

Setting Risk Priorities: A Formal Model

Jun Long¹ and Baruch Fischhoff^{1,2*}

This article presents a model designed to capture the major aspects of setting priorities among risks, a common task in government and industry. The model has both *design* features, under the control of the rankers (e.g., how success is evaluated), and *context* features, properties of the situations that they are trying to understand (e.g., how quickly uncertainty can be reduced). The model is demonstrated in terms of two extreme ranking strategies. The first, *sequential risk ranking*, devotes all its resources, in a given period, to learning more about a single risk, and its place in the overall ranking. This strategy characterizes the process for a society (or organization or individual) that throws itself completely into dealing with one risk after another. The other extreme strategy, *simultaneous risk ranking*, spreads available resources equally across all risks. It characterizes the most methodical of ranking exercises. Given ample ranking resources, simultaneous risk ranking will eventually provide an accurate set of priorities, whereas sequential ranking might never get to some risks. Resource constraints, however, may prevent simultaneous rankers from examining any risk very thoroughly. The model is intended to clarify the nature of ranking tasks, predict the efficacy of alternative strategies, and improve their design.

KEY WORDS: Risk ranking; muddling through; risk communication; comparative risk

1. INTRODUCTION

Scarce time and resources prevent individuals and societies from doing everything that they might to reduce risks to health, safety, and environment.^(1,2) When people face many risks, even evaluating the options for risk management can be overwhelming. One common strategy for coping with such overload is to rank risks in terms of their magnitude. Having done so, one can begin evaluating the options by starting with those directed at the largest risks.

This article provides a general analytic approach

to evaluating the efficacy of alternative risk-ranking strategies. The model can be used prescriptively, in order to design prioritization processes. It can also be used descriptively, in order to predict the efficacy of actual processes. Its parameters reflect both the goals that risk managers set for themselves and the situations that confront them. Once a situation has been characterized in the model's terms, one can, for example, compare the efficiency of devoting fixed resources to examining all members in a class of risks simultaneously or to learning sequentially about a series of focal risks (e.g., those nominated by the news media for the "risk-of-the-month club").

Setting priorities is, of course, as old (and as general) a process as making lists of personal worries. Ranking risks has gained prominence as a public policy tool, in part, through the U.S. Environmental Protection Agency's (EPA's) sustained efforts to evaluate its own resource allocations. In 1987, EPA published *Unfinished Business: A Comparative As-*

¹Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA.

²Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA.

*Address correspondence to Baruch Fischhoff, Department of Engineering & Public Policy, Carnegie Mellon University, Pittsburgh, PA 15213; baruch@cmu.edu

assessment of Environmental Problems.⁽³⁾ In it, EPA's senior scientific staff divided a wide range of environmental and ecological risks into 31 categories. Some categories reflected existing EPA programs or statutes; others were the responsibility of other agencies or of none at all. The ranking recognized the multi-attribute character of risks, by including health, ecological, and welfare effects. The technical feasibility and costs of controlling the risks were left for a later day. *Unfinished Business* propelled a public discussion of risk-based priority setting.⁽¹⁾ It was followed by the EPA Science Advisory Board's *Reducing Risks: Setting Priorities and Strategies for Environmental Protection.*⁽⁴⁾ In the ensuing decade, EPA sponsored a string of state, regional, and local comparative risk analyses.^(5,6) These efforts have generally been viewed as useful experiences, bringing together diverse individuals and reaching some degree of consensus on risk priorities. Given the need for public discourse about the meaning of "risk,"⁽⁷⁾ these encounters might have been socially valuable, whatever progress they made in determining the relative magnitude of risks. They are, in any case, but one representative of a process that plays itself out in many places, where regulatory agencies, industrial safety departments, public schools, hospitals, and the like decide where to focus their attentions.

If one accepts the need for setting risk priorities, then it is important to design the process effectively. Using rankers' time well should increase both the product of their labors and their willingness to invest in the process. This analysis begins by offering a model characterizing the fundamental structure of risk ranking. That model is then realized in the form of a simulation, designed to compute the efficacy of alternative ranking strategies. It is used here to examine several archetypal situations, chosen to capture variants on the extremes of sequential and simultaneous evaluation. The article concludes with a discussion of the data demands for applying the model to specific risk domains, as well as possible elaborations.

As characterized by Lindblom^(8,9) in the absence of systematic, simultaneous ranking, priorities change through some form of "muddling through"; As individuals or organizations, we face some current jumble of risks. Periodically, a specific hazard draws our attention. After investing some resources, we understand it better, possibly changing its place in the overall risk ranking. Then, we turn our attention to the next hazard, and the next. Over time, this sequential process should gradually improve the prioritiza-

tion of the whole set. How quickly that happens should depend on (1) the uncertainties in the situation we face, (2) what we hope to get out of it, and (3) how we allocate our resources. The same factors should determine our success, if we try to learn about several (or all) risks at once, but must spread the same learning resources over them.

The model presented here is designed to capture these three elements of risk-ranking situations. Its logic is as follows: At the beginning of a ranking period, beliefs about the magnitudes of risks are summarized in terms of subjective probability distributions (SPDs). Their spread reflects uncertainty about (1) the expected magnitude of the adverse effects that each hazard can cause, and (2) the weights to assign to those effects, when creating an aggregate measure of risk.^(10,11) The rankers decide how to allocate their resources in order to learn more about one, some, or all of these risks. After that learning period, the subjective probability distributions are updated as a function of how readily the uncertainties yield to such scrutiny. After completing each period, rankers evaluate the return on their investment, measured in terms of the reduction in their "confusion" about the relative magnitudes of the risks. The model is implemented in Analytica,³ and demonstrated here with hypothetical scenarios, meant to capture some archetypal situations. We believe, however, that it can clarify the nature of some risk-ranking tasks, even without pursuing full computational solutions.

2. MODEL DESCRIPTION

2.1. Assumptions

For the sake of simplification, the following assumptions are made in the current implementation of the model:

1. Risks are stable over the prioritization period. The ranking process may have been initiated by a perception that risks have changed (as well as by a perception that they have been misestimated). Whether or not that is the case, no further changes are allowed during the ranking process.
2. The risks can be represented along a single dimension, aggregating whatever attributes rankers deem relevant.⁽¹¹⁻¹⁵⁾ That dimension

³ Analytica Version 1.1.1, Lumina Decision Systems, Inc., 1997; <http://www.lumina.com/>.

has at least interval-scale properties. As mentioned, the overall uncertainty regarding the magnitude of a risk reflects the uncertainty about both how to weight the attributes of risk and how to characterize the risk on each relevant attribute. Thus, it is possible to know risks very well (in terms of the expected magnitudes of their effects, based on the different attributes), but still not know what to think about them (in terms of the different trade-offs that they present across the attributes).

2.2. Model Overview

The model represents each risk by an SPD over the risk measure. The risks are ranked according to a *ranking criterion*, representing some fractile of each SPD (e.g., the median, 0.99). Each round of the risk-ranking process involves devoting *learning resources* to reducing uncertainties about the risks. The updating process could be Bayesian (if one is designing an optimal process) or non-Bayesian (if one is predicting an imperfect one). When the initial SPDs are biased, the learning process will tend to correct them, perhaps increasing uncertainty. At the end of each round, the *residual confusion* (RC) in the ranking is measured by the overlap in the SPDs. The critical design decision is the *allocation rule* for spreading the learning resources available for a round, across the risks.

The initial SPDs are a state of nature, with which the rankers must contend; so is the *uncertainty reduction function* (URF), describing how quickly uncertainty decreases, as a function of the resources invested in understanding a risk. The rankers can control (1) how they rank the risks (given the SPDs), (2) how they allocate those learning resources, (3) how they update their beliefs after learning, and (4) how they evaluate their RC. The next section formalizes these two states of nature and describes four design choices.

2.3. Model Parameters

2.3.1. States of Nature

Initial Estimates. The model starts with a set of hazards, characterized in terms of SPDs, reflecting rankers' beliefs about the magnitudes of the risks. As mentioned, the uncertainties may reflect both questions of fact (how large each adverse effect is ex-

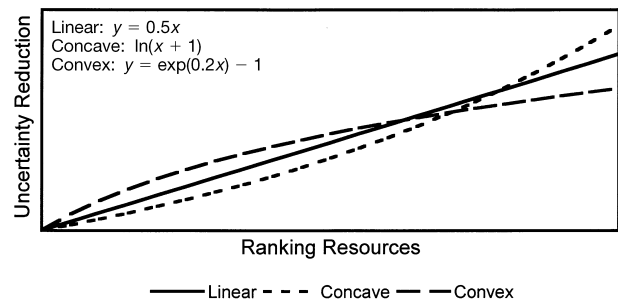


Fig. 1. Three possible uncertainty reduction functions. In the linear case, uncertainty decreases steadily as resources are invested; in the concave case, the initial investment has disproportionate marginal utility; in the convex case, one learns relatively little until a significant investment has been made.

pected to be) and questions of value (how the effects should be weighted). Analytical convenience favors characterizing SPDs in standard terms (e.g., a normal distribution, with specified mean and standard deviation). However, any form of distribution is possible in principle.

Initial SPDs can be biased by measurement errors, theoretical misconceptions, and erroneous perceptions of the facts.⁽¹⁶⁻¹⁸⁾ A model's specification includes whether bias is suspected in the initial SPDs.

Uncertainty Reduction. Barring systematic bias, the uncertainty about a risk will decrease as a function of the resources invested in learning about it. Figure 1 shows three such URFs. One is *linear*, meaning that uncertainty reduction is proportional to the ranking resources spent. The second is *concave*, reflecting cases where uncertainty reduction is initially very easy, but gets harder as time goes on. The third is *convex*, reflecting cases with little initial progress, until, after some significant investment, the risk quickly reveals itself. Other functions could be developed for specific risks, reflecting the nature of the relevant science and rankers' learning process.

2.3.2. Design Choices

Ranking Criterion. The SPDs are translated into rankings, by characterizing each by a common fractile. That might be the mean (as a "best guess" at the risk's value), an extreme high fractile (as a "conservative" estimate), or any other fractile that reflected the rankers' values.

Attention Allocation Rules (AARs). With simultaneous ranking, learning resources are divided equally across the risks. With more sequential processes, other rules allocate rankers' attention to clar-

ifying particular risks.⁴ A risk might be chosen at random (e.g., to make the selection process unpredictable to risk managers). Or, some special property might draw attention. For example, rankers might focus on risks with the largest means or the largest n th fractiles, or with the greatest uncertainty or coefficient of variation. Attention might depend on recent performance, such as focusing on risks that have had recent anomalous events.⁽¹⁹⁾ Slovic, Fischhoff, and Lichtenstein⁽²⁰⁾ discuss the attention drawn by events with high “signal value,” which observers may fear signals a change in the hazard, whose risk level needs to be reassessed. Risks with great uncertainty should be particularly likely to produce such unpleasant surprises (and less noticed pleasant ones).

Updating Process. Uncertainty reduction proceeds by combining new information with the initial SPDs in order to get posterior SPDs. Bayesian updating can be assumed, either when the ranking process is required to proceed that way or when it is expected to do so naturally (at least to a first approximation). Given the power of Bayesian inference, our model focuses on it and on the use of conjoint distributions, with which updating is computationally simple. However, individuals are not always Bayesians. Sometimes, tiny bits of new information swing opinions; at other times, people hold tenaciously to old beliefs, even in the face of great counterevidence. The model can accommodate such possibilities by the simple expedient of overweighting or underweighting new evidence, in a Bayesian calculation. Other updating rules are also possible.^(21–23)

Ranking Evaluation Criteria. Risk ranking is considered successful to the extent that it reduces the overlap among the SPDs describing different risks. The metric for the overlap between the SPDs for two risks, A and B, is the probability that a point drawn randomly from A’s SPD will be larger than a point drawn randomly from B’s SPD. If two SPDs overlap completely, then the $\text{Prob}[\text{Risk A} > \text{Risk B}] = 0.5$; if A dominates B (Fig. 2A), then $\text{Prob}[\text{Risk A} > \text{Risk B}] = 1$; if A is stochastically larger than B (Fig. 2B), then $0.5 < \text{Prob}[\text{Risk A} > \text{Risk B}] < 1$. We will use the absolute difference between $\text{Prob}[\text{Risk A} > \text{Risk B}]$ and 0.5 to reflect how much two distributions over-

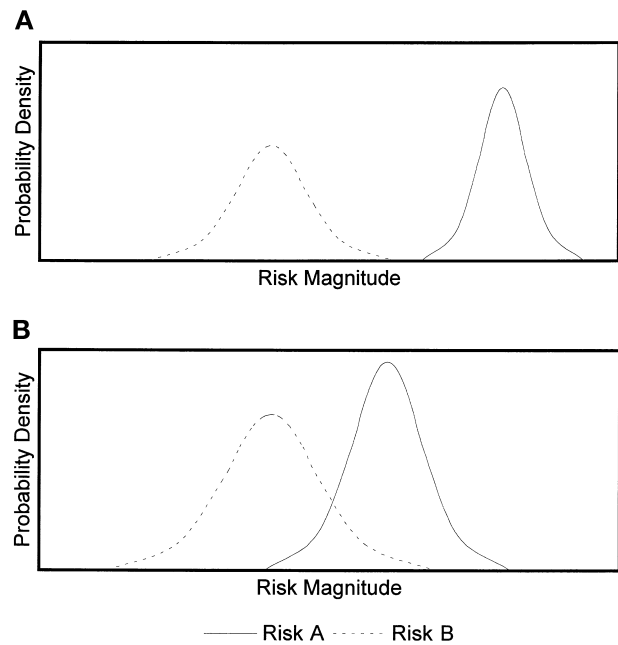


Fig. 2. Illustrations for calculation of residual confusion index: (A) Risk A dominates Risk B; (B) Risk A stochastically dominates Risk B.

lap. By this measure, larger values indicate less overlap and, hence, clearer priorities. Summing this metric over all N pairs of the n SPDs, and then subtracting from the maximum possible overlap (n coincidental distributions), produces a measure for the overall rankability of the risks, or the RC. It approaches zero as the SPDs become completely distinct.⁵

3. MODEL APPLICATION: UNBIASED INITIAL SPDs

This section illustrates the model by applying it to a simple case of ranking three risks. Different model features are then manipulated, in order to show the ranking process’s sensitivity to them. For expository purposes, it starts with the simplest (and uninteresting) case of very distinct initial SPDs. It proceeds to risks with stochastically dominant distributions, and then to ones without stochastically dom-

⁴ Given N risks, with total ranking resources RR , and URF $f(x)$: For sequential ranking, risk reduction for the focal risks is $y = f(RR)$, with no reduction for other risks. For simultaneous ranking, uncertainty reduction is $y = f(RR/N)$. If the risk distribution is normal, the standard deviation after the ranking process is $d_i = f(RR/N)$. If there is bias in the initial ranking, this is equivalent to applying $n = [d_i^2/(d_i - y)^2] - 1$.

⁵ This rule treats overlap among large risks the same as overlap among small ones. This would befit a situation in which, say, it was as important to decide that a risk no longer needed attention as to decide that another risk really needed it. The rule avoids assigning absolute-scale interpretations to risk values. Providing the measures could be assigned, it could be replaced by other rules (as could other discretionary features of the model).

inant SPDs. The initial SPDs are assumed here to have no systematic bias, in the sense of being centered on the true SPD. Thus, ranking should tighten, but not shift, the distributions. Section 4 allows for biased initial beliefs. In such cases, information gathering could increase uncertainty, even as it brings the rankings closer to their desired values.

3.1. Very Distinct Initial SPDs

This is the simplest case and an unlikely candidate for an actual risk ranking. Not only are the initial SPDs unbiased, but the uncertainties of the SPDs generate minimal overlap among the risks. Figure 3A shows such a case, with the three risks being $N(40, 2)$, $N(58, 3)$, and $N(80, 4)$, respectively. Here, the choice of ranking criterion (or fractile) has no effect on the initial prioritization. The choice of AAR will affect

which risks' uncertainties are reduced. The applicable URFs will determine how far that reduction proceeds. However, the ranking will be the same whether one SPD is tightened a lot or each is tightened some.

3.2. Stochastically Dominant Initial SPDs

Figure 3B shows SPDs for a more realistic case, with overlap among distributions, having $N(10, 7)$, $N(20, 6)$, and $N(30, 5)$. The stochastic dominance among these distributions means that the initial ranking is the same by any fractile. However, rankers might still be uncomfortable with the degree of overlap among the distributions. As a result, our simulations focus on how RC proceeds with different allocations of attention. We used $y = \ln(x + 1)$ as the URF, indicating cases where uncertainty initially reduces quickly.

Figure 4 shows how RC declines with the investment of ranking resources according to different AARs, either for 100 resource units (Fig. 4A) or just the first 5 units (Fig. 4B). These resources are invested either equally in all three risks simultaneously or exclusively in one risk. In this case, as more ranking resources are invested, the RC value decreases, both for the simultaneous strategy and for any of the three sequential ranking strategies (i.e., looking at just a single risk during this ranking period). However, the extent to which the overlap among the SPDs decreases depends on the strategy: With 4 or more resource units, RC is most efficiently reduced with simultaneous learning, and increasingly so. An intuitive account of this pattern is that with any reasonable level of resources, each risk receives enough attention to make a dent in its uncertainty. However, when resources are limited (Fig. 4B), sequential risk ranking can be more effective, especially when targeted at the risk causing the greatest confusion. In this case, that is Risk 2, which overlaps both of the other risks. For this reason, when resources are very limited, concentrating them all on Risk 2 is more efficient than concentrating on Risks 1 or 3 (or spreading them across the three risks equally). However, when more resources are available, they are wasted if spent just on that risk. By trial and error, we found an approximate optimum (not shown): devoting half the resources to Risk 2 and equal parts of the remainder to Risks 1 and 3. It is 5–10% more efficient than simultaneous learning for 5–40 resource units, then indistinguishable.

With simultaneous search, the risks retain their

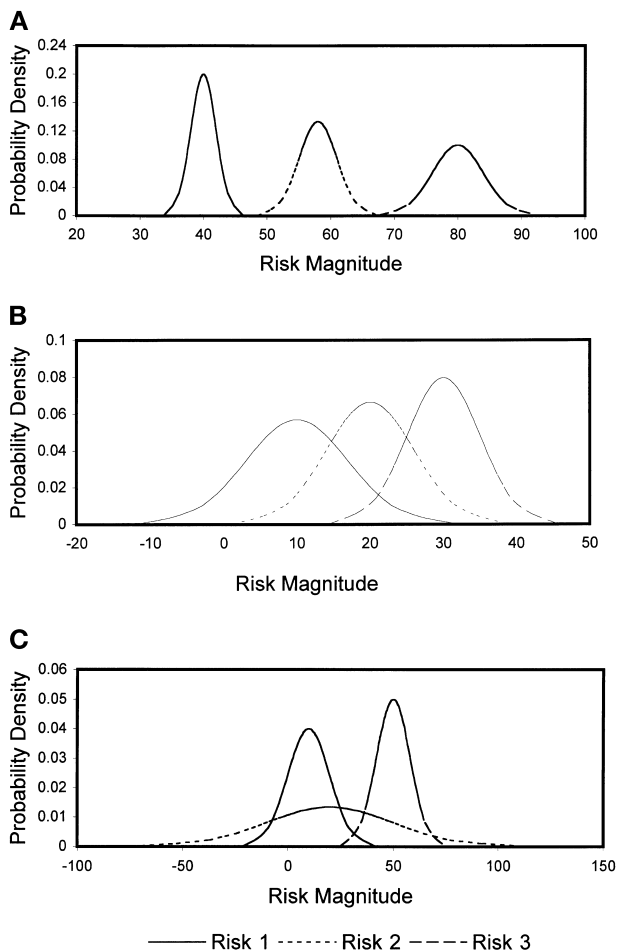


Fig. 3. Subjective probability distributions with (A) no overlap; (B) stochastic dominance; and (C) no stochastic dominance.

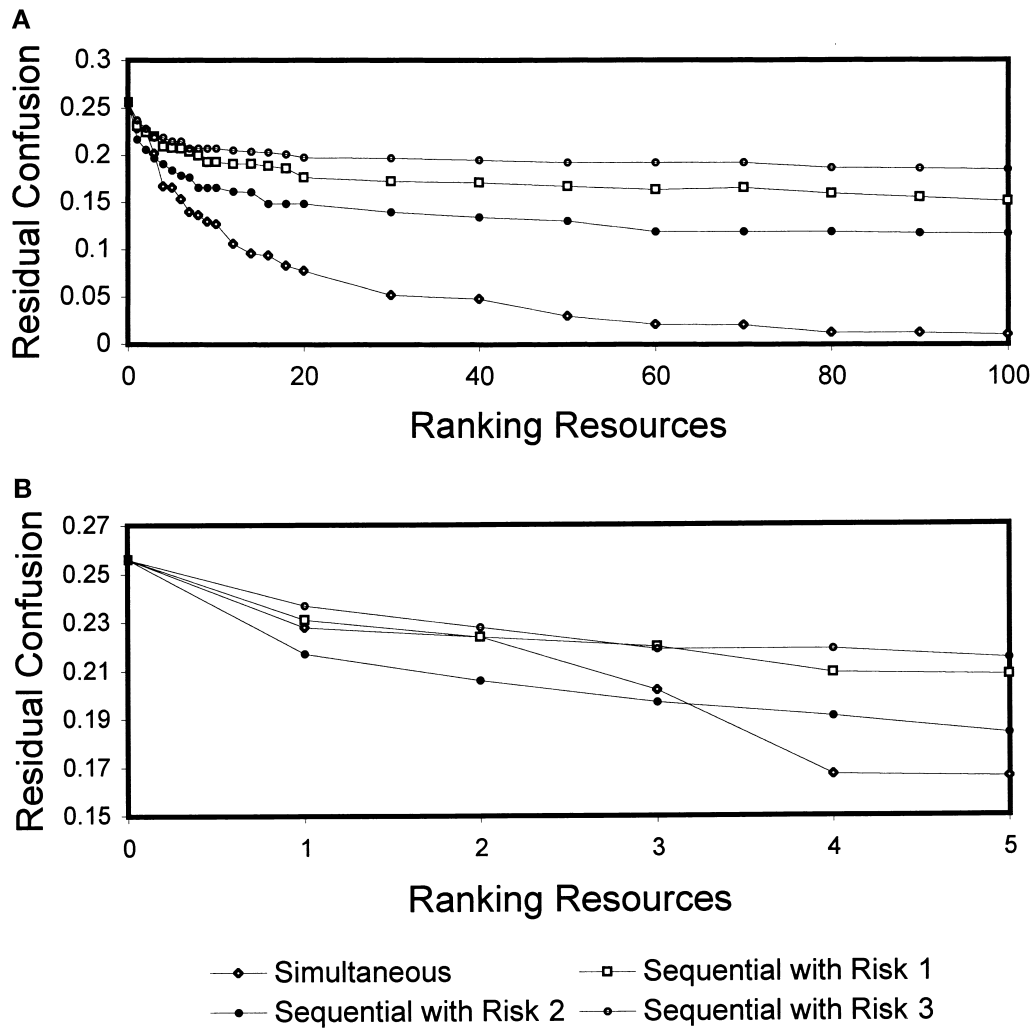


Fig. 4. Effects of the choice of attention allocation rule on change in residual confusion for three distributions with stochastic dominance (Fig. 3B). Simultaneous learning devotes resources equally to all three risks. The other three represent versions of sequential search devoting all resources to one risk for this round: (A) 100 resource units; (B) 5 resource units.

initial ranking whatever fractile is used. This follows from the absence of bias: Learning about all three risks reduces their uncertainty, while maintaining their relative ordering. Sequential learning can, however, change priorities. The risk that gets all the attention will shrink its distribution. At some point, that shrinkage will be so great that stochastic dominance no longer exists. In this example, with 30 resource units, focusing on Risk 2 changes the order from (1, 2, 3) to (2, 1, 3) for rankers using the 99th fractile—as a result of eliminating the possibility of very high values for Risk 2. By the same criterion, focusing on Risk 3 can change the order to (1, 3, 2), by reducing its assessed chance of having very high values.

Figure 5 shows the effects of assuming different URFs on RC, using the three URFs of Fig. 1. Figure 5A depicts the initial stages of simultaneous learning. The convex curve, $[\exp(0.2x) - 1]$, initially reduces RC rather little. The linear function ($y = 0.5x$) does much better⁶ and the concave URF $[\ln(x + 1)]$ better still. RC with the convex URF eventually reaches that obtained with the concave function. However, in this case, that occurs with a large expenditure of resources. As a result, where the state of nature provides a concave learning function, simultaneous learning is initially quite efficient.

⁶ The reduction is not strictly linear, or even monotonic, because of randomness in the simulation.

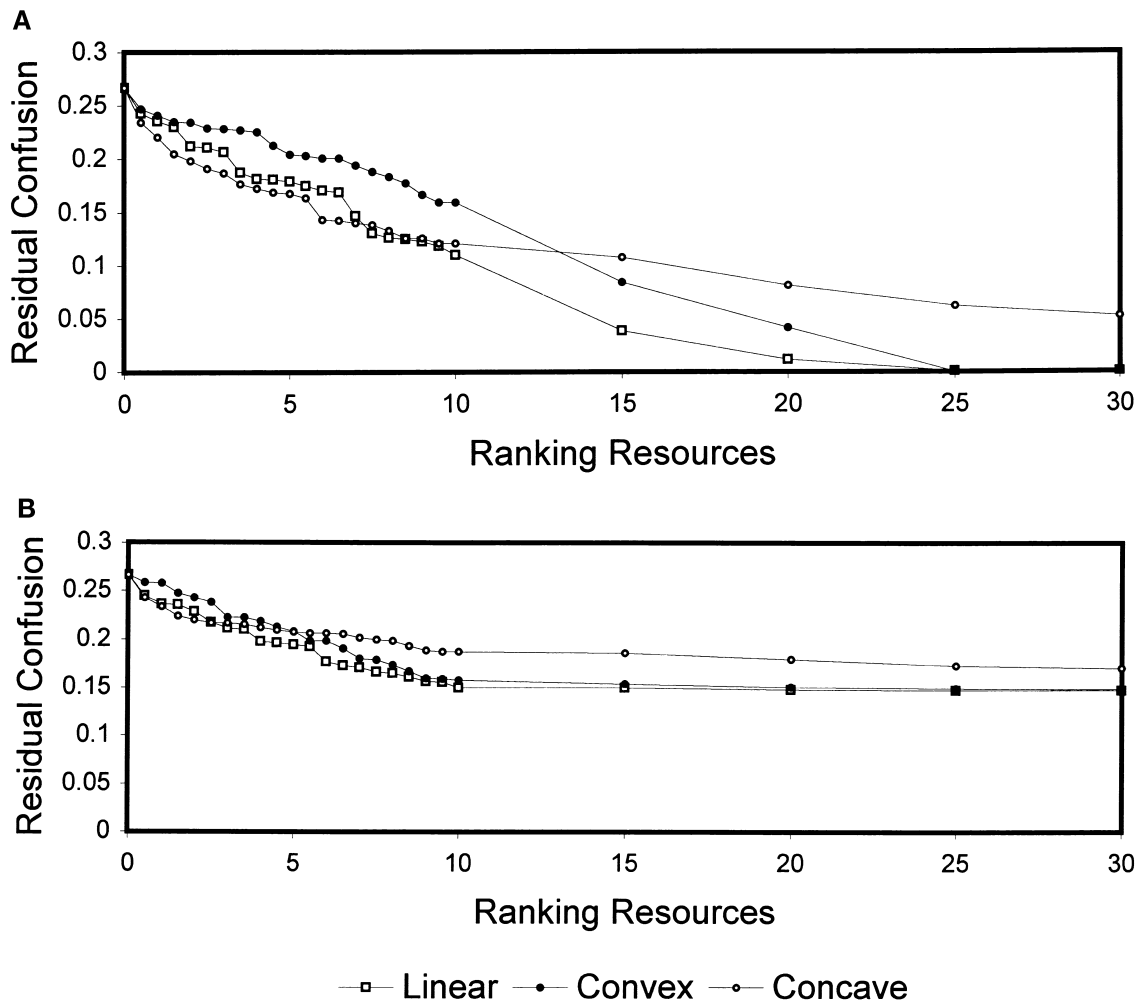


Fig. 5. Effect of the choice of uncertainty reduction function on residual confusion, with stochastically dominant initial SPDs (Fig. 3B): (A) simultaneous learning; (B) sequential ranking, focused exclusively on Risk 1.

Figure 5B provides another reflection on this situation. It shows the results of sequential learning, focused on Risk 1. The relative efficacy with the different URFs is in the same order as with simultaneous learning. The overall decline with each URF is less, however, because all attention is invested in Risk 1, initially suspected of having the least risk. (The picture would be similar with Risk 3, initially suspected of having the largest risk.) The asymptotic RC here is much higher than with simultaneous learning (Fig. 5A).

Thus, with simultaneous learning, the different URFs will reduce uncertainty to different extents, but won't affect the relative priority (as seen with the concave URF before). With sequential learning, however, rankings can change with any URF, provided

enough resources are invested in reducing the uncertainty about a given risk.⁷

3.3. Non-stochastically Dominant Distributions

Figure 3C shows such a case, with $N(10, 10)$, $N(20, 30)$, and $N(50, 8)$. It reflects one large risk (Risk 3) that is understood very well (as might happen if it had been studied heavily), along with two somewhat smaller risks (Risk 1 and Risk 2) that are understood less well. The same concave URF is used as before. As seen in Fig. 6, the AAR makes a critical difference

⁷ Think of the movement of an extreme fractile, if one distribution is tightened drastically, while the other two are left as uncertain as before.

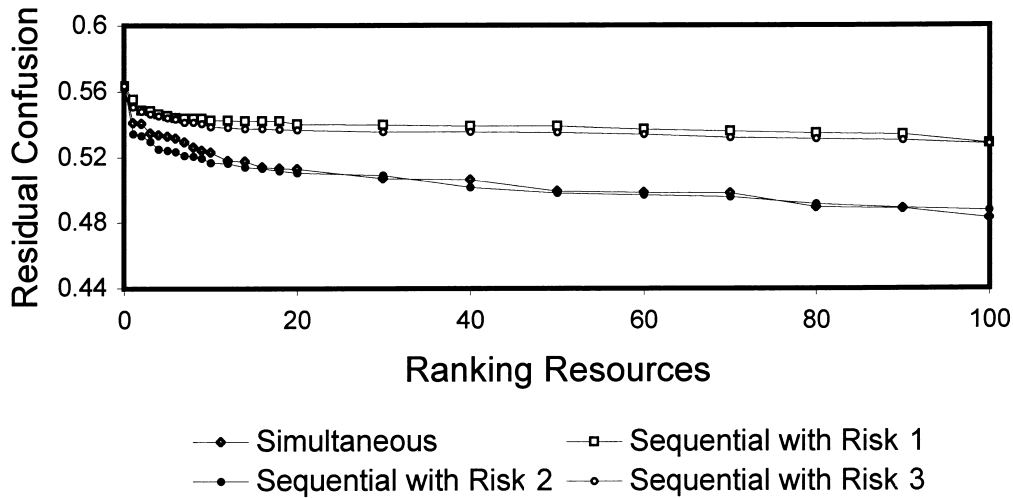


Fig. 6. Effects of the choice of attention allocation rules on residual confusion with initial subjective probability distributions lacking stochastic dominance (Fig. 3C).

in reducing RC. In this simulation, targeting Risk 1 or Risk 3 makes little difference, because each contributes little to the confusion in the set as a whole. Targeting Risk 2 is a much more effective strategy, even for small resource investments. As resources increase (to about 15 units), simultaneous ranking becomes about as good, by distinguishing Risks 1 and 3 from Risk 2. Comparing Fig. 6 with Fig. 4A, it can be seen that RC starts much higher here and declines less, in both relative and absolute terms. This is a property of the situation. Even investing significant resources in understanding Risks 1 and 3 will only distinguish them from one another, while still leaving them overlapping Risk 2 entirely. Thus, when performance is measured by RC, learning about Risk 2 is the clear “best buy.”

Table I shows rankings both before and after learning. In this case, the initial rankings depend on

the fractiles chosen, reflecting the overlap in the distributions. Given the great uncertainty about it, Risk 2 is worst for high fractiles. The one shaded cell shows the only case where priorities change with learning, compared with initial ones. Thus, in this example, despite the great uncertainty about the risks, the “action” in risk ranking is in the choice of fractile. Once that choice is made, the rankings are relatively set, even though learning can still reduce RC. The one reversal shows a case where learning a lot about a highly uncertain risk can reduce its ranking—in a case where its initial SPD is unbiased and ranking is by a high fractile.

If the state of nature pointed to a different URF, then it could be substituted for the one that we used in this simulation. If the URF were uncertain, then alternatives could be explored, as with the previous example (Section 3.2).

Table I. Effect of AAR on Priorities with Nondominating Initial SPDs

Fractiles	Initial priority	Priority after learning			
		Simultaneous	Sequential with Risk 1	Sequential with Risk 2	Sequential with Risk 3
1%	2 < 1 < 3	2 < 1 < 3	2 < 1 < 3	2 < 1 < 3	2 < 1 < 3
5%	2 < 1 < 3	2 < 1 < 3	2 < 1 < 3	2 < 1 < 3	2 < 1 < 3
50%	1 < 2 < 3	1 < 2 < 3	1 < 2 < 3	1 < 2 < 3	1 < 2 < 3
90%	1 < 2 < 3	1 < 2 < 3	1 < 2 < 3	1 < 3 < 2	1 < 2 < 3
99%	1 < 3 < 2	1 < 3 < 2	1 < 3 < 2	1 < 3 < 2	1 < 3 < 2

Note: 100 units of ranking resources.

4. MODEL APPLICATION: BIASED INITIAL SPDS

4.1. Systematic Bias

Sometimes, initial beliefs are suspected of being biased. For example, it may seem as though research-funding priorities have unduly focused on studying some physical processes, or that publication processes have unduly favored results creating a particular picture of risk. Those suspicions are, however, too diffuse to be captured in the SPDS.⁸ This section considers one extreme type of suspicion: *systematic biases*, which have the same direction and magnitude for all risks. Such a condition might arise, say, when people make the same mistake all the time (e.g., looking for trouble leads them to exaggerate all risks; their analyses share an erroneous parameter estimate). The next section briefly considers the other extreme, *random biases*.

As an example of systematic bias, we took the initial SPDS of Fig. 3C, but altered the true SPDS (by biasing the mean of the distribution upward by a factor of two), while leaving the standard deviation the same. In model terms, learning about a risk means sampling observations from the true SPDS, then combining those observations with the biased prior in a Bayesian manner. That process will tend to reveal (and reduce) the bias. Depending on the circumstances, the overall result might be to increase or decrease uncertainty. Figure 7 shows the effect of learning on estimates of the mean for Risk 1, as a function of investing resources in Risk 1 alone or in all three risks simultaneously. The same (concave) URF is used. The mean approaches the unbiased value (the lower dashed line) with either AAR, but does so more quickly when learning focuses on Risk 1 than when it is distributed over all three risks. A similar pattern emerges with the other two risks, with focused learning being moderately more efficient than simultaneous. Thus, for this case, in which all risks are biased in the same way, simultaneously learning seems advisable—in order to learn something about each risk—rather than to focus on any one. Of course, less progress would be made if the new information were also obtained from biased distributions (as might happen if rankers relied on the same faulty sources).

Figure 8 shows RC, investing the same resources

in sequential or simultaneous learning. Simultaneous learning shifts all distributions in the same direction. Given the common bias, that alone would leave RC relatively constant. However, a given amount of resources will update a distribution with small initial variance more than a distribution with large initial variance. As a result, Risk 2 moves toward its mean more slowly than do Risk 1 and Risk 3. That increases its overlap with Risk 3, as well as the overlap of Risk 1 and Risk 3—leading RC to *increase* with such simultaneous learning. Risk 3's initial mean is 50 and its true mean is 25, closer to Risk 2's initial mean. As a result, learning just about Risk 3 will move its SPD toward that for Risk 2, increasing their overlap (and RC). As mentioned, Risk 2's initial SPD has a mean of 20, while its true SPD has a mean of 10, the same as Risk 1's initial mean. Thus, learning about just Risk 2 will shift the mean of its SPD toward that of Risk 1, greatly increasing the contribution to RC of that overlap, while shifting its SPD away from Risk 3, somewhat decreasing the contribution to RC of that overlap. At the same time, reducing Risk 2's uncertainty will decrease its overlap with the other two SPDs, somewhat reducing RC. The net effect is the small overall reduction in RC shown in Fig. 8 (Sequential with Risk 2). Reducing the bias in Risk 1 still leaves it overlapping Risk 2; however, both the shift and the reduced uncertainty make it more distinct from Risk 3, slightly reducing RC.

Thus, learning about these risks does little to reduce rankers' confusion and may increase it—because they were less confused initially than they had a right to be. Where bias is suspected, RC does not provide a criterion for evaluating the efficacy of different strategies.

If these risks are ranked by the means of their SPDS, their initial order is (1, 2, 3). It remains that way whatever AAR is used. Simultaneous learning shifts all three means in tandem. None of the sequential learning rules pulls the focal mean past one of the others; as a result, their order stays the same. Initial ranking by the 95th fractile is (1, 3, 2) because of Risk 2's long right tail. The order will change to (1, 2, 3) if enough is learned about Risk 2 to pull in its tail and shift the distribution to the left.

4.2. Random Bias

Systematic bias describes one extreme. Another extreme involves biases that are completely random, in both direction and magnitude (over some range of

⁸ They might, however, be expressed in second-order distributions, expressing epistemic uncertainty⁽²⁴⁾—a possibility that we will not consider here.

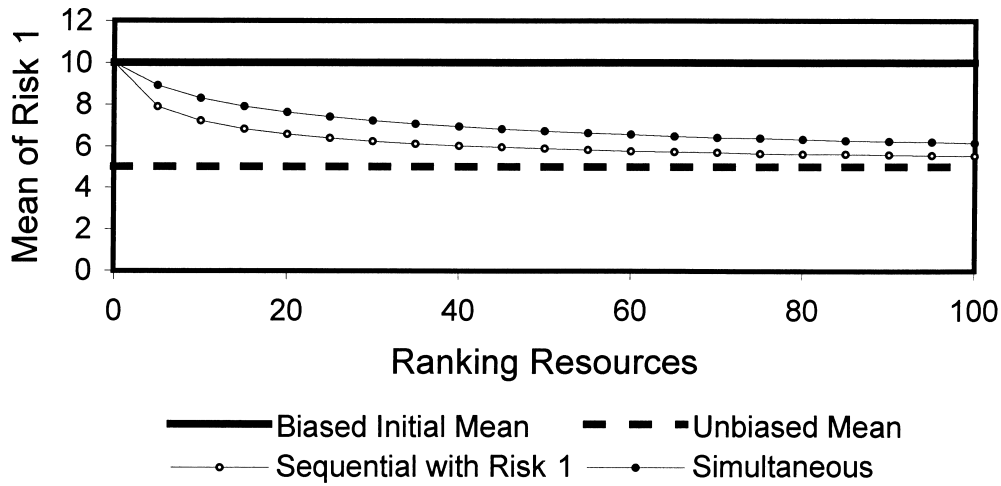


Fig. 7. Change in degree of bias, with simultaneous ranking and sequential ranking focused on learning about Risk 1.

possibilities). In this case, all the complications discussed above could happen, with the results of the simulations (and the ranking processes that they represent) being even less predictable. As before (Section 4.1), learning about the risks may appropriately increase uncertainty.

If bias is suspected, then it may pay to do some exploratory learning about the risks. If the width of the SPDs increases, then bias should be more strongly suspected. If the SPD shifts are in a common direction, then systematic bias is more likely. The model built for that situation could then incorporate

these assumptions when trying to design a ranking process or to predict its operation.

5. CONCLUSIONS AND POLICY IMPLICATIONS

Individuals, organizations, and societies often need priorities for addressing the myriad risks to their health, safety, and environment. Deciding on those priorities should help them to focus their search for ways to reduce risk. When risks are uncertain, so may be these priorities. Learning about risks may al-

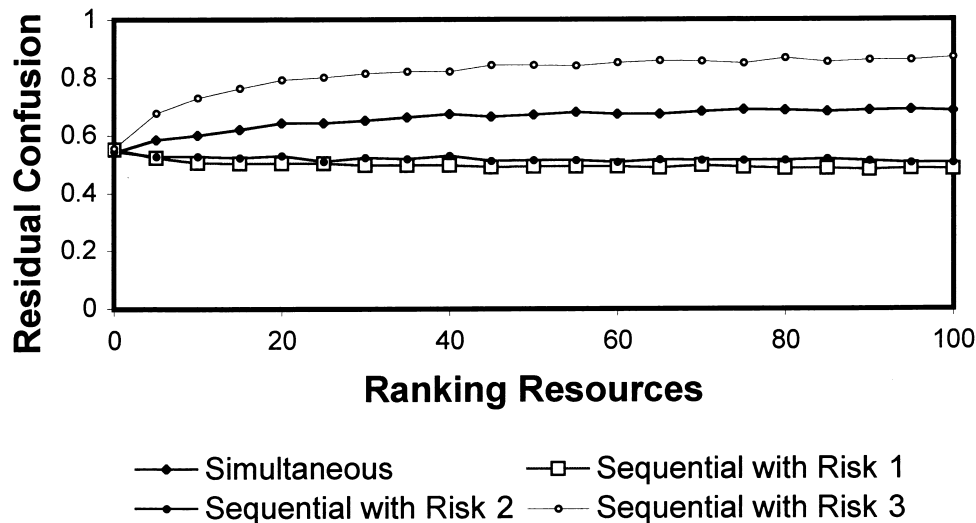


Fig. 8. Residual confusion with biased initial and an unbiased learning process, for simultaneous learning and sequential learning, focused on each risk.

low reducing residual uncertainty about both their individual magnitudes and their respective rankings.

We offer a model for such processes. It assumes an iterative process, in which limited resources are devoted to uncertainty reduction during successive learning periods, and rankings are revised in the light of what is learned. The model characterizes each round in terms of two states of nature, which people cannot control, and four design parameters, which they can. The states of nature are (1) the uncertainty, expressed in SPDs (which might be biased) and (2) the pace with which uncertainty shrinks, when resources are invested in learning about a risk (expressed in URFs). The design parameters are (1) the AAR, describing how learning resources are allocated across risks; (2) the *ranking criterion*, or fractile used to characterize the SPDs, for ranking purposes; (3) the Bayesian (or other) *updating rule* for combining new information about a risk with existing information; and (4) the RC measure for evaluating performance (which, however, is not directly interpretable when initial SPDs are suspected of bias).

Demonstrations of the model focus on two of the many possible ways in which attention can be allocated (during a round): simultaneous learning, where all risks receive equal attention, and sequential learning, where all learning resources are allocated to a single risk (for that round). The demonstrations consider the learning associated with three classes of the uncertainty reduction function: linear, concave, and convex.

The impact of these factors was examined in the context of five different initial conditions. Three of these assumed no underlying bias in the initial SPDs: (1) virtually no overlap among the distributions, (2) stochastic dominance, and (3) no stochastic dominance. The third of these situations was reconsidered, assuming bias in the initial SPDs that was either systematic or random.

In all cases, a Bayesian updating function was used to integrate what is learned with what was initially believed. For computational simplicity, we also restricted ourselves to one set of conjoint distributions (normal). A straightforward way to relax the optimality assumption is to use Bayesian updating, but over- or underweight the new information, thereby representing rankers who are too quick or slow to change.

Also for simplicity's sake, demonstrations have used only three risks. However, we hope that they illustrate the key features of the model and how it can illuminate the structure of risk-ranking tasks. In some

cases, just characterizing a ranking process in these terms may clarify how efficient, and appropriate, it is. For example, it may reveal whether the critical uncertainties are about facts or values. It might forestall large-scale data collection when there is little agreement about the risk metric, the ranking criterion, or the performance measure.

In other cases, though, it may be necessary actually to run the numbers. For example, in the example with stochastic dominance (Section 3.2), we found that the choice of AAR influenced both the final rankings and the efficiency of ranking strategies. With fewer resources, devoting them all to one risk could be more efficient than dividing them equally across the risks. With greater resources, however, simultaneous learning would be more efficient than sequential. In the example without stochastic dominance (Section 3.3) the choice of AAR greatly affected how far RC was reduced. The same was true when there was bias in the initial SPDs. Resources could be wasted if these issues were not sorted out before a ranking exercise was designed—or its results were interpreted. Even in our simple examples, these relationships would be hard to anticipate without explicit quantitative modeling.

We see value in characterizing stylized situations like those considered here. Doing so provides a way to think about the nature of risk ranking, including what one hopes to—and realistically can—gain from it. We also sought, however, to characterize the model clearly enough to allow its operationalization for specific settings. We now sketch how each of the model's parameters might be approached in an application, either designing a process or predicting its efficacy. These steps could be followed for the specific risks under consideration or for a small set of risks with properties like those in the full set (e.g., heterogeneous variance, no systematic bias, many risks clustered near the bottom of the set). Such archetypes might provide a useful feeling for the results of a full-fledged application.

States of Nature

SPDs: Follow accepted procedures for probability elicitation.^(18,25) Elicit suspicions of bias, perhaps in the form of second-order distributions.

URF: Consider the state of the science regarding each risk. How ripe is it for summarization?⁽²⁶⁾ What kinds of diagnostic tests are available (e.g., just rodent bioassays or also bacterial tests)? How quickly can novices be brought up to speed?

Design Choices

Ranking criterion: Elicit rankers' preferences for the fractile best suited to characterizing uncertain risks.

AAR: Start with a suite of stylized rules, including entirely systematic and sequential learning, along with promising hybrids. Translating these rules into URF terms may be challenging, even if the general shape of the URF is fairly clear. Once the other terms of the model have been set, backing out the best AARs is a logical way to optimize its application.

Performance measure: While simple, our RC metric treats distributional overlap the same, at all points on the risk continuum. It means that all confusion is equally troubling. Rankers might, however, choose to pay more attention to confusion among larger risks. Doing so will require a risk metric with more than interval-scale properties.

Updating: Although the designers of a ranking process might prescribe a Bayesian approach, other expectations might be more realistic. In that case, updating is a state of nature, knowledge of which will help in predicting how a ranking process will actually behave.

Although the model is fairly complex already, we can see several elaborations that could improve its fidelity, facilitate its application, and increase its design flexibility. One is to include variability as a source of the uncertainty in an SPD. Because variability is more directly estimated than uncertainties derived from other sources, considering it should refine the analysis. A second elaboration is to distinguish uncertainty about facts and about values. Part of a ranking process is determining what matters most, when integrating multiple attributes into a common measure of risk. Thinking hard about the ranking of some risks may teach useful lessons about the meaning of "risk" in general. Therefore, focused learning about a subset of risks might reduce confusion about the set as a whole. As a result, one hybrid strategy is to look closely at a few risks, in order to resolve the definition of "risk," then to think simultaneously about all risks in those terms.

In such ways, a model like this one can take advantage of experiences with risk ranking—and identify aspects of those processes that need to be better understood. Thus, one might look for structural prop-

erties of actual risk rankings that have been more or less satisfactory for participants. Are value issues brought to the fore, or left unarticulated beneath discussions of data? Has the technical staff provided cogent summaries of the issues, so that the rankers can learn a lot quickly—even if it would take them a very long time to acquire great mastery (making the URF more concave)? Is increased uncertainty an acceptable conclusion, when biased priors (and premature closure) are possibilities? Have the rankers been required to look at all the risks, when sequential evaluation of a few would have been more effective?

ACKNOWLEDGMENTS

The authors thank Mike DeKay, Paul Fischbeck, Keith Florig, Granger Morgan, Andy Parker, Mitchell Small, and two anonymous reviewers for their valuable comments on this work. This research was supported by National Science Foundation Grant SBR-9512023 and Environmental Protection Agency Grant CR824706-01-02; however, the views expressed here are those of the authors alone.

REFERENCES

1. Davies, J. C. (Ed.). (1996). *Comparing environmental risks*. Washington, DC: Resources for the Future.
2. Morgan, M. G., Fischhoff, B., Lave, L., & Fischbeck, P. (1996). A proposal for ranking risk within federal agencies. In C. Davies (Ed.), *Comparing environmental risks* (pp. 111–147). Washington, DC: Resources for the Future.
3. U.S. Environmental Protection Agency. (1987). *Unfinished business: A comparative assessment*. Washington, DC: Author.
4. U.S. Environmental Protection Agency. (1991). *Reducing risks: Setting priorities and strategies for environmental protection*. Washington, DC: Author.
5. Jones, K. (1997). *Ten years of comparative risk analysis*. <http://www.gmied.org> (June 9, 2000).
6. Minard, R. A., Jr. (1996). CRA and the states: History, politics, and results. In J. Clarence Davis (Ed.), *Comparing environmental risks—tools for setting government priorities*. Washington, DC: Resources for the Future.
7. National Research Council. (1996). *Understanding risk*. Washington, DC: Author.
8. Lindblom, C. (1959). The science of muddling through. *Public Administration Review*, 19, 79–88.
9. Bendor, J. (1995). A model of muddling through. *American Political Science Review*, 89, 819–840.
10. Crouch, E.A.C., & Wilson, R. (1981). *Risk/benefit analysis*. Cambridge, MA: Ballinger.
11. Fischhoff, B., Watson, S., & Hope, C. (1984). Defining risk. *Policy Sciences*, 17, 123–139.
12. Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sciences*, 8, 127–152.
13. Jenni, K. (1997). *Attributes for risk evaluation*. Unpublished

- doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA.
14. Slovic, P., Fischhoff, B., & Lichtenstein, S. (1979). Rating the risks. *Environment*, 21(4), 14–20, 36–39.
 15. Slovic, P., Fischhoff, B., & Lichtenstein, S. (1980). Facts and fears: Understanding perceived risk. In R. Schwing & W. A. Albers, Jr. (Eds.), *Societal risk assessment: How safe is safe enough?* (pp. 181–214). New York: Plenum.
 16. Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgments under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
 17. Kammen, D., & Hassenzahl, D. (1999). *Should we risk it?* Princeton, NJ: Princeton University Press.
 18. Morgan, M. G., & Henrion, M. (1991). *Uncertainty*. New York: Cambridge University Press.
 19. Kingdon, J. W. (1995). *Agendas, alternative and public policies*. Boston: Little, Brown and Company.
 20. Slovic, P., Fischhoff, B., & Lichtenstein, S. (1985). Characterizing perceived risk. In R. W. Kates, C. Hohenemser, & J. Kasperson (Eds.), *Perilous progress: Managing the hazards of technology* (pp. 91–125). Boulder, CO: Westview.
 21. Edwards, W., Lindman, H., & Savage, L. S. (1963). Bayesian statistical inference for psychological research. *Psychological Review*, 70, 193–242.
 22. Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review*, 90, 239–260.
 23. Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior & Human Performance*, 6, 649–744.
 24. Gärdenfors, P., & Sahlin, N. E. (1982). Unreliable probabilities, risk taking and decision making. *Synthese*, 53, 361–386.
 25. Clemen, R. (1991). *Making hard decisions: An introduction to decision analysis* (2nd ed.). Belmont, CA: Duxbury.
 26. Funtowicz, S., & Ravetz, J. (1990). *Uncertainty and quality in science for policy*. London: Kluwer.