment. Science vol. 242 pp. 44-49.

Halliday, D., & Resnick, R., (1988) Measurement. J. Merrill (Ed.) Fundamentals of Physics (pp.1-8). New York: John Wiley & Sons.

Hoberg, G. Jr., (1990). Risk science and politics: Alachlor regulation in the United States and Canada. Canadian Journal of Political Science, vol. 23 257-259.

Martz, H. F., and Bryson, M. C., Predicting low-probability/high consequence events. Walker, Ray A., and Covello V. T. (Eds.) Low Probability High Consequence Risk Analysis: Issues, Methods and Case Studies, New York:Plenum Press pp. 188-189.

Mendenhall, W., Wackerly, D. D., & Scheaffer, R.L., (1990). M. Payne (Ed.) Mathematical Statistics with Applications . Boston:PWS-Kent.

Otway, Harry, and Simms, D. L., (1987) Criteria for technological choice. Public Administration, vol. 65 pp. 131-143.

Perrow, C., (1984) Normal Accidents: Living with High-Risk Technologies, New York:Basic Books Inc. pp. 72-100

Regens, James L., Dietz, Thomas M., Rycroft, Robert W. (1983) Risk assessment in the policy-making process: environmental health and safety protection. Public Administration Review, vol. 43 pp. 137-145.

Rieth, Joseph P., and Starr, Thomas B. (1987) Upper-bound estimates of carcinogenic risk: More policy than science. Bonin, James J and Stevenson, Donald E. (Eds.) Risk Assessment in Setting National Priorities. New York: Plenum Press, pp. 223-229 .

Roberts, L., (1991) Dioxin risk revisited Science vol. 251 pp. 624-626

Rulis, Alan M. (1987) Establishing a threshold of regulation. Bonin, James J and Stevenson, Donald E. (Eds.) Risk Assessment in Setting National Priorities. New York: Plenum Press pp. 223-229.

Smith, Bruce L.R. (1990) American science policy since wwii Washington:The Brookings Institution. pp. 62-65.

Starr, C., (1969) Social benefit versus technological risk. Chalk, R. (Ed.) Science, Technology, and Society: Emerging Relationships Washington: American Association for the Advancement of Science. pp.166-172

Wildavsky, Aaron (1979) No risk is the highest risk of all. American Scientist vol. 67 pp. 32-35.

Wyatt, Gary (1989) Risk-taking and risk-avoiding behavior: the impact of some dispositional and situational variables. The Journal of Psychology, vol. 124 pp. 437-447.

Zeckhauser, R. J., & Viscusi, W. K. (1990) Risk within reason. Science, vol. 248 pp. 561-563.

