DRAFT

IMPROVING ACCURACY AND REDUCING COSTS OF ENVIRONMENTAL BENEFIT ASSESSMENTS

Risk Communication for Superfund Sites An Analysis of Problems and Objectives

by

Gary H. McClelland
William D. Schulze
Don L. Coursey
Brian Hurd
Julie R. Irwin
Rebecca R. Boyce

USEPA COOPERATIVE AGREEMENT #CR812054-02

September, 1987

PROJECT OFFICER

Dr. Alan Carlin
Office of Policy, Planning and Evaluation
U.S. Environmental Protection Agency
Washington, D.C. 20460

DISCLAIMER

Although prepared with EPA funding, this report has neither been reviewed nor approved by the U.S. Environmental Protection Agency for publication as an EPA report. The contents do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

TABLE OF CONTENTS

RISK COMMUNICATION FOR SUPERFUND SITES:
AN ANALYSIS OF PROBLEMS AND
OBJECTIVES

Page

CHAPTER 4: A FIELD STUDY OF FACTORS AFFECTING RISK BELIEFS

CHAPTER 5: RISK COMMUNICATION GUIDELINES

RISK COMMUNICATION FOR SUPERFUND SITES: AN ANALYSIS OF PROBLEMS AND OBJECTIVES

CHAPTER 1: INTRODUCTION AND OVERVIEW

1.1 The Problem

The Environmental Protection Agency has an extraordinarily difficult task in managing the Superfund Program. Although citizens are very concerned about Superfund sites, scientific estimates of the risks from most sites indicate that the hazards are very small. Thus, EPA faces public demands for extensive and expensive clean-up for many sites when scientific risk assessment fails to justify such efforts. Given these circumstances as well as the real fears of residents living near Superfund sites, the hypothesis underlying most research on risk communication has been that people just do not understand the "true" risks and that good communication of these risks will in and of itself solve or reduce the problem defined above. In other words, good risk communication is just good communication. This viewpoint, which is not supported in the research reported here, comes in great part from experiences with natural hazards such as tornadoes, hurricanes and floods where risk communication programs have shown considerable success. Sorensen and Mileti (1987), for example, argue that the social psychological model developed for risk communication about natural hazards provides many insights into good risk communication. We show in the research presented here that the argument that risk communication is just good communication is likely false for low probability risks such as those present at most Superfund sites. In fact, when the standard risk communication strategies from natural hazards research (which work well at high probabilities as Sorensen and Mileti suggest) were applied in a low probability situation, that of possible volcanic activity near the Mammoth Ski area in California, an utter fiasco resulted. Property values fell substantially after a warning was issued, even though the odds of a Mt. St. Helens type event were minute. Further, these odds were clearly described to residents and the news media (See Thayer et al., 1986). We have previously documented a similar substantial fall in property values near a Superfund site where scientific assessments have shown no evidence of significant risk (Schulze, McClelland and Hurd, 1986). Thus, both Superfund sites and natural hazards with low probabilities of harmful consequences have been shown to have the same potential for inappropriate levels of concern among populations exposed to or hearing about the risk. Consequently we agree with Slovic's (1986) assessment that, for low probabilities, we know almost nothing about risk communication.

The primary purpose of the research which we present here has been to attempt to determine what goes wrong with risk communication at low probabilities. The research initially attempts to answer two related questions. First, does something go wrong in the way people think about low probability hazards? Second, can it be conclusively shown that the individual and community response to low probabilities is inappropriate? The answers we provide to these two questions in turn raise some very serious policy issues for EPA with respect to risk communication at Superfund sites (which we discuss later in this chapter). Since we find that people do fail to behave appropriately at low probabilities <u>no matter how well they understand the risk,</u> we then analyze how factors that affect risk judgments may be employed as part of a risk communication strategy to help people better judge the risk from Superfund sites. That strategy in our view must incorporate an understanding of the inherent inability of the public to respond appropriately to even the clearest most understandable statements concerning low probability risks.

1.2 Summary and Implications of the Research

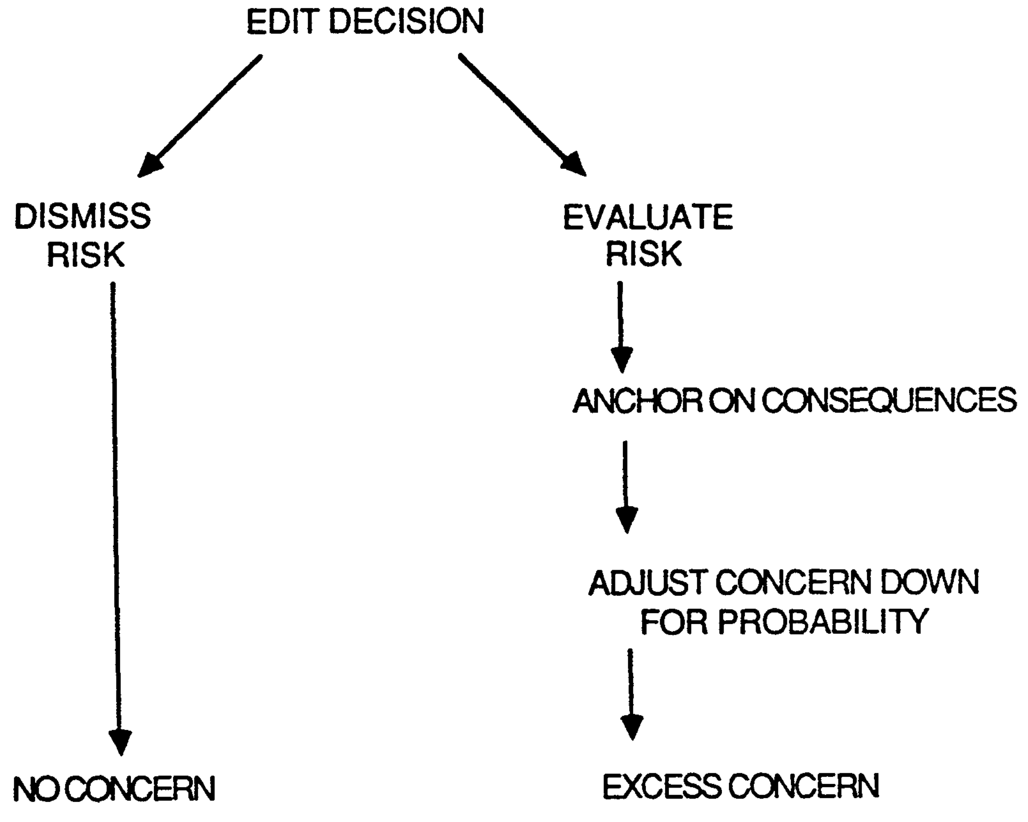
This section summarizes two research efforts used to develop specific suggestions for risk communication at Superfund sites. The first effort sought to expose people to a series of situations where the risks were made absolutely clear so the response to perfect risk communication could be examined. The probability was experimentally lowered to see what difference low as opposed to high probabilities made in the response and to gather data in a simple situation to help understand the source of any problems which might appear at low probabilities. The response chosen to measure concern for the risk in the experiment was to find out how much people would pay for insurance against a risk of loss. The loss was a fixed dollar amount. The probability of loss was obtained by putting red and white poker chips in a bag where subjects were first shown the chips and repeatedly told the number of each color before they saw them placed in the bag. Thus, people knew the exact dollar loss and exact number of red and white chips in the bag. They were then told if a red chip were drawn they would lose the specified amount unless they bought insurance which would prevent the loss if a red chip were drawn. Thus, they could buy their way out of being exposed to the risk, and their willingness to pay was a clear measure of their concern. The analogy to a hazardous waste site is as follows: the fraction of red chips in the bag corresponds to the odds of cancer or other health problems from the site, while the specified dollar loss corresponds to the consequences of cancer or other health problems if they occurred. The point is not that a monetary loss is like cancer or that drawing chips is actually like the risk of getting cancer; but rather, if the mental process generating concern for drawing a red chip fails to yield appropriate behavior at low probabilities in the very simple experimental situation, in the more complex and emotional real world, behavior will be much worse.

In our experiments, the value that people should place on insurance, at least to a first approximation, is the probability times the loss.[1] This amount is called the expected value of the loss. However, based on debriefing of subjects, none of them used this procedure to arrive at the value of insurance. Rather, as Hammond, et al., (1987) have noted, when people do not have access to an analytical mental process (calculation of expected value) to decide how concerned they should be, they use an intuitive mental process instead. In our experiments this intuitive process led to the following results. At higher probabilities of loss, such as four in ten, people intuitively valued and paid expected value for insurance. In other words, at higher probabilities the intuitive mental process yielded about the right level of concern for the loss. However, as we dropped the probability below one in ten the response began to split into two types. Either (1) the level of concern did not fall enough as the probability fell and people paid too much for insurance (many times expected value) or (2) people showed no concern at all, dismissed the risk, and bid zero. Figure 1.1 shows a model of how people judge how concerned to be at low probabilities based both on our debriefing of

[1]We demonstrate that risk aversion plays no role in our experiments in Chapter 3. However, in the real world situation of a Superfund site, risk aversion or risk seeking behavior may play a significant role. Risk preferences, however, have nothing to do with the cognitive errors which occur at low probabilities. These errors are not only clearly demonstrated in the research, described here but have been repeatedly demonstrated by psychologists. See for example research by Kahneman and Tversky, 1979, 1981, 1984, Lichtenstein, et al. 1978, and Combs and Slovic 1979.

FIGURE 1.1

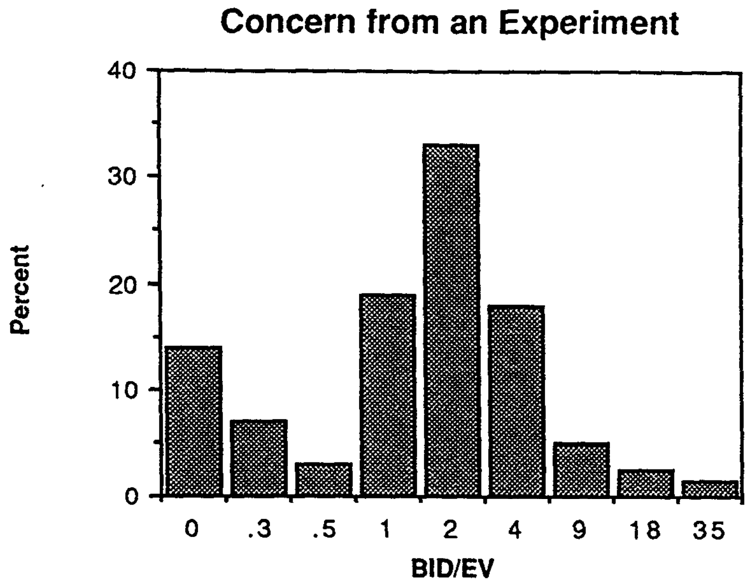
A MODEL OF RISK JUDGMENT

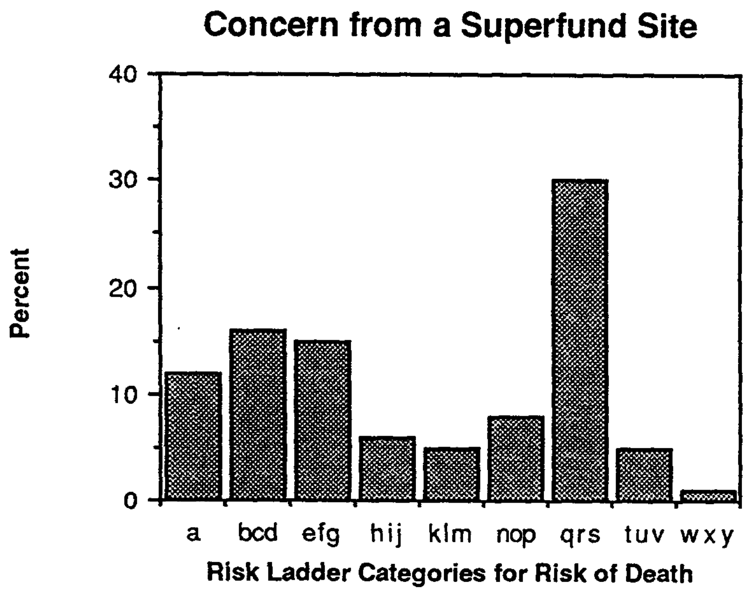


subjects and on the responses described above. Apparently when faced with a low probability people first try to decide if the risk is worth worrying about at all. People face innumerable small risks and evaluating all of them would be impossible. Thus, they first decide whether to evaluate the risk or just to dismiss it. If they edit the risk and dismiss it then they act as if they show no concern. On the other hand if they decide to evaluate the risk they then appear to go through the following process: first, they think about or anchor on the loss event. In the case of a Superfund site the loss event is likely to be cancer, a birth defect or other illness or disease possibly leading to death. In our experiments the consequence is just the loss of a sum of money. In either case people then take this level of concern, obviously large for a Superfund site, and attempt to adjust their concern downward to account for the fact that the consequence is not a certainty. People start out by thinking "wouldn't it be terrible if the site gave me or my family cancer"; and then think, "but maybe it won't, so I don't have to be quite so worried." Unfortunately psychologists have repeatedly demonstrated that when this intuitive anchoring and adjustment thought process is used, the adjustment almost always falls short. In other words, for low probability risks people do not adjust down enough and end up with excess concern. In the case of Superfund sites, since the potential consequences are so bad and the probability is so low, a lot of downward adjustment in concern is necessary. By falling short in the adjustment process people end up much too worried.

One way of showing the response to a low probability risk is to plot the frequency distribution of the amount of concern generated. The upper panel in Figure 1.2 plots the percent of subjects in an experiment offering to pay different amounts for insurance. The amounts shown on the horizontal axis are expressed as amount bid for insurance divided by expected value. Since the

FIGURE 1.2





Annual Risk of Death:
a = no risk
b = one in 9 million
f = one in 100 thousand
j = one in 10 thousand
n = one in one thousand
r = one in one hundred
v = one in ten

loss was $40 and the odds of loss were 1/100 in this experiment, expected value was 1/100 x $40 = $.40. Thus, people should all offer to pay about $.40 for insurance (shown as unity along the horizontal axis). However, the most frequent bid was 2 1/2 times expected value or $1.00 (which falls in the 2 BID/EV category). The second most frequent bid was zero. Note also that the horizontal axis is logarithmic (with the exception of the separate category for zero bids) to compress the scale since some bids were very large. The lower panel in Figure 1.2 shows the frequency distribution of concern for residents around a Superfund site in Monterey Park, California. The vertical axis again measures frequency as a percent of the population while the horizontal axis measures risk beliefs from a risk ladder, again on a logarithmic scale with a separate category for zero risk. The similarity between the two distributions is striking, showing a lower mode of individuals who edit the risk and show little concern, and an upper mode of people who <u>overestimate</u> the appropriate level of concern.

This last statement, that some people overestimate low probability risks, can only be justified on the basis of the experiment, not on the basis of the risk beliefs taken from the hazardous waste site. Who is to say that the upper mode group who believe that the Superfund site is very dangerous will not some day be proven correct, even if scientific assessments now show no risk? In contrast, under the controlled circumstances of the experiment, the true risk is known to all, subjects and experimenter alike. In this situation the fact that subjects behave inappropriately can be clearly demonstrated.

It is interesting to note just how badly subjects performed. We have run subjects at 1/100 odds of a $40 loss through more than 100 trials. Thus, a chip was drawn from a bag containing 99 white and one red chip more than 100 successive times. As noted above, subjects should bid $.40 for insurance on each trial. The logic of this approach is as follows: over 100 draws, one would on average expect one red chip to be drawn for one expected loss of $40. If insurance were purchased by a subject in every trial for $.40, $40 would be spent to avoid an expected $40 loss. In fact, most people in the experiment either obtained insurance for around $1.00 on each trial or bid $.00 and did not get insurance. People who got insurance all the time paid about $100 over one hundred trials to avoid one expected loss over that period of $40. People who always bid zero and never got insurance over one hundred trials typically suffered one loss of $40 but paid nothing for insurance and came out far ahead of those who bought insurance. In fact, over 100 plus trials, those who usually bought insurance often went broke.[2] These subjects left the experiment bewildered, realizing they had done something wrong, but not knowing what. They thought they were "playing it safe" by buying insurance to protect the initial money they were given but were in fact wasting their money and their concern on avoiding a low probability event which did not merit their attention. Therefore it can be seen that, for low probabilities, editing is better strategy when compared to overestimating the appropriate level of concern. This suggests that the actual goal of risk communication in analogous situations in the real world should be to encourage people to edit the risk.

People faced with such risks, in deciding what to do, can be thought of as trying to come up with the appropriate level of concern (expected value) using a defective electronic calculator. They punch in the loss and the numerically small probability. However, the calculator, rather than multiplying the two numbers, either comes back with an answer of zero or an answer which is much too high relative to expected value. At high probabilities (above .1) the calculator works much better, yielding numbers fairly close to expected values.

[2] These experiments used real money supplied by the Council on Research and Creative Work at the University of Colorado-Boulder.

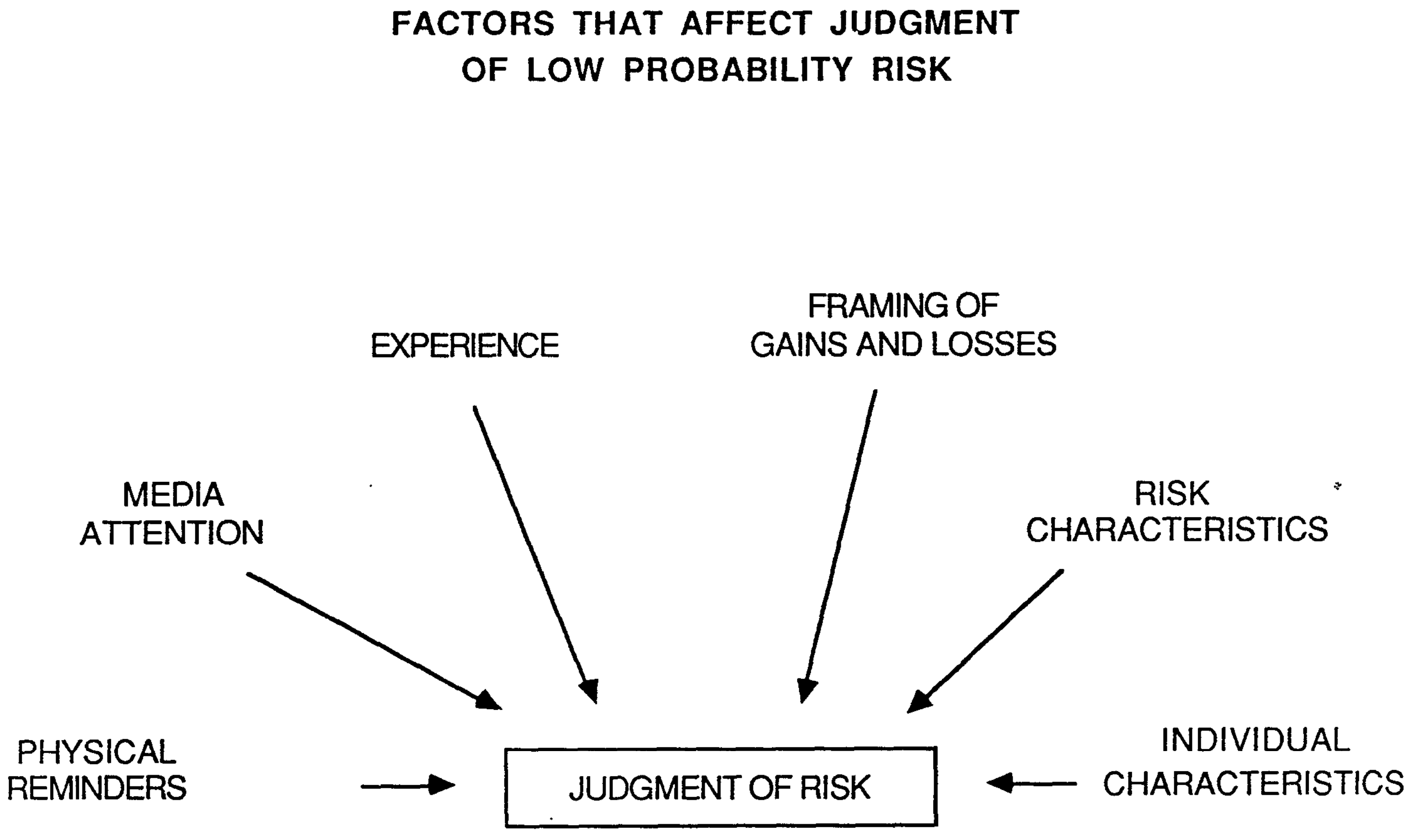
The fact that people are in some sense “lost” at low probabilities was also demonstrated in our experiments. Over many trials the concern people placed on the risk did not remain constant as it should have, since expected value was constant over trials. Rather, when long runs of white draws occurred (inevitable with 99 white and 1 red chip) the level of concern drifted downward. In contrast, after a red chip was drawn, concern would drift upward. Thus, these inappropriate cues influenced the level of concern. Slovic (1986) has extensively analyzed the real world factors which tend, often inappropriately to bias the level of concern. The remainder of this section describes the implications of such influences. Many of these factors, shown in Figure 1.3, play an important role in developing an appropriate risk communication strategy.

Based on our study of a Superfund site (summarized in Chapter 4), the factors shown in Figure 1.3 work to influence risk beliefs as follows:

Physical Reminders. In the absence of constant reminders people will tend to forget about, or edit, a risk. In contrast, tall fences around a Superfund site, warning signs, the noise of truck traffic associated with clean up of a site, views of workers wearing “space suits,” and odors or smoke from a site all remind people of the existence of risk and tend to cause excess concern. These reminders should be consistent with public safety but should not be exaggerated by a lack of awareness of how the pubic might misinterpret them.

Media Attention. Recent studies have shown that the media contribute substantially to excess concern (see Chapter 5 as well as Wilkins, 1987). Media attention provides reminders of risk when none may actually be present. The media, concerned with ratings or circulation, have an incentive to sensationalize issues especially where public controversy exists. Superfund sites will produce controversy between the no concern

FIGURE 1.3


FACTORS THAT AFFECT JUDGMENT
OF LOW PROBABILITY RISK
EXPERIENCE
FRAMING OF
GAINS AND LOSSES
MEDIA
ATTENTION
RISK
CHARACTERISTICS
PHYSICAL
REMINDERS
JUDGMENT OF RISK
INDIVIDUAL
CHARACTERISTICS

group and the group with excess concern described above. This controversy will attract media attention which provides reminders of the risk which are inappropriate cues reinforcing the beliefs of the group with excess concern. The media plays a central role in this process which has been termed social amplification. There is little we think EPA can do in this situation other than to avoid contributing to the amplification process by avoiding communicating risks in a way which leads to exaggerated concern.

Experience. Individual experience will strongly influence risk beliefs. If a neighbor has recently died of cancer, that death may well be falsely attributed to a nearby Superfund site by a worried family. If people have been aware of a risk for a very long time, i.e., have a lot of experience with the risk, they tend to adapt to the risk, view it as part of the status quo, and edit the risk. Thus, old risks such as coal fired power plants, garden chemicals, and driving are often edited while a new risk such as that from a Superfund site (which used to be viewed as relatively harmless) creates much concern.

Framing of Gains and Losses. People are much more concerned about losses than they are concerned about gains relative to the status quo. Thus, reducing an old risk (a gain relative to the status quo) is not very important to most people but finding out you have been exposed to a new risk (a loss from the status quo) gets people very concerned. Some people may become accustomed to a Superfund site and oppose clean up efforts because any new risk associated with clean up (a loss) increases concern more than removing the old risk from the site (a gain) decreases concern. Risk communication efforts should carefully consider how proposed actions will be stated and interpreted in terms of losses and gains.

Risk Characteristics. Some risks have characteristics which raise concern. For example, people tend to be more concerned if many people are killed at one time. Some modes of death such as cancer may be more feared than others such as accidental death. People may imagine that a Superfund site might explode, spreading a cloud of toxic gas, killing many people immediately, while leaving others to die years hence from cancer. Risk communication about the nature of consequences might do much to allay such fears.

Individual Characteristics. Families with children, younger people, and women all tend to be more fearful of Superfund sites. Education, income level and occupation seem to have little or no impact on risk beliefs. This individual information can be used to help predict where concerns are likely to appear over Superfund sites and over cleanup activities.

Many of the factors discussed above play a role in generating excess concern around Superfund sites. Unfortunately the role of these factors, which have been explored most extensively by Slovic, have been often misunderstood, as has much of Slovic’s research. Slovic is not arguing that these factors are *appropriate* determinants of risk beliefs. Rather, he has demonstrated that a rather long list of mostly irrelevant factors influences the level of concern shown for low probability risks. These irrelevant factors become influential because, as we have argued above, people are lost at low probabilities. Thus, since they have no analytical mental process to rely on, they substitute a very faulty intuitive process which is impacted by the factors shown in Figure 1.3.

Unfortunately, many non-psychologists writing in the area of risk communication, not realizing that the above factors lead to faulty judgments, have incorporated them into suggestions for risk communication. For example, dramatic illustrations which emphasize physical reminders may seem to constitute the elements of good communication, but instead can serve to unnecessarily frighten people. Much current advice on risk communication is ill advised for Superfund sites because the problem in the past has been to get people facing a high probability risk (from tornadoes, smoking, or driving a car) to stop editing the risk and do something (take shelter in a storm cellar, stop smoking, or buckle up, respectively). The problem at Superfund sites is just the opposite, to get the people who are overly concerned to edit the risk and not do anything (do not sell your home at a loss, do not be afraid, etc.).

1.3 Alternative Approaches for Risk Communication.

Given the current understanding of factors which affect risk beliefs, what strategies for dealing with the mismatch between the public's subjective beliefs and scientific estimates of risk at Superfund sites are available to EPA? Purely for purposes of defining issues and the range of options we briefly evaluate four alternative approaches: (1) Benign neglect, (2) Aggressive risk communication, (3) Conflict resolution, and (4) Complete site cleanup. These alternatives imply (moving from 1 to 4) increasingly expensive cleanup operations.

Benign Neglect. The discussion of the preceding section suggests that Superfund sites, if left completely alone, will eventually become viewed as part of the status quo just like risks from coal fired powerplants, automobiles and garden chemicals. As the new risks from Superfund sites become old risks, concern will subside. Obviously however, for the very reasons outlined above, public concern is now very intense at many sites and would likely take a very long time to subside under this scenario. Beyond the legal and political infeasibility of this approach, it is also bad policy in that people who believe Superfund sites pose great risk are actually suffering. The concern they express is genuine as demonstrated by large drops in property values. These concerns are not feined or insubstantial in spite of our arguments that they result from cognitive problems. Since economic damages are real (as expressed in property values) benefit cost analysis also suggests that action should be taken and that benign neglect is an inappropriate strategy.

Aggressive Risk Communication. Since the intense concern for Superfund sites is of ten inappropriate, one form of action would be to vigorously to manipulate all of the known factors which affect risk beliefs to attempt to decrease that concern and consequently avoid "unnecessary" expenditures for cleaning up sites which generate little actual risk. Two obvious problems are associated with such an approach. First, we do not know enough about risk communication to pursue such a policy successfully. Second, such a manipulative policy is offensive in a democratic society and smacks of policies pursued in closed societies. EPA has an obligation to provide accurate information to the public and, aside from ethical considerations, heavy handed attempts at minimizing risks will inevitably backfire, causing an even more intense over-response from the public.

Conflict Resolution. Since the pubic response to a problem Superfund site will be excess concern among one group of citizens and little or no concern among another group of citizens, conflict will inevitably arise over the appropriate course of action to take in cleaning up the site. This conflict provides the best opportunity for risk communication to provide scientific evidence to both sides in an open public debate. The natural tendency in a formal conflict resolution process (fostered by EPA) will be to reach a compromise. The resulting cleanup effort will likely be substantially less than the very concerned group would have desired. However, the concerned group will recognize, if the process is a fair one, that compromise was necessary to reach a solution given the diversity of public views. EPA should not take sides in this process (avoiding a no win situation) but rather provide technical information through a group of neutral scientists picked by the community to advise them in the conflict. What is critical in this process is that Superfund personnel avoid making the situation worse in the way the site is managed, the way cleanup operations are presented, etc. Chapter 5, which provides our guidelines for risk communication, follows this strategy.

Complete Site Cleanup. In the long run, fairly complete cleanup of Superfund sites, without much consideration of scientific assessments of the risks, in response to public pressure will solve the problem. However, this solution is not without severe disadvantages. First, it is unlikely that enough money could ever become available to satisfy concerned citizens near all Superfund sites. Second, it would obviously be preferable to use risk communication to allow more money to be spent on sites which have higher scientific estimates of risk. Third, cleanup activities themselves tend to frighten the public. The prospect, for example, of burning hazardous wastes on site will be regarded as a new risk and a loss from the current status quo. Both of these characteristics will generate intense concern. Thus, in the absence of a strategy such as conflict resolution cleanup efforts themselves may exacerbate problems.

It is our view that the only practical alternative available to EPA is to apply risk communication in the context of conflict resolution. In many respects, current EPA policy is consistent with this approach. However, it is our view that more effort should be devoted to conflict resolution within the. Superfund program. Current procedures do not provide an EPA sponsored framework for a formal process which includes, for example, the appointment of a community selected panel of scientists to review risks. An example of how risk communication blends with conflict resolution is described in a paper by Hammond, et al. (1976) which is included as the Appendix to this report.

In summary, we are very pessimistic about prospects for explicit risk communication at Superfund sites. If EPA inherits a site where people are not already concerned, attempts at risk communication may well provide inappropriate reminders and cause some people to move from an edit decision to become overly concerned. If EPA inherits a site where many people are already very concerned, attempts at risk communication can easily reinforce those concerns and falsely convince those who currently ignore the site to evaluate the risk. Since most people (experts included) when evaluating a low probability risk tend intuitively to overestimate how concerned to be, they are better off ignoring the risk. Implicit risk communication through management of perceptual cues such as physical reminders may, however, be a fruitful approach along with conflict resolution. Unfortunately most of what we know about explicit risk communication has come from attempts to warn people about important hazards which they are ignoring and so is inappropriate for Superfund sites. Consequently much of the advice on risk communication which EPA has received for these sites is, in our view, incorrect.

1.4 The Organization of the Report.

Readers who do not wish to examine the technical aspects of the research reported here may skip directly to Chapter 5, "Risk Communication Guidelines," which contains our specific recommendations for risk communication procedures at Superfund sites. Chapter 2 provides a technical synthesis of what is known about risk communication and the formation of low probability risk beliefs based both on our own research and the research of others. Chapter 3 presents the details of the core experiments described above. Chapter 4 summarizes our prior study of how risk beliefs were formed at a Superfund site.

CHAPTER 2: A TECHNICAL SYNTHESIS

In this chapter we review what is known about risk beliefs and risk communication from the research literature and especially from our own prior research in this field. This review provides necessary background for our recommendations which are presented in Chapter 5. Based both on the literature and our research we develop a model of risk judgments that leads directly to those recommendations.

2.1 Low Probabilities.

It is quite clear that it is low probability events that are especially problematic for risk judgment and communication. For example, Kahneman and Tversky's research (1979, 1984, and Tversky and Kahneman, 1981) supporting their Prospect Theory shows that outcomes with low probabilities receive a disproportional weight in the decision process. Lichtenstein, Slovic, Fischhoff, Layman and Combs (1978) and Combs & Slovic (1979) have found that many people seriously overestimate low probability events that receive disproportional coverage in the media (e.g., botulism poisoning).

Wallsten and his colleagues (see Wallsten, 1986) provide an interesting demonstration of the difficulties in communicating low probabilities. In one of their experiments, one member of a pair is shown a probability graphically (as an area of a pie chart). That person's task is to communicate that probability to the other member of the pair <u>without using numbers</u>. Instead of numbers the communicator can use phrases such as "very unlikely" or "rarely" or "probably." The person receiving the information then must estimate the probability being communicated. The result is that moderate probabilities (.15 to .85) are communicated rather accurately but more extreme probabilities (close to zero or one) are communicated quite poorly. The error is that the receiver of the communication overestimates low probabilities and underestimates high probabilities. These results suggest that non-numerical language for communicating very low and very high probabilities may simply not exist in English.

Our own laboratory research also dramatically illustrates risk judgment problems for low probabilities. In our studies, participants bid for insurance to protect them against a real economic loss that would occur with a known probability. We used an auction procedure (a multi-unit Vickery or competitive auction) known to have demand-revealing properties in order to get participants to bid what the insurance against the risk was really worth to them. Thus, their insurance bids can be considered a measure of their concern about the potential hazard. Figure 3.1 of Chapter 3 shows the mean values of the ratio of the bids (B) to the expected value (EV which equals probability of the loss times the magnitude of the loss) as a function of probability. If bids are consistent with standard assumptions of economic rationality, then B/EV should equal 1 for all probabilities. For probabilities of .2 and above, B/EV is very close to 1; however, B/EV is somewhat greater for $p = .1$ and very much so for $p = .01$ where insurance bids are about 2.5 times greater than EV. Thus, average concern for low probabilities is much greater relative to an EV model, than average concern for higher probabilities.

Camerer and Kunreuther (1987) report that in their experiments using double-oral auction markets insurance prices approach expected value for a large range of probabilities and loss amounts. They explicitly express surprise at our results reported in Chapter 3 because they obtain mean bids near expected value even for low probabilities. The explanation for the difference in the results between the two sets of studies is that

Camerer and Kunreuther recruited students from decision sciences and finance classes at the Wharton School and Camerer and Kunreuther themselves note that these students “were familiar with concepts of probability, expected value, and sometimes risk-aversion." While it is comforting to know that students explicitly taught expected value learn it and can apply it in these bidding experiments, it provides no evidence that typical citizens can understand and appropriately respond to low probability risks.

Thus, research is consistent in suggesting that low probabilities will be difficult for typical people to understand. By low, we mean probabilities of about .05 and below. Of course, many of the risk probabilities that are of interest to the Environmental Protection Agency, in general and to Super fund, in particular, are much less than that. It is not unreasonable to extrapolate from our data and expect that average concern about a risk will deviate even further from EV for very low probabilities such as 10^{-6}. It should be noted that we abstract from risk aversion in this discussion for clarity.

Figure 3.1 in Chapter 3 suggests that high probability risks (i.e., probabilities >.1) are likely to be well-understood. This is consistent with experience in the natural hazards area. While it is difficult to get people to plan appropriately for the 100-year flood which has a probability of .01, it is relatively easy to get people to respond to warnings of a higher probability such as alerts for imminent floods, tornadoes, or hurricanes. (Although in such cases it is sometimes difficult to communicate to people the appropriate action they should take.)

In summary, the problems of understanding and communicating low probability risks will be very different from those of high probability risks. It is therefore important to make this distinction in any discussion of risk judgments.

2.2 Bimodality.

The obvious next question is what goes wrong in judgments of risks at low probabilities. We believe our research is unique in providing explanations of the problems of judgments of low probability risks. Our explanations are based on natural cognitive processes. We were led to these explanations by discovering that the aggregate picture of Figure 3.1 is very misleading. Figure 3.2 shows the frequency distribution of B/EV for three probabilities ($p = .01$, $.2$, and $.9$) from our laboratory experiments. For $p = .9$, B/EV has approximately a normal distribution (note that the x-axis is a log scale) centered about 1, consistent with an EV model. The picture is similar for $p = .2$ although the distribution is more log-normal. However, the picture is very different for $p = .01$ where there is clear indication of bimodality. A sizeable proportion of people (about 25%) dismiss the risk out of hand (as indicated by bidding zero for the insurance) while another group of people bid at or substantially above EV. Thus, the high mean for B/EV for $p = .01$ in Figure 3.1 obscures the fact that many people are unconcerned about the risk. Somewhat paradoxically, for low probabilities more people are unconcerned about the risk, but those who do not dismiss the risk are much more concerned, relative to EV.

We think the apparent paradox can be explained by two cognitive processes: editing and anchoring and adjustment. Each individual is confronted by so many low probability risks that it would be paralyzing if one attempted to decide on an appropriate response to all of them. Thus, a useful strategy is to dismiss or "edit" risks that one considers to be below some threshold. For a fixed risk consequence (like the loss in our lab experiments), editing ought to increase as the probability of the risk decreases, or, conversely, the fraction of people concerned enough about the risk to bid for insurance ought to decrease as the probability decreases. For our data, f^+, the fraction of people bidding for insurance is modeled by

$$f^+ = .936 - .002p^{-1}$$

$$(152) \qquad (13.2)$$

$$DF = 4 \qquad R^2 = .98$$

The data and the model fit are plotted in Figure 3.3. Obviously, the fraction of positive bids falls sharply as the probability falls and the amount of editing increases.

If someone doesn't dismiss the risk and therefore is concerned about the risk, how does that person decide on an appropriate level of concern? In our experiments, deciding on an appropriate level of concern is equivalent to deciding how much to bid for insurance. We think an anchoring and adjustment process explains the mental steps involved in generating a number for a bid. To arrive at a bid, individuals focus or anchor on the loss and then adjust their bid downward to reflect that the loss will occur only some of the time. A consistent result from the cognitive psychology literature (Poulton, 1968; Tversky & Kahneman, 1974; Lichtenstein, et al. 1978) is that such adjustments are almost always underadjustments in that they stay too close to the original anchor. To model this we assume that people start with the anchor L (the amount to lose if the hazard occurs) and adjust this amount downward aiming at the target pL. We further assume a consistent proportional underadjustment error of β. If the total adjustment ought, according to an EV model, to be (L-pL) then the error will be β(L-pL). In other words, the bid will be given by

$$B = pL + \beta(L-pL).$$

Dividing both sides by EV = pL yields

$$B/EV = 1 + \beta(1-p)/p.$$

Fitting this model to the mean bids for those people not editing gives the following estimates:

$$B/EV = 1.1 + .023(1-p)/p$$

$$(13.3) \quad (11.5)$$

$$DF = 4 \qquad R^2 = .97$$

The intercept of 1.1 is not significantly different from the predicted value of 1 (t = 1.2). The data and the model fit are plotted in Figure 3.4.

Surprisingly, the model fit estimates the underadjustment error as only being between 2% and 3%. How can such a small error distort responses so much for low probabilities? For low probabilities the absolute size of the adjustment is large because the loss serves as the anchor and the distance between that anchor (L) and the target (pL) will be greatest for low probabilities. As the size of the adjustment required becomes bigger so too does the size of the adjustment error B(L-pL). This absolute adjustment error will be especially large when considered relative to pL, the EV target, which will be small for low probabilities. Hence, overbidding relative to EV will be especially pronounced, according to the anchoring and adjustment model, for low probabilities.

From subsequent laboratory experiments it appears that the anchoring and adjustment process may also apply to those who edit the risk, that is people who edit actually select O as their anchor and then adjust upward to reflect the probability and magnitude of the consequence. In such cases, the target is close to the anchors so adjustment errors may be relatively very small.

If bimodality were only found in the laboratory it would be of theoretical but not practical interest. However, we found the same bimodality of risk judgments in our field survey of residents living near the Operating Industries, Inc. (OII) landfill in California. Figure 4.3 of Chapter 4 shows the frequency distribution for resident's subjective health risk beliefs (again on a log scale). The bimodality is apparent: a large proportion of people essentially dismiss the risk by giving an estimate approximately equal to the probability of death from saccharin consumption while others give a much higher estimate approximating the risk of death from smoking a pack of cigarettes per day. We believe those equating the landfill risk with saccharin consumption are editing and then adjusting upward from a zero anchor while those equating it with lung cancer from smoking are adjusting downward from a cancer death anchor. Epidemiological studies and on-site monitoring of hazardous chemicals fail to reveal any significant risks at OII. Thus, those neighborhood residents in the low risk mode parallel those participants in our lab studies who essentially dismissed or edited the risk while those residents in the high risk mode parallel those lab participants who had an exaggerated concern for the low probability risk. We therefore think that the cognitive processes of editing and anchoring and adjustment are operating in the field sites.

When we have presented this research at professional meetings, other researchers have always come forward to tell us of similar bimodal results in the area of health risks which heretofore had been puzzling. We are convinced that it is a ubiquitous phenomenon.

An especially noteworthy aspect of the OII field study was the demonstration that the higher the proportion of residents in a neighborhood who were in the high risk mode the greater the decrease in property values. Thus, these subjective risk estimates had a real impact on economic behavior. Note also that bimodality has important implications for community conflict. It suggests there will be two distinct community perspectives towards low probability risks. One group complains that the risk is negligible and that all the fuss will only lower property values while the other group cannot understand why the former group only worries about money and is not concerned about the deadly risk confronting them all. Such conflicts are typical of low-risk Superfund sites. In summary, when deciding how concerned to be about a low probability, high consequence risk, the first decision is whether to give the risk any attention at all. If not, the risk is dismissed or edited and the person has virtually no concern for the risk. However, if the risk is not edited, in deciding how concerned to be, the person starts with his or her concern with the consequence if it were to occur for certain. Then the person adjusts that concern downwards to reflect the relative likelihood of the risk. This adjustment will almost always be an underadjustment. Given that the consequence will usually be very extreme (e.g., death from cancer), even a, slight underadjustment will leave the person with a relatively high level of concern. The operation of both editing and anchoring and adjustment processes produces a bimodal distribution of concern. As a shorthand in the following, we will refer to those people who dismiss a low probability risk as being in the “edit mode” and to those who are overly concerned as being in the “concern mode."

2.3 Two kinds of Low Probability Risks.

Given the bimodality and the difficulty that people have dealing with numerical risk estimates, especially for very low probabilities, we are doubtful that any attempts to communicate precise quantitative risk estimates will be successful. About the most that can be hoped for, in our opinion, is to get people to have levels of concern in the more appropriate mode, realizing that neither mode may be accurate. That is, if information about a new risk is being presented, the communication effort should be directed at getting people in either the edit mode or the concern mode depending on whether or not it is important to get people to take action. For existing risks for which there are some people in each mode, communication might be directed at moving people from one mode to the other.

The possibility of moving from one mode to another suggests that there is an important distinction between risk types. The first type is those risks for which it is appropriate to have communications attempting to move people from the concern mode to the edit mode. Examples might include Superfund sites where there is a great community concern but little or no scientific evidence of a risk. The second type is those risks for which it is appropriate to have communications attempting to move people from the edit mode to the concern mode. Examples might include efforts to get people to test their homes for radon gas accumulations. The communication strategies will likely be very different for the two kinds of low risk problems.

2.4 Determinants of Risk Mode.

Given the importance of the two risk modes, it is necessary to consider the factors which influence the mode in which a person’s risk judgments fall. Knowledge of those factors will automatically suggest strategies for moving people from one mode to another. In this chapter we try to answer why some low probability risks are generally ignored or edited while others generate great concern.

A dramatic example is presented by homeowners in New Jersey. Even though their homes are at high risk of having significant radon gas levels, it is extremely difficult to get homeowners in that area to test for radon, let alone make any building modifications to reduce radon gas accumulation. However, some of those same homeowners are greatly concerned about barrels of low-level radioactive wastes from the former manufacturer of radium watch dials in a nearby factory. Ironically, some of these homeowners would probably be less at risk if they were to live in a tent in the storage yard housing the barrels of low-level radioactive wastes than to sleep regularly in their basements.

In this subsection we consider three classes of factors that have been shown in our research and that of other’s to influence a person’s risk mode: perceptual cues, consequences, and experience.

2.4.1. Perceptual Cues.

Risk judgments are often partly based on perceptual cues. The more people are reminded of a risk, the more likely they are to be in the concern risk judgment mode. For example, residents near a landfill may form risk judgments based on such cues as foul odors from the site, heavy traffic to and from the site, and the presence of chain link fences. Combined with media attention indicating a possible offsite problem with cancer, these perceptual cues may move many residents to have a high concern about the risk of living near the site. In contrast, explosive concentrations of methane gas escaping from a landfill do not possess any perceptual cues to warn people of the dangers. The gas is odorless, colorless and tasteless. People depend totally on some form of official warning to take action. Fortunately, for low probability risks, many characteristics can be manipulated so that the public may evaluate such risks more accurately. In this context it is interesting to note that natural gas companies add an artificial odor to natural gas so that residents can easily detect the risk if natural gas is leaking in their homes.

Perceptual cues may be classified according to the five senses of sight, smell, taste, hearing and touch. Of these, visible cues are probably the most easily controlled. Often protective measures inadvertantly alarm the public to an excessive degree. For example, instead of institutional looking chain link fences, hedges could be planted to make a low-hazard area more aesthetic.

Other controllable perceptual cues include smell and taste. People who are bothered by odors emanating from a site such as a landfill or chemical facility are more apt to perceive greater risks. For example, in Chapter 4 we observe that residents near the OII landfill were made fearful by such odors. Another specific illustration is iron contamination in the Eagle River from the Eagle Mine which resulted in discoloration and poor taste of drinking water in Vail, Colorado. Local residents were greatly alarmed and feared that their water supply had been poisoned even though the iron in the water was itself non-toxic.

The media also provide many perceptual cues through written or spoken statements and still or moving photographic images. Our research at OII indicated that frequent exposure to media reports about OII increased the likelihood that someone would be in the concern mode.

In summary, the presence of strong perceptual cues moves risk judgments to the concern mode while the absence of perceptual cues allows risks to be edited more easily. Clearly, a key component of any risk communication strategy would be to add or reduce perceptual cues, depending on whether it was appropriate, respectively, to increase or decrease concern about a particular risk.

2.4.2. Consequences.

The nature of the consequences, or more accurately, a person's beliefs about and characterization of those consequences, are obviously important in determining the risk mode. Clearly, the more serious and dramatic the consequences of a risk, the higher will be the anchor in the anchoring and adjustment process so the final level of concern will be higher. Slovic, Fischhoff, Lichtenstein (1980) have identified a number of important characteristics of risks that cause people to under- or overestimate risks, or, in our terms, to be in the edit mode or the concern mode. Many of these characteristics pertain to the magnitude and imagibility of the consequence. Dreaded risks that are believed to have the potential of killing many people at one time in a dramatic event will usually have risk estimates in the concern mode.

One very important characteristic of a risk's consequence that will strongly affect attempts at risk communication is whether people believe the change in risk is a loss or a gain relative to the status quo. The differentiation of loss and gain effects is based on Kahneman and Tversky's prospect theory. As a rough rule of thumb, a loss relative to the status quo will have about three to ten times the psychological impact as an equivalent gain. Going from thinking you are safe to believing you are unsafe makes people a lot more unhappy than going from unsafe to safe makes people happy. In the context of hazardous wastes, for example, informing residents about an old chemical waste dump in their neighborhood about which they had no awareness will create a great deal of unhappiness because it is a loss relative to the status quo. Conversely, telling people who have worried about a known waste site for many years that it is in fact safe, even if they were to believe it, would not increase their happiness much because it is a gain relative to the status quo. This means that informing people about new risks and hazards must be done very carefully and that informing people about old risks is not likely to have much impact.

We believe this distinction between gains and losses partly explains the usual disparity between willingness-to-pay (WTP) and willingness-to-accept (WTA) measures of concern about risks. Fisher, McClelland, and Schulze (1986) review a number of studies comparing WTP and WTA responses. Of particular interest is our OII study (summarized in Chapter 4) which found that for closing the site if it were open (again) people were willing to pay (WTP) only about a tenth as much as they demanded (WTA) to allow the site to reopen (a loss) if it were closed.

It is important to note that there are at least two important ways in which a risk could be viewed as a loss relative to the status quo. First, people may learn of the probability or possibility of a risk that they previously did not believe they were exposed to. Second, people may learn that the consequences of a risk are more severe than they realized. Either or both of these will create an exaggerated concern about the risk.

2.4.3. Experience.

The amount and nature of prior experience with the risk is an important determinant of whether a person will be in the edit or concern risk mode. Research by Slovic et al. (1980) and others has shown that risks that are familiar, well-known to science, and with which we have had lots of benign experience are more likely to be edited while risks that are unfamiliar, not well understood by science, and with which we don't think we have had benign experiences are more likely to generate high levels of concern. For example, almost everyone, even those who have had automobile accidents, has had numerous benign driving experiences. Hence, the risks of automobile driving tend to be underestimated or edited, resulting in the underuse of seat belts. On the other hand, people who have not flown much, if any, have not had benign experience with airplanes and therefore tend to overestimate the dangers of air travel.

In our laboratory studies we have been able to take a close look at the effect of experience on risk mode. In several studies we have looked at the concern (as reflected in bids for insurance) about a low probability risk when people are exposed to many rounds of the risk (either 50 or 150 rounds). Figure 3.10 in Chapter 3 shows a 50 round experiment where the concern mode steadily decreases with benign experience until the risk occurs on the 33rd round. At that point, there is a sharp drop in the proportion in the concern mode reflecting the gambler's fallacy that a low probability event is less likely to occur on the next round because it occurred on the previous round. But then for succeeding rounds the number of people in the concern mode grows as fewer and fewer people feel comfortable in editing the risk. Note that there is not a corresponding sharp drop in the average bids for those bidding after the first risk event at round 33. This suggests that the effect of experiential variables such as benign experience and the gambler's fallacy are more important in determining whether someone edits or not than in determining the actual level of their concern.

The three factors that influence whether a person is in the edit or concern mode can explain the ironic, paradoxical response of those homeowners in New Jersey. They have all had lots of benign experience living in their homes; it is hard for them to imagine that they will be harmed in their own homes and that it will be their home that does the harm. Deaths from the radon gas risk will be undramatic and difficult to attribute to radon. Also, there are no perceptual cues because radon gas is invisible, odorless, and tasteless. Hence, the radon gas risk is underestimated, probably seriously underestimated because so many people edit the risk. On the other hand, residents are likely to believe, falsely, that the radioactive barrels might explode and wipe out the neighborhood in a dramatic event; the well-known radioactivity sign provides a perceptual cue; residents are also unaware of their benign experience living with this risk; and they don't think that the radioactivity is well understood by scientists. Hence, it is a risk that will be overestimated.

Summary. Our basic model is that perceptual cues to the risk, the risk consequences, and experience with the risk determine whether people dismiss or edit the risk. If they edit the risk, they will show little or no concern for the risk. If they don't edit, then an anchoring and adjustment model describes how people arrive at their level of concern for the risk. Anchors will generally be the consequence of loss, and adjustments downward from this anchor to reflect the low probability will almost always be underadjustments resulting in an exaggerated level of concern. The operation of the editing and anchoring and adjustment cognitive processes results in a bimodal distribution of concern levels for low probability, high consequence risks. The best that may be hoped for in risk communication is to move people to either the edit or concern mode, depending on which mode is closer to the appropriate level of concern for that risk.

2.5 Changing Responses to Low Probabilities.

The literature review and especially our own research suggests that the primary problems in risk communication involve low probability, high consequence situations, which are typical of many Superfund sites. In this subsection we describe some preliminary research which suggests some strategies for changing responses to low probabilities. The studies reported here are indeed preliminary and do not yet warrant their own separate technical reports or chapters in this report.

2.5.1 Changing Low to High Probabilities.

Our research and that of others suggests that although people have difficulty understanding the implications of low probabilities for appropriate behavior, they have little or no difficulty with moderate to high probabilities. One way to change an apparent low probability of risk on annual basis to a moderate probability is to change the focus from an annual basis to a longer period such as a lifetime basis. For example, for an annual risk of .01, the risk for a 70-year lifetime would be $1-(1-.01)^{70}$ = .51. Our research presented in Chapter 3 suggests that while people would have difficulty understanding the .01 annual risk they might be better prepared to understand the .51 lifetime risk. Slovic, Fischoff, and Lichtenstein (1978) report suggestive evidence that such a lifetime focus increases willingness to use automobile seat belts.

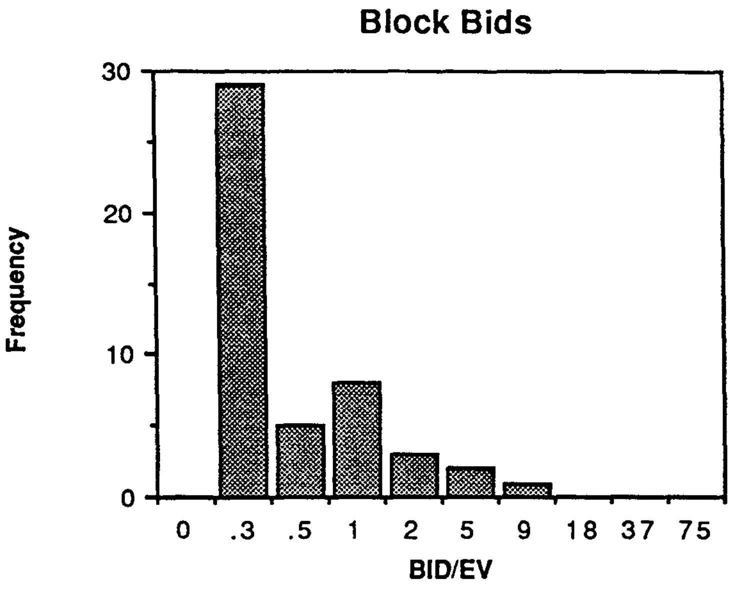
To test this idea more rigorously, we have modified the experimental procedure used in Chapter 3. Participants were told that the probability of loss on any given round was .01 but that the probability of at least one loss during 25 rounds was approximately .22. Participants were then given the opportunity to bid for an insurance policy that would protect them against a loss for the entire 25 rounds. After the first block of

25 rounds (in which no loss was experienced), they had the opportunity to bid for a second insurance policy that would protect them against a loss for the second block of 25 rounds. For comparison, the round-by-round bids of a different group of participants were summed across each block of 25 trials. In the first block of 25 rounds, for those bidding round-by-round mean BID/EV = 6.5, consistent with our earlier studies; however, for those bidding for a block of 25 rounds mean BID/EV was only 0.86. For the second block of 25 rounds the respective means were 8.5 and 0.63. Thus, even though the objective risk situation was identical for both groups of participants, changing the focus from round-to-round to blocks of 25 rounds dramatically changed behavior, as expected from our model, in a direction towards EV. Somewhat surprisingly, many participants were substantially below EV indicating that the proportion editing (or more likely anchoring on zero) increased with the block-of-25 focus. Figures 2.1 and 2.2 display the comparison of the frequency distributions of BID/EV. The block-of-25 focus shifts the entire distribution downward, note especially that extremely high bids (relative to EV) are eliminated by the block-of-25 focus.

Our results , consistent with Slovic, et al. (1978) suggest that an effective strategy in communicating low probability risks is to use a long enough focus so that the probability will be at an understandable level. For annual risks on the order of 10^{-2} and maybe even 10^{-3}, a lifetime focus might be useful or other time frames such as the "length of your mortgage." Unfortunately, this strategy will not work for risks of 10^{-4} or lower because even the lifetime risk is not in a range of probabilities that most people will be able to understand. One possibility might be also to change the focus from the individual to the family or community. For example, for

FIGURE 2.1

FREQUENCY DISTRIBUTIONS OF BID/EV FOR ROUNDS 1-25



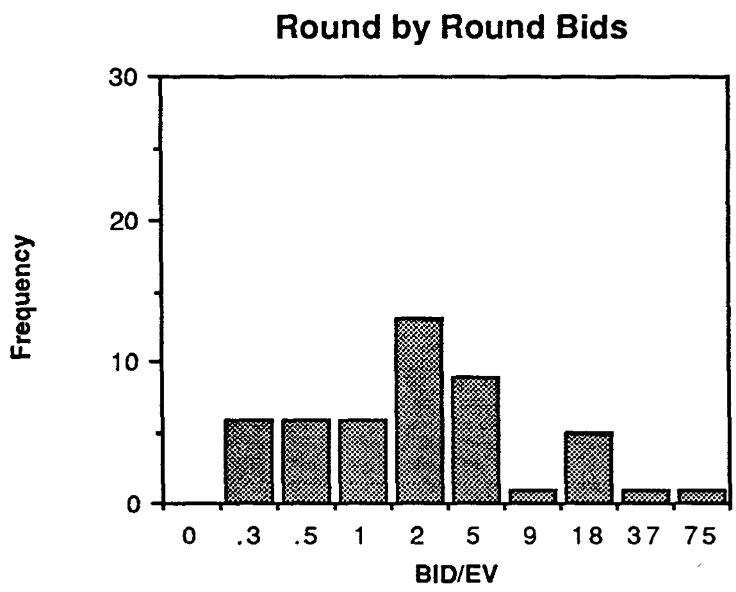


FIGURE 2.2

FREQUENCY DISTRIBUTIONS OF BID/EV FOR
ROUNDS 26-50

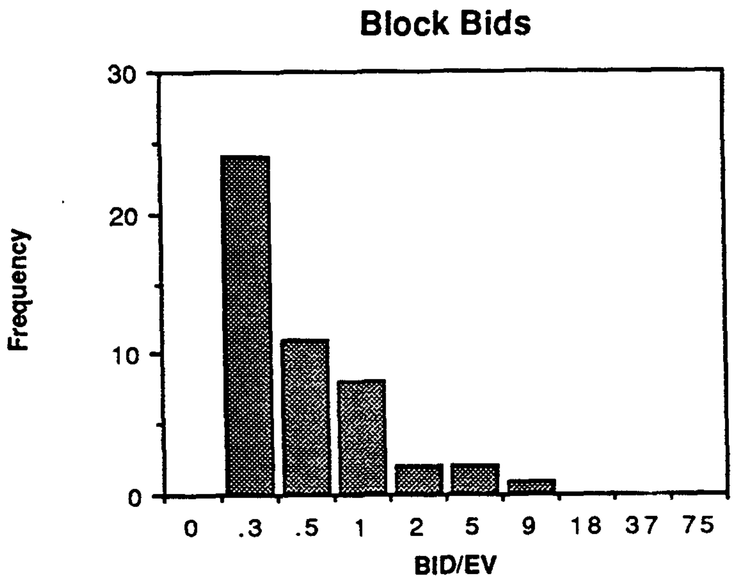
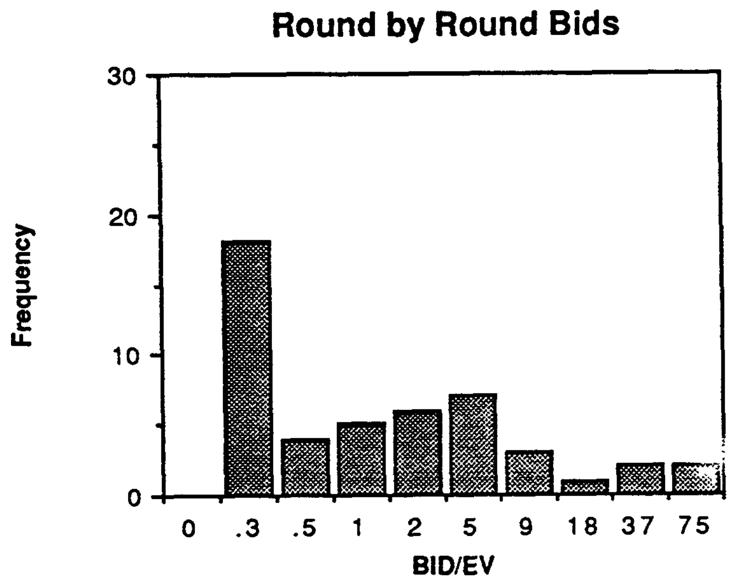


FIGURE 2.2
(CONTINUED)

FREQUENCY DISTRIBUTIONS OF BID/EV FOR
ROUNDS 26-60



an annual risk of 10^{-4}, the probability that over a lifetime at least one member of a family of four will suffer the risk is much more understandable than the annual risk. Similarly, in a neighborhood of 100 people, the probability that at least one person would suffer the risk over a lifetime would be approximately .5. The neighborhood should be able to understand the .5 probability and respond appropriately. However, this is a speculative conclusion and should not be implemented until more research on the effects of changing focus have been conducted.

Section 2.4 reviewed the findings on determinants of risk mode--those factors which influence whether people dismiss particular low probability risks or develop a high level of concern about them. It is natural to look to those factors for potential mechanisms for changing responses to low probability, high consequence events. We consider each of the three types of determinants in turn.

2.5.2 Perceptual Cues.

The perceptual cues about a hazard such as a Superfund site are just as important a part of the risk communication process as any of the formal intended messages sent to neighborhood residents by public officials and the media. Thus, it is critically important that as much as possible the perceptual cues are consistent with the actual level of risk. Too often, at sites with very low risk levels, perceptual cues such as high chain link fences, danger signs, and odors send a message inconsistent with the actual risk. On the other hand, if people are not taking sufficient precautions near a Superfund site, it would be advisable to increase the perceptual reminders about the hazard. It would be interesting to test these notions about changing perceptual cues in the laboratory but we have not yet had the opportunity to do so.

Cutter (1987) has analyzed the case histories of the evacuation response for several large accidental airborne toxic releases such as the derailment and chlorine release at Mississauga, Ontario in 1979. She too points to the important role of perceptual cues in citizen judgments of risk. Those residents who have perceptual cues about an accident (see smoke, hear an explosion, smell unusual odors, see and hear emergency equipment, etc.) are easy to evacuate; in fact, those residents often evacuate before asked to do so. On the other hand, those residents who live close enough to be in danger but far enough away so that perceptual cues are weak or nonexistent are often very reluctant to evacuate even when ordered to do so by uniformed policy officers. Cutter worries that sometimes residents over-respond to perceptual cues that might cause panic and suggests that efforts be made to bring perceptual cues into line with the actual risk as much as possible.

2.5.3 Consequences.

How citizens view the potential consequences of a risk event is an important determinant of the magnitude of their response. As noted above, of particular importance is whether a change in risk is viewed as a gain or a loss relative to the status quo. Sometimes it is possible to present risk information in a form that suggests a reference frame. For example, McNeil, Pauker, Sox, and Tversky (1982) asked both physicians and patients to choose between two treatments: an operation involving some risk of perioperative death and a less effective treatment not involving an operation. For half the people the probability was expressed in terms of dying (e.g., 5 out of 100 will die); for the other half the probability was expressed in terms of living (e.g., 95 out of 100 will live). Even though the risk was the same, both physicians <u>and</u> patients more often selected the risky operation when its risk was expressed in terms of the number living instead of the number dying. In a similar study involving the risks of oral contraceptives, Halpern, Blackman, and Salzman (1986) presented women with the same risk information expressed in these two forms: In the age group 15 to 34 it has been estimated that, in general, about 99991.7 out of every 100,000 women who use oral contraceptives will not die of circulatory disorders each year; whereas, for nonusers the rate is about 99998 out of 100,000 each year who will not die of circulatory disorders. In the age group 15 to 34 it has been estimated that, in general, about 8.3 out of every 100,000 women who use oral contraceptives die per year; whereas, for nonusers the rate is about 2 out of every 100,000 per year. Women who received the first communication judged that oral contraceptives were less risky than did the women who received the second communication. These studies suggest, therefore, that expressing risks at a Superfund site in terms of the probability that there will not be an accident or the probability that no one will be hurt by the site is likely to generate lower levels of concern than expressing those risks in terms of the probability that there will be be accident or that someone will be hurt by the site.

Some preliminary laboratory experiments we have conducted suggest that reference frames (viewing something as a loss or a gain) may change with time. We made a slight modification of our standard procedure (see Chapter 3, for a complete description of the standard procedure) so that instead of buying insurance for a low probability, high consequence risk, participants were given insurance with the opportunity to sell it back to the experimenter. That is, the procedure was changed from willingness-to-pay (WTP) to willingness-to-accept (WTA) or compensation demanded (CD). On early rounds some participants demand very high prices for their insurance but in later rounds suddenly switch to much lower prices more consistent with the expected value of the insurance. Our hypothesis is that they initially adopt the frame that they have the insurance policy, a valuable commodity, and that to give it up will be a loss. Hence, they demand a high amount for their insurance. But we buy back the insurance only from the four (out of eight) lowest bidders so those demanding high amounts keep their insurance. They then come to view the money that others are getting for selling their insurance as money that they themselves could have had; hence, the money not obtained becomes viewed as a loss. From this new frame or perspective, participants then quickly lower their offering prices so as not to miss out on the money from the experimenters.

It is possible that reframing of gains and losses similar to our laboratory studies might take place in the context of a Superfund site. For example, consider people living near a site who are so concerned about possible water contamination that they use bottled water exclusively. As they see that no problems are occurring to their neighbors who are not engaging in such averting behaviors, they may come to see their own extra efforts as a loss. Similarly, if a lot of money were spent on a detection and warning equipment by a municipality and a problem never materialized, that money and any future expenditures would likely be viewed as a loss. However, this remains an interesting speculation that ought to be tested empirically.

2.5.4 Experience.

With respect to exaggerated concern about a risk and community conflict about the level of a risk, the aphorism that "time heals all wounds" does appear to apply. With time, as long as the risk remains a potential risk and nothing happens, concern about the risk will generally decrease. The operating mechanism is of course not time; instead, the reduction in risk judgments is probably due to increasing familiarity. If so, this suggests that an amelioration strategy is to increase community familiarity with the Superfund site. Thus, in those cases where it was safe to do so, someone ought to conduct tours of a site so that people can become more familiar with the situation. In such cases, people's worst imaginations are almost always worse than the actual fact so becoming more familiar with the risk almost always reduces their level of concern.

CHAPTER 3: AN EXPERIMENTAL STUDY OF FACTORS AFFECTING RISK BELIEFS

3.1 Introduction

Psychologists have documented many systematic deviations in behavior from that predicted by the expected utility model. Much of this evidence has been generated in experiments in which subjects have been asked what their behavior would be in response to hypothetical situations (see, for example, Lichtenstein and Slovic, 1971; Slovic et al., 1977; Kahneman and Tversky, 1979; Tversky and Kahneman, 1974, 1981; Abelson and Levi, 1985). Based on these experiments, psychologists have argued that errors in decision making under uncertainty arise from the improper application of intuition or simplifying rules of thumb (heuristics), from the improper consideration of factors irrelevant to the decision (framing or context effects), and from errors in reasoning about probabilities. Such errors may play a dominant role under some circumstances such as those found at Superfund sites.

Economists have also conducted laboratory experiments exploring behavior under uncertainty. Results of these experiments, while in part confirming deviations from the expected utility model (e.g., Grether and Plott, 1979), suggest that when individuals make repeated choices in a market environment the expected utility model is "not universally misleading” (Plott and Sunder, 1982, p. 692). Economic experiments generally use actual cash payments, induce values (control the value to the subject of the commodity used in the experiment so it is known with certainty to the experimenter (Smith, 1976), and employ many repeated trials to allow individuals to practice and become familiar with the market institution (e.g., Coppinger, Smith and Titus, 1980; Smith, Williams, Bratton and Vannoni, 1982; and Coursey, Hovis and Schulze, 1986).

One principal focus of experimental economics has been an examination of the efficiency and Pareto optimality properties of market institutions. Since Pareto optimality by definition is an idealized rational outcome, experimental economists have been concerned with finding institutions which tend to produce rational behavior. This focus contrasts substantially with the objective of many experiments, conducted by psychologists, that have as their objective the detection of situations where deviations from rational behavior will occur.

This chapter presents results of two experiments that attempt to integrate these separate lines of research conducted by economists and psychologists to understand behavior at low probabilities. To this end, our experimental design and our analysis follow procedures and employ concepts drawn from both cognitive psychology and experimental economics.

The aim is to collect a body of evidence which might help in interpreting the empirical study of a Superfund site presented in Chapter 4. This study suggests that for low probability, high loss events, large deviations from rational behavior are likely to occur. For example, past studies of flood and earthquake insurance (Kunreuther, et al., 1978) and of the value of avoiding exposure to hazardous substances (Burness et al., 1978 and Smith and Desvouges, 1986) all suggest deviations from rationality. Such studies document a difficult and as yet unresolved policy dilemma. In some cases, such as hazardous wastes, many individuals seem to place inexplicably large values on avoiding risks. Yet in other cases, such as floods or earthquakes, many individuals refuse to insure against objectively similar or even greater risks. While it is difficult or impossible to replicate the high loss nature of such events in the laboratory, it is possible to explore a range of risk to see if behavior at relatively lower probabilities is in some way different from behavior at relatively higher probabilities.

Finally, considerable controversy has surrounded the use of hypothetical as opposed to actual responses from individuals. Thus, the experiments were also designed to collect both hypothetical and actual data involving cash purchases of insurance. Hypothetical values were obtained both before and after individuals had actual repeated market-like experience so that the effect of experience on the accuracy of hypothetical responses could be assessed.

In interpreting results such as those reported here, there is always some question about how well responses to laboratory risks generalize to real risks posed by a Superfund site. Although the precise responses might not generalize there is good reason to expect that the cognitive processes underlying the responses will generalize. That is, while it may not be possible to predict from laboratory studies the proportion of people who edit or the amount by which people underadjust from an anchor or the degree to which they are susceptible to the gambler’s fallacy, it is almost surely true that the cognitive processes of editing, anchoring and adjustment, and the gambler’s fallacy will be evident in people’s responses to a Superfund site. It is unreasonable to presume that people in emotionally-charged situations attempting to deal with real risks will suddenly have access to cognitive processes to aid them in their decisions that they do not have access to in the relative calmness of the laboratory. It is our view that, laboratory. It is our view that, if anything, the decision making problems identified in the studies reported here are likely to be exacerbated when people confront risks outside the laboratory.

3.2. Experimental Design

3.2.1 Theoretical Issues

This section develops the theoretical basis for the detailed experimental design presented below. In contrasting expected utility theory (EUT) with models from cognitive psychology, we draw strongly on the formalized theoretical structure developed by Kahneman and Tversky (1979) which they term prospect theory (PT). PT has been evolving over the last decade and we apply the label broadly to include several extensions of the model.

In general, our experiments were conducted as follows: each subject is given the opportunity to make a bid of B dollars for insurance against a possible loss of L dollars that occurs if a red chip is drawn. The probability of drawing a red chip is given as p. If a white chip with a stated probability of 1-p is drawn, each subject is rewarded with a gain of G dollars. The gain is included in part to finance successive trials. If a subject has an initial wealth of Y° dollars and utility is a function U(Y) of wealth Y, then, according to EUT the expected utility of the situation described above without purchase of insurance is

(3.2.1) $pU(Y^{\circ}-L>+(1-p)U(Y^{\circ}+G)$

and the expected utility with purchase of insurance is

(3.2.2) $pU(Y^{\circ}-B)+(1-p)U(Y^{\circ}+G-B)$.

The most that an individual should pay for insurance can be obtained by setting (3.2.2) equal to (3.2.1) and solving for the bid, B. The notion here is that individuals will only be willing to increase the bid to the point that the expected utility with insurance falls to the level of of expected utility without insurance. Since the loss and gain used in the first experiment ($4 and $1, respectively) are small relative to wealth, EUT would imply that it is reasonable to suppose that changes in wealth are constrained to an approximately linear segment of the utility function. Thus, a linear approximate utility function

(3.2.3) $U(Y)=U(Y^o)+U'(Y^o)\cdot\Delta Y$ where $\Delta Y=Y-Y^o$

may be substituted into (3.2.1) and (3.2.2) without loss of generality. If (3.2.1) and (3.2.2) are then set equal, the bid for insurance solves as

(3.2.4) $B=p\cdot L.$

Thus, the bid is equal to the expected value of the loss (EV). Since, as noted above, Vickrey auctions have been shown to be strongly demand revealing, we would expect bids to be equal to EV or at least normally distributed around EV for a large range of probabilities if EUT is a good predictor of behavior.

While maintaining the linear weighting of EUT, prospect theory makes two modifications. First, the utility function is replaced with a rather different value function. Second, the probabilities are replaced by a weighting function which depends on the probabilities.

PT postulates that individuals are assumed to care only about relative changes from their current wealth position and to dislike a loss in wealth much more than they enjoy an equivalent gain. Thus, according to PT the value function is not an argument of wealth, but rather of changes in wealth, AY. Further, the value function $\nu(\Delta Y)$ has the properties that $\nu(o)=0$, the left hand derivative $\nu'(o)^-$ exceeds the right hand derivative $\nu'(o)^+$ at the origin, and that both derivatives are positive, so $\nu'(o)^->\nu'(o)^+>0$. As we show below, the value function likely plays no role in the structure of our experiment, but it has been introduced by cognitive psychologists because many individuals seem to make errors in judgment because they reason in relative rather than absolute terms and show intense aversion to perceived losses.

The weighting function $\pi(p)$ of PT overweights small probabilities ($\pi(p)>p$), underweights large probabilities ($\pi(p)<p$) and shows subcertainty ($\pi(p)+\pi(1-P)<1$). The subcertainty feature implies that when a certain outcome is compared to an uncertain prospect, the prospect will be underweighted relative to the certain outcome. This attribute of the model adjusts for the observation drawn from psychology experiments that individuals seem to be biased towards certainty. Similar probability weighting functions have been proposed by Handa (1977) and Karmarkar (1979).

Given PT as described above, the value of the prospect posed by the experimental situation without insurance would be given by

(3.2.5) $\pi(p)\nu(-L)+\pi(1-p)\nu(G)$

and the value of the situation with insurance would be given by

(3.2.6) $\nu(-B)+\pi(p)\nu(o)+\pi(1-p)\nu(G)$.

Note that (3.2.6) is not written as $\pi(p)\nu(-B)+\pi(1-p)\nu(G-B)$. This is because subjects must first pay for insurance, a <u>certain</u> loss which is valued as $\nu(-B)$ and implicitly weighted with unity. After this adjustment, subjects face a modified prospect of $\pi(p)\nu(o)+\pi(1-p)\nu(G)$ which is underweighted since $\pi(p)+\pi(1-p)<1$, reflecting a bias against uncertainty central to PT.

To obtain the bid for insurance, the two expressions (3.2.5) and (3.2.6) are set equal. This algebraic manipulation is specifically legitimized by cognitive psychology in the following way. The model presented here can be interpreted as a mental representation that

that describes how individuals decide how much to bid for insurance. Thus, subjects in the experiment will note that the gain of G dollars will occur with or without purchase of insurance. This implies that $\pi(1-p)\nu(G)$ may be cancelled from (3.2.5) and (3.2.6), that is, the gain can be ignored in the decision process. If an individual has insurance, a red draw causes no loss, so the term $\pi(p)\nu(o)$ may be dropped from (3.2.6) since $\nu(o)=0$. 'his leaves a comparison of the certain loss associated with purchasing insurance which is valued as $\nu(-B)$ with the uncertain loss associated with drawing a red chip which is valued as $\pi(p)\nu(-L)$. Thus, we arrive at

(3.2.7) $\nu(-B)=\pi(p)\nu(-L)$.

Since the value functions on both sides of (3.2.7) evaluate small decreases in income, -B and -L, respectively, a linear approximation of the value function is appropriate so, for decreases in income $(\Delta Y<0)$ we have

(3.2.8) $\nu(\Delta Y)=\nu(o)+\nu'(o)^{-}\cdot\Delta Y=\dot{\nu}'(o)^{-}\cdot\Delta Y$

since $\nu(o)=0$. Substituting (2.8) into (2.7) yields

(3.2.9) $B = \pi(p)L$

and therefore the bid is equal to the weighting function times the loss. Thus, bids for insurance against a small loss will, according to PT, involve the weighting function but not the value function. Individuals can be thought of as recognizing that they must choose between two small losses: a sure one of B dollars and an unsure one of L dollars. We will discuss a possible mental process for arriving at this bid later.

In analyzing the data from the experiments we can evaluate the predictions of of relative to those of EUT by dividing the actual bids obtained in the experiment by EV which is a known constant, pL, for any stated probability, P, and loss, L. If PT is taken as the basis of analysis, dividing (3.2.9) by EV yields

(3.2.10) $B/EV=\pi(p)/p.$

Given the assumptions on the weighting function (relative overweighting of low probabilities) B/EV should be greater than unity for small probabilities, and B/EV should be less than unity for larger probabilities. Thus, our experimental design focuses on the values of B/EV over alternative probability levels. If the frequency distribution of individual values of B/EV at all probability levels is normally distributed around unity, then bids should closely correspond to EV and EUT would be supported by the data. However, if the frequency distribution of individual values of B/EV is not normally distributed around unity, some alternative, such as PT, is likely to be the more appropriate theoretical structure.

3.2.2 The Structure of the Experiment

Each experimental session employed eight student volunteers recruited from undergraduate economics classes at the University of Colorado. Five experimental sessions (total of 40 participants) provide data at probabilities of .01, .1, .2 and .4 while three experimental sessions (total of 24 participants) provide data at probabilities of .6 and .9. No student participated in more than one session. Subjects received a $5 guaranteed payment for participating. In addition, they were given a $10 stake at the beginning of the five lower probability experimental sessions and a stake of $65 at the start of the three higher probability experimental sessions. They were allowed to keep any of the stake remaining and any gains at the end of the experiment. Subjects were assured that even if they lost all their stake, they would still receive the $5 payment.

Overview. In the course of the five lower probability experimental sessions, each participant made a total of 51 bids to purchase insurance in the following risky situation which was fully described to the participants: A chip is to be drawn from a bag containing R red chips and W = 100-R white chips. If a white chip is drawn, each participant receives $1. If a red chip is drawn, those having insurance lose nothing but those without insurance lose $4. Before being placed in the bag, the stacks of chips were displayed on a table in front of the participants so they would have a more concrete representation of the specific probability levels. The four values of R used in each session were 1, 10, 20, and 40 corresponding to p = .01, .1, .2, and .4, respectively. The particular value of R being used was always made explicit before each bid. The total of 51 bids consisted of two basic types: hypothetical bids (7) and Vickrey auction bids (44). The method used for obtaining each bid type is described separately below and then the sequence of the bid types is described. In the three higher probability sessions an identical situation was employed where, however, R was equal to 60 and 90 corresponding to probabilities of .6 and .9, respectively. Three hypothetical and 20 actual bids were collected from each subject.

Hypothetical Bids. Two types of hypothetical bids were collected: inexperienced and experienced. For the inexperienced hypothetical bids, the risky situation was described to subjects as hypothetical and they were asked how much they would hypothetically pay for an “insurance policy”, which would offer full protection against the $4 loss associated with the draw of a red chip. Subjects wrote their bids on paper. These inexperienced hypothetical bids were meant to be comparable to the types of responses obtained in many psychology experiments (for example, Slovic & Liechtenstein, 1968).

To obtain the experienced hypothetical bids, subjects were asked the same hypothetical question after they had had experience with the Vickrey auction and with the drawing of chips for other probability levels. Subjects entered their bids on computer terminals in the same manner as described below for the Vickrey auction.

Vickrey Auction Bids. A Vickrey (1961) auction determined who received insurance on each round. Subjects read written instructions, heard an oral explanation of the auction procedure, and were given an opportunity to ask questions. After the appropriate number of chips were displayed and placed in the bag, the eight subjects in each session entered bids on a computer terminal for one of four insurance policies sold in each round. The terminal also displayed the current composition of the chip bag. The computer accepted bids between, inclusively, 0 and the subject's current balance in units of one cent. After everyone had entered a bid, the computer rank ordered the bids from highest to lowest and displayed the "reigning price"--the fifth highest bid for insurance--on each subject's terminal screen. Only the four subjects with bids above the reigning price received insurance. In the case of ties for the fourth highest bid, remaining insurance policies were randomly allocated among those with tied bids. Those receiving insurance were only required to pay the reigning price. This represents the key feature of the Vickrey auction and is intended to eliminate the incentives for strategic behavior that are present in auctions in which individuals must pay exactly what they bid. After each auction, the computer displayed the original balance, the reigning price, whether or not insurance had been received, adjustments to the balance, if any, and the new balance. Other than the reigning price, subjects received no information about the bids of other subjects. Terminals were arranged so that no subject could see the terminal of any other subject and subjects were not allowed to talk with each other. At the beginning of the experiment subjects participated in four practice bidding rounds which did not affect their balances in order to familiarize them with the procedures used in the Vickrey auction.

Great care was taken to avoid the use of any judgmental words in the written and oral instructions. This is in contrast to some previous experiments using the Vickrey auction which have used “winners” to designate those who have received insurance. The use of such words might artificially increase the subjective value of holding insurance above its value as protection against the loss associated with the draw of a red chip.

Risky Event. After the auction and distribution of insurance, the experimenter reached into the bag of chips, stirred the chips noisily to reinforce beliefs of randomness, and drew a chip from the bag so that all subjects could see its color. Another experimenter entered the color of this chip at a control terminal so that the appropriate adjustments--$1 to all if a white chip was drawn and $4 loss to those without insurance if a red chip was drawn--could be made to the subjects’ balances and displayed on their terminals. To allow pooling of data across sessions and to ensure that all subjects received the same probabilistic experience, the drawing was controlled (the different colors of the chips were distinguishable by texture as in Phillips and Edwards, 1966, and many similar psychology experiments) according to the following sequences: [1]

---

[1] Subjects were chosen and sessions were arranged so that communication between subjects participating in different sessions outside of the laboratory was unlikely. In fact, the supposedly random draws were never questioned by subjects. Rather, subjects were suspicious that the computer run auction was rigged.

Experimental Sessions	Probability Level	Sequence of Chips (W= white R= Red)
Lower Probabilities	p = .01 p = .10 p = .20 p = .40	W W W W W W W W W W W W R W W W W W W W W W R W W W R W W W W R R W W W R W W W
Higher Probabilities	p = .60 p = .90	R W R R W W R W R R R R R W R R R R R R

Sequence. The different components of the experiment were presented and data were obtained in the following fixed order in the lower probability experimental sessions:

Inexperienced Hypothetical Bids at p = .2, .1, .01, and .4

Vickrey Auction Practice Bids, 4 rounds at p = .2

Vickrey Auction Binding Bids, 10 rounds at p = .2

Experienced Hypothetical Bids at p = .1

Vickrey Auction Binding Bids, 10 rounds at p = .1

Experienced Hypothetical Bids at p = .01

Vickrey Auction Binding Bids, 10 rounds at p = .01

Experienced Hypothetical Bids at p = .4.

Vickrey Auction Binding Bids, 10 rounds at p = .4.

In the higher probability experimental sessions the following fixed order was used:

Inexperienced Hypothetical Bids at p = .6 and .9

Vickrey Auction Practice Bids, 4 rounds at p = .6

Vickrey Auction Binding Bids, 10 rounds at p = .6

Experienced Hypothetical Bids at p = .9

Vickrey Auction Binding Bids, 10 rounds at p = .9

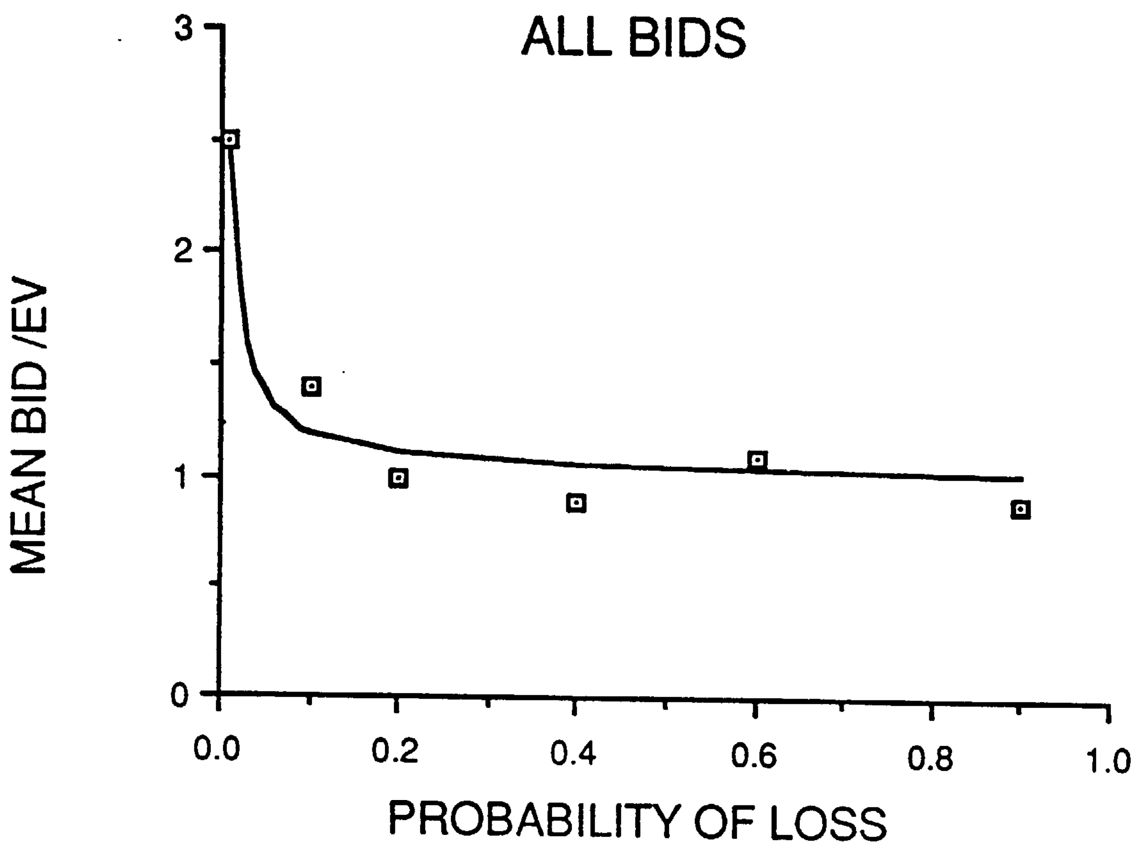
The fixed order of probabilities makes it impossible to have experienced hypothetical bids for p = .2 and p = .6 because these were always the first probability levels presented in the sequence of actual auctions.

3.3 Results

3.3.1 Overview

Summary statistics describing results of the experiment are presented in Figure 3.1. This figure depicts means of bids pooled across rounds divided by expected value, B/EV, plotted against probability of loss. As noted in Section 3.2.1, we normalize bids for insurance by dividing by expected values so we can directly compare results at different probability levels with each other and with the predictions of EUT. According to EUT we would, of course, expect mean measures of B/EV to equal unity. Note that, at probabilities of loss of .2 and above, mean B/EV is close to unity. However, at the lower probabilities of .1 and .01, EUT fails to predict observed values. The mean bid rises to about two and one-half times EV at a probability of loss of .01. Thus, on average, individuals overbid for insurance at low probabilities. This result at low probabilities is entirely consistent with the predictions of PT and can be interpreted as a direct consequence of the weighting function. From equation (3.2.10), PT predicts B/EV = $\pi(p)/p$ which should exceed unity for small p since it is assumed that $\pi(p)>p$ in this case. Mean auction values do not necessarily support PT at the higher probabilities (.2 and above) since PT argues that $\pi(p)< P$ for large P which implies B/EV<1. However, it should be noted that PT only predicts small underbidding at higher probabilities for the specific weighting functions typically proposed, so

FIGURE 3.1



we doubt that these data support a rejection of PT at higher probabilities. Rather, EUT and PT are similar in their predictions at higher probabilities for the case of insurance against loss.

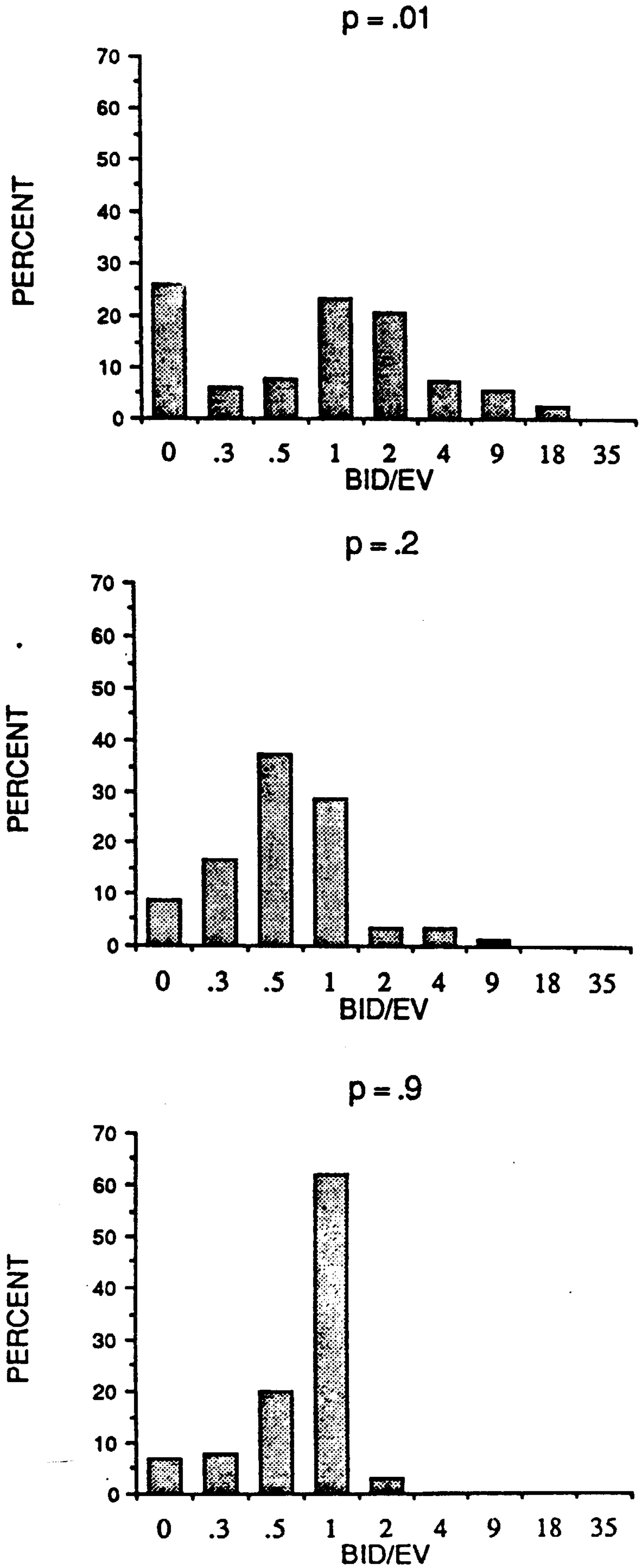
To attempt to understand the source of the large deviation from EUT that apparently occurs at low probabilities, we turn to a detailed analysis of the frequency distribution of B/EV. Figure 3.2 presents frequency distributions pooled across trials for auction values of B/EV at probabilities of loss of .9, .2 and .01. Since the frequency distributions for B/EV at probabilities of .1 and .4 are similar to that shown for .2, and since the distribution for .6 is similar to .9, only three distributions are presented. Also, since the variance increases greatly at lower probabilities a logarithmic horizontal axis is used to allow comparisons across probabilities. The approximate midpoint value of B/EV for each bin is shown under the bar representing the frequency of bids falling within the bin.[2]

The most striking feature in the top panel of Figure 3.2 is the pronounced bimodality of the distribution of bids which occurs at a probability of loss of .01. More than 25 percent of the bids in the sample are equal to zero, forming a lower mode. The distribution of positive bids on the logarithmic scale is approximately normal, thereby implying a log-normal distribution of the positive bids. The two modes suggest that two different cognitive processes may be operating at low probabilities.

---

[2] Bins were chosen as follows: The largest values of B/EV obtained in the experiment were equal to 50 and occurred at p=.01. A logarithmic scale was created by successive halving of this value. Thus, bins were created for values of B/EV <50 and >25, <25 and >12.5, ≤12.5 and >6.25, <6.25 and >3.125, <3.125 and >1.5625, <1.5625 and >.78125, ≤.78125 and >.390625, ≤.1953 and >zero. A separate bin was provided for zero bids. The rounded geometric means of the end points of each of the bins are shown along the horizontal axis of Figure 3.2.

FIGURE 3.2



p = .01
PERCENT
70
60
50
40
30
20
10
0
0
.3
.5
1
2
4
9
18
35
BID/EV
p = .2
PERCENT
70
60
50
40
30
20
10
0
0
.3
.5
1
2
4
9
18
35
BID/EV
p = .9
PERCENT
70
60
50
40
30
20
10
0
0
.3
.5
1
2
4
9
18
35
BID/EV

The first process, editing, occurs when individuals dismiss the risk and bid zero. The editing process is generally necessary since no decisionmaker can explicitly consider every possible risk, no matter how small, givdecision maker a finite time constraint on the decision process. Some rule of thumb or heuristic must be used to decide which risks are worth considering more carefully. Examination of the lower two panels in Figure 3.2 shows that the number of zero bids falls sharply as the probability of loss increases from .01 to .2 and .9. Thus, editing seems to depend on probability of loss in this experiment (where the loss is held constant).

The second process, anchoring and adjustment, attempts to explain the mental steps that individuals use to generate positive bids. The first step is, of course, the judgment not to edit, i.e., that the risk is worth considering. Second, individuals focus (anchor) on the loss, $4.00, and attempt to adjust the loss downward to account for the fact that the loss will occur only some of the time. Thus, for example, with a probability of loss of .1, subjects may be viewed as going through the following mental iterations:

<u>Example</u>

> “Should I bid $4.00? No, the loss will not occur all the time so insurance is not worth that much. Should I bid $2.00? No, this still seems to be too high a proportion of $4.00. Should I bid $1.00? Maybe. Should I bid $.50? Maybe. I think $.50 is probably closer than $1.00 to the proportion of $4.00 which represents the value of the risk of loss so I guess that will be my bid.”

Note that EV is $.40 in this case and, in the example above, the adjustment process has generated a bid which is quite appropriate. However, many subjects are likely to “guess" $1.00 since the intuitive process used in the example above is not highly accurate. That is, individuals may not engage in formal mathematical calculations in arriving at their bids. Further , since the process starts from the loss and, at least for the average individual, works downward, any error is likely to produce an upward bias in bids in that the adjustment process is likely to fall short as has been demonstrated in many studies of the anchoring and adjustment process. As is evident in examining the bottom panel in Figure 3.2, although the mean bid at p=.9 is near EV (B/EV = 1 in the figure) the variance is very large. Thus, some individuals adjust too far down while others mistakenly adjust upwards from the loss producing bids greater than the loss. Also, the distribution of bids is essentially normal (as opposed to log-normal) at p=.9, possibly reflecting more of a two way adjustment process either up or down from the loss anchor.

3.3.2 Models for Editing and Anchoring and Adjustment

We propose and test the following formal models to explain the data from this experiment. Define the fraction of zero bids as f° and the fraction of positive bids as $f^+=1-f^0$. Figure 3.2 suggests that f^+ will be a function of p, $f^+(p)$. Figure 3.3 plots the fraction of positive bids versus probability of loss. The fitted curve shown in the figure labeled "model" is estimated using data pooled across trials for the six probabilities as

(3.3.1) $$f+ = \underset{(152)}{.936} - \underset{(13.2)}{.00}\frac{1}{p} .$$

$$DF = 4 \quad R^2 = .98$$

where t-statistics, testing whether the coefficients differ from zero, are shown in parentheses. Obviously the fraction of positive bids fails sharply as the probability falls and the amount of editing increases. The functional form used was chosen on the basis of fit. For example, replacing the l/p term with an exponential in p lowered the R^2 to a value near .4.

To model the anchoring and adjustment process, we focus on explaining

FIGURE 3.3

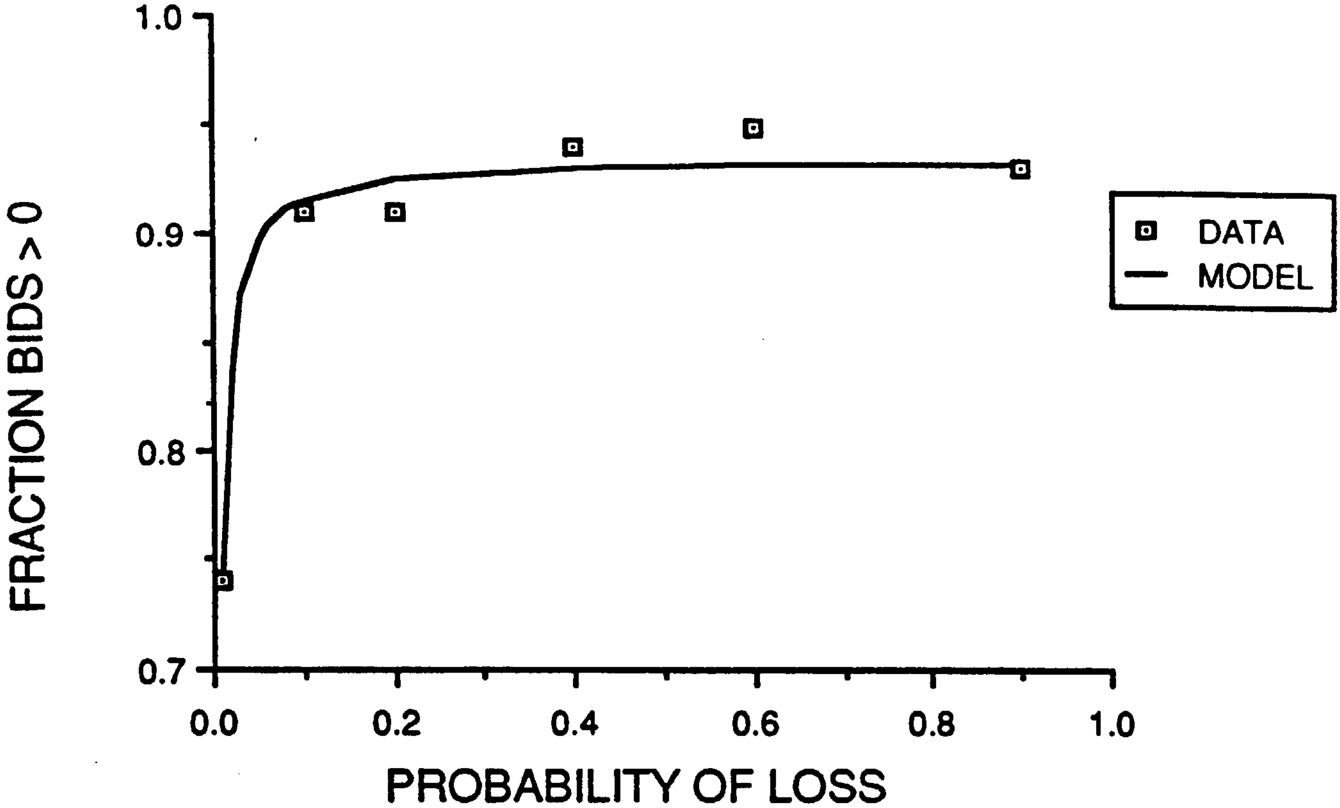
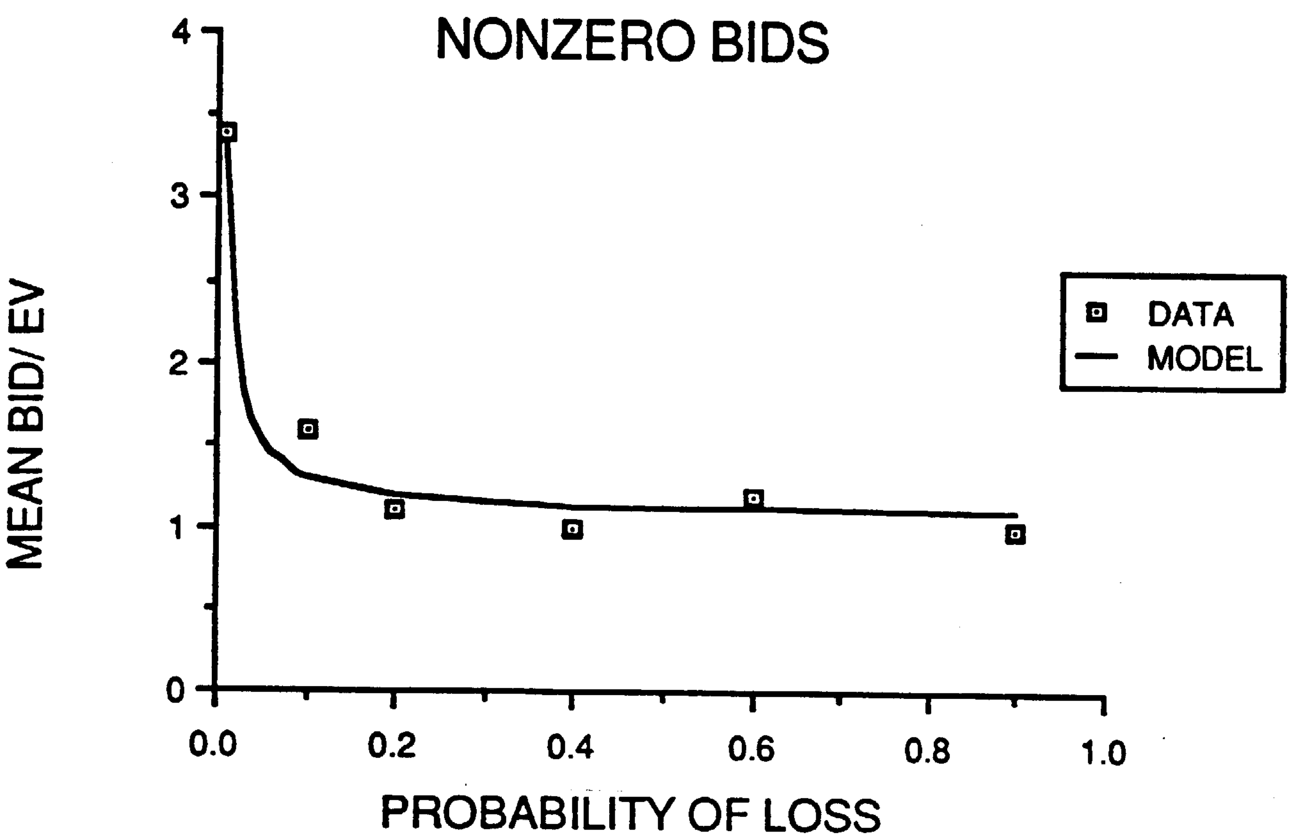


FIGURE 3.4



the mean value of positive bids, $\overline{B}^+$. The mean of all bids $\overline{B}$ is then equal to $f^+ \cdot \overline{B}^+$. Thus, the relationship developed for f^+ above which describes the editing process and the relationship developed below for $\overline{B}^+$ together yield the curve labeled "model" in Figure 3.1 that explains the average bid $\overline{B}$ in terms of p.

The average positive bidder is assumed to start from the loss, L, and move towards EV = pL. Thus, the distance over which the adjustment occurs is L-pL. Individual positive bids, B^+, can be viewed as being equal to EV plus an error term, ε, so

(3.3.2) $$B^+ = pL + \varepsilon.$$

The error term ε is assumed to have a distribution $g(\varepsilon, L-pL)$ which is shifted by the distance, L-pL, over which the adjustment process occurs. Further it is assumed that the mean of the error term, $\overline{\varepsilon}$ has the following properties

(3.3.3) $$\overline{\varepsilon}(L-pL) = \int_{-\infty}^{+\infty} g(\varepsilon, L-pL)\varepsilon d\varepsilon \begin{cases} =0 \text{ for } L-pL=0 \\ >0 \text{ for } L-pL>0 \end{cases}$$

so the greater the distance over which the adjustment must occur, the more the mean error exceeds zero.[3] We use a first order Taylor series linear approximation of (3.3.3) to obtain

(3.3*4) $$\overline{\varepsilon} \approx \overline{\varepsilon}(0) + \overline{\varepsilon}'(0) \cdot (L-pL)$$

where $\overline{\varepsilon}(0) = 0$ and $\overline{\varepsilon}'(0)>0$ by (3.3.3). Thus, if we assume that adjustment is a linear process we can substitute (3.3.4) into (3.3.2) to obtain the

---

[3] We note that the anchoring and adjustment process can also possibly produce zero bids. Some fraction of the population may adjust to zero or beyond to a negative bid which presumably appears as a zero bid. The fraction of zero bids generated by anchoring and adjustment would, based on our model be

$$\int_{-\infty}^{-pL} g(\varepsilon, L-pL) d\varepsilon.$$

We do not attempt to account for this fraction in our statistical analysis, rather assigning all zero values as edits. The error introduced by this assumption is presumed to be small.

mean positive bid as:

$$\bar{B}^{+} = pL + \bar{\varepsilon}'(0)\cdot(L-pL). \qquad (3.3.5)$$

$\bar{\varepsilon}'(o)$ can be interpreted as a parameter which estimates the fraction of the distance L-pL that the average individual falls short in the attempt to adjust the bid from the anchor, L, to pL.[4] The anchoring and adjustment process is an example of a framing effect in which consideration of the loss as an anchor biases the estimation of EV. A non-random error has thus been introduced by the intuitive thought process used to estimate EV. It is the non-random nature of this type of error which makes analysis of cognitive processes of importance in understanding economic behavior in which low probabilities are involved.

To obtain a functional form for statistical estimation and testing of hypotheses we divide (3.3.5) by EV to obtain

$$\frac{\bar{B}^{+}}{EV} = 1 + \bar{\varepsilon}'(o) \quad (\frac{1-p}{p}). \qquad (3.3.6)$$

This relationship is estimated using the six observations on mean positive bids pooled across trials for the six alternative probabilities as

$$\frac{\bar{B}^{+}}{EV} = \underset{(13.3)}{1.1} + \underset{(11.5)}{.023} \left(\frac{1-p}{p}\right) \qquad (3.3.7)$$

$$DF = 4 \qquad R^2 = .97$$

where the constant is free (not forced equal to unity). This estimated relationship is plotted in Figure 3.4 along with the data points. Using (3.3.7) we can test two hypotheses: The constant is not significantly different from unity (t(4)=1.2) but the coefficient on (1-p)/p is reliability different from, zero (t(4)=11.5). Thus, we can conclude that the probability level affects $\bar{B}^{+}/EV$.

---

[4] An anchoring and adjustment model which has some similarities to the one presented here has been successfully tested by Johnson and Schkade (1986) for a rather different experimental situation involving uncertainty. Their experiment did not involve the lower probabilities examined here and focused on having individuals provide hypothetical estimates of probability and certainty equivalents.

$\overline{B}^+/EV$ should, of course, simply equal unity and be unaffected by p if EUT holds. Thus, we must in general reject EUT. However, from (3.3.7), $(1-p)/p \to 0$ as $p \to 1$ and as a result $B^+/EV \to 1.1$, the constant, which is not significantly different from unity. Thus, we do not reject the hypothesis that EUT applies asymptotically for high probabilities; in Figure 3.4, $\overline{B}^+/EV$ is not very different from unity for p>.2. Finally, the data are consistent with the anchoring-and-adjustment model as expressed in (3.3.6) because the constant is similar to unity as predicted by the model and $\overline{\varepsilon}'(o)$ (estimated as .023) is significantly different from zero. This interpretation suggests that the average individual adjusts 97.7 percent of the distance from the loss to the expected value. The 2.3 percent shortfall in the adjustment process only leads to a large error in estimating EV (as a proportion of EV) at smaller probabilities as the distance between L and pL increases.

3.3.3 Comparison to other Models and Experiments

Returning to the original specification of PT we can now suggest an explanation for the weighting function $\pi(p)$, which can be interpreted as a decision weight on the loss L. The mean $\overline{B}/EV$ of the entire population from PT should be $\pi(p)/p$ from the analysis of section 3.2.1. Thus, using (3.3.6)

$$\frac{\pi(p)}{p} = \frac{\overline{B}}{EV} = \frac{f^+\overline{B}^+}{EV} = f^+(p)\cdot(1+\overline{\varepsilon}'(o)\frac{1-p}{p}). \tag{3.3.8}$$

so

$$\pi(p) = f^+(p)(p+\overline{\varepsilon}'(o)(1-p)). \tag{3.3.9}$$

The weighting function can thus be interpreted as an artifact of the editing process (described by the $f^+(p)$ relationship) and the anchoring-and-adjustment process (captured in the $\overline{\varepsilon}'(o)$ parameter). Unfortunately, at low probabilities where the divergence of $\pi(p)$ from p becomes important, the decision weight does not describe the behavior of any "typical" individual, but rather is the average of two divergent behaviors. In one behavior individuals bid zero. In the other, the mean individual bids well above EV, This bimodality has several implications for analyzing low-probability, high-loss events. For example, the intense conflict which often arises over technological risks such as those from nuclear power might be explained as a conflict between individuals from upper and lower modes similar to those apparent in the top panel of Figure 3.2. Such conflict cannot be explained in terms of the weighting function of PT which likely represents the average of a bimodal distribution. Rather, consistent with the spirit of PT and as an extension of PT we would propose the weighting function be discarded in favor of explicit modeling of the editing and the anchoring and adjustment processes. One way to formalize these notions is to refocus PT on the determinants of the fraction of positive bids, f^+, and on the determinants of the mean positive bid - expected value ratio, $\bar{B}^+/EV$.[5]

[5] Studies by Hershey and Shoemaker (1982), Shoemaker and Kunreuther (1979), and Slovic, et al. (1977) have also investigated insurance preferences. However, those studies are difficult to compare to the present study because those earlier studies did not involve market pressures, observed only hypothetical responses, and used very large losses (e.g., $100,000). The major difference in those studies is that respondents did not bid for insurance but only indicated whether they would accept or reject insurance offered at an actuarially fair price. In terms of our analysis, for lower probabilities these studies were essentially tracking f^+, the proportion of the sample in the upper mode. These studies are not consistent with each other in terms of their implicit modeling of f^+ as a function of p and L and it should be noted that we have offered only an atheoretical empirical model of f+ as a function of p. Hershey and Shoemaker suggest that some of the differences between studies are due to context or framing effects. We offer the further suggestion that comparisons between these studies will be difficult because the actuarially fair price will be dividing a bimodal distribution between modes for some combinations of p and L and a unimodel distribution (at approximately the mean) for other distributions. Clearly, much work remains to be done to understand the factors determining whether people edit or anchor-and-adjust from the loss.

3.3.4 Trial Dynamics

We now turn to an examination of trial dynamics - how experience over trials affects bids for insurance. Figure 3.5 shows mean auction values of B/EV (including zero bids) across rounds or trials. The means in Figure 3.5 remain constant and near unity across rounds for the higher probabilities of .2 and .9 shown (.4 and .6 are similar) but show a slight upward drift at .1 and a large upward movement at .01 across rounds. We interpret the upward drift over rounds of B/EV at the lower probabilities to be the result of gambler’s fallacy. That is, if a run of successive white chips is drawn, subjects become falsely convinced that the subjective probability of drawing a red chip has increased. This effect is not apparent at higher probabilities because when a red chip is drawn, subjects either “reset” their subjective probability close to the objective probability or assume that the odds of drawing another red chip have gone down. Thus, gambler’s fallacy appears to be self cancelling when subjects experience fairly frequent draws of a red chip. Of course at low probabilities, long runs of successive draws of white chips are likely and the cumulative effect of gambler’s fallacy will be apparent. When examining Figure 3.5, it is important to note that no red chips were drawn across the ten rounds at a probability level of .01. Also, at the probability level of .1 only one red was drawn (on the third round).

To analyze the mechanics of gambler’s fallacy we separate the data at P=.01 by again analyzing the fraction of positive bids and the mean B/EV of positive bids which are, on a round by round basis:

FIGURE 3.5

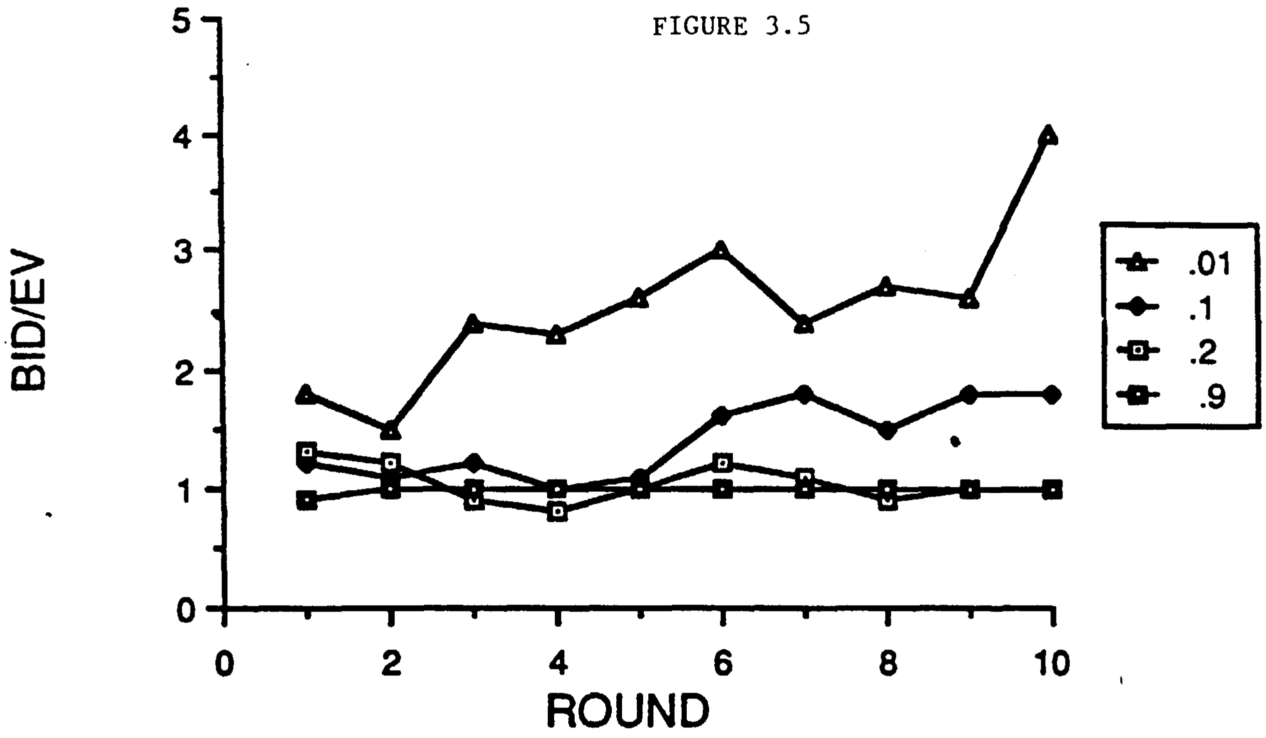
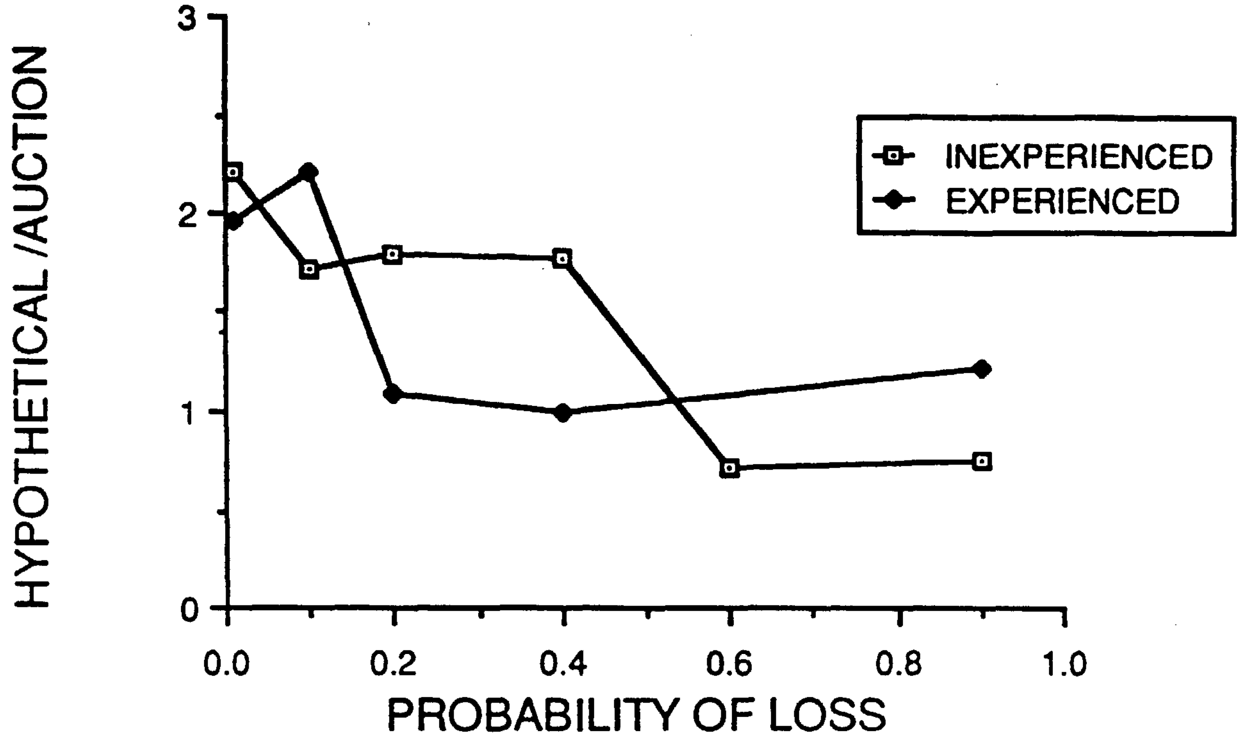


FIGURE 3.6



Round	1	2	3	4	5	6	7	8	9	10
f^+	.60	.62	.68	.75	.78	.78	.72	.85	.82	.80
$\overline{B}^+/EV$	3.0	2.3	3.5	3.0	3.4	3.9	3.3	3.2	3.2	5.0

These data suggest two observations. First, the fraction of positive bids shows a relatively steady increase across rounds. Thus, mode switching from the zero mode to a mode centered around B/EV=3.2 (the mean of B^+/EV for rounds 1 to 9) seems to be the source of most of the gambler's fallacy effect apparent in Figure 3.5. Second with the exception of the last round (which shows a strong end effect) $\overline{B}^+/EV$ is relatively stable across rounds. A linear regression of f^+ as a function of round yields an estimated equation (using only the first 9 rounds)

(3.3.10) $$f^+ = .59 + .028 \cdot \text{round}$$
(20.9) (5.6)

DF = 7 $R^2 = .82$

which contrasts sharply with a linear regression of $\overline{B}^+/EV$ on round,

(3.3.11) $$\overline{B}^+/EV = 2.89 + .065 \cdot \text{round}$$
(9.6) (1.2)

DF = 7 $R^2 = .17$

which has a coefficient not significantly different from zero for the round variable. However, we are not prepared to dismiss the hypothesis that successive rounds affect the mean positive" bid for two reasons.

First, the amount of adjustment which occurs may depend on experience; for example $\overline{\varepsilon}'(o)$ may decrease with more experience to produce bids close to the target pL. Since subjects had much experience at other probabilities prior to purchasing insurance at p = .01, the effect of experience on valuing a new risk may not be apparent here. Second, it is likely that gambler's fallacy affects the subjective probability belief of individuals. Both f^+ and B^+/EV seem to be functions of p. If we replace p with s, we should see an effect of gambler's fallacy on both f^+ and $\overline{B}^+/EV$. The next section describes an experiment at p=.01 in which subjects have no prior experience and in which the number of successive trials is raised to 50. This second experiment was specifically structured to further explore trial dynamics. In any case, the analysis above confirms another cognitive source of deviation from EUT, gambler's fallacy, which again, in a market like auction environment, seems to occur only as a problem at lower probabilities.

3.3.5 Hypothetical Behavior

As noted in the introduction of this chapter, psychology experiments of risky decision making have often used hypothetical bids and risks. In contrast, experimental economics traditionally employs actual financial transactions. The obvious question is whether using real monetary consequences differs from using hypothetical amounts. Figure 3.6 shows how means of hypothetical bids collected in our experiment compare to means of actual auction bids pooled across trials. Hypothetical mean bid divided by actual auction mean bid is shown on the vertical axis and probability of loss is shown on the horizontal axis. The inexperienced hypothetical bids collected at the start of the experiment clearly overestimate actual auction bids at low probabilities (since the ratio shown in Figure 3.6 is greater than one) and underestimate actual auction bids at high probabilities (since the ratio is less than one). The single deviation from the predictions of PT apparent in our auction results was that at high probabilities bids were close to EV, that is, we did not see the underweighting predicted by PT. We have no explanation as to why inexperienced hypothetical bids at high probabilities show underweighting and actual auction bids do not. In contrast, however, experienced hypothetical bids, which were collected after actual auction experience at other probabilities, were good predictors of auction bids at probabilities of .2 and above. It should be noted that the experienced hypothetical data point shown for .2 was not taken from the experiment described herein but from a pilot study where the order of probabilities was different so that an experienced hypothetical value could be obtained for p=.2.

Both inexperienced and experienced hypothetical bids are about twice actual auction bids at p=.1 and .01. We conjecture that the overestimation of hypothetical bids which occurs at low probabilities may be due to an incomplete adjustment process. In other words, since individuals start with the loss and work downward in deriving bids and since the distance between the loss and EV is great at low probabilities, practice may increase the amount of downward adjustment which occurs, bringing bids closer to EV. At the lower probabilities more adjustment is required and both inexperienced and experienced hypothetical bids may represent the first iteration in the adjustment process. In the experiment described in the next section subjects begin an actual auction at p=.01 with no prior laboratory experience of the auction procedure or this type of risk. If this hypothesis is correct, actual auction bids should start at very high values. From Figure 3.1, actual auction bids are about 2.5 times EV at p=.01. From Figure 3.6, hypothetical bids are about 2 times actual auction bids at p=.01. Thus, we conjecture that completely inexperienced actual auction bids might be 5 times EV. If this is the case, then hypothetical bids might be good predictors of completely inexperienced auction behavior. Note in this context that all of the auction behavior in the experiment described above was of the experienced type because we began the auctions with four non-binding practice trials.

3.4 A Laboratory Simulation of the Response to a ‘New” Risk

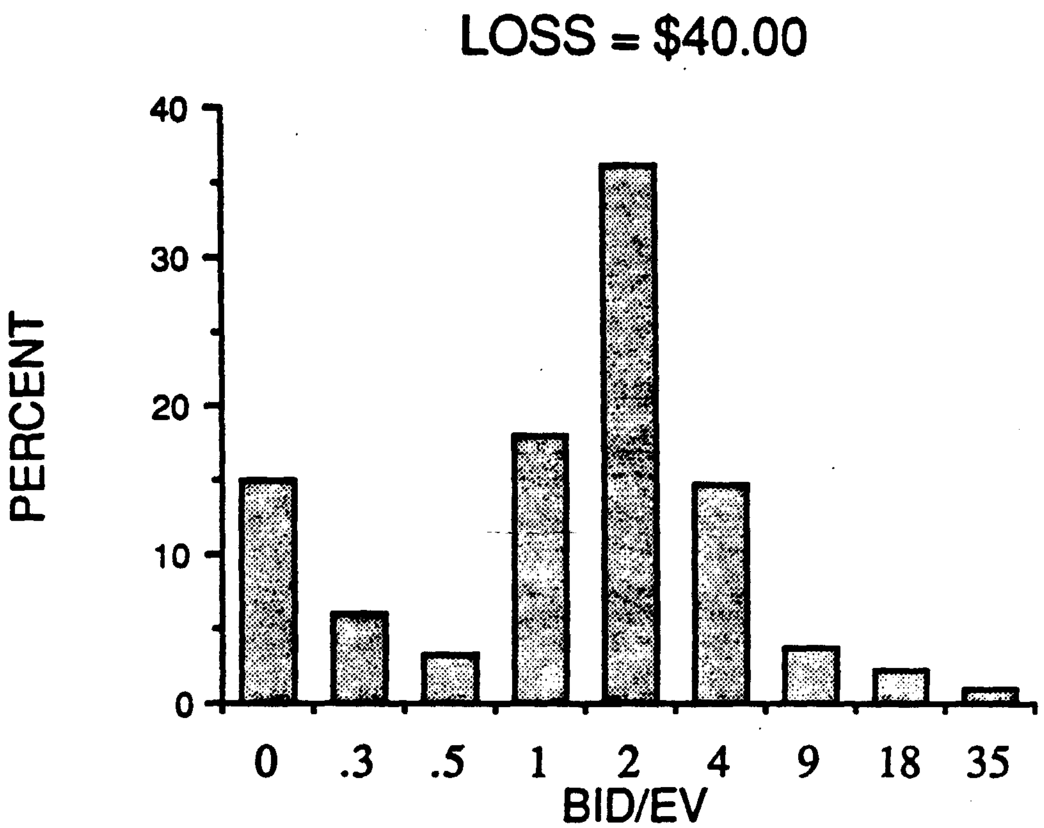
3.4.1 Overview

Given the bimodality and dynamic instability of values obtained at p=.01 in the experiment described above, a second experiment was conducted to explore further these phenomena at p=.01. A number of specific questions motivated the design of the new experiment. First, the editing phenomenon might have resulted from the relatively small $4 loss employed. Would editing still occur at similar frequencies for a much higher loss? Will bimodality still characterize the distribution of bids? To address such questions, we raised the loss to $40. Second, subjects in the experiment described above faced odds of loss of .01 after they had obtained a considerable amount of experience both with the Vickrey auction for insurance and with other probabilities. Many real world policy problems are associated with the response to new risks. Populations are informed that the landfill near which they have lived for a decade is leaking toxic substances or that the long dormant volcano or fault now poses a threat. New technologies are often resisted because they are viewed as dangerous, but later become accepted. Thus, to explore these issues, the auction was begun with no practice trials and the number of rounds was increased to 50 so that the lengthy recurrence intervals between events (draws of a red chip) characteristic of real world low probability hazards could be simulated in the laboratory. A red chip was drawn on round 33 so that subjects could accumulate benign experience in the form of a lengthy sequence of white chips before the event occurred. Seventeen remaining rounds were then available to examine behavior after the event occurred. As in the previous experiment, each of the six sessions employed eight subjects drawn from undergraduate economics classes, a Vickrey auction was used to sell four insurance policies in each round, and each subject was given one dollar if a white chip was drawn to help finance successive trials. The risky situation, initial balance and Vickrey auction were described to subjects both in written instructions and in an oral explanation which allowed questions. Subjects were shown 99 white chips and 1 red chip as they were placed in a bag. The sequence of 50 binding actual auctions then began immediately. A chip was drawn and replaced following each auction.

3.4.2 Results

Figure 3.7 shows the frequency distribution of B/EV pooled across all 50 trials. Generally, the distribution of B/EV for the $40 loss looks remarkably similar to the frequency distribution shown for the $4 loss in the top panel of Figure 3.2. Both distributions are strongly bimodal with

FIGURE 3.7



one mode at zero and another above EV (EV is shown as B/EV=1 in the figures). Since the horizontal axis is logarithmic, the upper modes in both cases appear to center on approximately log-normal distributions. Two minor differences are also apparent. First the upper mode for the $40 loss is shifted slightly to the right compared to the $4 loss. As we show below when we examine trial dynamics, inexperienced bids for the case of the $40 loss were very high in the early rounds. Thus, the difference in the initial amount of experience between the two experiments likely explains this shift. Second, in the $40 loss experiment some bidders seem to be adjusting upwards from a zero anchor creating a descending step pattern (moving to the right) for the zero, .3 and .5 B/EV bins in Figure 3.7. This suggests that individuals who edit are in reality choosing a zero anchor as opposed to the loss anchor as the basis for an upward as opposed to downward adjustment process. This leads to the conjecture that a lower mode just above zero, made up of individuals who edit might evolve under some circumstances.[6] In any case, the pronounced bimodality of the

---

[6] To understand these circumstances we need to consider why we do not see evidence of upward adjustment in the case of the $4 loss shown in the top panel of Figure 3.2. Subjects demonstrated a strong tendency to submit bids in round monetary values such as $.00, $.05, $.10, $.25, $.50, $.75, $1.00, $1.50, $2.00, $3.00, $5.00, $10.00 and so on. This monetary anchoring has often been tested in psychological studies of decision making (e.g., Combs, Bezenbinder, & Good, 1967) and in survey research. Since EV was $.04 in the $4 loss experiment, no strong monetary anchors fell between $.05 and $.00 so any upward adjustment from zero likely fell in the B/N=1 bin. In contrast, with a $40 loss, EV=$.40, and monetary anchors of $.10 and $.25 were available in the B/EV=.3 and the B/EV=.5 bins respectively. Thus, the larger the loss, the more apparent upward adjustment from the zero. anchor will be in experiments of this sort. For very low probabilities and very large losses a mode just above zero might then become apparent since monetary anchors will be available just above zero.

earlier experiment is present at the higher loss, consistent with the editing and the anchoring and adjustment models developed in the previous section.

Trial dynamics are shown in Figures 3.8 and 3.9. Figure 3.8 shows mean bid divided by EV. Early bids for insurance averaged about five times expected value, which is consistent with our conjecture of the previous section, but decreased to about two times expected value just before a red chip was drawn on round 33. Reigning price (shown in Figure 3.9), after an initial rise, remained constant at about 2 1/2 times EV until, following the draw of the red chip, a sharp drop in reigning price occurred in round 34. Both mean bid and reigning price then increased to the conclusion of the experiment at round 50. We conjecture that bids fell in early rounds both because individuals gained experience (i.e., learned to adjust more completely) and because benign experience may work in the opposite direction from gambler's fallacy by reducing the subjective probability of loss. Note that in the 10 round experiment subjects both had experience in forming values at other probabilities prior to facing odds of loss of .01 and had actually experienced the loss of $4 on the draw of a red chip. After the draw of the red chip in the $40 loss experiment bids also rose over following rounds as in the $4 loss experiment. Thus, we suspect that, in the absence of the experience of loss, draws of white chips may convince some that they should dismiss the risk and bid lower or bid zero for insurance. Experience with loss, however, seems to reverse this process. Convinced by experience that the loss can occur, some subjects seemingly felt that successive draws of white increased the need for insurance. The actual odds of drawing red remain constant over trials since the drawn chip

FIGURE 3.8

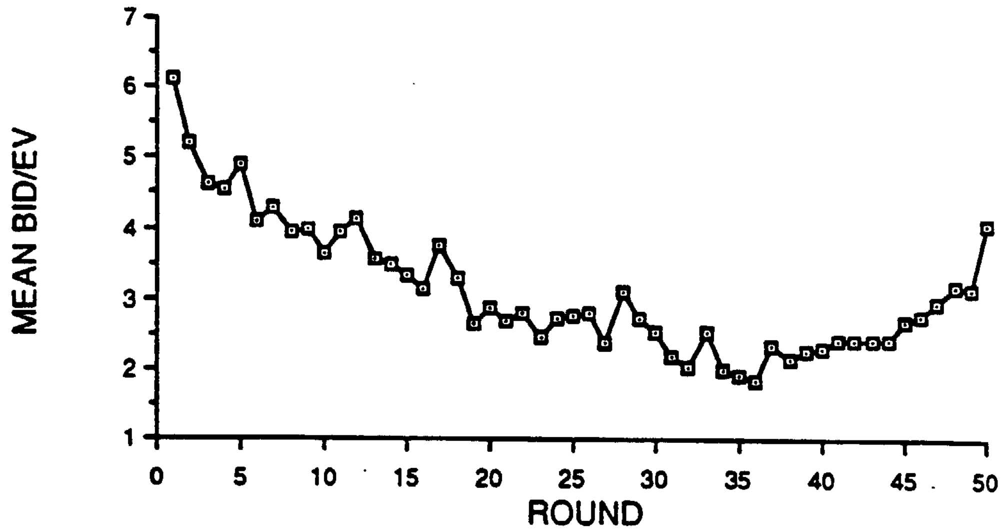
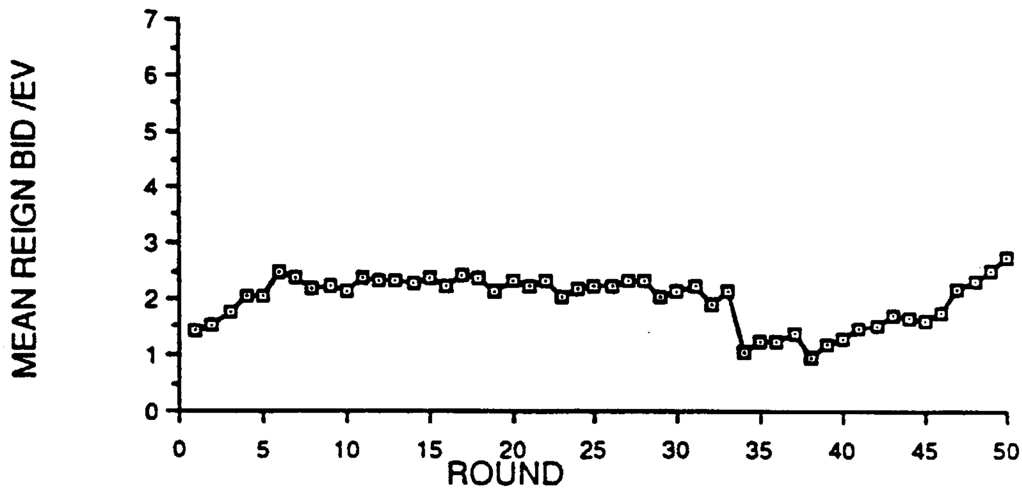


FIGURE 3.9



was replaced before the next trial. Thus, the possible benign experience and the gambler's fallacy effects on subjective probability of loss represent simple logical errors concerning probability.

To test these hypotheses we again split the data into two portions. One portion consists of the edit bids and the other contains all those bids presumed to be generated by anchoring on the loss L = $40 and then adjusting downwards towards pL. This time we count as edit bids not only the zero bids but also those bids that are slightly above zero but still in the lower mode of the frequency distribution of Figure 3.7. The category for B/EV = .5 appears to be the boundary between the two groups of bids so we use the midpoint of that category as the dividing line. The fraction of bids in the upper portion is now f^+ and the mean of the bids divided by expected value in the upper portion is B^+/EV. Both f^+ and B^+/EV are calculated for each round and plotted in Figures 3.10 and 3.11 respectively. To develop a statistical model of the effects of benign experience and gambler's fallacy on f^+ and $\overline{B}^+/EV$ we define the following variables on the basis of round:

	Round								
Variable	1,	2,	. . . . .,	33,	34,	35,	36,	. . . . .,	50
Benign	-33,	-32,	. . . . .,	-1,	0,	0,	0,	. . . . .,	0
After Red	0,	0,	. . . . .,	0,	1,	1,	1,	. . . . .,	1
Since Red	0,	0,	. . . . .,	0,	0,	1,	2,	. . . . .,	16

---

[7] If the prior definition of edit bids as only zero bids is used, the functional form of the analysis which follows is essentially unchanged. The only difference is that f+is much less stable when editing is restricted to zero bids. This suggests that those editing sometimes switch between zero and very low bids in the .3 and .5 B/EV categories of Figure 3.7.

FIGURE 3.10

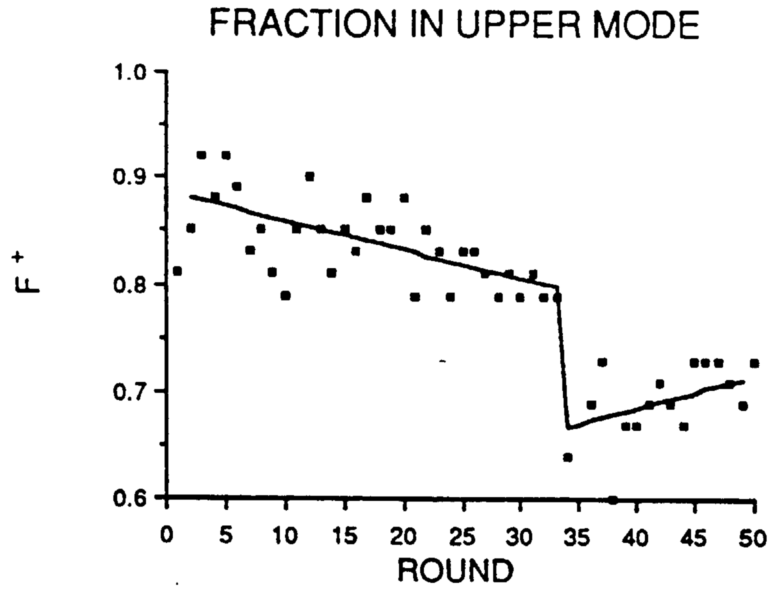
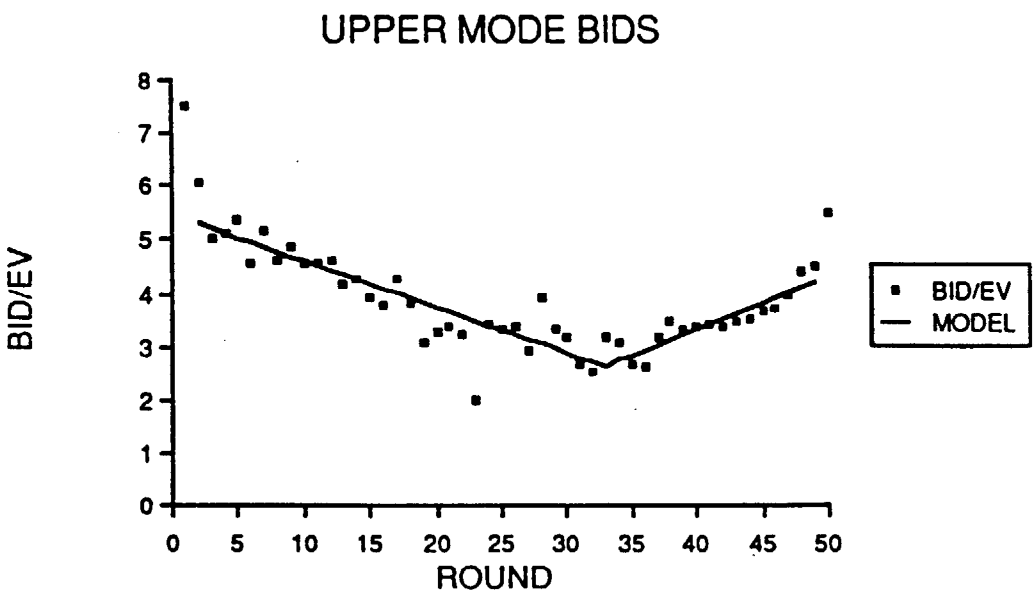


FIGURE 3.11



Linear regression estimates (excluding data from the first and last rounds, which are obvious obvious from Figure 3.11) are:

$$(3.4.1)\quad f^{+} = \underset{(73)}{.793} - \underset{(-4.6)}{.0027}\,(\text{Benign}) - \underset{(-7.1)}{.128}\,(\text{After Red}) + \underset{(1.1)}{.003}\,(\text{Since Red})$$

DF = 44 $R^2 = .87$

and

$$(3.4.2)\quad \frac{\overline{B}^{+}}{EV} = \underset{(22)}{2.54} - \underset{(-14)}{.086}\,(\text{Benign}) + \underset{(-1.1)}{.21}\,(\text{After Red}) + \underset{(5.6)}{.099}\,(\text{Since Red})$$

DF = 44 $R^2 = .88$

We can interpret the data shown in Figures 3.10 and 3.11 using the statistical model above as follows: The fraction of bids in the upper mode (those presumed to be adjustments from the anchor L=$40), f^{+}, begins (after only one round of prior experience) at approximately .88 and decreases significantly with benign experience (t(44)= -4.6) to approximately .80 just before a red chip is drawn on round 33. The draw of a red chip causes a sudden and statistically significant (t(44) = 7.1) drop in f+ (measured by the “After Red” intercept shifter) in round 34 to approximately .67. Gambler’s fallacy then appears to explain the increase in f^{+} to .71 by the next to the last round of the experiment but the “Since Red” variable does not have statistical significance (t(44) = 1.1).

The statistical model for the mean upper mode bid divided by expected value indicates that $\overline{B}^{+}/EV$ begins at approximately 5.3 and falls steadily to about 2.5 just before a red chip is drawn on round 33. This decline, associated with benign experience, is significant (t(44) = -14) and is possibly caused by an improvement in the adjustment process which occurs with experience or by a negative effect of successive draws of a white chip on the subjective probability of loss. In contrast to the f+ relationship, the draw of a red chip had a negligible immediate impact on $\overline{B}^+/EV$ as measured by the "After Red" intercept shifter (t(44) = -1.1). This difference supports a model of the decision structure which separates the editing decision from the process used to derive a positive bid. Note in this context that the drop in reigning price which follows the draw of a red chip on round 33 as shown in Figure 3.8 is the result of the drop in f^+ which reflects a sharp increase in editing due to a gambler's fallacy effect (a draw of red reduces the chance of red on the next round). Positive bids, as measured by $\overline{B}^+/EV$, are however affected by successive draws of white after a red in a manner consistent with gambler's fallacy. As measured by the "Since Red" variable, gambler's fallacy is significant (t(44) = 5.6) in apparently increasing the subjective probability beliefs of positive bidders so that B^+/EV climbs to approximately 4.2 by the next to last round. Gambler's fallacy could be motivated in part by an end effect in that subjects increasingly attempt to defend their balances as the last round approaches.

The dynamic pattern shown in Figures 3.8-3.11 is broadly consistent with the conventional wisdom concerning subjective risk beliefs about natural and man-made hazards. Such risks are viewed as being overestimated both when people first become aware of the possibility of a catastrophe as well as in some period following the occurrence of a catastrophe. Risk beliefs are viewed as being underestimated following long periods of benign experience that inevitably occur given the long recurrence intervals of low probability events. We were not able to demonstrate fully this last supposed attribute because, although $\overline{B}/EV$ declined steadily with benign experience, its value did not fall below unity nor did f^+ approach zero in 33 rounds. Obviously, the asymptotic properties of both f^+ and $\overline{B}^+/EV$ need to be explored in future experiments involving more trials.

The bimodality present in these experiments suggests that actual insurance markets for disasters such as floods, earthquakes, etc. are likely to be peculiar. If insurance is provided by competitive suppliers, in the long run the offered price of insurance should be equal to expected value of the loss plus a minimum of administrative and transaction costs. Thus, insurance should be offered just above expected value. For low probability hazards, the offered price of insurance is likely to fall between the upper and lower modes of the frequency distribution of bids for disaster insurance. Thus, nearly all of the individuals in the upper mode (who anchor on the loss) will likely purchase insurance, while all of the individuals in the lower mode (who edit) are not likely to purchase insurance. In other words, the editing decision will completely control the number of insurance policies sold. Since benign experience appears to decrease f^+, which can be interpreted as the fraction of the population at risk likely to buy insurance, people may be reluctant to insure against disasters which have not occurred recently. On the other hand, sales of such insurance might well increase in a period following an event.

3.5 Can Risk Aversion Account for the Results?

Intuition suggests that the relative impact of risk aversion on bids for insurance should increase as the probability falls. Thus, in our first experiment with a loss of $4, one would expect little evidence of risk aversion at a probability of .9 and our results do show mean bids equal to expected value. However, at odds of .01 of a $4 loss, enormous risk aversion <u>could</u> raise mean bids to 2.5 times expected value as our data indicate. This possibility is excluded by the results from our second experiment with a loss of $40 as follows: To assume risk aversion is sufficient to explain our results for a $4 loss in wealth implies a very highly curved utility function in the neighborhood of the current level of wealth. If the size of the loss is increased from the $4 level to $40 level, as is done in our second experiment for odds of loss of .01, then given such a highly curved utility function, the effect of risk aversion should be to dramatically increase the ratio of bids to expected value above the 2.5 obtained in the $4 loss experiment. This does not happen. Rather, Figure 3.7 looks almost identical in pattern to the top panel of Figure 3.2. In other words, the frequency distribution of bids divided by expected value is about the same in the $4 loss case as it is in the $40 loss case. This is very strong evidence that relative risk aversion plays almost no role in our experiments. The enormous degree of risk aversion necessary to use expected utility theory to explain the 2.5 ratio of mean bid to expected value obtained at a .01 probability with a $4 loss implies that the frequency distribution shown in Figure 3.7 ($40 loss) should be far to the right compared to the frequency distribution shown in the top panel of Figure 3.2 ($4 loss). In fact, the slight rightward shift shown in Figure 3.7 compared to Figure 3.2 is much more likely attributable to the fact that we began our second $40 loss experiment with no prior auction experience for subjects to simulate what would happen when people faced a new risk. In contrast, subjects facing the $4 loss at .01 odds had prior experience with 4 practice trials plus 10 binding trials at odds of .2 plus 10 binding trials at odds of .1. In other words, learning, not risk aversion, likely explains any increase in bids relative to expected value in the $40 versus $4 loss experiments (both at .01 odds).

A second demonstration of the inability of risk aversion to explain our results comes from empirical studies of the coefficient of relative risk aversion. Following the notation used above, utility is U(Y) where Y is wealth. If we assume risk aversion, then $U''<0$. The coefficient of relative risk aversion is defined as $c=(-U''/U')\cdot Y$, a positive number. The empirical evidence on the coefficient of relative risk aversion has recently been summarized as follows:

> In particular, Cohn et al. (1975) found evidence that the coefficient of relative risk aversion is slightly decreasing in wealth. Friend and Blume (1975) found that "if there is any tendency for increasing or decreasing proportional risk aversion, the tendency is so slight that for many purposes the assumption of constant proportional risk aversion is not a bad first approximation" (p. 915). More recently, Morin and Suarez (1983) found the coefficient to be slightly decreasing for wealth levels up to $100,000, after which it becomes approximately constant. Furthermore, Friend and Blume estimated the market price of risk to determine a value for the coefficient, which they argue is greater than one and may be as high as two.[8]

In contrast to the field studies cited above, what value for c is implied by our experiments if risk aversion is to account for the observed increase in mean bid for insurance relative to expected value at a probability of loss of .01? Using the same notation as before, where we defined

$Y°$ = initial wealth,

P = odds of loss (red chip),

L = size of monetary loss,

B = bid for insurance against loss,

(1-p) = odds of gain (white chip),

and G = size of monetary gain,

---

[8]"A Test of the Expected Utility Model: evidence from Earthquake Risks," by Brookshire, Thayer, Tschirhart and Schulze, JPE, 1985 Vol. 93:2, p. 381.

the true bid for insurance can be obtained in an expected utility framework by setting the expected utility of paying and obtaining insurance for $B equal to expected utility without insurance:

$$p\,U(Y^\circ-B)+(1-p)U(Y^\circ+G-B)=pU(Y^\circ-L)+(1-p)U(Y^\circ+G).$$

To incorporate the coefficient of relative risk aversion, we substitute a second order Taylor series approximation of U(Y) into the expression above and obtain (where expected value = $EV=p\cdot L$) an expression for bid divided by expected value (B/EV):

$$B/EV=1+\frac{c}{Y^\circ}\left[\frac{1}{2}(1-p(B/EV)^2)L+(1-p)(B/EV)G\right].$$

Note, that if no risk aversion is present, c=0, and B/EV=1. This is the assumption used in our prior analysis and, for losses of $4 and $40, can be justified as follows: The largest known value from field studies for c is about 2. In these studies, Y° = $100,000. Thus, the term in square brackets above is multiplied by 2/$100,000. In our $40 loss experiment at odds of .01 (where the effect of risk aversion should be greatest in our experiments) the relevant observations on p,L, G and B/EV are .01, $40, $1, and 2.5 respectively. Using these values in the term in brackets above yields a value for that term of about 21.2. Multiplying c/Y° times this in the formula above gives a B/EV of 1.0004. Obviously this value of B/EV is inconsistent with our experimental results and with our use of B/EV=2.5 in the r.h.s. of the formula above. The quadratic formula can be used to solve for B/EV assuming c=2 and gives values of B/EV negligibly different from unity. Clearly, existing field evidence on risk aversion justifies our assumption of risk neutrality for losses of $4 and $40 used in our experiments.

Another approach is to ask what could c have to be to explain our results.

The expression above can readily be solved for c in terms of Y° B/EV, p, L and G. Using data from our $4 loss experiment at .01 odds gives c=33,333 (assuming Y°=$100,000, B/EV=2.5, p=.01 L=$4 and G=$1). For the $40 loss experiment c=6,667 (assuming Y°=$100,000, B/EV=2.5, p=.01, L=$40 and G=$1). Thus, not only must risk aversion take on absurd levels, but the degree of risk aversion must be much larger for small losses than for large losses. As noted above, field studies of risk aversion have shown c to be relatively constant for large changes in wealth. In summary the empirical evidence on the coefficient of relative risk aversion suggests that risk aversion plays no role in our experiments. This is consistent with our earlier argument that the increase in loss from $4 to $40, did not shift the frequency distribution of B/EV to the right.

3.6 Conclusion

The principal objective of the experiments reported in this chapter was to explore insurance behavior in a laboratory market-like environment where the probability of loss was varied. Thus, the predictions of expected utility theory as well as models from psychology could be compared against actual behavior at both higher and lower probabilities of loss. Additionally, repeated trials were included in the experiments so that the effect of experience on decision making could be determined. The results of the experiments suggest that although expected utility theory is an adequate explanation of behavior at higher probabilities of loss, at lower probabilities a much more complex model is required to explain observed behavior. This complex model has been evolving within psychology principally under the guise of prospect theory and includes features such as the editing phenomenon and the anchoring and adjustment process documented in our results. Additional results of our experiment include, at low probabilities, a large gambler's fallacy effect and strong bimodality. These results are consistent with the direction and spirit of prospect theory. Further, they serve to reinforce our general conclusion that models which arise from psychology and which consequently focus on the mental processes and possible errors in those processes are central to any explanation of economic behavior motivated by low probability events.

Although it can be argued that markets themselves seem to promote behavior consistent with expected utility, they do not seem to help very much for low probability, uncertain situations, at least within the range of experience observed in our experiments. This implies that decision making at low probabilities is likely to be subject to error even in a market context. Individual responses to threats from low probability hazards such as Superfund sites are likely to suffer from the entire litany of cognitive difficulties identified above.

Although behavior differs from predictions of expected utility theory due to these cognitive difficulties, it is not appropriate to call behavior at low probabilities irrational. Anyone attempting to consider seriously the myriad of low-probability natural and technological hazards would quickly be overwhelmed and paralyzed with indecision. Therefore, it is rational to edit away many hazards that appear to be unlikely and to concentrate only on those that appear to be somewhat more likely. For those risks that are worth considering, the anchoring and adjustment process may produce estimates that are in many cases "close enough" in the sense that additional cognitive effort would not generally be worth its cost. This viewpoint is similar to that expressed in the recent paper by Russell and Thaler (1985). However, for very low probabilities and very high losses the intuitive reasoning that leads to the bimodality shown in our results has important implications for public policy.

For example, consider a controversy about whether a landfill containing toxic materials needs to be cleaned up when the scientific estimate of the risk is low. An application of the results from this study would lead us to expect that some residents living near the landfill would edit and therefore dismiss the risk. Other residents would consider their response to the situation by anchoring on the losses, which could be extreme such as cancer and birth defects, and then adjusting downward. Given such extreme anchors, the judged levels after insufficient adjustment, even if the percentage misadjustment factor is small, are likely to be quite high. This produces two groups of residents who disagree strongly about what needs to be done. One group complains that the risk is negligible and that all the fuss will only lower property values while the other group cannot understand why the former group is not concerned about the deadly risk confronting them all.[9] Chapter 4 shows that a drop in property values near a hazardous waste site seems to be associated with a bimodal distribution of risk beliefs very similar to that shown in the laboratory results presented here.

---

[9] Note that we obtained this bimodality in the laboratory where subjects were able to perceive the risk directly by actually viewing the number of red and white chips put into the bag. For a real risk where such direct perception of the risk is not possible, there is likely to be wide individual variation in the estimation of the risk probability which will likely exacerbate the difficulties.

CHAPTER 4:

A FIELD STUDY OF FACTORS AFFECTING RISK BELIEFS

4.1 The Policy Dilemma of Superfund Sites.

Superfund sites have created intense public concern both for residents living near such sites and for the general public. For people who live near a Superfund site, fears of cancer or other health problems are very real, and such fears genuinely reduce the quality of life. Many residents living near a Superfund site judge that the health risk is substantial. In contrast, experts often judge the risks from particular sites to be very small, or at least significantly smaller than they are in the judgment of the residents.

The discrepancy between the large subjective risk that the public believes is at issue and the small risk experts believe is scientifically founded creates a policy dilemma for institutions concerned with risk management. Should large sums of money be spent cleaning up Superfund sites when experts' judgment of the risk imposed by the site to be small? Is the harm to residents near such sites in some sense real even if health is not actually affected adversely? If a Superfund site is not cleaned up because the expert assessments of risk indicate only a small risk and the local population still believes the site to be harmful, has a disservice been done? Can a community's beliefs change to better reflect what is actually known about possible risks?

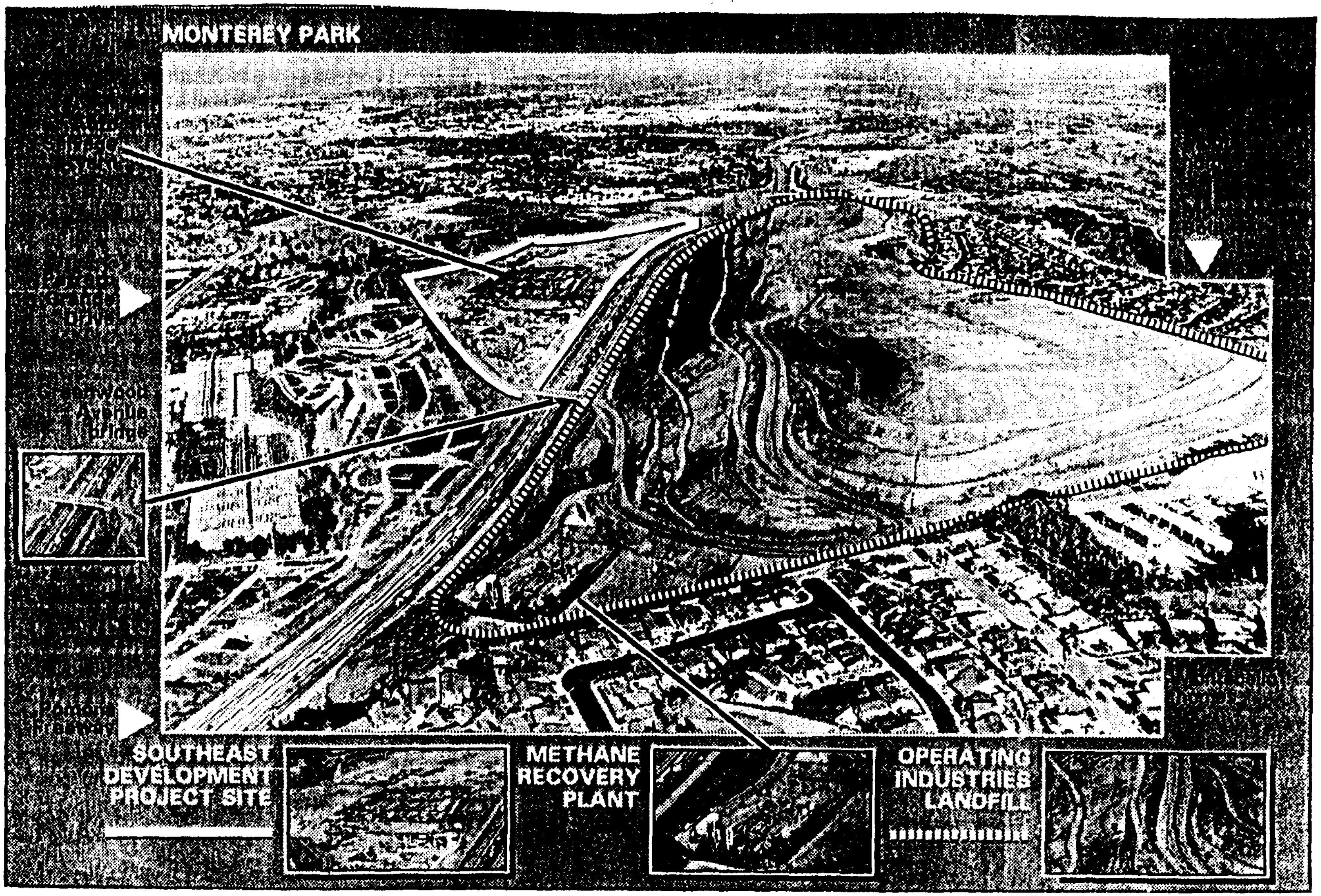
The study summarized here explores these issues in the context of a particular landfill site located near a large number of homes. Expert estimates of the risks associated with this site were low, but nearby residents were quite concerned about the effects of the landfill on their health. We modeled housing prices to determine whether the residents' concerns were measurable in terms of economic damages, and we surveyed residents to identify the sources of these economic and psychological damages.

4.2 The Site.

As background for our survey and study of property values, we briefly describe the site and its history. The Operating Industries Inc. (OII) Landfill is located between the communities of Montebello and Monterey Park in the Los Angeles, California metropolitan area (See Figure 4.1). OII opened in 1948 as a municipal landfill and began accepting hazardous wastes in 1976. 011 stopped accepting hazardous materials in January, 1983; and in October, 1984 the landfill reached its capacity and was closed. At that time the OII Landfill was proposed for inclusion on the National Priorities List for "Superfund" monies. The landfill covers 190 acres and contains approximately 30 million cubic yards of refuse, which is generating sufficient landfill gas (methane) to be commercially extractable.

Several land use and policy changes have affected the site and the surrounding area. During the early 1970s, the city of Montebello approved development plans for residential housing along the southern edge of the landfill. Original plans were to reclaim the landfill area and to build a golf course and park. The housing development coincided with several other land use changes in the area, including the construction of the Pomona Freeway, which disects the OII Landfill. Construction of the freeway restricted activities at the landfill to the area of the site south of the freeway. As compensation for this loss of area, the height restrictions at the landfill were relaxed. This increase in the height limitation has been linked to increased erosion problems including slope failure and mudslides,

FIGURE 4.1



KEN LUBAS and MICHAEL HALL / Los Angeles Times

Aerial view eastward along Pomona Freeway through Monterey Park outlines landfill site nominated for federal Superfund list.

which have exposed decaying refuse.

Soon after residents occupied the newly constructed homes in the mid 1970s, they began to complain of odors to the office of the South Coast Air Quality Management District. Complaints of rodents and leachate pooling off-site accompanied the odor problems. Additional wells for collection of landfill gas and better leachate control systems have been installed since 1983 to mitigate odors and reduce risks. In early 1985, the U.S. Environmental Protection Agency began feasibility studies of further remedial measures.

In 1979, some residents of the immediate area formed a group called Homeowners to Eliminate Landfill Problems (HELP) to organize their efforts to eliminate odor and health and safety problems emanating from the OII Landfill. HELP has a membership of approximately 460 dues-paying families. Issues on the HELP agenda include possible health problems associated with the site, leachate disposition, migrating gas, landfill use after closure, and property devaluation. Media attention at the site has been intense over the past several years. Television, radio and regional newspaper coverage has accompanied local coverage from newspapers, community meetings and an EPA newsletter, The OII Update. The nomination of the OII for the National Priorities List has also been a significant catalyst for media attention.

4.3 Expert Judgments of the Health Risk.

It is not our purpose either to defend or to criticize the scientific studies of the OII Landfill or the expert judgment of its risk to health. In this section we simply want to document what that judgment is so that we can compare it to the judgments of the residents living near the OII site.

The scientific studies and expert judgment have sought to answer two questions about the health risks of OII: Are there any known hazardous chemicals emanating from the site that can be detected in the surrounding neighborhoods? Are there any demonstrable ways in which the health of current residents differs from the health of people living in nearby control communities? We consider each question in turn.

The Regional Water Quality Control board has monitored the ground water supply continually since 1976 and has found no evidence of contamination. In April, 1983 the off-site level of vinyl chloride, a carcinogen, was measured at 19 ppb, which exceeds the California regulatory level of 10 ppb. However, workers experiencing exposures 170 to 500 times these levels have not experienced health problems and more recent random samples of air within homes showed no detectable levels of vinyl chloride gas (above 2 ppb). No other hazardous chemicals have been detected in appreciable quantities in off-site air monitoring. Thus, Satin, Huie, and Croen (1986) in a study conducted by the California Department of Health concluded that "the recent environmental monitoring of the area indicates that with the levels of chemicals found, long-term (health) problems would not be expected to occur."

The one potentially serious carcinogen detected off-site is vinyl chloride. Calculations of the cancer risk from exposure to vinyl chloride in the highest amounts detected therefore provide an upper bound to the risk. The USEPA (1985) Carcinogen Assessment Group estimated that the unit risk for exposure to vinyl chloride over a 70-year lifetime at a concentration of 1 mg/m^3 is $2.6x10^{-6}$. The concentration of 1 mg/m^3 corresponds approximately to a concentration of .38 ppb, which, if adjusted linearly from a 70-year exposure to a one-year exposure, is equivalent to 26.9 ppb, 140 percent of the highest monitored level of OII. Dividing the EPA unit risk for exposure by 1.4 gives an estimated annual risk of 1.86 x 10^{-6} for residents exposed at 19 ppb for one year. However, residents nearest OII have lived there as long as nine years (since 1977). The maximum cumulative risk might then be as much as nine times higher or 1.67×10^{-5}. It should be noted that these risk calculations are likely to be serious overestimates because the 19 ppb level represents the highest 24-hour average value ever obtained near the site and because vinyl chloride concentrations have been below detection since then. Thus, our assumed exposure of 19 ppb cannot be characterized as typical for any individual living near the site but rather represents an extreme upper bound for possible exposure.

A second approach to assess the health risk has been to compare health status of residents living near OII to others living in the Los Angeles metropolitan area. A study conducted by the Los Angeles County Department of Health Services in 1983 concluded that no consistent pattern of absences from school had occurred around the landfill. Nearby residents had not suffered excess mortality, nor had they experienced more adverse outcomes of conception than had residents in other parts of Los Angeles County. Of course, current epidemiological studies may not indicate serious health effects that may arise in the future because of, for example, the long latency periods for many types of cancer.

The California Department of Health has conducted a survey of residents living near the OII Landfill and residents of comparable control communities approximately ten miles away. There were no statistically significant differences between the OII area and the control communities in terms of mortality or increased incidence of adverse pregnancy outcomes, cancer, and liver disease. There was a statistically significant difference in self-reports of headache, sore throats, sleeping problems, eye and skin irritation, and feeling tired (see California Department of Health Services, 1986). These reported health problems were greater in those neighborhoods near OII where odor was more frequently a problem. However, toothaches were also more frequently reported in neighborhoods located near OII. Because there is no known biological mechanism for toothache involving any of the possible toxic chemicals at the landfill site, this finding suggests that residents may have simply monitored their health more carefully or just remembered these minor health problems better because they were aware of the possible association with OII.

In summary, although the OII Landfill is not a pleasant place, there is no indication that it has caused serious health problems, nor is there reason to believe, based on water and air monitoring, that there are likely to be major health problems in the future. The possibility does remain that there is some as yet undetected toxic chemical associated with one odor from the landfill.

4.4 Residents’ Judgments of the Health Risk.

In the fall of 1985, we conducted a mail survey to gather judgments of health risk from people living near the OII site. From maps, reverse telephone books provided by Pacific Bell, and records of real estate transactions, an address list of 1912 residences near OII was constructed. Surveys were mailed to all 1912 known addresses. Using standard follow-up reminders (Dillman, 1978), we obtained responses from 768 residents, which after adjustment for bad addresses represent 45 percent of the original sample.

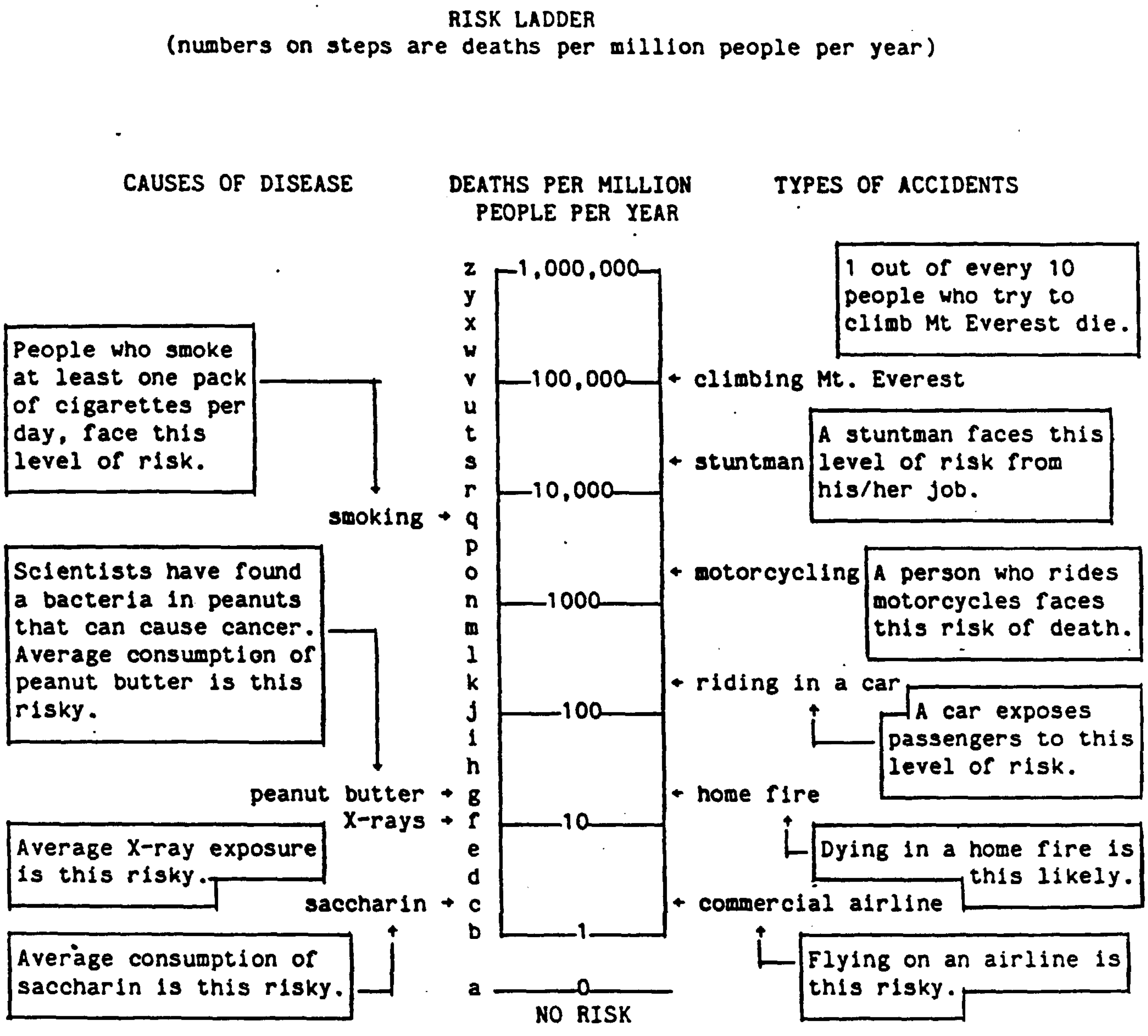
The survey questionnaire assessed residents' beliefs about health and safety risks, odor problems, sources of information about the site, and attitudes towards local, state and federal officials, the news media and landfill operators. The questionnaire also included standard sociodemographic questions.

On a "risk ladder" (see Figure 4.2) respondents matched their belief about the risks they faced from the OII Landfill to specific levels of risk defined in terms of the probability of death. Respondents reported retrospectively their belief about risk before site closure as well as their current belief about risk after site closure.

Figure 4.3 shows on log scales the frequency distributions of subjective health risk both before and after closure of the 011 Landfill. There are two striking features of the frequency distribution of subjective health risk before closure of the site. First, there is a wide diversity of opinion; every category on the risk ladder received responses. Second, the distribution is bimodal with a sizable proportion of the respondents estimating the risk around 10^{-3} and 10^{-2}, approximately the risk of smoking at least one pack of cigerattes per day and another segment of the sample clustering around estimates of the risk between 10^{-5} and 10^{-6}, approximately the risk from the average consumption of saccharin. In other words, some residents believed the risk to be very large whereas others judged the risk to be very small.

In Chapter 3 we obtained a similar bimodal distribution of responses in a laboratory study of risk decision making with low probability risks. The distribution from the laboratory experiment was very similar to the

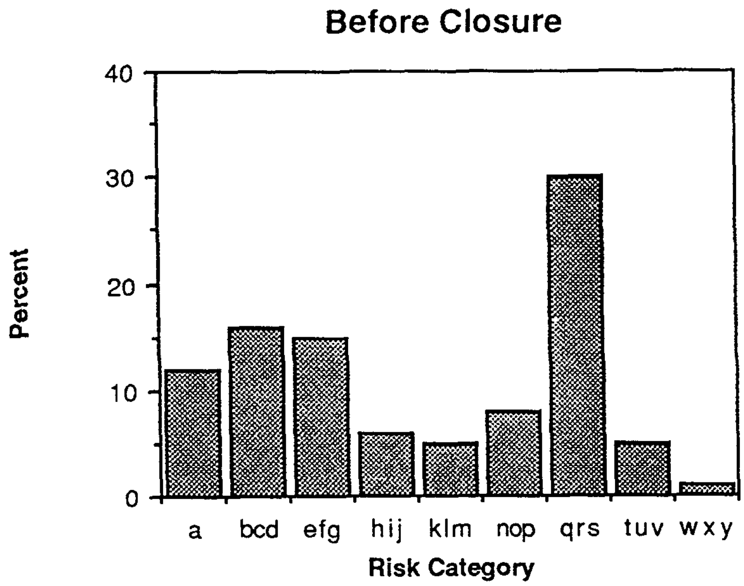
FIGURE 4.2

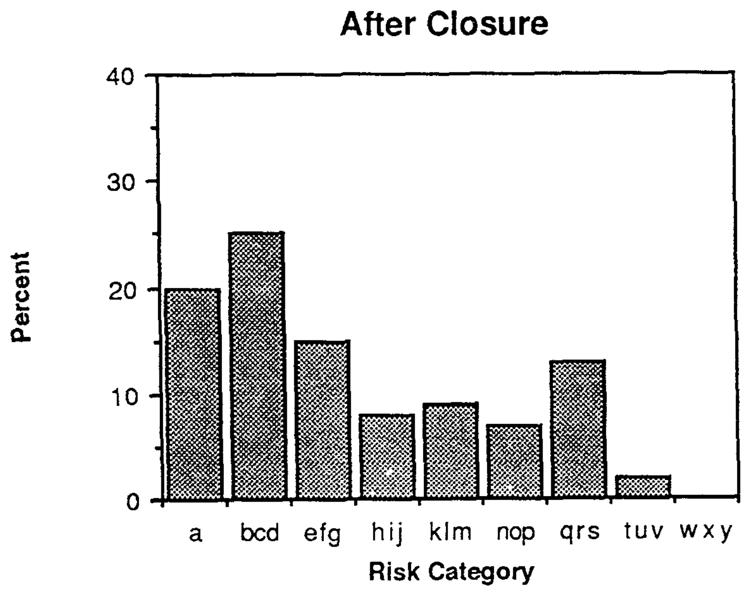


RISK LADDER
(numbers on steps are deaths per million people per year)
CAUSES OF DISEASE
DEATHS PER MILLION PEOPLE PER YEAR
TYPES OF ACCIDENTS
z y x w v u t s r q p o n m l k j i h g f e d c b a
1,000,000
100,000
10,000
1000
100
10
1
0
NO RISK
1 out of every 10 people who try to climb Mt Everest die.
climbing Mt. Everest
A stuntman faces this level of risk from his/her job.
stuntman
motorcycling
A person who rides motorcycles faces this risk of death.
riding in a car
A car exposes passengers to this level of risk.
home fire
Dying in a home fire is this likely.
commercial airline
Flying on an airline is this risky.
People who smoke at least one pack of cigarettes per day, face this level of risk.
smoking
Scientists have found a bacteria in peanuts that can cause cancer. Average consumption of peanut butter is this risky.
peanut butter
X-rays
Average X-ray exposure is this risky.
saccharin
Average consumption of saccharin is this risky.

FIGURE 4.3

SUBJECTIVE HEALTH RISK





Annual Risk of Death:
a = no risk
b = one in 9 million
f = one in 100 thousand
j = one in 10 thousand
n = one in one thousand
r = one in one hundred
v = one in ten

distribution in Figure 4.3(a). Other researchers have also observed a similar bimodality. As Slovic, Fischhoff, and Lichtenstein (1981) have noted, "people often attempt to reduce the anxiety generated in the face of uncertainty by denying the uncertainty, thus making the risk seem so small it can safely be ignored or so large that it clearly should be avoided." It therefore appears that bimodality of risk judgments characterizes responses to low probability risks: some people "edit" the risk away while others may exaggerate its importance.

Figure 4.3(b) shows the frequency distribution of beliefs about risk after closure of the site. The bimodality, although present, is much less pronounced than for judgments before closure, and the judgments of risk are in general lower.

For purposes of subsequent analysis, we constructed a new binary health belief variable that indicated the mode for each respondent's judgment of risk. The dividing line between the two modes is the letter L on the risk ladder, approximately $5x10^{-4}$ per year. Approximately 51 percent of the sample was in the high health risk mode before closure.

4.5 Comparison of Expert and Resident Health Risk Beliefs.

The epidemiological studies have found virtually no health risk, so from that perspective almost all the respondents believe the health risk to be higher than the expert estimate. If we use the calculated risk for the one extreme recorded exposure of vinyl chloride as an upper bound for the scientific risk ($1.67x10^{-5}$) then the half of the sample in the upper health risk mode (the cut point equals $5x10^{-4}$) overestimates the health risk by at least one order of magnitude. The bimodality also implies that whatever the true risk, approximately half the respondents seriously misestimate that risk. Either those in the high mode are greatly overestimating the true risk or those in the low mode are greatly underestimating the true risk. For all these reasons it is reasonable to include that the subjective health risk beliefs for many respondents differ substantially from the expert judgments.

A comparison in Figure 4.3 of the frequency distributions of beliefs of risk before and after closure of the site also suggests another way in which the beliefs of residents are inaccurate. There was a substantial believed reduction in health risk as a function of site closure. However, at the time of closure OII had already stopped accepting hazardous material and whatever toxic chemicals may have emanated from the site before closure would be just as likely or even more likely to emanate from the site after closure. Thus the aggregate reduction in subjective judgments of the health risk cannot be accurate.

4.6 A Model of Subjective Beliefs About Risk.

The great variation in estimates of subjective health risk suggests that those judgments must be due in part of psychological and sociological factors other than a perception of the true health risk. It is therefore interesting to model subjective health risk judgments using psychological and sociodemographic variables assessed in the survey. Potential variables for inclusion in the model are described below in conceptual groups.

Experiential Variables. The more that experience has made one aware of the potential health problems from the landfill, the higher one’s estimate of the health risk is likely to be. Thus, the model includes variables which assess awareness of the potential problem through several sources. In particular, the model includes respondent awareness of media attention to the problem and perception of odor from the site as experiential variables. Also included is geographic distance to the site as a proxy variable for other experiential effects. Presumably, those respondents who live near the landfill will have had more visual reminders of the potential health hazards.

Sociodemographic Variables. Judgments of health risk may vary as a function of various sociodemographic variables. For example, older respondents will have necessarily survived a number of hazards and may therefore place the present landfill risk in a different context than a younger respondent who is raising children. Although we do not have specific hypotheses about the risk effects of these variables, we examine income, education, age, gender, number of children living at home, occupation, and ethnicity as possible components in a model of health risk judgments. It is possible to examine ethnicity because of the high proportion of Asian-Americans in this sample.

Site Closure Variable. As already noted, the mean judgment of the health risk was lower after closure of the site to further dumping. We therefore include an indicator variable to mark whether the estimated health risk is for before or after closure of the site.

Health Risk Dependent Variable. The strong bimodality in the distribution of health risk judgments suggests that the error from any model of those judgments would be unlikely to meet the usual distributional assumptions necessary for statistical tests. Also, we are more interested in the correlates of which mode a respondent is in father than the minor variation within each mode. So, the subjective health risk scores were recoded to reflect mode: those in the upper mode received a score of 1 and those in the lower mode received a score of 0. This recoding does not solve all the problems with the error structure because ordinary least squares (OLS) analysis of binary data can be problematic. We therefore perform both OLS and PROBIT analysis. Computer limitations constrained the number of variables possible in the probit analysis with this many observations, so we used OLS to screen variables for inclusion in the probit analysis.

4.6.1 Model for Health Risk Judgments.

Table 4.1 gives the partial regression coefficients and their associated t statistics for both the OLS and PROBIT analyses. Both analyses produced exactly the same conclusions. We therefore discuss the results in terms of the OLS regression because it is generally easier to understand. It should be remembered that the statistical tests are for partial regression coefficients. That is, the test asks whether the given variable reliably explains a portion of the variation in health risk after controlling for all the other variables included in the model. With covariation among the predictor variables this can produce conservative conclusions about the importance of a variable.

As expected, the site closure variable is a statistically significant component of the model even after controlling for all the other variables. All three experiential variables had significant coefficients. Odor in particular stands out as an important predictor of subjective health risk. Distance from the site was also a significant predictor after controlling for odor perceptions. Thus, there must be other perceptions or concerns associated with distance besides the perception of odor which affect judgments of health risk. Frequency of exposure to media attention about the site also predicted increased health risk judgments.

TABLE 4.1

Regressions Explaining Subjective Health Risk
Before Closure of the Operating Industries Inc. Landfill

Variable Name	Mean	Std. Dev.	Estimated Coefficients (t in parentheses) OLS	Probit
DEPENDENT VAR.				
Subjective Health Risk (1 if in upper mode 0 if in lower mode)	0.415	0.49		
INDEPENDENT VAR.				
Constant			0.57 (1.78)	-0.67 (-2.26)
Closure Dummy Var. (1 before closure) (0 after closure)	0.52	0.50	0.094 (2.77)	0.29 (2.681
Experiential Var.				
Frequency of hearing or reading about OII problems.	4.11	0.96	0.037 (2.16)	0.14 (2.61)
Perceived odor problems	16.45	14.35	0.013) (9.83)	0.040 (9.44)
Distance from site (blocks)	11.50	7.07	-0.0083 (-3.53)	-0.028 (-3.78)
Socio-Economic Var.				
Number of people under 18 living in house	0.91	1.05	0.047 (2.64)	0.12 (2.27)
Age of respondent	48.48	12.63	-0.0035 (-1.98)	-0.0097 (-2.14)
Income	47.631	22,038	0.354E-6 (0.45)	-
Sex of respondent (0 female) (1 male)	0.79	0.41	-0.12 (-2.91)	-0.31 (-2.52)
level of education (1-9)	6.34	1.91	0.0019 (0.18)	-
Occupation Var.				
(Sales or Managerial = 1; service. Repair. Labor, or Farm/Fishery = -1; Retired = 0)	0.39	0.84	-0.00078 (-0.038)	-
Ethnic Var.				
#1 (Caucasian = 2; Asian or Hisp. = -1)	-0.17	1.28	0.00076 (0.056)	-
#2 (Caucasian = 0; Hispanic = -1; Asian = 1)	0.22	0.79	0.030 (1.45)	-
Sample Size			762	
R^2			0.282	
Likelihood Ratio Test				238.87

It is important to recognize that a cross-sectional survey such as this must necessarily suffer from causal ambiguity. For example, we have included frequency of exposure to media attention as a predictor of health risk judgments. However, it might be the case that someone who becomes concerned about the health risks will pay more attention to and seek out media reports about the problem. Similarly, someone who is concerned about the health risk may be more alert for the odor problem and hence report having experienced it a greater number of times.

It is interesting to ask whether sociodemographic variables can explain variation in health judgments over and beyond the variation attributable to the more direct experiential and perceptual variables. Having statistically controlled for the experiential variables, any effects of sociodemographic variables represent largely attitudinal effects. The two socioeconomic status variables of income and education had inconsequential effects. Thus, it is not true that those who had more to lose economically were more concerned about the risk. However, the number of children living at home was a significant predictor so in that sense those who had more to lose were more concerned about the risk. Age of respondent is obviously correlated with having children living at home but age predicted variation over and above that variable. The direction of the effect is that younger people thought the hazards of the site were more risky. Gender also made a significant difference with females believing the site is more risky than did males. A coded variable contrasting managers and sales people against service, labor, and repair occupations (those in the latter group are presumably exposed to more on-the-job risks) indicated no differences in risk judgments. Similarly, two variables coding ethnic group (one contrasting Caucasians with Asian-Americans and Hispanics and one contrasting Asian-Americans with Hispanics) yielded no significant differences. There are, therefore, no suggestions in this sample of any occupational or cultural differences in the evaluation of risk.

A reasonable model of judgments of the health risks associated with the OII Landfill site includes the following components: site closure, media exposure, odor, distance to site, number of children living at home, age, and gender. This model accounts for approximately 28 percent of the variation in the coded health risk variable. This is substantial for a model of this type, especially given that the dependent variable is binary. What does the model mean? First, the importance of the perceptual odor variable above and beyond the other variables is striking. It is easy to speculate that without vivid, perceptual cues from the site, risk judgments would be greatly reduced. More important than the specific pattern of significant coefficients, however. are the following conclusions: (a) there is great variability and bimodality in judgments of health risks; (b) many respondents have inaccurate beliefs about the extent of the health risk; and (c) the variation in health risk judgments is not random but can be related to systematic differences among respondents.

4.7 Real Estate Markets Around OII.

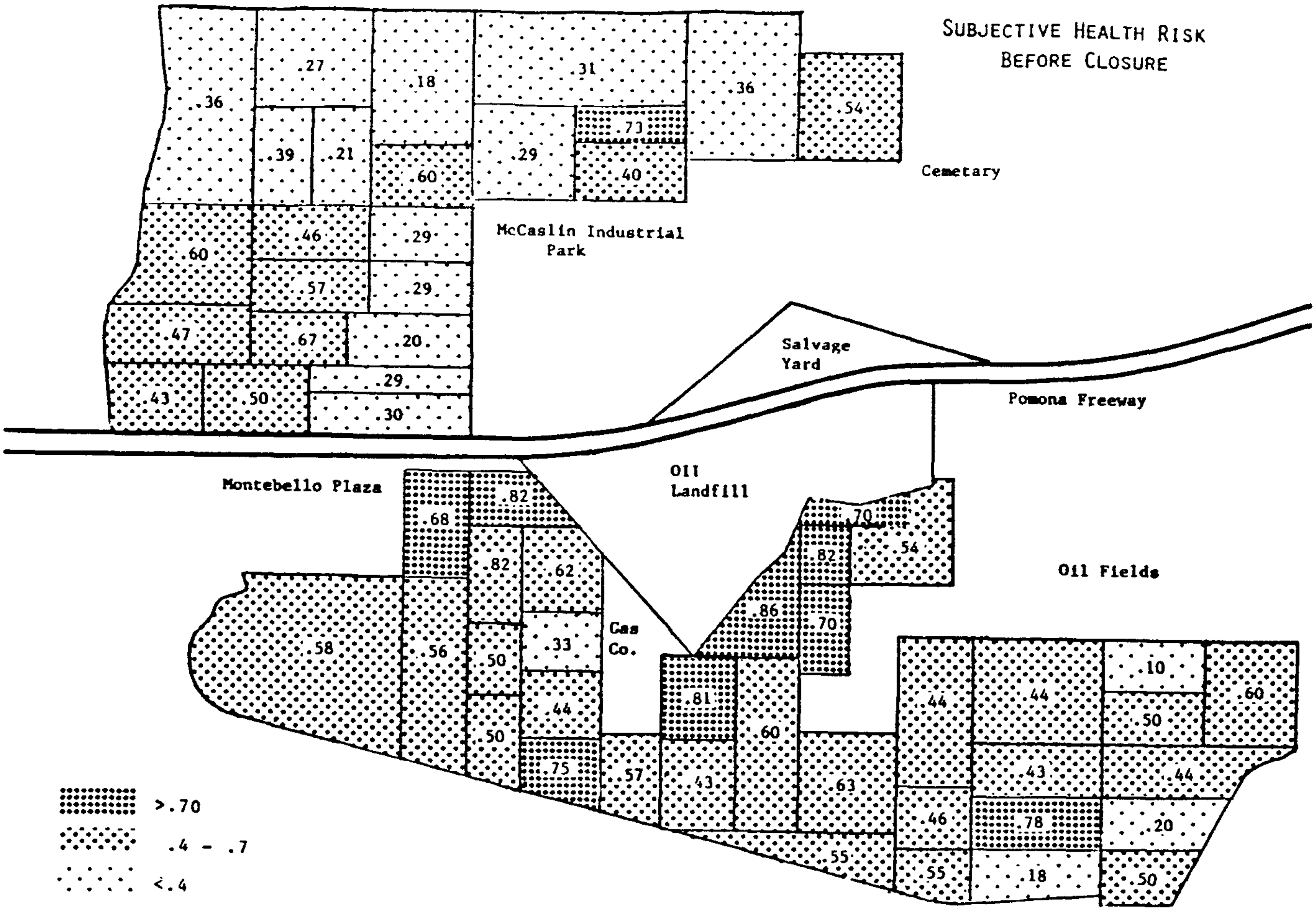
In this section we analyze the role of perception and risk judgments on the real estate market in the area around the OII Landfill. However, in analyzing the real estate market near the OII Landfill, individual perceptions and attitudes are of less importance than the collective perceptions and attitudes of individuals residing in neighborhoods in the vicinity of the waste site.

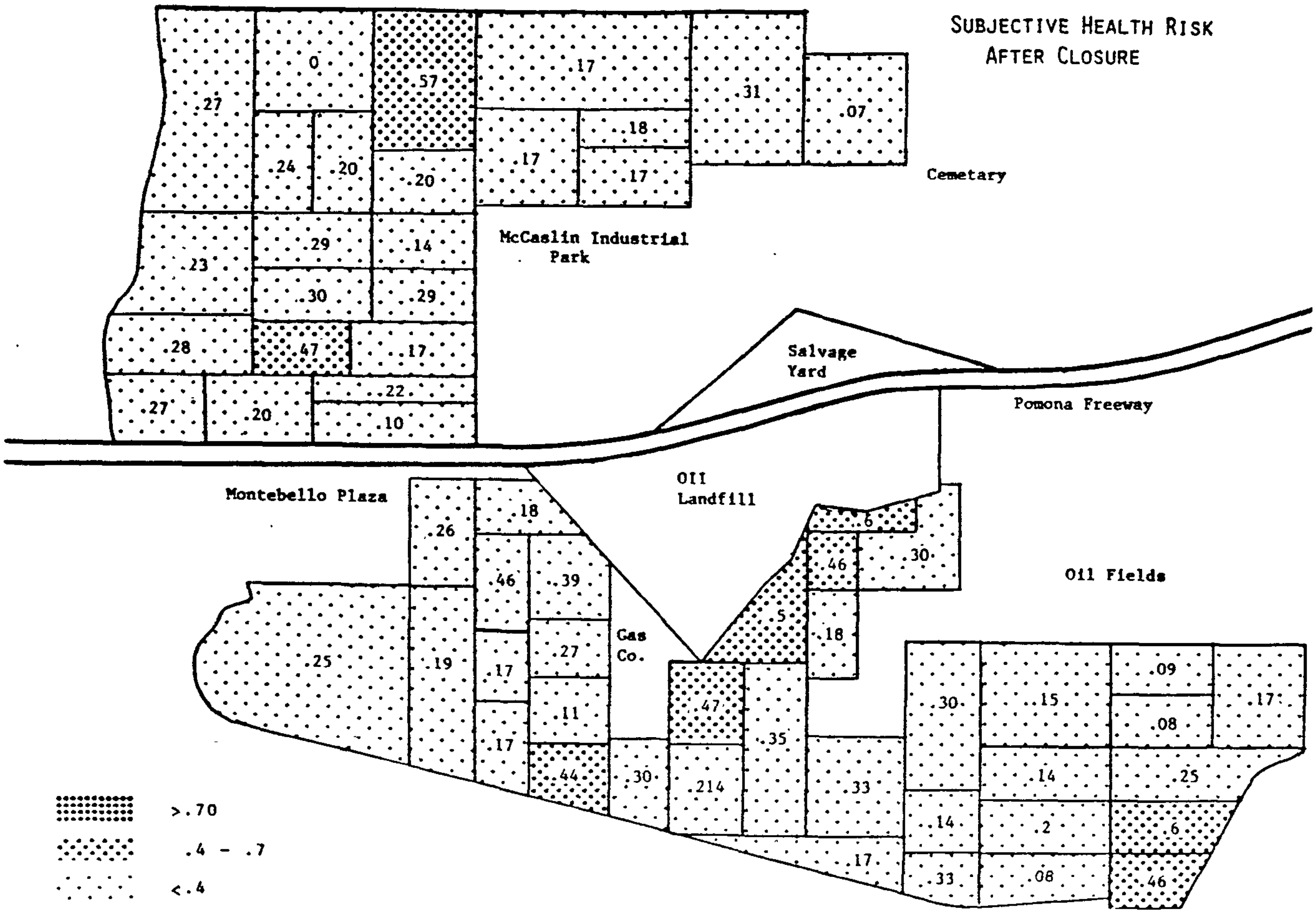
Although residents may well be willing to sell at a price adjusted downward by their willingness to pay to avoid any subjective risk associated with proximity to the OII Landfill, they are likely to list homes, after consulting a realtor, at the "going" market rate. Thus sellers will attempt to obtain a price higher than their actual willingness to sell. In effect, sellers will try to obtain some consumer surplus as is normal in all competitive markets. In fact, in a neighborhood the supply curve will be shifted to the right to the extent that homeowners within a neighborhood feel that the OII poses a risk. Thus, the greater the percentage of homeowners in a neighborhood who feel that the OII Landfill poses a threat, the further the supply curve will be shifted to the right relative to an initial supply curve that assumes no homeowners in the neighborhood feel threatened by the site. Thus, the observed price for homes in a particular neighborhood will fall as more homeowners in a neighborhood feel threatened. Unfortunately, we have no information on the subjective risk beliefs held by potential purchasers who make up the demand curve; but note that sixty-two percent of recent purchasers were not aware of the site when they bought their homes, despite local requirements for information disclosure to new buyers. Those that were aware may, of course, have lowered their offered bids, shifting the demand curve downward to the left, causing a further decline in observed prices. Since we have no data on subjective risks by neighborhood for prospective purchasers, we must assume that the subjective risk of residents measured for each neighborhood around the OII Landfill can proxy for that of purchasers in our reduced form estimated property value equation. Thus, we focus on the development of neighborhood rather than individual measures of perceived odor problems, subjective health risk and explosion risk. In order to provide the spatial distribution of the key variables for the property value study, we plotted households responding to the survey on an aerial photograph of the area. Using the aerial photograph, the area around the site was divided into neighborhoods with about 10 to 15 data points in each neighborhood. Having identified responses within a given neighborhood, perception characteristics can be attributed to homes sold in neighborhoods and used as independent variables in the property value study. Figure 4.4 shows how judgment of health risk is spatially distributed around the site both before and after closure of the landfill. The number used for each neighborhood represents subjective health risk as the fraction of residents who lie in the upper mode of the bimodal distribution of risk perception. Therefore, values of the subjective risk measure will fall between 0 and 1, with neighborhoods having a high number of upper mode residents approaching 1 and neighborhoods with a low number of upper mode residents falling near 0. The Figures generally show that in neighborhoods closer to the landfill, the fraction of residents with a high level of health risk perception is larger. In the discussion that follows, the effects of perceptions and subjective judgments on property values is explored.

4.7.1 Property Values Near the OII Landfill.

Residents in the vicinity are troubled by a decline in the value of their property that they believe is caused by the location, size and the presence of hazardous wastes at the OII Landfill. The effects on property values are further aggravated by intensive media coverage that has tended to focus on the possible risks and the presence of odor problems, which has appeared to have strongly influenced perceptions and subjective judgments within the area.

FIGURE 4.4


SUBJECTIVE HEALTH RISK
BEFORE CLOSURE
Cemetary
McCaslin Industrial Park
Salvage Yard
Pomona Freeway
Montebello Plaza
Oil Landfill
Oil Fields
Gas Co.
.36 .27 .18 .31 .36 .54 .73 .39 .21 .60 .29 .40 .60 .46 .29 .57 .29 .47 .67 .20 .43 .50 .29 .30
.68 .82 .70 .82 .62 .54 .82 .58 .56 .50 .33 .86 .70 .81 .44 .50 .60 .44 .44 .10 .60 .50 .75 .57 .43 .63 .43 .44 .46 .78 .20 .55 .55 .18 .50
>.70
.4 - .7
<.4


SUBJECTIVE HEALTH RISK
AFTER CLOSURE
Cemetary
McCaslin Industrial Park
Salvage Yard
Pomona Freeway
Montebello Plaza
Oil Landfill
Oil Fields
Gas Co.
.27 0 .57 .17 .31 .07 .18 .24 .20 .20 .17 .17 .23 .29 .14 .30 .29 .28 .47 .17 .27 .20 .22 .10
.26 .18 .46 .39 .30 .46 .25 .19 .17 .27 .5 .18 .47 .11 .35 .30 .15 .09 .17 .08 .17 .44 .30 .214 .33 .14 .25 .14 .2 .6 .17 .33 .08 .46
>.70
.4 - .7
<.4

The Hedonic Price Method (HPM) attempts to value certain environmental amenities (or disamenities) by studying markets in which an environmental attribute may be captured (See Rosen 1974). In this case, the value that people hold for avoiding hazardous waste problems may be proxied by relative declines in the real estate market near the hazardous waste site. The model postulates that the value of a home is a function of the quantity and quality of certain physical attributes of the home and neighborhood including perceived environmental conditions. By estimating a reduced form property value equation, the relative role of each of the factors can be determined, including the relative importance of perceived environmental conditions in determining the value of homes.

We obtained property value data through a real estate information network. These data included home sales information and characteristics from August 1983 through November 1985 (which spans the closing of the OII Landfill late in 1984). Combining current property sales data from secondary sources with current perception and subjective judgment data from the survey has made it possible to construct a hedonic model to explore how perceptions and subjective judgments affect property values. As discussed above, subjective risk and perceived odor data were grouped into neighborhood variables.

Neighborhood subjective risk and perceived odor data are available for both before and after closure of the OII Landfill. Therefore, there arises a question about the timing of the shift from before to after closure risk judgments and perceptions. It was hypothesized that a lagged effect would be present and that before closure perceptions would persist (at least in terms of buyers moving to the area) past the date that the OII

Landfill actually closed. A six month lag was used, evenly splitting the period between the two points in time for which subjective risk and odor perception information was obtained. The OII Landfill officially ceased accepting additional wastes on the last day of October 1984, but home sales during the first 6 months following the closure were assigned the neighborhood subjective risk and perceived odor values that were present before closure. A linear functional form was used in specifying the equations because of the ease in interpreting the coefficients and because results obtained from alternative log forms were not significantly different.

4.7.2 Property Value Model.

In the secondary data set, 179 home sales were identified within the area near the OII Landfill during the 28 month period. The data was pooled in order that information on both before and after closure could be included in the analysis. Table 4.2 shows the results of four model specifications corresponding to the inclusion of subjective health risk, subjective risk from explosion, perceived odor and all three, respectively. The results suggest that subjective health risk may be the primary factor causing a decline in property values. With a coefficient of $-13,719 and a t-value of -1.80, it appears that the effect of subjective health risk is both significant and non-trivial. Neither subjective explosion risk nor perceived odor appears to be significantly contributing to the fall of property values. Considering the change in the size of the coefficient on subjective health risk from the first specification to the fourth, it appears that the multicollinearity between subjective health risk, subjective explosion risk and perceived odor is sufficient to cause

TABLE 4.2

HEDONIC PROPERTY VALUE REGRESSION
For Homes Near the Operating Industries Inc.
Landfill in Monterey Park. California

Variable Name	Mean	Std. Dev.	Estimated Coefficients (t in parentheses) 1	2	3	4
Dependent Var. Sale Price ($)	135,863	35,253				
Independent Var. constant			96,231.1 (8.26)	90,674.9 (7.72)	95,711.9 (7.65)	95,560.0 (7.70)
Subjective Health Risk[1]	0.41	0.20	-13719.8 (-1.80)	-	-	-22051.7 (-2.07)
Subjective Risk From[2] Explosion	8.43	3.26		-5.66 (-0.014)		865.8 (1.53)
Perceived Odor[3]	17.43	7.20	-	-	-184.1 (-0.95)	-88.9 (-0.35)
Date of Home Sale by month (08/83 - 1;08/85 - 25)	15.1	7.7	491.8 (2.70)	647.2 (3.83)	581.0 (3.29)	464.7 (2.52)
Area of Home (ft^2) (X - SqFt)	-0.041	475.5	50.63 (9.04)	49.61 (8.81)	50.61 (8.87)	51.09 (9.00)
" " " $(ft^2)^2$	224,807.3	262,400.7	0.021 (3.83)	0.0194 (3.61)	0.0191 (3.56)	0.019 (3.68)
Number of bathrooms	2.0	0.64	488.0 (0.12)	1,653.6 (0.41)	1,062.7 (0.27)	538.5 (0.13)
Year Home Built (i.e., 77, 84, 56)	58.8	9.8	523.4 (2.82)	454.0 (2.44)	457.7 (2.51)	499.3 (2.66)
Swimming Pool (0 if no pool; 1 if pool)	0.17	0.38	13,354.0 (4.00)	12,564.4 (3.76)	12,614.2 (3.79)	13,153.0 (4.19)
Scenic View From Home (0 if no view; 1 if view)	0.07	0.26	1 ,554.3 (0.31)	1,636.8 (0.33)	1 ,633.6 (0.33)	1,145.4 (0.23)
Fireplace in Home (0 if no fireplace; 1 if fireplace)	0.45	0.50	-603.4 (-0.21)	-1,219.5 (-0.42)	-883.9 (-0.30)	-502.2 (-0.17)
Proximity to Highway (1 if within 2 blocks; O otherwise)	0.06	0.24	-12.173.8 (-2.35)	-10.831.3 (-2.09)	-10,776.1 (-2.09)	-12,331.5 (-2.36)
R^2			0.802	0.798	0.799	0.805
sample size	179					

[1]This variable represents the fraction of respondents within a neighborhood who responded to survey Question 12 with a subjective health risk greater than 5 deaths in 10.000 (atop L). Homes sold prior to May 1985 were assigned a value corresponding to before closure subjective risk, and home sold after May 1985 were assigned the Corresponding risk value for <u>after closure</u> subjective risk.

[2]This variable represents a logarithmic scale from 1 (no risk) to 26 (certain risk) taken from responses to question 13 of the survey. Each neighborhood was assigned the mean value of responses within that neighborhood with home sold prior to May 1985 receiving the mean before closure value and home sold after May 1985 receiving the mean of the after closure value.

[3]This variable represents the product of frequency and intensity of perceived odor problems from responses to Question 11 in the survey. The resulting scale goes from 1 (very small problem) to 50 (very great problem) with homes sold prior to May 1985 receiving the mean neighborhood value before closure and homes sold later receiving the mean neighborhood value after closure.

sign changes in the coefficient on subjective explosion risk and to alter the coefficients on odor and subjective health risk. However, it is clear from the individual specifications that odor and risk from explosion are much less significant in explaining the observed property value decline. Other significant variables in the model include the date of home sale, the area of the home, the year the home was built, presence of a swimming pool, and the proximity of the house to the Pomona freeway.

4.7.3 Assessment of Total Subjective Damages Around the Site.

The coefficient on the effect of subjective health risk on property values, as identified in the econometric model, is $-13,719. To arrive at a total assessment of property value damage for the area, the total number of homes in each neighborhood cell was identified from an aerial photograph. This number was multiplied by the fraction of homes with a high subjective risk judgment in each neighborhood and by the coefficient on subjective health risk ($-13,719) and then summed over the sixty neighborhoods. This same procedure was followed using the after closure fraction of residents in the upper mode of subjective risk judgment to arrive at an after closure assessment of damages. The subjective benefits of closing the landfill amount to the difference between the before and after subjective damage assessments. The before closure estimate of subjective damages amounted to over $27 million for the 4100 homes near the site. After closure subjective damages amounted to $13 million resulting in a subjective benefit of closing amounting to $14 million.

These figures represent the magnitude of the real economic damages that residents in the area must bear because of property devaluation in the area of the OII Landfill. These figures also indicate the effect that closing the site may have had on property values and also suggest the magnitude of the potential benefits of better risk communication if, residents and potential home buyers could be convinced that the risk is truly small.

4.8 Changing Subjective Health Risk Judgments.

The evidence suggests that although the damages that have occurred to property values are real, the damages depend on subjective health risk beliefs which may change in response to factors other than objective risks. With effective risk communication measures and the further reduction of negative perceptual cues, property values may show a further recovery from these subjective damages. The relevant question becomes: Does mitigation of subjective damages require a complete and costly site cleanup or can other measures such as attempts to communicate objective risks along with more limited action to clean up the site provide a satisfactory solution?

It appears that large benefits can be obtained by changing subjective risk beliefs by communicating objective risk information to the public living near Superfund sites, and that these benefits may substantially exceed those from even eliminating objective health risks that may exist. In fact, community agreement that the problem has even been adequately addressed seems unlikely as long as current subjective risk judgments prevail. We concur with the conclusion of Covello, Von Winterfeldt and Slovic (1986) who state

> ... the literature specifically focused on risk communication is relatively small. Substantial progress has been made on some topics, such as psychological research on public perceptions of risk, but large gaps exist in our understanding of virtually every issue relevant to risk communication.

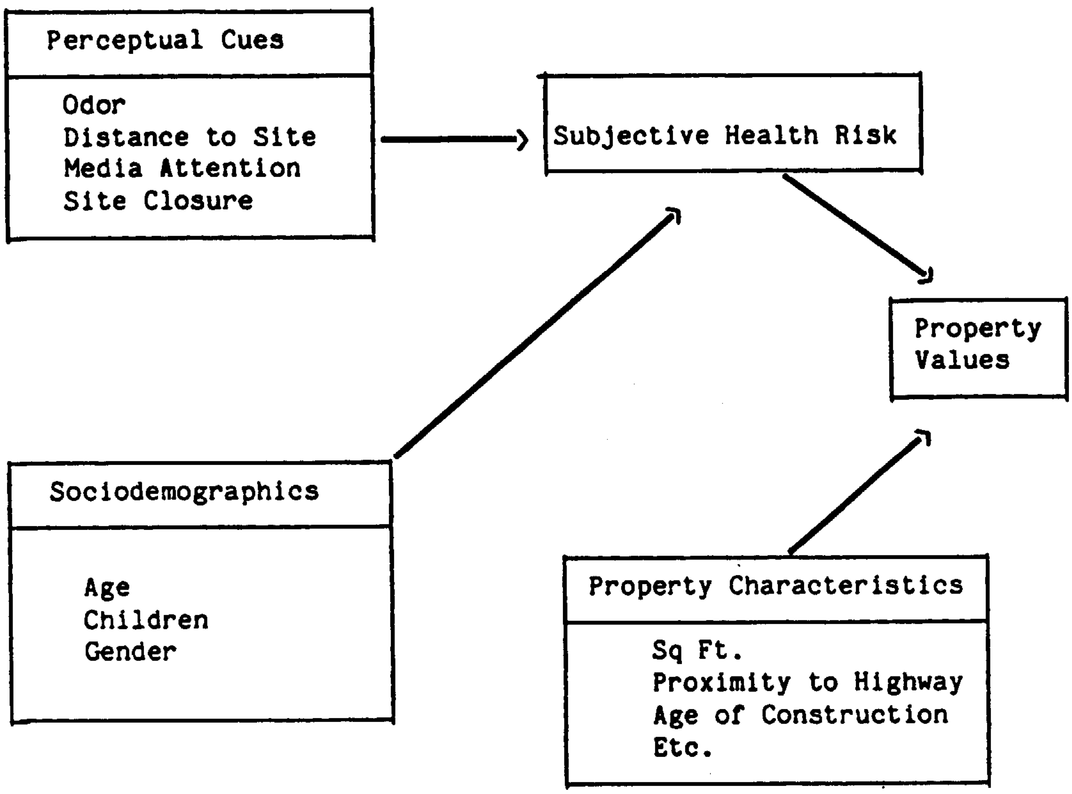
The importance of better risk communication is well understood but the methods are lacking. In a study of public perception and response to EPA warnings concerning the risks of ethylene dibromide (EDB), Sharlin (1986) analyzed and compared what EPA was trying to tell the public about the risks of EDB to the information the public actually received through the media about these risks. He found vivid contrasts between the public’s view of the health risks and the EPA’s aggregate statistics on health risks. The extent and nature of this contrast is an area that needs further exploration.

Two main conclusions emerge from the OII study results: (a) subjective health risks are likely to be overestimates of the objective risks and (b) the overestimated subjective health risks are associated with significant property value losses. In many respects it is similar to the situation described in Chapter 1 where a warning was issued for possible volcanic activity. In several instances the overreaction to such warnings has resulted in economic losses due to property devaluations that far exceeded the expected economic losses. When, as in the case of the OII Landfill, total damages from the overestimates of risk are on the order of $27 million, a program designed to change subjective estimates of health risks can easily be cost effective.

Figure 4.5 illustrates a schematic framework that integrates the model of subjective health risk with the model of property values. The left side of the Figure represents a model for subjective health risk estimates of individual survey respondents. The right side of the Figure shows the factors impinging on property values. (The property value modeling is necessarily an aggregate analysis because property value changes could be linked with subjective health risk variables only at the neighborhood

FIGURE 4.5

A MODEL OF SUBJECTIVE HEALTH RISK AND PROPERTY VALUES



level.)

The modeling of subjective health risk judgment points to two components for possible intervention: perceptual cues and attitudes associated with sociodemographic variables. Of the two, psychological research shows that perceptual cues are much easier to change than attitudes. Managing the perceptual cues which serve to remind people about the risk can be very effective in reducing risk estimates to more appropriate levels. The management of perceptual cues would involve such things as reducing odor, reducing visibility of the site using plantings or screening, reducing activity at the site (e.g., reducing number of trucks entering and leaving), and reducing sensational media coverage of the site. These are not necessarily easy to implement. Some of these strategies such as reduced media coverage can only be recommended, not mandated. Others such as reducing odor and reducing activity are difficult or impossible to implement short of closing the site. However, if such reductions can be obtained, the management of perceptual cues can have dramatic effects. If subjective health risks for a hazardous site are overestimates of the objective risk, then the perceptual cues about the risk should be managed as extensively as possible. The economic savings obtained by correcting and/or avoiding inappropriate property devaluations are likely to be large.

After major changes in the perceptual cues associated with closing the site, many people maintained high risk estimates. These high risk estimates translate via the property value equation into an estimated remaining loss of about $13 million. This residual loss is due partly to perceptual cues that cannot be easily modified (visibility of the site and the methane plant) and to risk attitudes. Given that further modifications of perceptual cues are probably impossible, further reductions in subjective health risks and their associated effects on property values could only be achieved by credible, effective communications about the objective risk.

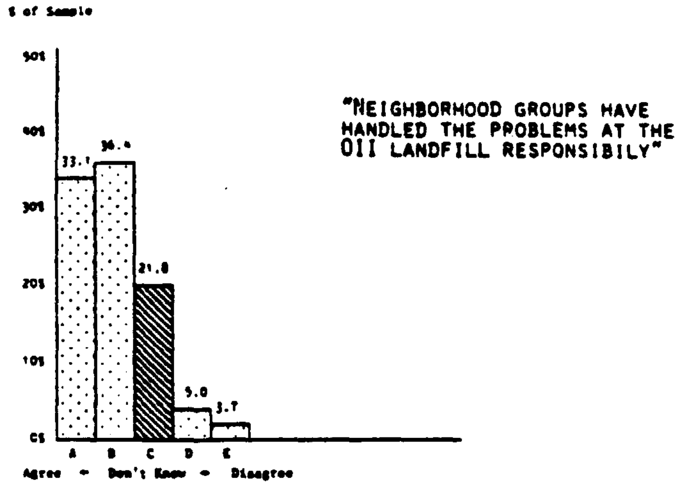
Risk attitudes and beliefs should be changed if health risks are truly small. Changing attitudes is notoriously difficult and there are several factors which compound the problem in this context.

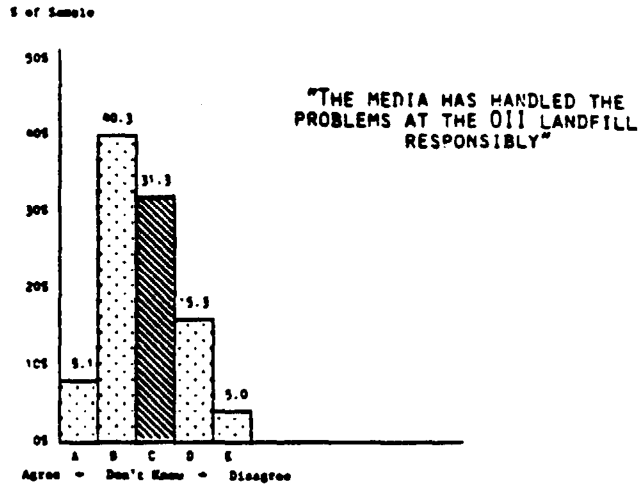
First, many psychological studies (see Tversky and Kahneman 1974; Slovic, Fischhoff, and Lichtenstein 1977) have shown that most people have trouble understanding probabilistic information in general and expert assessments of risk in particular. To be understood, expert assessments are best communicated by comparing new risks to better known risks such as smoking and X-rays rather than presenting technical measures such as mortality rates for a given exposure. No information of the appropriate type of risks has been provided to residents near the OII Landfill.

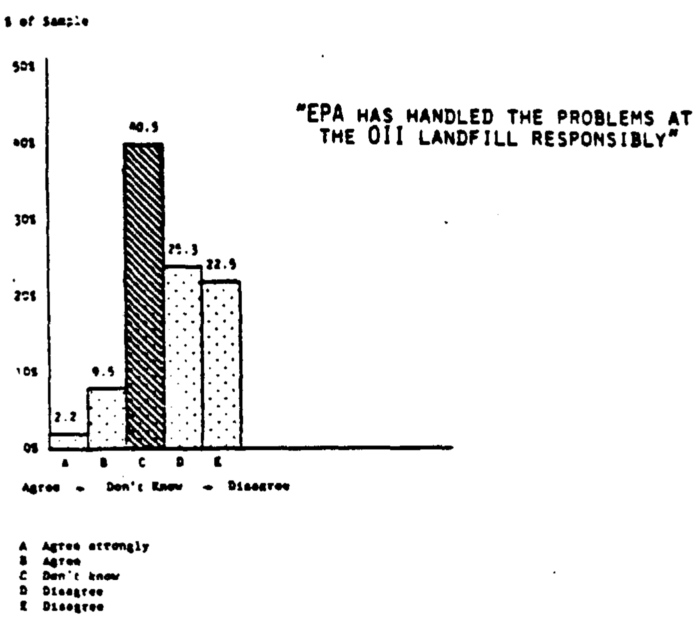
Second, to be effective, risk communication must come from credible sources. Figure 4.6 shows how credibility is perceived among a few of the important actors at the OII Landfill. Residents in the area perceive that neighborhood groups have acted the most responsibly with the media also receiving a favorable response. The EPA, however, was not as well perceived, and is now unlikely to be viewed as a credible source since residents ranked EPA nearly as low as the operators of the OII Landfill in terms of how "responsibly" the agency had dealt with problems at the site.

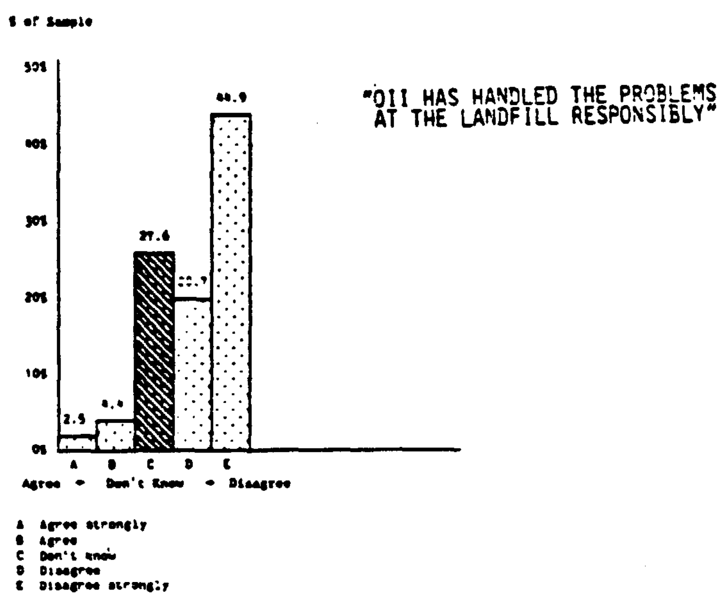
Third, even though it has not been especially effective, much more is known about increasing subjective risk judgments (e.g., risks of smoking,

FIGURE 4.6


"NEIGHBORHOOD GROUPS HAVE HANDLED THE PROBLEMS AT THE OII LANDFILL RESPONSIBILY"
3.7
A B C D E
Agree – Don't Know – Disagree


"THE MEDIA HAS HANDLED THE PROBLEMS AT THE OII LANDFILL RESPONSIBLY"
5.0
A B C D E
Agree – Don't Know – Disagree


"EPA HAS HANDLED THE PROBLEMS AT THE OII LANDFILL RESPONSIBLY"
2.2
9.5
22.5
A B C D E
Agree – Don't Know – Disagree
A Agree strongly
B Agree
C Don't know
D Disagree
E Disagree


"OII HAS HANDLED THE PROBLEMS AT THE LANDFILL RESPONSIBLY"
2.5
27.6
A B C D E
Agree – Don't Know – Disagree
A Agree strongly
B Agree
C Don't know
D Disagree
E Disagree strongly

risks of not using seat belts) than about decreasing risk judgments.

Fourth, communications about issues with a high affective component (e.g., the emotionality surrounding a landfill hazard issue) are often misinterpreted and misunderstood. For these and other reasons a quick fix via risk communications for the attitudinal inflation of risk estimates is improbable. The potential elimination of approximately $13 million in property value losses would, however, justify considerable efforts to change subjective risk estimates to more realistic levels.

4.9 Conclusions.

While changing risk attitudes will not be easy, there are several studies which suggest some optimism. Hammond and his colleagues at the University of Colorado (see Hammond and Adelman, 1976; Hammond et al. 1984) have been successful in reducing disagreements about risk among experts and then communicating the resulting judgment about the risk to the public. Examples include public concern about a new police handgun bullet and about possible plutonium pollution from a nearby facility. Characteristics of these successful efforts to reduce overestimated risks share the following attributes.

First, a citizen panel (such as the HELP group) selects a group of independent scientists to evaluate the risk. The danger at this stage is that, all too often, the citizen’s panel will want to become technical experts themselves in order to make their own risk judgments. Their proper role is representing community values and the procedure generally works best if they stick to that.

Second, the group of scientists uses standard scientific and scholarly procedures (e.g., references to referred journal articles, development and defense of mathematical equations producing the risk estimate) to resolve their differences. Also of use in this stage are psychological techniques for studying judgments and techniques that help identify issues of disagreement that need resolution. Contrary to the danger in the first stage, the danger here is that the scientific experts will make action recommendations for the community. Such recommendations necessarily are based on both risk judgments, which the technical experts should make, and assumptions about community values, which the technical experts should not make.

Third, once agreement on the magnitude of the risk is obtained (and surprisingly such agreement is almost always obtained), the results are communicated to the public via the local media. What is communicated to the public is the experts' conclusion that the risk is either low or high and a comparison of the risk to known, widely-accepted risks. For example, comparing the danger of plutonium emissions to smoking or hospital X-rays.

Although the above approach is not a panacea, it does offer a reasonably inexpensive means for attempting to reduce subjective health judgments, which due to attitudes, overestimate the true risk. Given the magnitude of potential benefits, the past success and relatively small cost of such procedures justifies their use in an attempt to change subjective health risks.

CHAPTER 5: RISK COMMUNICATION GUIDELINES

5.1 Introduction.

Proper communication of the level of risk for a particular Superfund site is a crucial component of successful community relations. Poor risk communication leads to confusion among community members as to the appropriate level of concern about the hazards of a site and can also produce unnecessary disagreement within the community. Thus, poor or inadequate communications about risk can make an already difficult community relations problem even more difficult. This chapter describes characteristics of good risk communication; all EPA and contractor personnel should always have these characteristics in mind when developing any communication about the hazards of a site. This chapter also makes the important point that there are many indirect ways--for example, the appearance of the site to community members--in which risk is communicated. EPA personnel must carefully monitor those indirect communications and make sure that technical contractors do not inadvertantly communicate inappropriate messages about risk levels through their actions. The chapter concludes with a series of specific recommendations for communicating information about the risks of Superfund sites.

There are numerous situations in which the level of risk must be communicated to a community in which a Superfund site is located. Successful risk communication cannot be accomplished unless all aspects of community relations are handled with care. Thus, all the guidelines from Community Relations in Superfund-A Handbook apply to risk communication as well. We do not repeat those guidelines here but instead describe some principles for presenting information about levels of risk, principles that apply to all community relations situations.

It cannot be overemphasized that good risk communication, as all good communication, is a two-way process. Community relations personnel must learn from members of the community their concerns and their beliefs about the risk associated with a particular site. Community relations personnel have an important role to play between the technical experts and the community. Risk assessments developed by off-site experts sometimes respond only to technical characteristics of the site while unintentionally ignoring some community concerns. For example, consider a site from which emits an unpleasant odor. A technical expert might know immediately that the odor was not harmful and so ignore it completely in a technical report, instead concentrating on chemicals, the names of which are probably unfamiliar to the general public, that might be leaking from the site. The community would obviously be aware of and concerned about the possible health consequences of the odor so a risk communication that did not explicitly address the odor problem would be inadequate and unacceptable to the community. Thus, risk communication must address all the concerns of a community , whether or not the technical experts see those concerns as important.

It should be stressed that guidelines in this chapter apply to all risk communication situations. Research has so far shown that the best ways of communicating risk are the same for experts and community leaders as for the general public. Thus, the same communications should be given to everyone. Not only is this more efficient, but it also avoids the potential problems that could be created by attempting to give different information to different groups in the same community.

5.2 Credibility.

No communication enterprise can be expected to be successful if the source of the information has not established credibility. If the source of risk information is not credible, then it will have little chance of being accepted. Community concerns may in fact get exaggerated because the dubious attitude toward the source will extend to the risk communication. Obviously, this situation could escalate, causing permanent damage to community relations. There are several steps that can help in establishing credibility:

1. Neutral, well-regarded experts should be appointed to assess the risk. Scientists in the appropriate fields are especially good candidates.
2. These experts should report to a credible level of government or to a credible group of government officials. Often the most local level is most credible to the community, but this will vary with the site. If no credible level of government exists, a citizen committee of community leaders can be used.
3. The credible government officials or citizen committee should release the risk information to the general community, including the press.

Information should be released in a consistent manner. Care must be taken with the news media, especially, who can inadvertently cause distortion in the risk perception by presenting inconsistent information and distorted perceptions (please see Section 5.5).

The remaining sections of this chapter will address how the level of risk of a site should be expressed to achieve the best understanding as well as some principles of risk perception that are important for Superfund and contractor personnel to know when preparing information releases, developing a community relations plan, or planning on-site activities.

5.3 Overview of Risk Communication Principles.

There are a number of important factors that need to be considered in communicating risk; each factor will be addressed in detail in the remaining sections of the chapter. The major components in the formation of community risk beliefs for a Superfund site are:

(5.4) Physical reminders around the site,

(5.5) News media presentation of the risk,

(5.6) Community characteristics,

(5.7) Reaction to low-level risks,

(5.8) Characteristics of the risk, and

(5.9) The framing of losses and gains

The Chapter concludes with

(5.10) Use of example risks, and

(5.11) Recommendations.

5.4 Physical Reminders.

Physical reminders of the site provide an indirect but very powerful and important means of communicating levels of concern to the community. Heavy truck traffic to and from the site, heavy chain-link fences with large imposing warning signs, odors emitted from the site, discoloration of water, and on site workers wearing protective "space" suits are all examples of physical reminders that implicitly send a message to the community about the level of risk at the site.

Community relations personnel should make a careful inventory of all the physical reminders at a particular site. For those physical reminders that are modifiable, the community relations officer should try to ensure that each physical reminder is appropriate for the actual level of risk. Regional contractors as well need to be aware of physical reminders. For example, if being at or near the site is not actually harmful then a more discrete fence such as hedging may be preferred to a heavy chain-link fence with large red warning signs. Conversely, if the site is an old, familiar spot in the community and the citizens have become inappropriately careless about the site, then an imposing fence with many warning signs will be critically important.

For those physical reminders that are not modifiable, the community relations officer must ensure that the community has the proper information necessary to understand the meaning of the physical reminder in terms of the level of risk. Many physical reminders are very misleading indicators of the true level of risk. As an example, in Colorado, iron contamination from an old mine that is now a Superfund site caused a poor taste and red color in the drinking water of several downstream communities. Although the poor taste and discoloration were undesirable and clearly needed to be remedied, drinking the water was, in fact, not likely to be harmful. However, the taste and color served as physical reminders of the iron contamination and made many residents fearful about drinking the water. A successful community relations effort would have provided residents with the necessary information to understand that the taste and discoloration were undesirable but not harmful.

The community relations officer must also be particularly careful to alert the community to any changes in the physical reminders. Unannounced changes in the physical reminders almost always send a more extreme message about the level of risk than is appropriate. For example, if decontamination workers in protective clothing appear at a site unannounced, then the level of community concern will immediately increase to very high levels. Even if the arrival of such workers is announced, levels of community concern will increase but the level of concern will generally be more consistent with the actual level of risk than if no advanced preparation had been done.

5.5 News Media Presentation of Risk.

The news media provide much information concerning the risk level of a site. For many citizens the newspaper, radio, and television are their primary sources of information about the site. As in any good community relations plan, it is essential that the news media have advance announcement of all activities at the site. Also, the community relations officer must provide them with any background information that might be necessary to help them and the public to interpret the reports from technical experts.

The community relations officer should be alert to a problem that often arises in communicating risks to the public through the news media. It often happens that someone very concerned about the site who is not a technical expert becomes an unofficial spokesperson for the group of citizens who are very concerned. This person is always available for interviews with the news media and is often eager to make statements about what they believe about the risk posed by the site. To balance the sometimes exaggerated risk estimates that may result, consistent information from credible sources (see Section 5.2) should be regularly available to the news media. It is of course important that the credible source be independent of the control of the Environmental Protection Agency.

5.6 Community Characteristics.

High levels of concern about risks are often associated with certain community and individual characteristics. Knowledge of these characteristics can help Superfund personnel and contractors anticipate and prepare for difficult risk communication situations, situations in which unwarranted conflict may arise. This section identifies a number of personal and demographic characteristics that are often associated with high levels of concern.

An obvious characteristic of great importance is proximity to the site. Residents living near a hazardous waste site will have higher levels of concern and will be more skeptical of attempts to communicate levels of risk.

Older residents tend to have lower levels of concern about risks; conversely, young adults tend to be more sensitive to risks, especially new ones. There are several reasons for this general association with age and lower levels of concern. Older residents living near a site have likely been exposed to the risk for a long time without experiencing any significant consequences; they may consider it to be benign. Also, some older individuals may feel that they have less time left in which to suffer the consequences of long-term exposure.

People with families, especially those with young children, tend to have higher levels of concern. People with children may have higher levels of concern because they feel that they have “more to lose.” Thus, many parents with young children will likely be very concerned about a Superfund site. Note that this very concerned group, because of family responsibilities and child care arrangements, is unlikely to attend community meetings or participate in other activities where risk information would be communicated. Thus, community relations personnel must be creative in finding ways to communicate to this very concerned group.

Occupation, income, ethnic background, and level of education have generally not been found to have an appreciable effect on levels of concern.

5.7 Reaction to Low-Level Risks.

In planning risk communication, it is important to understand how people interpret and respond to information about risks, especially risks with low probabilities but serious consequences. Virtually everyone--citizens, the news media, community relation officers, and even many technical experts--have difficulty with interpreting and responding appropriately to risks with low probabilities. Low probabilities are those with annual odds on the order of 1 in 100 or less. It is quite common for the serious risks associated with a Superfund site to be this low or lower so the communication problem will be quite difficult. In this section we describe the likely responses to communication about low probability risks.

When receiving new information about a low probability risk most people make one of two judgments. Some judge the risk to be a serious threat to , them, their families, or their property and so they have a high level of concern. Others decide that the chances of the risk are so remote that they dismiss the risk and act as if the probability of the hazard is zero.

It is often difficult to find grades of concern between these extremes for low probability risks; either people are very concerned or they are not concerned at all. Thus, people often do not make distinctions between low level risks, especially when risk levels are presented in terms of powers of ten such as 10^{-5}. Although a risk of 10-5 is 1000 times more likely than a risk of 10^{-8} and although the level of concern should be about 1000 times greater, most people will make very little distinction between such risks. Either they will be very concerned or they will dismiss the risk as too unlikely to worry about.

The disparity in the two types of responses to information about low probability risks has obvious implication for community conflict about a Superfund site. One group of citizens will be very concerned and so will be motivated to attend public hearings, write letters to the newspaper, circulate petitions, file lawsuits, etc. Another group of citizens will be unconcerned about the risk of the site but will become quite concerned about the activities of the other group. They will see the other group as needlessly “stirring up trouble” which, in their view, will result in unwarranted publicity and decreases in property values. For example, at a Superfund site in California, a concerned group of citizens formed an organization named "HELP" ("Homeowners to Eliminate Landfill Problems"). This group was in turn opposed by another group of homeowners who thought that HELP would only succeed in reducing property values. To reduce community conflict about a site, the community relations officer must be alert for possibilities to help each group understand the other group’s concerns using techniques for conflict resolution such as those presented in the Appendix.

The community relations officer should also be aware that in some sense the response of neither group is appropriate. We make that statement cautiously because citizens are of course free to have whatever level of concern they believe appropriate. However, it is often the case that the responses of those in the very concerned group are inconsistent with their responses to other risks with similar probabilities and consequences. For example, use of certain household and garden chemicals or activities such as smoking may expose them to risks as large or larger than the risks posed by the Superfund site yet they have little or no concern for those other risks. It has been shown that experts do no better in their personal lives dealing with low level risks than the general public, so no individual or group is immune to the problem of over or under reacting to low level risks. Some of our suggestions presented in Section 5.10 for communicating risk levels involve comparing the risk of the superfund site to comparable risks which are more familiar.

Just as some people are overly concerned, there are others who are completely unconcerned but sometimes ought to be more concerned. People who are unconcerned will be less motivated to heed warnings and to take precautionary actions that may be temporarily necessary for those living near a Superfund site. For example, an unconcerned resident might ignore a warning sign and trespass on a Superfund site because of good hunting or might not follow precautions about contaminated drinking water.

Given this information about the tendency for people either to be over- or under-concerned, the role of the community relations officer is to provide the best and most appropriate information so that people can make their own decisions. However, it will be most helpful if the credible source and the news media are provided with suggested actions (e.g., treatment of drinking water, avoidance of the site) that are appropriate for the level of the risk. But in the end, the community relations officer must be prepared for the fact that some people will be concerned that the actions are not enough and others will believe that they are too much.

It is also important for the community relations officer to understand that when thinking about risk and unfortunate events, people often make attributions about responsibility, attributions that are sometimes unwarranted. Blame for bad outcomes is usually attributed to something specific even if there is no evidence to justify such blame. This tendency to blame something in particular is especially true if there is already a community “bad guy.” For example, the Superfund site or the former ‘operator of the site may be labelled as this bad guy so that the blame for many random events actually unassociated with the site will be attributed to the site. Thus, the cause of an otherwise unexplained cancer death of a resident living near the site or of a former worker at the site will almost surely be attributed to the site. Adverse pregnancy outcomes, the real cause of which is often difficult to determine, will also be attributed to the site.

An interesting example of false blame is provided by reactions at a Superfund site in California. An epidemiological survey by the state health department found an increased reporting of toothaches from residents living near the site. Even though there is no known biological mechanism by which anything at the site could cause toothaches, residents had a heightened awareness of any adverse event and were quick to attribute it to the neighborhood bad guy--the Superfund site.

When expressing levels of risk, complicated mathematical expressions (such as 10^{-5}) should be avoided. The community relations officer should also avoid the temptation to express the probability level in many different but equivalent ways. For example, a risk could be phrased in terms of how many people would be harmed per 10,000 people exposed (say, 2 in 10,000) or in terms of how many would not be harmed per 10,000 people exposed (9,998 in 10,000). Doing so may seem more complete than presenting only one expression, but so much information at the same time leads to confusions that can be fatal to a risk communication enterprise. There have been cases of such information being misconstrued and reported erroneously.

In summary, many problems may arise because people have difficulty in understanding low probability risks. Community relations personnel must try to provide information about risks in a manner that will not increase community conflict. An effective way to accomplish this is to compare the low probability risk to other familiar example risks. Precisely how to do this is described in Section 5.10.

5.8 Characteristics of the Risk.

There are a number of characteristics about a risk or hazard other than its probability or the seriousness of its consequences that influence response to the risk. In order to anticipate risk communication problems it is important for community relations personnel to identify the relevant risk characteristics associated with a particular Superfund site. It is usually not possible to change these characteristics, but just as knowledge of community characteristics is important in predicting response to risk information, so too is knowledge of the risk's characteristics. This section describes a few of the important characteristics of risks which help determine the level of concern.

The general effect of risk characteristics is to raise or lower the level of concern relative to comparable risks. Risks which tend to generate relatively lower levels of community concern are familiar, well-known to science, and undramatic. Conversely risks which tend to generate relatively high levels of community concern are unfamiliar, not well-understood by science, dramatic (in that many people might be killed or injured in a single event), and contain an element of dread.

Let's consider these characteristics in the context of typical Superfund sites. Familiarity is often a very important issue. Some sites are well known to community residents; residents will have had many experiences with the site such as just driving by without experiencing any effects. With such familiarity the level of concern will be much lower than for a site which poses a comparable risk but which has been covered up for many years and so the risk was unknown to the community. Similarly, if the risk is well-known or familiar to science, it will usually generate lower levels of concern. For example, if the risks at a site are due to chemical or toxic materials that are commonly found and monitored in workplace settings and for which acceptable levels have previously been established, concern will be relatively lower than for toxics that are not well-understood and for which scientists are uncertain about acceptable levels. Whether or not the consequences of the risk are dramatic is also important. Some sites have risks where exposure now will result in a fatal illness many years hence while other sites have risks which are more immediate. Finally, there are some risks that people simply dread more than others. Radioactivity and cancer-inducing toxics are especially dreaded. So, even if the scientific estimates of the probability of harm were the same at two sites, the site with even small amounts of radioactivity or cancer-inducing toxics would produce a much higher level of concern in the community.

The task of risk communication therefore depends somewhat on the risk characteristics. If a site has characteristics which produce relatively high levels of concern, then the community relations officer will need to provide information that will help residents place the risk in its proper context. On the other hand, if a site has those characteristics which produce relatively low levels of concern, then the community relations officer must take special care to alert residents that a real risk does exist so that they will be motivated to take necessary precautionary actions.

5.9 The Framing of Gains and Losses.

In preparing communications about risks associated with a Superfund site, it is important to determine whether the level of risk being communicated will be viewed by residents as an increase or decrease in the level of risk. That is, each resident will have some prior belief about the riskiness of the site; any risk communication will be responded to differently depending on whether the riskiness being communicated is higher or lower than that prior belief.

This is important because a perceived increase in risk will have between three and ten times the psychological impact on the level of concern because it is perceived as a loss than an equivalent reduction in risk which is perceived as a gain. Going from thinking you are safe to believing you are unsafe makes people a lot more unhappy than going from unsafe to safe makes people happy.

In the context of Superfund sites, for example, informing residents about an old waste site in their neighborhood about which they had no awareness will create a great deal of unhappiness because it is an increase in perceived risk and viewed as a loss. Conversely, telling people who have worried a lot about a known site for many years that the site is in fact very safe, even if they were to believe it, would not increase their happiness a great deal because it is a decrease in risk and viewed as a gain. This means that informing people about new risks and hazards must be done very carefully and that informing people about reductions in old risks is not likely to have much impact.

There are two important ways in which a risk level can be viewed as a change relative to the current perceived level: (a) the probability or likelihood of a hazard event may increase and/or (b) the consequences or severity of a risk may increase. Even slight increases in either or both will cause high levels of community concern while moderate or even large decreases in either or both will only slightly reduce levels of community concern.

5.10 Use of Example Risks.

The previous sections have outlined the problems that must be addressed in any risk communication enterprise. Risk communication is broadly defined as both the physical reminders that are present at any Superfund site and the communication enterprise undertaken by the community relations personnel. It is important that both types of communication be monitored; too often only the latter is given careful consideration. As noted in Section 5.6, community relations personnel must also be alert to characteristics of the community and to characteristics of the risk that will make risk communication more or less difficult. With that context of the difficulties of risk communication, this section describes a useful strategy for the accurate and understandable communication of levels of risk.

As noted in Section 5.8, some risks are better understood than others. Well-understood risks can be used to great advantage as a tool to help people understand new risks such as those posed by a Superfund site. This is most effectively done in a compare/contrast manner; for example, a low-level lung cancer risk posed by the site is presented with and compared to a high-level example risk, such as smoking. Explaining how much lower the site risk is than smoking helps the concerned citizens place the site risk in perspective, thus helping promote an appropriate level of concern. Familiar risks with comparable levels of risk can also be presented. For example, low-level example risks such as x-rays or saccharin may be presented in conjunction with analogous low-level site risks. Naturally, care must be taken that comparisons between risks are scientifically valid, and it is especially helpful if the example risks have many characteristics (see Section 5.8) in common with the site risk (such as the lung cancer example, above).

A particularly good procedure for using example risks is the risk ladder. Instead of a few example risks, the ladder contains many types and levels of risk, arranged on a risk scale. An example of a ladder that has been particularly useful for these situations is given in Figure 5.1. The risks on the ladder are generally well-understood risks, presented with their actual risk level. The site risk can be added to the ladder, in its. proper place in comparison to the example risks, and the ladder can be presented to concerned citizens. This procedure is much more effective than presenting the risk by itself because context has been established. Thus, citizens can view the risk in a more natural, real-world context.

Two very important guidelines must be followed when example risks and ladders are used. First, the example risks themselves must be well-understood risks that do not cause inappropriate levels of concern on the part of the public. Thus, risks that are considered unknown or dreadful (see 9.7) are inappropriate as example risks. It would be unwise, for example, to use nuclear war or AIDS as example risks. Better risks are those given on the risk ladder (Figure 5.1).

A second important guideline concerns the choice of ladders versus one or two example risks. It should be remembered that the ladder, while effective, takes time to read and understand, and would be cumbersome for the media, for example, to use effectively. It is better to present one or two appropriate example risks, explaining in detail the nature of the comparison and repeating the risk levels to make sure they have been understood. Otherwise, it is possible for a misunderstanding to arise, and the high-level contrast example risk (e.g., smoking) could be construed as analogous to the site risk, thus confusing and alarming citizens.

Example risks, then are an effective tool for communicating risks. A risk ladder of many risks is especially useful, but only for situations in which it would not be cumbersome. Finally, example risks must be chosen carefully, or risk perception will be further confused.

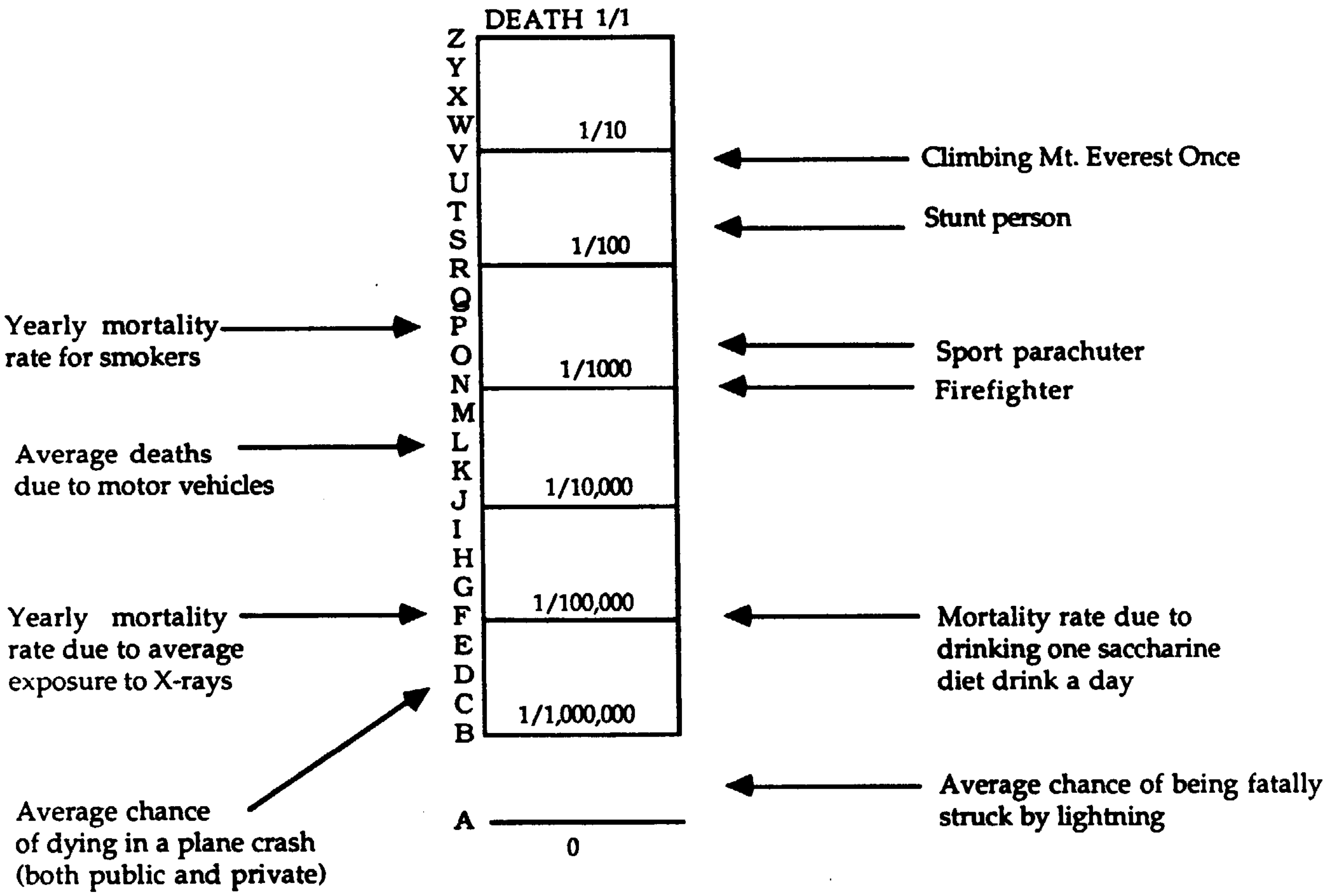
5.11 Recommendations.

This section will refer to the guidelines already established in previous chapters. The purpose of these recommendations is to give some examples of risk communication situations, and to suggest effective ways of performing risk communication in those situations.

FIGURE 5.1

ANNUAL CHANCE OF DEATH

(Unless otherwise specified, risk is for one year of exposure)



1. The neutral, well-respected group of risk experts (scientists, usually) that is chosen to assess the site should report to a credible level of government. Sometimes, however, trust has deteriorated, and no government body in the area is considered credible. In this case, a citizen committee should be formed. The risk experts then report to these respected citizens, and the citizens communicate the risk to the community at large.
2. Certain communities or neighborhoods, especially those with young families, will be more likely to have very high levels of concern. These areas should be targeted for careful communications, remembering that these families are the least able to participate in traditional public hearings.
3. Information about levels of risk should be expressed in concrete terms, using example risks or the risk ladder. Also, physical reminders should be monitored so that they communicate a level of risk consistent with the actual level of risk.
4. The news media easily can inadvertantly become a catalyst for inappropriately high levels of concern. To prevent this, complete and consistent information should be provided to the news media, starting as early as possible before the news media’s beliefs have been formed and reported. If the news media nevertheless begins to escalate concern on the part of citizens, the risk communication officials must increase the effort to counteract misleading information. For example, if non-experts are interviewed, then experts should become available to the news media, providing accurate information and addressing the non-experts’ information. Example risks should also be provided, although too many example risks (as in a risk ladder) could

be counter productive given the brief coverage typically provided by the news media. In all risk communication enterprises, probability
5. phrases should be utilized with caution. Low probabilities are especially problematic; probabilities at or below one in one hundred are poorly understood. Expressing long-range probabilities (the risk over a period of many years, for example) is often helpful because it brings the probability level to a better-understood range. Also, pairing probabilistic information with example risks and risk ladders helps citizens put the numbers in context.

6. Complicated mathematical expressions (such as 10^{-5}) should be avoided, as should too many permutations of numeric information.

REFERENCES

Abelson, R. P., and A. Levi. 1985. "Decision Making and Decision Theory" Chapter 5 in J. Lindzey and E. Aronson (eds.), Handbook of Social Psychology, Vol I: Theory & Method, New York, Random House, pp. 231-309.

Burness, H. S., R. C. Cummings, A.F. Mehr, and M.S. Walbert. 1983. "Valuing Policies Which Reduce Environmental Risk," Vol. 23, Natural Resources Journal, pp. 675-682, July.

Camerer, C., and H. Kunreuther. 1987. Experimental Markets for Insurance. (Draft technical report. Center for Risk and Decision Processes. Department of Decision Sciences. The Wharton School, University of Pennsylvania, Philadelphia, PA 19104.)

Cohen, R.A., W.G. Lewellen, R.C. Lease, and G.G. Schlarbaum. 1975. "Individual Investor Risk Aversion and Investment Portfolio Composition," J. Finance, 30:605-20.

Coombs, C.H., T.G. Bezembinder, and F.M. Goode. 1967. "Testing Theories of Decision Making Without Measuring Utility or Subjective Probability." Journal of Mathematical Psychology, 4, pp. 72-103.

Coppinger, V. M., V. L. Smith, and J.A. Titus. 1980. "Incentives and Behavior in English, Dutch and Sealed-Bid Auctions," Economic Inquiry, (XVIII(I) pp. 1-22, January.

Coursey, D.L., J. J. Hovis, W. D. Schulze. 1986. "On the Disparity Between Willingness to Accept and Willingness to Pay Measures of Value,n Quarterly Journal Of Economics.

Combs, B. and P. Slovic. 1979. "Causes of Death: Biased Newspaper Coverage and Biased Judgments," Journalism Quarterly, 56:837-843.

Covello, V., D. Von Winterfeldt, and P. Slovic. 1986. Risk Communication: An Assessment of the Literature on Communicating Information About Health, Safety and Environmental Risks," Draft Report to USEPA, Institute of Safety and Systems Management, University of Southern California, Los Angeles.

Cutter, S.L. 1987. Airborne toxic Releases: Are Communities Prepared? Environment, 29(6), 12-17, 28-31.

Dillman, D., 1978. Mail and Telephone Surveys: The Total Design Method, John Wiley & Sons, New York, N.Y.

Einhorn, H., and R. M. Hogarth. 1985. "Ambiguity and Uncertainty in Probabilistic Inference." Psychological Review, 92, pp. 433-461.

Friend, I. and M.E. Blume. 1975. "The Demand for Risky Assets," American Economic Review, 65:900-922.

Grether, D. and C. Plott. 1979. "Economic Theory of Choice and Preference Reversal Phenomenon," American Economic Review, Vol. 69, pp. 623-638.

Halpern, D.F., S. Blackman, and B. Salzman. 1986. "The Use of Statistical Risk Information to Decide if Oral Contraceptives are Safe." Paper presented by the annual meetings of the Psychonomic Society, New Orleans, November 1986.

Hammond, K.R., R.M. Harem, J. Grassia and T. Pearson. 1987. "Direct Comparison of the Efficacy of Intuitive and Analytical Cognition in Expert Judgment," Manuscript, Center for Research on Judgment and policy, University of Colorado-Boulder.

Hammond, K., B. Anderson, J. Sutherland and B. Marvin, 1984. "Improving Scientists' Judgement of Risk," Risk Analysis, Vol. 4.

Hammond, K.R., and L. Adelman, 1976. "Science, Values and Human Judgment," Science, 194, 389-396.

Handa, J. 1977. "Risk, Probabilities, and a New Theory of Cardinal Utility." Journal of Political Economy, 56, pp, 279-304.

Hershey, J.C., and P.J.H. Shoemaker. 1980. "Risk Taking and Problem Context in the Domain of Losses: An Expected Utility Analysis." The Journal of Risk and Insurance, 47, pp. 111-132.

Johnson, E.J., and D. A. Schkade. 1986. "Anchoring, Adjustment and Bias in Utility Assessments." (Working Paper #12-86-87, Carnegie-Mellon University, Pittsburgh).

Kahneman, D. and A. Tversky. 1984. "Choices, Values, and Frames," American Psychologist 39:341-350.

Kahneman, D., and A. Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk," Econometrics, Vol. 47, No.2, pp.263-291, March.

Kahneman, D., and A. Tversky. 1972. "Subjective Probability: A Judgment of Representativeness," Cognitive Psychology, Vol 3, pp. 430-454.

Karmarkar, U.S., 1978. "Subjectively weighted Utility: A Descriptive Extension of the Expected Utility Mode." Organizational Behavior and Human Performance, 21 , pp. 62-72, 1978.

Kunreuther, H., R. Ginsberg, L. Miller, p. Sagi, p. Slovic, B. Borkan, and N. Katz. 1978. Disaster Insurance Protection: Public Policy Lessons, Wiley.

Liechtenstein, S., and P. Slovic. 1971. "Reversal of Preferences between Bids and Choices in Gambling Decisions," Journal of Experimental Psychology, Vol 89, pp. 46-55.

Liechtenstein, S., P. Slovic, B. Fischhoff, M. Layman and B. Combs. 1978. "Judged Frequence Of Lethal Events," Journal of Experimental Psychology: Human Learning and Memory 4:551-578.

Los Angeles County Department of Health Services, 1983. "Conception outcomes and Mortality Study of the Operating Industries Landfill Area," unpublished Report, L.A., California.

McClelland, G.H and W.D. Schulze. 1986. "Measures of Willingness to Pay Versus Willingness to Accept: Evidence, Explanations, and Potential Reconciliation," Paper presented at the U.S. Forest Service Workshop on Integrating Economic and Psychological Knowledge in Valuations of Public Amenity Resources, Estes Park, Colorado. (May).

McClelland, G.H., W. Schulze and D. Coursey. 1986. "Valuing Risk: A Comparison of Expected Utility With Models From Cognitive Psychology Submitted to American Economic Review.

McClelland, G.H., and B. H. Hackenberg. 1978. "Subjective Probabilities for Sex of Next Child: U.S. College Students and Philippine Villagers," Journal of Population, Vol. 1(2), 132-147.

Morin, R.A. and A. Suarez Fernandez. 1983. "Risk Aversion Revisited," J. Finance 38:1201-16, September.

McNeil, B.J., S.G. Pauker, H.C. Sex, and A. Tversky. 1982. On the Elicitation of Preferences for Alternative Therapies. New England Journal of Medicine, 306, 1259-1262.

Poulton, D.C. 1968. "The New Psychophysics: Six models for Magnitude Estimation." Psychological Bulletin 69: 1-19.

Phillips, L. D., and W. Edwards. 1966. "Conservatism in a Simple Probability Inference Task,n Journal of Experimental Psychology, Vol. 72, pp. 346-357.

Plott, C. R., and S. Sunder. 1982. "Efficiency of Experimental Security Markets With Insider Information: An Application of Rational-Expectations Models," Journal of Political Economy, Vol. 90, No. 4, pp. 663-698.

Rosen, S., 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," Journal of Political Economy, 82, pp. 34-55.

Russell, T., and R. Thaler. 1985. "The Relevance of Quasi Rationality in Competitive Markets," American Economic Review, Vol. 75, No. 5. pp. 1071-1082.

Shoemaker, P.J.H. and H.C. Kunreuther. 1979. “An Experimental Study of Insurance Decisions.” The Journal of Risk and Insurance, Vol. 46, pp. 603-618.

Schulze, W.D., G. H. McClelland and B. Hurd. 1986. "A Case Study of a Hazardous Waste Site: Perspectives From Economics and Psychology,” Manuscript, University of Colorado, Center for Economic Analysis.

Sharlin, H.I., 1985. "EDB: A case study in the Communication of Health Risk,n Report to USEPA, HIS Associates, Washington, D.C.

Slovic, P. 1967. "Influence of the Response Mode Upon Relative Importance of Probabilities and Payoffs in Risk Taking.” Proceedings of the 75th Annual Convention of the American Psychological Association, pp. 33-34.

Slovic, P., B. Fischhoff, S. Liechtenstein, B. Corrigan, and B. Combs. 1977. "Preference for Insuring Against Probable Small Losses: Insurance Implications,” Journal of Risk and Insurance, Vol. 44, pp. 237-258.

Slovic, P., and S. Liechtenstein. 1968. “The Relative Importance of Probabilities and Payoffs in Risk-taking.” Journal of Experimental Psychology, Vol. 78 (3, pt. 2) 1-18.

Slovic, P. 1986. "Informing and Educating the Public, About Risk,” Risk Analysis, Vol. 6:403-415

Slovic, P., B. Fischhoff, and S. Liechtenstein. 1978. “Accident Probabilities and Seat Belt Useage: A Psychological Perspective.” Accident Analysis and Prevention, 10, 281-285.e

Slovic, P., B. Fischhoff, and S. Liechtenstein, 1977. “Behavioral Decision Theory,” Annual Review of Psychology, 28, 1-39.

Smith, V. K. and W. H. Desvousges. 1986. “Asymmetries in the Valuation of Risk and the Siting of Hazardous Waste Facilities,” Proceedings, American Economic Review, (forthcoming).

Smith, V. L. 1976. “Experimental Economics: Induced Value Theory,” American Economic Review, Vol. 66, pp. 274-279.

Smith, V. L., A. W. Williams, W. K. Bratton, and M. G. Vannoni. 1982. "Competitive Market Institutions: Double Auctions v. Sealed Bid-Offer Auctions,” American Economic Review, Vol. 72, pp. 58-77.

Sorensen, J.H., and D. Mileti. 1987. Public Warning Needs, in Proceeding of Conference XL, A Workshop on “The U.S. Geological Survey’s Role in Hazards Warnings," U.S. Department of the Interior Geological Surveys Denver, Colorado, February 2-3, 1987.

Tverskey, A. and Kahneman, D. 1981. "The Framing of Decisions and the Psychology of Choice," Science, Vol. 211, pp. 453-458, January.

Tverskey, A., and D. Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Biases," Science, Vol. 185, pp. 1124-1131.

United States Environmental Protection Agency, Carcinogen Assessment Group, 1985. Air Toxics Problem in the United States: An Analysis of Cancer Risks for Selected Pollutants, USEPA Office of Policy, Planning and Evaluation, EPA 450/1-85-001, Washington, DC.

Vickrey, W. 1961. "Counterspeculation, Auctions, and Competitive Sealed Tenders," Journal of Finance, Vol 16, pp. 8-37.

Wilkins, L. 1987. Media Coverage of the Chernobyl Incident: A Preliminary Assessment," Presentation to Institute of Behavioral Science Program, Spring, 1987. University of Colorado-Boulder.

Reprinted from
22 October 1976, Volume 194, pp. 389-396

SCIENCE

APPENDIX

Science, Values, and Human Judgment

Kenneth R. Hammond and Leonard Adelman

Copyright© 1976 by the American Association for the Advancement of Science

Science, Values, and Human Judgment

Integration of facts and values requires the scientific study of human judgment.

Kenneth R. Hammond and Leonard Adelman

Scientists and policy-makers are uncertain how scientific facts are to be integrated with social values. For their part. scientists are uncertain whether their contributions should be restricted to presenting the facts. thereby leaving the policy judgment entirely to the political decision-makers. or whether they should also advise politicians which course the scientist believes to be best. And politicians. for their part are uncertain how much scientific information they are supposed to absorb. and how much dependence they should place on scientists for guidance in reaching a judgment about policy (1). As a result. "the scientific community continues its seemingly endless debate about the role of science and scientists in the body politic" (2).

One principal reason for the "endless debate" is that scientific progress has increasingly come to he judged in the context of human values. These judgments find their ultimate expression in the forming of public policy because it is during that process that the products of science and technology are integrated. or aligned. with human values: it is during that process that scientific and technological answers to questions of what can be done are judged in the content of what ought to k done.

The key element, therefore, in the process of integrating social values and scientific facts is human judgment—a cognitive activity not directly observable and generally assumed to he recoverable

Dr. Hammond is professor of psychology and director of the Research Program on Human Judgment and social Interaction. Institute of Behavioral Science. University of Colorado, Boulder 80309. Dr. Adelman is a research associate in the Research Program on Human Judgment and Social Interaction.

only by (fallible) introspection and "self-report." These characteristics. among others. have led to the general belief that human judgment is beyond scientific analysis and therefore little has been learned about the cognitive activity that produces crucial decisions. The integration of social values and scientific information in the effort to form public policy remains largely a mystery.

The fact that an essential element in the policy formation process remains a mystery has serious consequences. one of which is a search for safeguards. Means must be found to avoid both poor judgments and self-serving judgments. TWO general methods have been recommended by scientists for these purposes: (i) the adversary method. in which scientists with differing judgments are pitted against one another in front of a judge or jury, or both, and (ii) the search for and use of scientists who have somehow gained a reputation for wisdom in the exercise of their judgment. Neither of these methods provide enlightenment with regard to the judgment process that produces the ultimate decision. Consequently. we reject both methods because they are "ascientific": they leave the body politic at the mercy of a cognitive activity which remains as much a mystery as ever.

We contend that policy judgments can be brought under scientific study and, as a result a process that is now poorly understood can be examined, understood. assisted, and thereby improved. To support this contention we describe a scientific framework for integrating (i) scientific information (the province of scientists) and (ii) social value judgments (the province of the electorate and their representatives) in a manner that is scientifically. socially, and ethically defensible. and offer an example of its use. First, however, we briefly consider two contrasting viewpoints concerning the role of science and scientists in the body politic.

Contrasting Viewpoints of the Role of the scientist

There arc two main viewpoints: one is that scientists should merely present unbiased information, while the other is that scientists should provide advice with regard to the implications of scientific information. The first view can be illustrated by the comments of Phillip Handler, president of the National Academy of Sciences (NAS), in an interview with Otten, of the *Wall Street Journal.* Often (3) writes; "Once the scientific community has presented the facts. however, it must leave final decisions to the policy-makers and the public. Mr. Handler asserts. "Science can contribute much 10 enhancing agricultural production. but American policy with respect to food aid is not intrinsically a scientific question." Similarly. science can study whether energy independence is technically feasible or whether Soviet underground nuclear tests can be detected, but [Handler] insists. [scientists] must then let regular policy-makers decide whether to try for energy independence or just what arms control proposals to put to the Russians." Often concluded that "Both science and government seem well served by this reasonable man."

Handler's viewpoint as represented in the above quotation is exactly in accord with the two Executive Orders (1918. 1956) concerning the role of the National Research Council. These documents indicate that scientists are to render information to those who are entitled to receive it. but they do not imply that scientists should offer their judgment as to what public policy should follow from their studies.

In practice it may be impossible not to offer such judgments. With the ever-increasing reliance of society on science and technology it is difficult to imagine how modern scientific information could be conveyed to nonscientists without providing such judgments. In a recent editorial in *Science,* Boulding (4) argued that if policy judgments were not offered by scientists. they would be demanded by politicians.

> Every decision involves the selection among an agenda of alternative images of the future. a selection that is guided by some system of values. The values are traditionally supposed to be the cherished preserve of the political decision-maker, but the agenda. which involves fact or at least a projection into the future of what are presumably factual systems, should be very much in the domain of science. . . . [But] if the decision-maker simply does not know what the results d alternative actions will be, it is difficult to evaluate unknown results. *The decision-mob wants to know what are the choices from which he must choose* [italics ours].

Toulmin (5. pp. 102-103) goes further than Boulding. Whereas Boulding notes that politicians may demand policy judgments from scientists, Toulmin argues that it may be part of the scientists. responsibility to offer policy judgments before such judgments are requested by political decision-makers. Thus. "In the early days. the picture was always of the politician as the man who *first* formulated for himself questions about the political options. about the choices he had to make: on this view, he *subsequently* turned to people called "technical advisors and asked them how to do this or that, how much each option would cost. and so on. A 101 of people still see the relationship between the scientist or technologist and the politicians on this model. . . .'" But. Toulmin observes. " . . even during [World War II] scientists were being transformed into people who could very often see a fresh range of policy options *before* the politicians could." Significantly. Toulmin notes that "To some extent, the institutional relationships between politics and science have not yet caught up with this change."

Thus, Toulmin points out that the decision-maker not only wants to know "the choices from which he must choose." as Boulding put it, but he also wants to know which choice the scientist thinks he should choose. Senator Muskie's call for a "one-armed scientist" (one who would not qualify his advice with "on the other hand") exemplifies the politician's demand for an unequivocal answer to the question of what ought to be done as well as to that of what can be done.

This situation has not escaped the attention of students of the role of scientists in the formation of public policy. The presence of, the demand for and the exercise of value judgments has led to a sharp focus on the values and thus on the motives of the scientists who participate in the preparation of NAS reports that affect public policy.

The Focus on Scientists and Their Motives

In his book The *Brain Bank of America* (6. p. 54) Boffey attributes self-serving motives to scientist who provide information and advice to the government within the framework of NAS committees, and thus questions their objectivity and honesty. For example:

> The Academy claims that the most distinctive feature of its committees is that they are independent of any pressure of special interests. . . . But the Academy 's record in recent years suggests that its protestations of Supreme Court impartiality should not be taken at face value. In actual practice many of the Academy's reports have been influenced by powerful interests that have a stake in the questions under investigation.

Boffey admits, however, that "We found no cases of direct, personal conflict of interests at the Academy—no cases, for example. where a committee member profited financially as a direct result of the advice he rendered" (*6*. p. 54). The charge that "many of the Academy's reports have been influenced by powerful interests'" is directed toward the broader social and political motives which he claims influence scientists' judgments.

The NAS has already accepted the principle that the motives of scientists must be examined. Boffey (*6*, p. 87) notes with approval that the NAS demands a "bias statement" from the scientists who provide information to the government. a report that is intended to reveal one's true interests, as may be inferred from a list of "all jobs. consultantships, and directorships held for the past 10 years. all current financial interests whose market value exceeds $10,000, or 10 percent of the individuals holdings: all sources of research support for the past five years. and any other information." such as public stands on an issue which 'might appear to other reasortable individuals as compromising of your independence of judgment. " Thus the NAS has already fallen victim to the ethic of the lawyer (and the journalist). Trust no one. is the rule. unless they can offer this negative proof: I am not now. nor have I ever been. under the control of any incentive to lie. cheat. or otherwise compromise my judgment. Whereas this approach may begin with a request for a "self-report"' on sources of bias. it seldom ends there, as scientists know all too well. Investigation is undertaken by others. and by other means. precisely because the focus has been successfully turned away from methods to persons and their motives.

The results of the focus on persons and their motives can be seen in Poisby's review (7) of Boffey"s book. Poisby indicates what the results might have been had he taken ii similar approach in his review by raising suspicions about Boffey's impartiality and thus his motives. That is. by using "Boffey's own primary method of demonstration: a glance at somebody's background gives a 'motive" for selected characteristics of his performance." Poisby finds that "Boffey"s employer for the writing of this book was Ralph Nader (identified as "consumer champion Ralph Nader" on p. 186). who of late has gotten rather heavily into the business of sponsoring exposes of establishment-type establishments. . . . Under these circumstances of employment. could Boffey have done other than to produce an attack. no matter how flimsily founded. on the Academy?" (7. P. 666).

Poisby's review shows the customary result of such mutual destruction. Boffey's approach, he concludes, "is only good for so much mileage. . . . Arbitrarily imposing the symmetrical assumption . . . that Boffey and the Academy arc both fatally incapacitated by conflict of interest has the effect of condemning both the Academy and the book out of hand" (7. p. 666). In short. because neither the critic nor those criticized can be trusted. the reader, the consumer, and the public remain buried in doubt as to where the truth lies. Thus. Poisby acknowledge that, "After reading The Brain Bank of America I do not know what to think about the Academy as an organization for evaluating the state of scientific knowledge" (7. p. 666). In all likelihood. Poisby is not the only reader of Boffey"s book who no longer knows what to think about the Academy.

It is precisely because scientists have learned that it is not only fruitless, but harmful. to focus on persons and their motives that they have learned to ignore them in their work as scientists. When scientists look for the truth and the troth appears to be in doubt. neither scientific work nor the scientific ethic requires the investigation of the characteristics of the person working on the problem: instead. they require the analysis of the method by which the results are produced. Unfortunately. in the confusion of t he "endless debate" there has been a tendency to forget the scientific procedure and its associated ethics. The focus on persons and their motives . has led not only to the filing of bias statements but to the advocacy of the adversary method for the settlement of disputes about the truth--a method which is ascientific not only in its procedure. hut in its greater commitment to victory rather than to truth.

Scientists as Adversaries

The concept of a "'science court" reached Congress several years ago when Kantrowitz (8} urged that members of Congress "appoint a science advocate for (each) side of the story. . . ."He further suggested that a procedure be worked out which would be "modeled on the judicial procedure for proceeding in the presence of scientific controversy." The final judgment would be exercised by a group of scientific judges who would cross-examine each other and challenge each other's position. Kantrowitz's argument is currently y being given serious consideration by members of the scientific community. *Physics Today* (published by the American Institute of Physics) recently indicated that a science court was worth trying, as did H. Guyford Stever, director of the National Science Foundation (9).

Members of the scientific community are not unanimous. however, in their appraisal of the value of the adversary system. as the following interchange between Platt. Dror. and Waddington in a Ciba symposium indicates (*10*. p. 210):

> PLATT: In the U.S. . . . we are beginning to have something called "adversary science." where scientists speak on public issues. doing their best like lawyers for a particular side. and then in a later case perhaps doing their best for the opposite side. The hope is that in this kind of open confrontation as in a court of law one comes closer to the truth than by having just l accidents of committee structure or unanswered polemics decide the matter.
>
> WADDINGTON: I would strongly oppose that way of advancing science.
>
> PLATT: But somebody should make the total case for a nuclear plant. and somebody should make the total case against the plant for environmental reasons. so that we can see all of both sides before we decide.
>
> DROR: Why shouldn't the two sides make two balanced presentations for and against? Why total . . . ?
>
> PLATT: Do you know a better system?
>
> DROR: Yes, reliance on professional judges in courts and careful policy analysis on television for the public.
>
> PLAIT: Who judges the judges?
>
> DROR: Who judges the juries?
>
> WADDINGTON: That is a piece of politics, not a piece of learning. Learning is not advanced by legal procedures.

The above interchange not only indicates a divergence in viewpoint with regard to a science court and illustrates the morass (Who judges the judges? Who judges the juries?) into which scientists can be drawn because of the focus on persons but it also points to the unproductiveness of the effort Even if the concept of a science court were to be accepted by scientists and even if scientists could be persuaded to make the "total case for (say) a nuclear plant" (*10*. p. 201). the adversary procedure would indicate only who had been judged to be the winner in the arena of competing scientific facts and Scientific judgments. integration of scientific judgments with social values would remain buried in the minds of the judges and the juries (and their judges): the "endless debate" would not be terminated.

It remains to be seen whether a science court with its judges and juries and its ascientific adversary proceedings in which one scientist is pitted against another will be accepted by scientists. In any event. scientists not advocating the adversary method recommend a different ascientific method the person-oriented approach.

Scientists' Advocacy of the Person-Oriented Approach

When scientists have addressed themselves to the function of human judgment in policy formation they have treated the unexamined intuitive abilities of persons as t bough they were somehow superior to the scientific method. For example. in its report on technology assessment to the House Committee on Science and Astronautics. the Committee on Public Engineering Policy (COPEP) of the National Academy of Engineering observed *(11*. p. 17) that “applying only cause-effect [i.e.. scientific] methods to technology-initiated studies produces a mass of data but few broad conclusions.” Apparently assuming that it had no other recourse the committee called for “. . . contributions of talented individuals or groups who can intuitively perform analysis and evaluations . . . ." an approach which “demands an integrated combination of information and value judgments that cannot always be formulated explicitly.”

Not only does the COPEP report illustrate the advocacy of a person-oriented approach to the combination of “information and value judgments" that appeals to the mysterious as a substitute for the scientific method. it provides a clear case of the failure to recognize that it is precisely such person-oriented “combinations of information and value judgments that cannot always be formulated explicitly” that are defenseless against charges of self-serving bias.

Skolnikoff and Brooks *(12)* were critical of the NAS study of science and public policy-making because it suggested that persons who provide science advice should have personal qualities of “intelligence wisdom judgment humanity and perspective” on the ground that "These qualities are so obviously desirable for anybody in a high position that they are hardly helpful criteria.” Yet they are as willing as COPEP or the NAS committee to let the process of combining tacts and values remain subject to the unexamined vagaries of human judgment. For example (12. p. 38):

> Judgment on both technical and nontechnical issues l and on their interaction is thus required [on policy issues]: a logically reasoned single answer is not possible. Judgment is necessarily affected by biases policy preferences. ignorance differing estimates of the nontechnical factors and other vagaries. There is nothing wrong with this: it is unavoidable.

But there is something wrong with this. and this situation is avoidable. What is wrong is that both solutions indicated above focus on persons rather than on method. and both confuse scientific and valuative judgments. That is bad practice: it is bad for scientists. bad for leaders in government. and bad for the public that both are trying to serve. It is bad because it condones and encourages confusion of thought and function. substitutes an appeal to the unknown in place of the knowable. and makes scientists easy targets for charges of self-serving bias. The argument advanced by Skolnikoff and Brooks merely puts a brave face on a bad situation. for they imply that because scientific and valuative judgments cannot be separated there is nothing wrong with confusing them. That argument suggests that if such judgments could be separated. it would be wrong to confuse them. We argue that. from the point of view of science. it is not impossible in principle or in practice to achieve such a separation *(13)*.

A scientific approach toward the role of judgment would be quite different from the person-oriented approach that is embedded in the adversary system, A scientific approach would emphasize that judgment is a human cognitive activity and is therefore subject to scientific analysis. as arc all natural phenomena. The premises of a scientific approach to the relation of science to public policy are: (i) human judgment is a critical part of the policy-making process: (ii) it is a part of the process that remains poorly understood: and (iii) it might well be improved through scientific study. Rather than searching for persons who possess mysterious talents, or indicating that the present situation is unavoidable. the scientific approach to this problem would be similar to the scientific approach to all problems: carry out theoretical and empirical analyses of the process in a manner that is subject to criticism and that provides cumulative knowledge.

The remainder of this article (i) provides an example that illustrates the social costs of employing the adversary system and the person-oriented appreach and (ii) outlines a scientific framework for integrating scientific information and social values in the formation of public policy *(14)*.

An Example of Contrasting Approaches

In 1974, the Denver Police Department (DPD), as well as other police departments throughout the country. decided to change its handgun ammunition. The principal reason offered by the police was that the conventional round-nosed bullet provided insufficient “stopping effectiveness” (that is. the ability to incapacitate and thus to prevent the per. son shot from firing back at a police officer or others). The DPD chief recommended (as did other police chiefs) the conventional bullet be replaced by a hollow-point bullet. Such bullets it was contended. flattened on impact. thus decreasing penetration. increasing stopping effectiveness. and decreasing ricochet potential.

The suggested change was challenged by the American Civil liberties Union. minority groups. and others. Opponents of the change claimed that the new bullets were nothing more than outlawed “dum-dum” bullets. that they created far more injury than the round-nosed bullet. and should. therefore. be barred from use. As is customary judgments on this matter were formed privately and then defended publicly with enthusiasm and tenacity, and the usual public hear. ings were held. Both sides turned to ballistics experts for scientific information and support.

Adversary, Person-Oriented Approach

From the beginning both sides focused on the question of which bullet was best for the community. As a result of focusing on bullets and their technical ballistics characteristics. legislators and city councilmen never described the social policy that should control the use of force and injury in enforcing the law: they never . specified the relative e importance of the societal characteristics of bullets (injury stopping effectiveness or ricochet). Instead. the ballistics experts assumed that function. When the legislators requested their judgment as to which bullet was “best.”. the ballistics experts implicitly indicated the social policy that should k employed. That is in recommending the use of a specific bullet. they not only implicitly recommended specific degrees of injury. stopping effectiveness. and ricochet but also recommended a social policy regarding the relative importance of these factors. In short. the legislators. function was usurped try the ballistics experts who thus became incompetent and unauthorized legislators-incompetent because of their lack of information about the social and political context in which a choice would be made: unauthorized because they assumed a function for which the! had not been elected.

In parallel fashion. the ballistics experts turned their scientific-technical function over to those who should have formed social policy-the legislators. When the experts presented scientific information to policy-makers about various bullets, they found themselves disputing ballistics data with legislators who preferred a different type of bullet. Thus. the legislators, none of whom were ballistics experts in their turn served as incompetent ballistics experts in the hearings.

When legislators and scientists accept the adversary system with its concomitant person-oriented approach as the primary means for integrating science and social values, they may expect to find a reversal of roles, and when scientists accept the person-oriented approach they may expect to be confronted by challenges to their objectivity y *(15)*. The outcome is well represented by the comment of one legislator who said to an opponent *(16)*: “You have your expert and we have ours. . . .”

A Scientific Approach

We now consider. by way of an example, a scientific method for integrating scientific information and social values that is scientifically. socially. and ethically defensible. This method was employed in solving the dispute about handgun ammunition for the police as described above. A broad outline of the method is presented *(17)*.

The general framework of the method as it was applied to the above problem is shown in Fig. 1. Basic to any policy involving scientific information are objectively measurable variables (Fig. 1. left). Scientific judgments regarding the potential effects of technological alternative . are also required (Fig. 1. middle). Finally. social value judgments by policy-makers or community representatives are necessary (Fig. 1. right). The overall acceptability of an alternative is determined by how closely its potential effects satisfy the social values of the community.

Application of this framework to the bullet dispute involved three phases: (i) externalization of social value judgments: (ii) externalization of scientific judgments: and (iii) integration of social values and scientific judgments. Each phase is discussed in turn.

Phase 1: Externalizing Social Value Judgments

The participants in phase included the mayor and city council other elected officials representatives of the DPD (including the chief). and official representatives of community organizations. including minority groups and members of the general public. Each person was asked 10 make judgments concerning the relative desirability of hypothetical bullets, described in terms of their (i) stopping effectiveness, (ii) seventy of injury, and (iii) threat to bystanders. These value judgments were made at the console

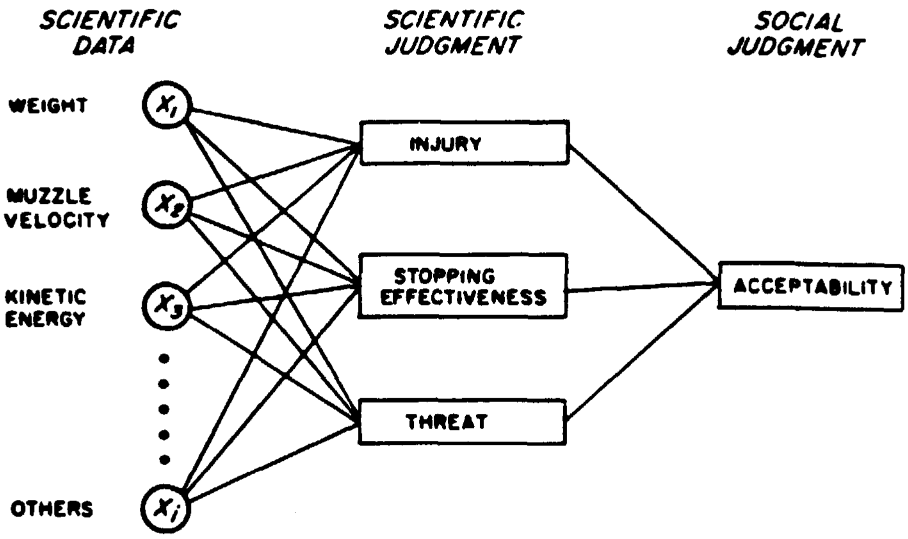


Fig. 1. A pictorial representation of a framework that combines scientific facts with social values.

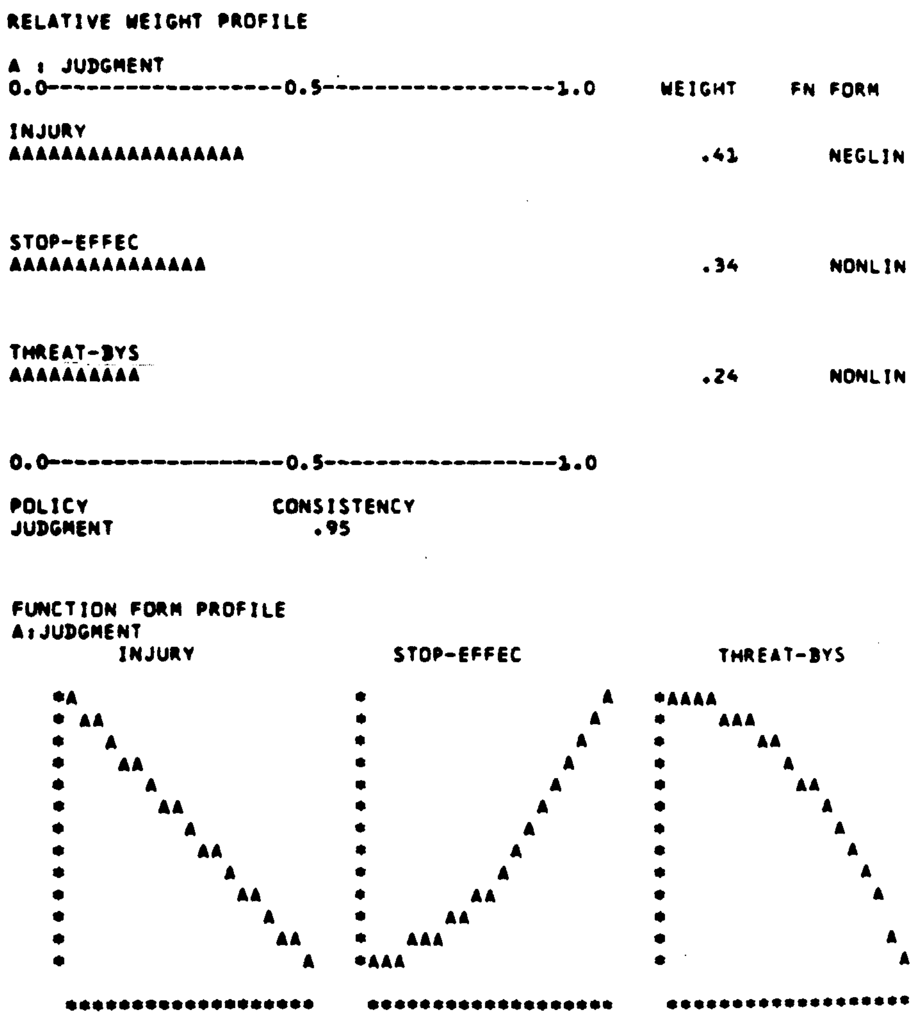


Fig. 2. A reproduction of a participant's interactive computer display of relative weights and functional relations, *FN.* function: *NEGLIN,* negative linear, *NONLIN,* nonlinear.

of an interactive computer terminal. After their judgments were made the participants were immediately shown the relative importance they gave to each of these three functional characteristics of bullets. That is, a statistical analysis was carried out on the data and the results were then displayed at the terminal for the participant to observe *(18)*. In addition. each participant was shown the form of the relation (linear curvilinear) between his or her judgment and each of the three characteristics mentioned above. in this way, each participant saw the relative importance he or she attached to stopping effectiveness. injury, and threat to bystanders, as well as the optimal point for each (a typical display is shown in Fig. 2).

After viewing the display, the participants were asked if the results reflected their considered judgment. The data corrected when necessary, were then stored, and a cluster analysis was carried out in order to discover whether different groups held different judgment policies. Widely differing policies with regard to the relative importance of each characteristic were found, although the functional relations between bullet characteristics and judgments were all found to be approximately linear in form.

The above procedure presides objective, visible data not otherwise available. The same procedure was used to externalize the required scientific judgments.

Phase 2: Externalizing Scientific Judgments

A panel was assembled that included one firearms expert, one ballistics expert and three medical experts in wound ballistics. The judgments of these experts provided scientific information regarding the stopping effectiveness. severity of injury, and threat to by standers of 80 bullets. The data for these bullets were obtained from the National Bureau of Standards. Each dimension (stopping effectiveness, injury and threat to bystanders) was judged separately for each of the 80 bullets: agreement among the experts was found to be quite high *(19)*. Only the results for stopping effective- - ness and injury are summarized here. as these were the central factors in the controversy.

Three factors were found to be important in judgments of slopping effectiveness: (i) The maximum diameter of the temporary wound cavity: (ii) the amount of kinetic energy lost by the bullet in the target: anti (iii) the muzzle velocity of the bullet. The close but not perfect relalion between stopping effectiveness and injury (shown in Fig. 3) is reflected in the fact that independent judgments of potential injury were positively related to the amount of kinetic energy lost maximum diameter of the temporary cavity and degree of penetration.

The data in Fig. 3 are important because they suggest that contrary to previous. unexamined assumption there is not a perfect relation between stopping effectiveness and injury: increasing one does not necessarily increase the other. These data illustrate the value of scientific information by indicating the possibility of finding a bullet that increases stopping effectiveness without increasing injury *(20)*.

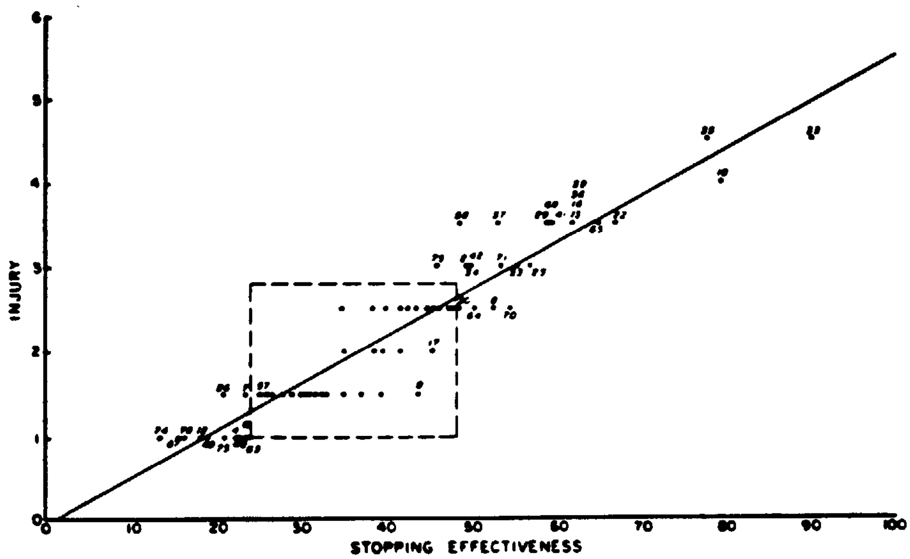


Fig. 3. the average rating of stopping effectiveness and injury are plotted above. Each point on the graph represents a bullet. The diagonal line determined by linear regression analysis. indicates the average value of injury for bullets with a specific level of stopping effectiveness. Bullets above the line produce more injury than the average bullet with the same stopping effectiveness bullets below the line produce less injury.

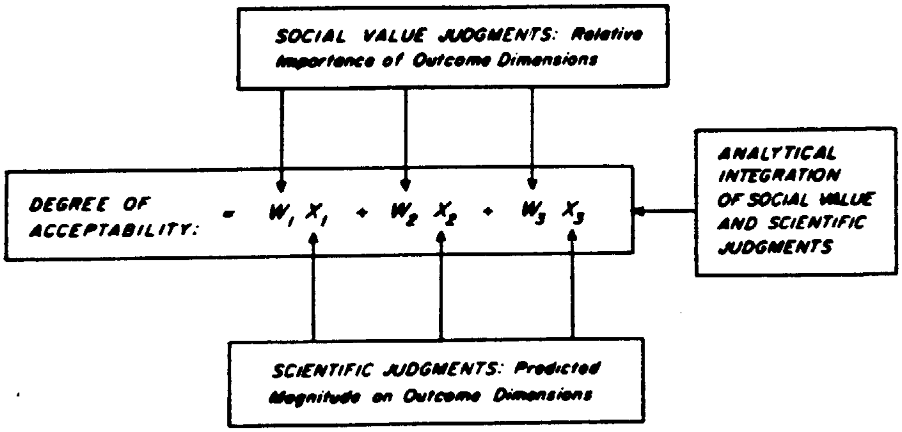


Fig. 4. A schematic representation of the analytical combination of scientific facts and social values.

Phase 3: Integrating Social Values and Scientific Information

Social value judgments and scientific judgments were combined by means of the equation in Fig. 4. where the separation and combination of the judgments of policy-makers and scientist-technologists may be seen. We used the following algebraic form of this equation

$$Y_i = W_1X_1, + W_2X_2 + W_3X_3$$

where Y, is the overall acceptability of a bullet: W_j, j = 1.3. indicates the weight. or relative importance policy-makers placed on stopping effectiveness. injury. and threat to bystanders: and X_j, j = 1. 3 are the experts judgments regarding stopping effectiveness. injury and threat to by standers.

Because phase I resulted in a variety of different weights on stopping effectiveness, injury, and threat to by standers.

the city council took all three factors into consideration by placing l equal weight on each. As a result, when considering stopping effectiveness and severity of injury only, the appropriate bullet is one which lies farthest from the line of average relation in Fig. 3. this distance from the line being measured perpendicularly from the point to the line. Bullet 9 in Fig. 3 satisfies this criterion. It has greater stopping effectiveness and is less apt to cause injury (and less apt to threaten bystanders) than the standard bullet then in use by the DPD (bullet 57). In addition, bullet 9 (a hollow-point bullet) is less apt to cause injury than is bullet 17, the hollow-point bullet recommended by the DPD. Bullet 9 was accepted by the city council and all other parties concerned. and is now being used by the DPD *(21).*

Finally. three points should be mentioned with regard to the application of judgment analysis to the above problem,

1) Intense political and social conflict existed prior to our participation in the project. During the controversy a Denver police officer was killed by a hollow-point bullet: as a result, hundreds of policemen staged a march that ended in demands on both the police chief and the governor that the police be permitted to use hollow-point bullets. Members of the city council and others seemed convinced that the usual adversary methods had failed, and that they faced a dangerous impasse. The fact that the above procedures were used in these circumstances indicates that elected officials and special interest groups can accept a scientific approach to critical social problems, even when they have become immersed in sharp political dispute. Moreover, interviews with members of the city council and others not only indicated a high degree of satisfaction with the procedure but appreciation of its impersonal approach as well.

2) The procedures were applied to complex technical judgments. As far as we could determine at the time of the research no standard quantifiable definition of severity of injury (with regard to handgun ammunition) had ever been developed. Moreover, in developing such a definition, and in making their judgments, the ballistics experts considered **11** distinct characteristics of handgun ammunition.

3) The procedure is general in nature. Despite the apparent simplicity of the framework presented in Fig. 4. judgment analysis can be applied to a variety of complex problems involving value judgments and scientific judgments by differentiating the elements in Fig. 4 in a hierarchical fashion (22).

Scientific Defensibility

The above method is scientifically defensible not because it is flawless (it isn't), but because it is readily subject to scientific criticism. It is vulnerable to such criticism (i) because its aim is to meet appropriate standards regarding replication. quantification and logic for the problem under study (an aim all scientific efforts share) and (ii) because the procedure for achieving that aim is public (as all scientific effort must be). The locus and degree of imperfection in method and procedure are thus available for public inspection and subsequent improvement. In short, the process provides the opportunity for cumulative knowledge, as scientific efforts should.

Social Responsibility

The above method is socially responsible because it provides a public framework for (i) separating technical. scientific judgment from social value judgments and (ii) integrating them analytically, not judgmentally. The separation phase permits elected representatives to function exclusively as policy-makers, and scientists to function exclusively as scientists. Neither role is confused or exchanged because policy-makers are not forced to become amateur scientists, nor are scientists required to make judgments on public policy. The intergrative phase provides an overt, rather than covert, process for combining facts and values. Because the social values in the community are identified before the decision is implemented, the decision process is not seen to be a mere defense of a predetermined choice: rather it can be evaluated in terms of its rational basis before the final choice is made.

Ethical Standards

Ethical and scientific standards converge in the process of combining facts and values because both scientific ethics and public ethics require controls against bias. Scientific control against bias is illustrated by the use of the double-blind control in experiments: in the above procedure public control against bias is carried out by a similar blindness. That is, the method described above has the advantage of situating all parties (policy-makers, scientists, and the public) behind what Rawls (23. p. 136) calls "a veil of ignorance." It fits Rawls" requirement that the participants should not "know how the various alternatives [would] affect their own particular case and they are obliged to evaluate principles solely on the basis of general considerations," in the approach described above, the technical experts were not aware of the relative importance the policy-makers placed on the three societal characteristics of bullets, nor were the policy-makers aware of the technical judgments made by the scientists-technologists in regard to specific bullets. In short by implementing Rawls" veil of ignorance, both scientific and ethical standards were met.

Conclusion

Current efforts to integrate scientific information and social values in the forming of public policy are confused and defeated by the widespread use of ascientific methods--the adversary system and the person-oriented approach. The adversary system suffers from an ascientific commitment to victory rather than truth: the person-oriented approach suffers from an ascientific focus on persons and their motives rather than on the adequacy of methods. The reason for the widespread use of both lies in the failure to recognize that human judgment can be brought under scientific, rather than ad hominem, analysis. The argument advanced here is that a scientifically socially, and ethically defensible means for integrating science and human values can be achieved.

References and Notes

1. See. for example. Public Law 92-484 which established the Office of Technology Assessment.
2. J. W. Curlin. *Science* 190. 839 (1975).
3. A. L. Otten. *Wall Street Journal*. 3 April 1975. p. 12.
4. K. E. Boulding. *Science* 190. 423 (1975).
5. S. E. Toulmin. in *Civilization and Science: In Conflict or Collaboration?*. Ciba Foundation Symposium 1 (Elsevier. Amsterdam. 1972).
6. P. Boffey. *The Brain Bank of America* (McGraw-Hill. New York. 1975).
7. N. W. Polsby. *Science* 190. 665 (1975).
8. Hearings before the House Committee on Rules and Administration (1971).
9. See the article by J. N. Wilford. *New York Times*. 19 February 1976. p. 3.
10. *The Future as an Academic Discipline*. Ciba Foundation Symposium 36. (Elsevier. Amsterdam. 1975).
11. Committee on Public Engineering Policy. National Academy of Engineering. *A Study of Technology Assessment* (Government Printing Office. Washington. D.C.. 1969).
12. E. B. Skolnikoff and H. Brooks. *Science* 187. 35 (1975).
13. There are clear indications that scientists are beginning to acknowledge the need for explicit methods for decision-making in areas where science and the public interest intersect. Two recent NAS committee reports [*Environmental Impact of Stratospheric Flight* (1975): *Decision Making for Regulating Chemicals in the Environment* (1975)] as well as others mentioned in the latter describe the application of normative decision theory to such problems. Although these efforts represent a clear step forward through their insistence on the use of an explicit framework for decisions. they do not indicate how such decisions might be assisted or improved through the study of human judgment.
14. For a general review of current research on

judgment and decision-making see [P. Slovic, B. Fischhoff, S. Lichtenstein, in *The Annual Review of Psychology* (Annual Reviews, Palo Alto, Calif., in press), vol. 28]. See also M. Kaplan and S. Schwartz, Eds., *Human Judgment and Decision Processes* (Academic Press, New York, 1975); W. Edwards, M. Guttentag, K. Snapper, in *Handbook of Evaluation Research*, E. L. Struening and M. Guttentag, Eds. (Sage, Beverly Hills, Calif., 1975), vol. 1; R. A. Howard, in *Proceedings of the Fourth International Conference on Operational Research* (Wiley-Interscience, New York, 1966); H. Raiffa, *Decision Analysis: Introductory Lectures on Choices Under Uncertainty* (Addison-Wesley, Reading, Mass., 1968).

15. Can the adversary system produce this confusion of roles at the national level, and does it have similar negative effects? Apparently it can, and does. For example, in Polsby's review of Boffey's book, Polsby (7, p. 666) states: "Boffey notes, in criticizing a National Academy of Engineering committee on pollution abatement, that it was no more qualified than any other group of citizens to judge what should be 'wise' public policy." (In this instance, Boffey argues that scientists overstepped their bounds and should have confined their role to presenting the facts.) "Sound doctrine," observes Polsby, "and yet Boffey criticizes another of the Academy's committees for taking on an assignment pertinent to a naval communications project that did not include evaluating its 'desirability,' and for not venturing to raise 'questions as to the basic worth' of the space shuttle program." (In this instance, Boffey argues that scientists failed to help form social policy and thus failed in their responsibility to the public.) Thus, concludes Polsby, "the Academy is damned if it does pronounce on the overall wisdom of public policies, and damned if it doesn't."

16. Public Broadcasting Service, "Black Horizons," 16 February 1975.

17. K. R. Hammond, T. R. Stewart, L. Adelman, N. Wascoe, *Report to the Denver City Council and Mayor Regarding the Choice of Handgun Ammunition for the Denver Police Department* (Report No. 179, University of Colorado, Institute of Behavioral Science, Program of Research on Human Judgment and Social Interaction, Boulder, 1975).

18. To determine the relative importance a person places on each characteristic, linear multiple regression analysis was performed to obtain the beta weights on each of the three judgment dimensions, or factors. The absolute value of the beta weight for a factor was then divided by the sum of the absolute values of the beta weights over all factors to determine the relative weight, or importance placed on each factor. The relative weights were displayed on the computer console. For technical details on the procedure see [K. R. Hammond, T. R. Stewart, B. Brehmer, D. O. Steinmann, in *Human Judgment and Decision Processes*, M. Kaplan and S. Schwartz, Eds. (Academic Press, New York, 1975)].

19. The judgment dimensions were defined as follows. (i) *Stopping effectiveness*: the probability that a 20- to 40-year-old man of average height (5'10") and weight (175 lbs) shot in the torso would be incapacitated and rendered incapable of returning fire. Judgments ranged from 0 to 100, indicating, on the average, how many men out of 100 would be stopped by a given bullet. (ii) *Severity of injury*: the probability that a man, as described above, shot in the torso would die within 2 weeks of being shot. (iii) *Threat to bystanders*: penetration was defined as the probability that a bullet would pose a hazard to others after passing through a person shot in the torso at a distance of 21 feet. Ricochet was defined as the probability that a bullet would pose a hazard after missing the intended target at a distance of 21 feet.

20. The separation of stopping effectiveness from injury that is indicated in the graph for bullet 9 was not due to inconsistencies and inaccuracies in the experts' ratings. The three medical experts agreed that the shape of the temporary cavity is an indicator of differences in severity of injury for bullets with the same stopping effectiveness. More severe wounds are produced by bullets that have a long, wide temporary cavity; less severe wounds localize the maximum diameter of their temporary cavity and do not penetrate deeply. According to all three experts, a temporary cavity that reaches a maximum diameter of 10 to 15 cm at 5 to 7 cm from the surface, and does not penetrate more than 15 cm, would provide the best compromise between stopping effectiveness and survivability.

21. The time, manpower, and cost of the handgun study were as follows. (i) The project was completed in 6 weeks and (ii) research personnel included four people of whom one worked full time. Total cost, including salaries of the project staff, did not exceed $6000; an additional $3500 was required to pay the travel and consulting costs of the ballistics experts.

22. For examples of the application of a hierarchical framework, see K. R. Hammond, J. Rohrbaugh, J. Mumpower, L. Adelman, in *Human Judgment and Decision Processes: Applications in Problem Settings*, M. F. Kaplan and S. Schwartz, Eds. (Academic Press, New York, 1976).

23. J. Rawls, *A Theory of Justice* (Harvard Univ. Press, Cambridge, Mass., 1971).

24. Supported by National Institute of Mental Health grant MH-16437. We thank S. Cook, D. Deane, and B. Fischhoff, among many others, for their help.